2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

TC-LEC: Targetless Calibration for LiDAR-Event Camera Systems

Ying Yang[1], Jianing Li[2*], Jiangming Shi[1], and Yanyun Qu[1*]

***Abstract*— LiDAR-event camera integration has shown considerable promise and is gaining traction across various perception applications. Event cameras offer high temporal resolution and wide dynamic range but suffer from noise sensitivity and lack depth information. LiDAR complements these capabilities by providing absolute scale and robustness, yet accurate calibration between the two sensors remains a significant challenge. This paper presents targetless calibration framework for LiDAR–event camera systems that removes dependence on dedicated calibration targets and strong initial assumptions. The method estimates the event camera angular velocity by analyzing the timestamp and spatial changes of per-pixel, enabling precise detection of natural edges. Calibration proceeds in two stages: (i) motion-based initialization, where Canonical Correlation Analysis (CCA) on rotational estimates from the event camera and LiDAR jointly recovers the temporal offset and rotation; (ii) nonlinear refinement of the extrinsics via cross-modal alignment of natural edge features. Experiments on physical platforms and public datasets demonstrate robust performance and high calibration accuracy across diverse scenarios. This work provides a solid foundation for further development and application of LiDAR-event camera fusion.**

I. INTRODUCTION

Event cameras, owing to their unique sensing principle, have recently attracted increasing attention in a wide range of perception tasks [50], consistently demonstrating remarkable performance in simultaneous localization and mapping (SLAM) [1], [2], [27], 3D reconstruction [3], [4], [5], robust object detection [6], [7], [8], and accurate optical flow estimation [9], [10]. Meanwhile, LiDAR [15], [16], [17] provides reliable absolute scale and strong environmental robustness, thereby effectively complementing and mitigating the inherent limitations of event cameras [12], [13], [14].

The complementary characteristics of event cameras and LiDAR enable significant advancements in perception. Their fusion facilitates depth interpolation, point cloud densification, and temporal enhancement [18].Building on event streams and LiDAR sparse depth hints, a "hallucinated events" strategy is introduced to effectively mitigate the sparsity problem[21]. In challenging lighting or subterranean environments, event–LiDAR integration achieves robust human and object tracking [20]. For autonomous driving[19], such fusion supports low-latency dynamic object perception and real-time detection. Combining event cameras with LiDAR not only overcomes the limitations of individual sensors but also provides profound significance for high-precision, robust environmental perception in intelligent systems.

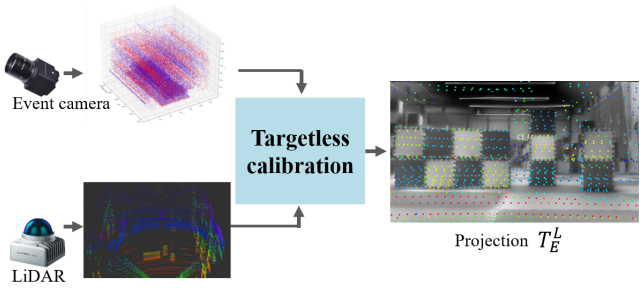


Fig. 1: Overview of the extrinsic calibration method that parameters (T_L^E) of the event camera with LiDAR.

However, conventional calibration methods for LiDAR–event camera systems [23], [28] typically rely on physical calibration objects such as chessboards or specialized patterns, or depend on controlled variations in ambient illumination for sensor data alignment. Such approaches are not only labor-intensive and inflexible, but also struggle to adapt to complex and dynamic real-world environments.

Although targetless LiDAR-event camera calibration methods [29], [32], [43] have been proposed to enhance deployment flexibility by eliminating the need for dedicated targets, they often depend on specific sensor configurations or constrained environments and are sensitive to initialization. Some methods [44], [45] recover the pose or velocity of each sensor and align the motion information using external parameters but are limited to estimating only the rotational parameters. Accurate calibration between the two sensors remains a significant challenge.

To address the aforementioned issues, this paper presents a targetless extrinsic calibration method for LiDAR–event camera systems, which fully leverages event cameras the high temporal resolution and asynchronous nature. Specifically, due to the asynchronous sampling nature of the two sensors, there exists a temporal offset between them. we explicitly estimate this offset and align both data streams to a common time base. Calibration is then initialized by combining angular-velocity measurements and cross-sensor rotational correlations to obtain a coarse pose. Building on this initialization, we perform a nonlinear optimization that aligns natural edge features extracted from both modalities, thereby refining the extrinsic parameters. This design reduces reliance on dedicated targets and precise initialization, and yields accurate, robust calibration across diverse operating

*This work was supported by Science and Technology on Sonar Laboratory under grant 2024-JCJQ-LB-32/07 and State Grid Corporation Headquarters Technology Project (52120025004V-463-FGS).

[1] Key Laboratory of Multimedia Trusted Perception and Efficient Computing, Ministry of Education, Institute of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen, Fujian, China, yangyingrs@stu.xmu.edu.cn.

[2]School of Computer Science, Peking University, Beijing, China.

*Corresponding authors.

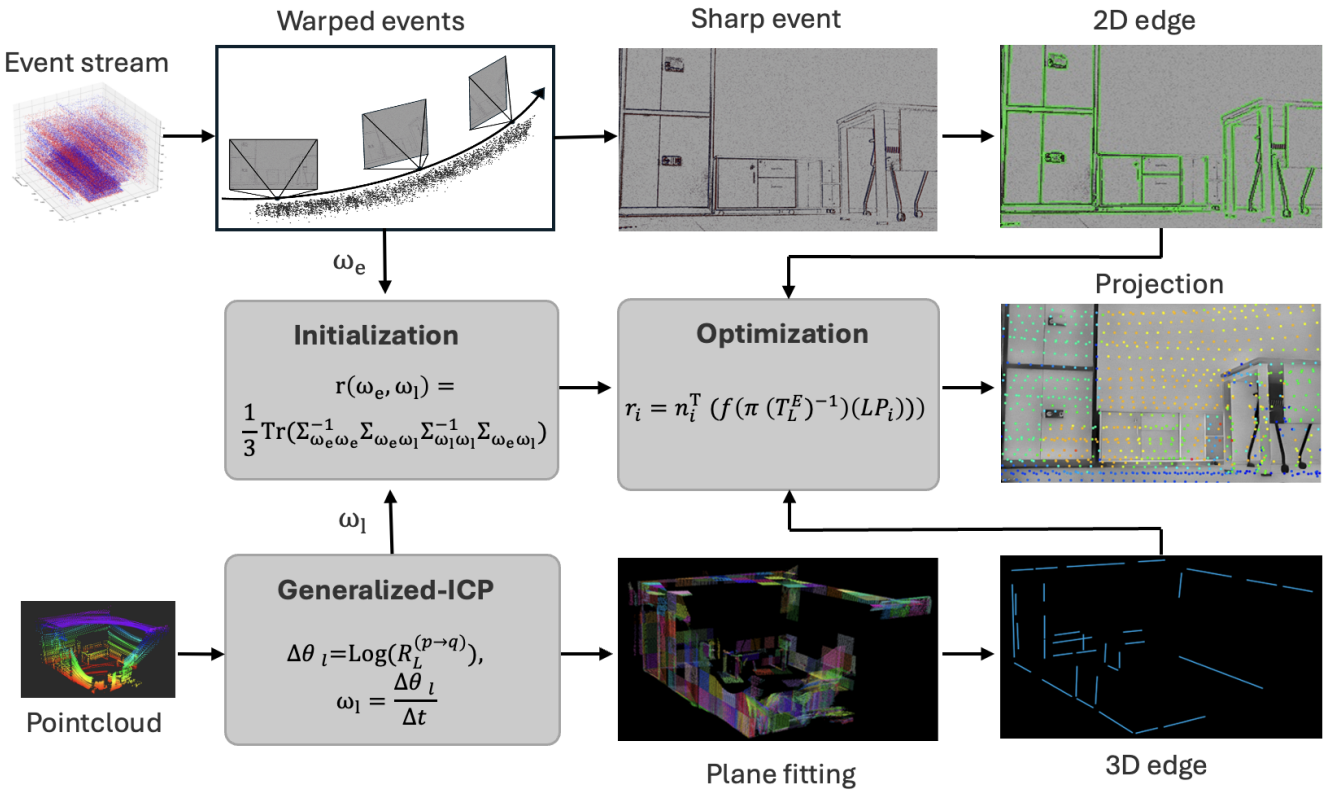


Fig. 2: Calibration pipeline. Leveraging the event sensor microsecond-level temporal resolution and asynchronous operation, the camera angular velocity ω_e is estimated via an event-warping consistency objective while natural edge features are simultaneously extracted. The LiDAR angular velocity ω_l is obtained using Generalized-ICP (G-ICP). Applying CCA to the two rotational estimates, the temporal offset and an initial rotation estimate are jointly recovered, yielding a coarse alignment. With this initialization, a cross-modal nonlinear optimization aligns natural edge features and refines the extrinsics T_E^L. Finally, the LiDAR point cloud is projected onto the event-domain visualization to assess calibration accuracy.

conditions. Our contributions can be summarized as follows:

1) A targetless calibration framework for LiDAR–event camera systems that eliminates dedicated calibration targets and tight hardware synchronization, exploiting the event camera high temporal resolution and asynchronous operation.

2) We propose a two-stage calibration procedure that first jointly estimates the temporal offset and rotation from cross-sensor rotational cues under a unified time base, then refines the extrinsics via cross-modal nonlinear alignment of natural edge features.

3) Extensive evaluations on public datasets and physical platforms, including UAV and ground-vehicle deployments, demonstrate robust performance and high calibration accuracy across diverse scenarios.

II. RELATED WORK

Extrinsic calibration techniques for frame-based cameras and LiDAR systems are relatively mature and widely adopted [22], [24], [25], [26], [34], [35], [36], [37], [33]. In contrast, extrinsic calibration for LiDAR–event camera systems remains comparatively underexplored. Existing approaches can be broadly categorized as target-based and targetless, with the main distinction being in how cross-modal correspondences are defined and extracted. Target-based methods rely on calibration artifacts with known geometry and dimensions to provide accurate correspondences for recovering extrinsic parameters. However, targetless methods directly estimate the extrinsics from naturally occurring structures, such as edges, corners, and high-reflectance returns, which are jointly observable by both sensors in natural scenes.

Target-based Calibration. Due to the operational characteristics of event cameras, they do not generate events in a steady state with constant illumination. Some intrinsic calibration techniques [38] use active markers, such as blinking LEDs or dynamic displays, to enable event generation through controlled illumination changes while maintaining fixed sensor positions. Similar principles of active stimulation have been extended to LiDAR-event camera extrinsic calibration. For example, several studies have used changes in LED light brightness [39] or flickering screens [23], [40] for calibration. Jiao et al [28] employed automatic feature extraction and chessboard tracking, using E2VID [41] for event camera reconstruction to achieve calibration. However, these methods require pre-prepared calibration markers and manual operations, which may not be feasible or practical in changing environments.

Targetless Calibration. To reduce reliance on specific

devices, some studies have started to leverage the event camera's ability to capture dynamic features for calibration [30], [31]. Xing et al. [19] proposed a method that entirely eliminates the need for external calibration targets. This approach utilizes the event camera's capability to extract geometric edges during motion and associates these edges with those extracted from LiDAR point clouds. In contrast, Ta et al. [42] impose stringent hardware requirements, requiring spectral overlap between LiDAR emissions and the camera's infrared response range, as well as sufficient emission intensity to reliably trigger events. These conditions are not always met in typical sensor configurations. MULi-Ev [43] is a deep learning framework for online calibration of event cameras and LiDAR, but it is optimized for specific models of event cameras and LiDAR. Additionally, a motion calibration framework [45] for event-centric multi-sensor systems estimates angular velocity from normal flow observations and refines both temporal and rotational parameters.

Building upon existing research, our method takes a targetless route that directly leverages the differential and asynchronous characteristics of raw event streams to recover event camera motion without frame reconstruction, improving flexibility in real-world deployment. We then align natural cross-modal edge features through nonlinear optimization to refine the extrinsic parameters , thereby reducing reliance on dedicated targets, tight synchronization, and precise initialization.

III. METHODOLOGY

This section introduces our extrinsic calibration framework for an event camera and a LiDAR. Event-camera motion and edge features are extracted (Section. III-A), while LiDAR motion and edge features are recovered in parallel (Section. III-B). An initial extrinsic estimate is obtained using CCA (Section. III-C) and subsequently refined through nonlinear edge-based optimization (Section. III-D). The complete pipeline is illustrated in Fig. 2.

A. Event Camera-Based Angular Velocity Estimation

In order to fully leverage the high temporal resolution and asynchronous trigger characteristics of event cameras, this method estimates the angular velocity by maximizing the contrast warped-event accumulation image [46], [47], [11] Given events in a time window:

$$\varepsilon = \{(x_i, y_i, t_i)\}_{i=1}^{N}. \tag{1}$$

and a short-time pure-rotation model with reference time t_e, each event at pixel $\mathbf{u}_i = [x_i, y_i]^\top$ and time t_i is warped to the reference frame by the motion model $W(\mathbf{u}_i, t_i; t_e, \omega_e)$, which is derived from the camera intrinsics and the rotation induced by the angular velocity ω_e. The warped-event accumulation image is defined as:

$$E_{\omega_e}(\mathbf{u}) = \sum_{i=1}^{N} \kappa(\mathbf{u} - W(\mathbf{u}_i, t_i; t_e, \omega_e)), \tag{2}$$

where $\kappa(\cdot)$ is a bilinear or Gaussian kernel for differentiability and $\mathbf{u} \in \Omega$ spans the image domain. The contrast is defined as the variance over pixels:

$$\begin{aligned} \mathrm{Var}_\Omega(E_{\omega_e}) &= \frac{1}{|\Omega|} \sum_{\mathbf{u}\in\Omega} (E_{\omega_e}(\mathbf{u}) - \bar{E}_{\omega_e})^2, \\ \bar{E}_{\omega_e} &= \frac{1}{|\Omega|} \sum_{\mathbf{u}\in\Omega} E_{\omega_e}(\mathbf{u}). \end{aligned} \tag{3}$$

The optimal angular velocity is obtained by contrast maximization,

$$\omega_e^\star = \arg\max_{\omega_e} \ \mathrm{Var}_\Omega(E_{\omega_e}). \tag{4}$$

The resulting image is then normalized and thresholded, after which natural edge features are extracted. Edge features (e.g., those obtained using the Canny detector) represent salient scene boundaries, which are then employed for calibration.

B. LiDAR Angular Velocity Estimation and edge Feature

To determine the LiDAR rotation from one scan to the next, we employ GICP algorithm [48], which combines point-to-point and point-to-plane metrics within a probabilistic formulation. First, to correct for motion distortion, we project each sampled point in the scan to the current LiDAR frame's scan-end using the previously estimated motion. Then, we downsample the point cloud using the open-source small-GICP library [49] and match the source and target point clouds to recover the relative transform. To improve robustness to outliers and low-structure regions, we estimate the pose increment by minimizing the following robust cost:

$$\min_{\xi\in\mathfrak{se}(3)} \sum_i \rho\left(\left\| \mathbf{n}_i^\top (R(\xi)\,\mathbf{p}_i + \mathbf{t}(\xi) - \mathbf{q}_i) \right\|^2 \right), \tag{5}$$

where $\rho(\cdot)$ is a robust loss, $\mathbf{p}_i$ and $\mathbf{q}_i$ are a correspondence pair, and $\mathbf{n}_i$ is the associated surface normal.

The transformation matrix $T_l^{(p\to q)}$ represents the rotation and translation from source scan p to target scan q, where the rotation is described by the rotation matrix $R_l^{(p\to q)}$ and the translation vector $\mathbf{t}_l^{(p\to q)}$, with $R_L^{(p\to q)} \in SO(3)$. By performing a logarithmic map on the rotation matrix, we obtain the rotational increment $\Delta\theta_l$, and the angular velocity ω_l is calculated from the rotational increment and the time interval t_l:

$$\Delta\theta_l = \mathrm{Log}\big(R_L^{(p\to q)}\big), \qquad \omega_l = \frac{\Delta\theta_l}{t_l}, \tag{6}$$

where $\mathrm{Log}(\cdot)$ is the matrix logarithm and t_l is the inter-scan time.

The relative positioning between the LiDAR and event camera, together with scene occlusions, can cause zero-value and multi-value issues [51] when the point cloud is directly projected onto the image plane. To address this problem, we extract edge features directly from the point cloud. The extraction of edge features depends on the structural characteristics of depth values [52], [53]. Different types of edges differ significantly in physical properties and extraction methods, thus requiring different processing approaches.

Discontinuous edge features refer to abrupt depth changes between the foreground and background, which are prone to

noise interference, impacting extraction accuracy In contrast, the extraction of continuous edge features relies on plane fitting. This method involves dividing the point cloud into small voxels, using the Random Sample Consensus (RANSAC) algorithm to fit planes, and then extracting the intersection lines of the planes as edges. The edges extracted in this manner are more stable and precise, significantly improving the accuracy of edge extraction.

C. Rough Calibration

The event camera and LiDAR measurements are subject to a temporal offset, caused by transmission and processing delays, as illustrated in Fig. 3. To compensate for this offset, we evaluate the correlation of angular velocities between the two sensors by computing their covariance and cross-covariance matrices. Specifically, the calculation of cross-covariance is a key step to measure the correlation of angular velocities between the event camera and LiDAR. We define the $\omega_e(t_i)$ and $\omega_l(t_i)$ are the angular velocities of the event camera and LiDAR at time t_i, $\bar{\omega}_e$ and $\bar{\omega}_l$ are the mean angular velocities of the event camera and LiDAR. Then, The covariance matrix for the event camera:

$$\Sigma_{\omega_e\omega_e} = \frac{1}{N-1}\sum_{i=0}^{N-1}(\omega_e(t_i)-\bar{\omega}_e)(\omega_e(t_i)-\bar{\omega}_e)^T \tag{7}$$

Similarly, the same formula is used to obtain the covariance matrix for LiDAR: $\Sigma_{\omega_l\omega_l}$, and the cross-covariance matrix between the event camera and LiDAR: $\Sigma_{\omega_e\omega_l}$.

These covariance matrices are used to evaluate the relationship between the event camera and LiDAR, especially the synchronization of their angular velocities. The key step in calculating the time offset is to maximize the trace correlation between the event camera and LiDAR. The trace correlation formula is:

$$r(\omega_e,\omega_l) = \frac{1}{3}\mathrm{Tr}\left(\Sigma^{-1}_{\omega_e\omega_e}\Sigma_{\omega_e\omega_l}\Sigma^{-1}_{\omega_l\omega_l}\Sigma_{\omega_l\omega_e}\right) \tag{8}$$

where Tr represents the matrix trace operation, $\Sigma^{-1}_{\omega_e\omega_e}$ and $\Sigma^{-1}_{\omega_l\omega_l}$ are the inverse covariance matrices, and $\Sigma_{\omega_e\omega_l}$ and $\Sigma_{\omega_l\omega_e}$ are the cross-covariance matrices.

The external rotation matrix R^L_E can be computed using the cross-covariance matrices. The objective of this step is to calculate the external rotation using SVD. The formula for calculating the external rotation matrix is:

$$R^L_E = U\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \det(UV^T)\end{pmatrix}V^T \tag{9}$$

where V and U are the left and right matrices obtained from the SVD of the covariance matrix $\Sigma^{-1}_{\omega_l\omega_l}\Sigma_{\omega_l\omega_e}$. $\det(UV^T)$ represents the determinant of the matrix UV^T. This formula represents the rotation matrix obtained, which is used to align the coordinate systems of the event camera and LiDAR.

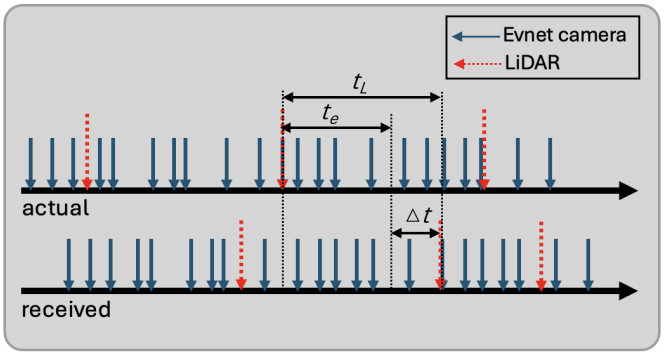


Fig. 3: Temporal offset.

D. Optimization

We consider a LiDAR point $LP_i \in \mathbb{R}^3$ expressed in the LiDAR frame L. Using the current extrinsic estimate between the event camera frame E and the LiDAR frame L, the point is mapped into the event-camera coordinates as:

$$E_{P_i} = (T^E_L)^{-1}(LP_i) = R^L_E LP_i + t^L_E. \tag{10}$$

where $T^E_L \in SE(3)$ denotes the rigid transformation, and $(T^E_L)^{-1} = (R^L_E, t^L_E)$ is its inverse. Here, $R^L_E \in SO(3)$ represents the rotation matrix and $t^L_E \in \mathbb{R}^3$ denotes the translation vector.

In this optimization process, the reprojection error is used as the optimization objective, which directly measures the geometric alignment between LiDAR and camera. By minimizing the reprojection error, the optimization process ensures pixel-level calibration precision, thus improving the alignment between the LiDAR point cloud and the image. The goal of optimization is to minimize the reprojection error. For edge point LP_i, the residual is formulated as:

$$r_i = n_i^T\left(f\left(\pi\left((T^E_L)^{-1}(LP_i)\right)\right) - (q_i + I_{wi})\right). \tag{11}$$

where n_i is the normal vector of the edge in the image, q_i is the corresponding edge point in the image, I_{wi} is the image noise, and r_i is the residual for the edge point.

Next, weighted least squares is applied to optimize the extrinsic parameters E_LT. The noise model accounts for measurement noise in both LiDAR and image sensors. LiDAR noise is typically modeled using a Gaussian distribution, which allows the quantification of the impact of different types of noise on the measurement results via a covariance matrix. During the optimization process, the noise is handled by the weighted least squares method, reducing the influence of noise on the calibration result and enhancing the robustness of the optimization.

$$\min_{\delta T}\sum_i (r_i + J_i\delta T)^T(J_w\Sigma J_w^T)^{-1}(r_i + J_i\delta T) \tag{12}$$

where $J_w\Sigma J_w^T$ is the covariance matrix of the residuals, and Σ contains the noise models for both LiDAR and image. This weighted optimization effectively mitigates the impact of noise on the calibration result, ensuring stable and robust

calibration outcomes. To minimize the objective function, the update to the extrinsic parameters δT is computed as:

$$\delta T^* = -\left(J_T^T (J_w \Sigma J_w^T)^{-1} J_T\right)^{-1} J_T^T (J_w \Sigma J_w^T)^{-1} r \tag{13}$$

The extrinsic parameters are updated as:

$$T_E^L \leftarrow \exp\left(\widehat{\delta T^*}\right) T_E^L, \tag{14}$$

with $\widehat{(\cdot)}$ the $\mathfrak{se}(3)$ hat operator. This cross-modal, edge-alignment refinement recovers accurate full extrinsics and improves calibration robustness. Through this optimization process, we obtain the optimal extrinsic parameters, effectively reducing the impact of noise on calibration uncertainty via noise modeling and weighted optimization, thus ensuring the stability and robustness of the calibration process.

IV. EXPERIMENTS

A. Experiment Setup and Evaluation Metrics

Experiments on both real-world and open-source datasets are conducted to validate the proposed method. The system employs a Prophesee EVK4-HD event camera with a resolution of 1280×720 and a Livox MID-360 LiDAR using Livox's unique rotating mirror solid-state technology. A compact multi-sensor platform with LiDAR and event camera is deployed on mobile robots and drones. Experimental datasets are collected from this platform as shown in Fig. 4. To further validate the versatility of our calibration system, we utilized open-source datasets. The open-source data is based on the DSEC dataset[54], which includes scenarios from urban, suburban, and rural environments. This platform is equipped with the Prophesee Gen3.1 event camera, offering a resolution of 640×480, and the Velodyne VLP-16 LiDAR.

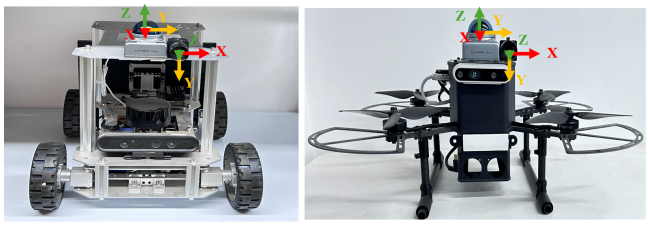


Fig. 4: Physical platform. The left side is the Livox MID-360 LiDAR, and the right side is the Prophesee EVK4-HD event camera. Data is collected using two mobile platforms: a ground-based mobile robot and an aerial drone.

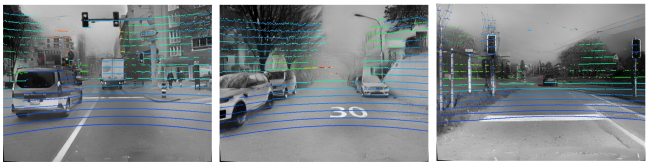
Mean Absolute Error (MAE) is used for both translation and rotation parameters [43]. The MAE of the translation component is defined as the average of the absolute difference between the translation vector and the actual translation vector. The error for each component is given by:

$$T_{MAE} = \frac{1}{N} \sum_{i=1}^{N} \|t_{p,i} - t_{g,i}\|_2 \tag{15}$$

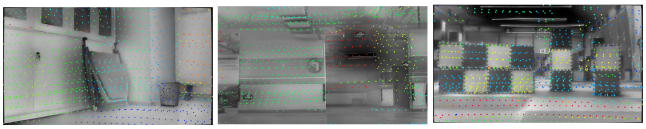
where $t_{p,i}$ and $t_{g,i}$ represent the predicted and ground truth translation vectors for the i-th sample, respectively, and N denotes the number of test samples. The same applies to the rotation R_{MAE} as well.

B. Calibration Results Assessment

We conducted a series of experiments to validate the effectiveness and robustness of the proposed extrinsic calibration method, with the projection results shown in Fig. 5. The figure displays the results of our outdoor tests using the DSEC dataset, where data from the left event camera and LiDAR sensor were used. The test scenarios include urban, suburban, and rural environments. Additionally, the figure also presents data collected in indoor environments using a physical platform, covering scenes such as offices, laboratories, and underground parking garages. Our algorithm accurately aligns the point clouds with the event camera and demonstrates high adaptability. The method is not limited by specific sensor types or environmental conditions and can effectively handle and calibrate various sensor types, ensuring consistent and accurate results across different scenes.



(a) Outdoor scenes



(b) Intdoor scenes

Fig. 5: Qualitative results. (a) Data collected by dsec in an outdoor autonomous driving scenario. (b) Data collected by the physical platform in an outdoor environment. Both (a) and (b) are images depicting LiDAR point clouds projected onto the event frame.

To evaluate the robustness of our method, we randomly multiple initial values in different scenarios for the extrinsic calibration (90 trials). The Fig. 6 shows the optimized extrinsic calibration parameters, including translation and rotation, with the initial seed values marked by gray dots and the calibration results represented by green markers. The results demonstrate that our method produces consistent calibration outcomes across different initial values and scenarios.

C. Comparison Experiments

We compared our method with existing event camera and LiDAR calibration methods, including L2E[29], LEC-Calib [28], and MULi-Ev [43], using the publicly available DSEC dataset. As shown in Table **??**, the translation error is reduced to an average of 0.67 cm, and the rotation error is reduced to 0.12°. Although our method has a slightly higher rotation error than MULi-Ev, the difference is very small, and the overall performance is excellent. The estimation of rotation error relies on the directional information of features, which

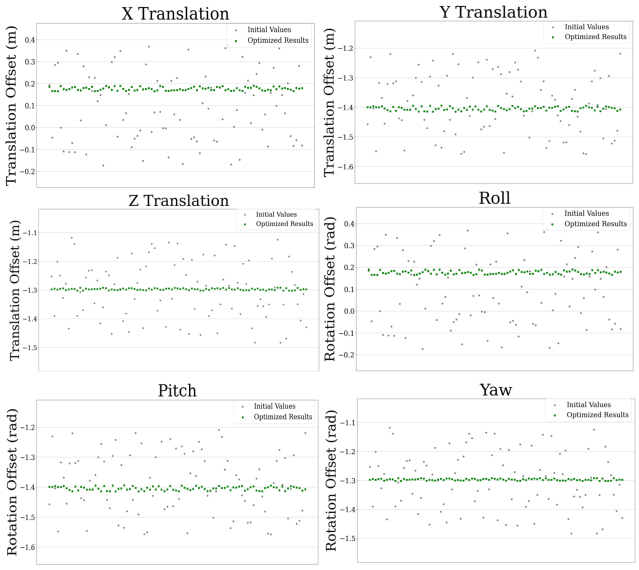


Fig. 6: The external parameters are presented, where grey represents the random initial values, and green represents the optimized solutions.

TABLE I: Comparison results with other methods.

Method	Translation Error (cm)	Rotation Error (deg)	Online	Target
L2E[29]	N/A	N/A	No	No
LCE-Calib[28]	1.5	0.3	No	Yes
MULi-Ev[43]	0.81	**0.10**	Yes	No
TC-LEC	**0.67**	0.12	No	No

is more susceptible to noise, changes in lighting, or scene complexity. These factors can lead to such results.

To further evaluate the robustness of our method across different types of scenes, we measured the various sequences in different scenes from the DSEC dataset. We analyzed the errors using a box plot, which displays the first quartile, third quartile, and the mean of the data. As shown in the Fig. 6. It is evident from the figure that in Zurich, both the translation and rotation errors are the smallest, achieving the best calibration results. This result is primarily because there are more sequences from Zurich, with a total of 35 sequences. The Zurich scenes come from an urban environment, containing rich edge features such as roads, buildings, and vehicles. Despite the Zurich scenes including night scenes, the best accuracy was still achieved, indicating that our edge feature extraction method is robust under different lighting conditions. However, the Interlaken scenes are mostly rural, lacking the abundant edge features found in Zurich, and have fewer data sequences, with only 5 sequences. Overall, the test results across the three scenarios demonstrate stable performance and high accuracy.

D. Application

we equipped a mobile robot with the LiDAR and event camera and had it traverse the experimental site. By improving the SLAM algorithm [55], an accurate and detailed environmental map, along with the platform motion trajectory, was successfully created through the fusion of LiDAR and the event camera, as shown in Fig. 7. This map demonstrates the accuracy and effectiveness of the calibration of the event camera and LiDAR parameters on the physical platform. It also highlights the significant potential of LiDAR and event camera fusion in practical applications, providing strong support for the development of advanced environmental perception.

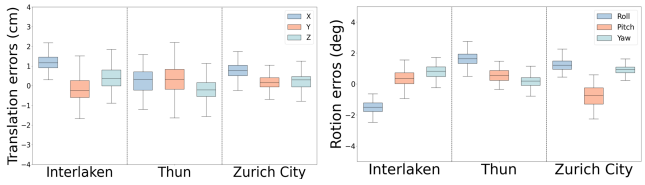
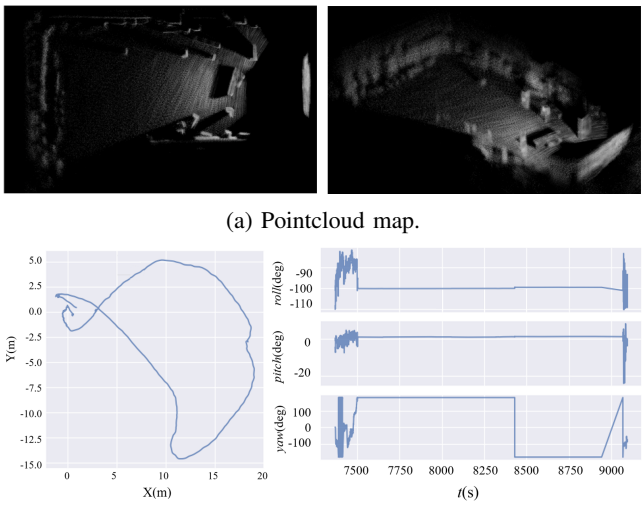


Fig. 7: Box plots of translation and rotation errors on the test DSEC set across different scenes.



(a) Pointcloud map.

(b) Trajectory Plot.

Fig. 8: Map and trajectory. (a) Different views of the mapping results, where the structures and features are clearly defined and closely aligned with the actual scene. (b) The movement trajectory of the mobile platform.

V. CONCLUSIONS

In this work, a reliable method is developed to calibrate the external parameters of Event-LiDAR systems. It performs effectively in general environments without controlled lighting or specialized equipment. The proposed approach is not limited by sensor models, and provides higher calibration accuracy and operational flexibility. It not only simplifies the calibration process, but also offers a reliable solution for deploying small autonomous systems in various environments. In the future, we plan to extend the framework by integrating multi-sensor data to further improve perception accuracy and robustness in complex scenes, enabling reliable decision-making in more challenging scenarios.

REFERENCES

[1] Guo S, Gallego G. "CMax-SLAM: Event-based rotational-motion bundle adjustment and SLAM system using contrast maximization," IEEE Transactions on Robotics, vol. 40, pp. 2442-2461, 2024.

[2] Niu J, Zhong S, Lu X, et al. "ESVO2: Direct visual-inertial odometry with stereo event cameras," IEEE Transactions on Robotics, 2025.

[3] Wang J, He J, Zhang Z, et al. "Physical priors augmented event-based 3D reconstruction," in Proc. 2024 IEEE Int. Conf. Robotics and Automation (ICRA), pp. 16810-16817.

[4] Scheerlinck C, Rebecq H, Gehrig D, et al. "Fast image reconstruction with an event camera," in Proc. IEEE/CVF Winter Conf. Applications of Computer Vision, 2020, pp. 156–163.

[5] S. Ghosh, V. Cavinato, and G. Gallego, "ES-PTAM: Event-based stereo parallel tracking and mapping," in *Proc. Eur. Conf. Comput. Vis. Workshops (ECCVW)*, 2024.

[6] Gehrig M, Scaramuzza D. "Recurrent vision transformers for object detection with event cameras," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2023, pp. 13884–13893.

[7] Mitrokhin A, Fermüller C, Parameshwara C, et al. "Event-based moving object detection and tracking," in *Proc. 2018 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 1-9.

[8] Afshar S, Nicholson A P, Van Schaik A, et al. "Event-based object detection and tracking for space situational awareness," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 15117-15132, 2020.

[9] Ding Z, Zhao R, Zhang J, et al. "Spatio-temporal recurrent networks for event-based optical flow estimation," in Proc. AAAI Conf. Artificial Intelligence, vol. 36, no. 1, 2022, pp. 525–533.

[10] Luo X, Luo A, Wang Z, et al. "Efficient meshflow and optical flow estimation from event cameras," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, 2024, pp. 19198–19207.

[11] Gallego G, Rebecq H, Scaramuzza D. "A unifying contrast maximization framework for event cameras, with applications to motion, depth, and optical flow estimation," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2018, pp. 3867–3876.

[12] K. Tang, X. Lang, Y. Ma, Y. Huang, L. Li, Y. Liu, and J. Lv, "Monocular event-inertial odometry with adaptive decay-based time surface and polarity-aware tracking," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 12544-12551, October 2024.

[13] V. Vasco, A. Glover, E. Mueggler, et al., "Independent motion detection with event-driven cameras," in *2017 18th International Conference on Advanced Robotics (ICAR)*, pp. 530-536, July 2017.

[14] H. Rebecq, R. Ranftl, V. Koltun, et al., "Events-to-video: Bringing modern computer vision to event cameras," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3857-3866, June 2019.

[15] Y. Li and J. Ibanez-Guzman, "Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 50-61, July 2020.

[16] S. Royo and M. Ballesta-Garcia, "An overview of lidar imaging systems for autonomous vehicles," *Applied Sciences*, vol. 9, no. 19, p. 4093, 2019.

[17] R. Ishikawa, T. Oishi, and K. Ikeuchi, "Lidar and camera calibration using motions estimated by sensor fusion odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 7342-7349, October 2018.

[18] Li B, Meng H, Zhu Y, et al. Enhancing 3-D LiDAR point clouds with event-based camera[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-12.

[19] Gehrig D, Scaramuzza D. Low-latency automotive vision with event cameras[J]. Nature, 2024, 629(8014): 1034-1040.

[20] Saucedo M A V, Patel A, Sawlekar R, et al. Event camera and lidar based human tracking for adverse lighting conditions in subterranean environments[J]. IFAC-PapersOnLine, 2023, 56(2): 9257-9262.

[21] Bartolomei L, Poggi M, Conti A, et al. Lidar-event stereo fusion with hallucinations[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 125-145.

[22] A. Dhall, K. Chelani, V. Radhakrishnan, and K. M. Krishna, "LiDAR-camera calibration using 3D-3D point correspondences," *arXiv preprint arXiv:1705.09785*, 2017.

[23] R. Song, Z. Jiang, Y. Li, Y. Shan, and K. Huang, "Calibration of event-based camera and 3d lidar," in *2018 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*, 2018, pp. 289–295.

[24] T. Tóth, Z. Pusztai, and L. Hajder, "Automatic LiDAR-camera calibration of extrinsic parameters using a spherical target," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 8580–8586.

[25] X. Li, Y. Xiao, B. Wang, H. Ren, Y. Zhang, and J. Ji, "Automatic targetless LiDAR–camera calibration: a survey," *Artificial Intelligence Review*, vol. 56, no. 9, pp. 9949–9987, 2023.

[26] H. Yu, W. Zhen, W. Yang, and S. Scherer, "Line-based 2-D–3-D registration and camera localization in structured environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8962–8972, 2020.

[27] E. Mueggler, H. Rebecq, G. Gallego, et al., "The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and SLAM," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 142-149, 2017.

[28] J. Jiao, F. Chen, H. Wei, J. Wu, and M. Liu, "Lce-calib: automatic lidar-frame/event camera extrinsic calibration with a globally optimal solution," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 5, pp. 2988–2999, 2023.

[29] K. Ta, D. Bruggemann, T. Brödermann, C. Sakaridis, and L. Van Gool, "L2E: Lasers to events for 6-DoF extrinsic calibration of lidars and event cameras," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 11425–11431.

[30] N. Ou, H. Cai, and J. Wang, "Targetless LiDAR-camera calibration via cross-modality structure consistency," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 2636-2648, 2023.

[31] C. Park, P. Moghadam, S. Kim, et al., "Spatiotemporal camera-LiDAR calibration: A targetless and structureless approach," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1556-1563, April 2020.

[32] B. Elnashef and S. Filin, "Target-free calibration of flat refractive imaging systems using two-view geometry," *Optics and Lasers in Engineering*, vol. 150, p. 106856, 2022.

[33] L. Li, H. Li, X. Liu, et al., "Joint intrinsic and extrinsic LiDAR-camera calibration in targetless environments using plane-constrained bundle adjustment," *arXiv preprint arXiv:2308.12629*, 2023.

[34] C. Yuan, X. Liu, X. Hong, and F. Zhang, "Pixel-level extrinsic self calibration of high resolution lidar and camera in targetless environments," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7517–7524, 2021.

[35] J. Kümmerle and T. Kühner, "Unified intrinsic and extrinsic camera and LiDAR calibration under uncertainties," in *2020 IEEE International conference on Robotics and Automation (ICRA)*, 2020, pp. 6028–6034.

[36] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, "Automatic extrinsic calibration of vision and lidar by maximizing mutual information," *Journal of Field Robotics*, vol. 32, no. 5, pp. 696–722, 2015.

[37] Z. Taylor and J. Nieto, "A mutual information approach to automatic calibration of camera and lidar in natural environments," in *Australian Conference on Robotics and Automation*, 2012, pp. 3–5.

[38] M. J. Dominguez-Morales, A. Jimenez-Fernandez, G. Jimenez-Moreno, C. Conde, E. Cabello, and A. Linares-Barranco, "Bio-inspired stereo vision calibration for dynamic vision sensors," *IEEE Access*, vol. 7, pp. 138415–138425, 2019.

[39] M. Muglikar, M. Gehrig, D. Gehrig, and D. Scaramuzza, "How to calibrate your event camera," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 1403–1409.

[40] E. Mueggler, B. Huber, and D. Scaramuzza, "Event-based, 6-dof pose tracking for high-speed maneuvers," in 2014 IEEE/RSJ International Confere

[41] Rebecq H, Ranftl R, Koltun V, et al. High speed and high dynamic range video with an event camera[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(6): 1964-1980.

[42] K. Ta, D. Bruggemann, T. Brodermann, C. Sakaridis, and L. Van Gool, "L2e: Lasers to events for 6-dof extrinsic calibration of lidars and event cameras," 2022.

[43] M. Cocheteux, J. Moreau, and F. Davoine, "MULi-Ev: Maintaining Unperturbed LiDAR-Event Calibration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 4579–4586.

[44] X. Li, Y. Zhou, R. Guo, X. Peng, Z. Zhou, and H. Lu, "Spatio-temporal calibration for omni-directional vehicle-mounted event cameras," IEEE Robotics and Automation Letters, 2024.

[45] Mai J, Lu X, Dai K, et al. Temporal and Rotational Calibration for Event-Centric Multi-Sensor Systems[J]. arXiv preprint arXiv:2508.12564, 2025.
[46] G. Gallego, M. Gehrig, and D. Scaramuzza, "Focus is all you need: Loss functions for event-based vision," in IEEE Conf. Comput. Vis. Pattern Recog. (CVPR), 2019, pp. 12 272–12 281.
[47] T. Stoffregen and L. Kleeman, "Event cameras, contrast maximization and reward functions: an analysis," in IEEE Conf. Comput. Vis. Pattern Recog. (CVPR), 2019, pp. 12 292–12 300.
[48] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in Robotics: science and systems, vol. 2, no. 4. Seattle, WA, 2009, p. 435.
[49] K. Koide, "small-gicp: Efficient and parallel algorithms for point cloud registration," Journal of Open Source Software, vol. 9, no. 100, p. 6948, Aug. 2024.
[50] H. Rebecq, R. Ranftl, V. Koltun, and D. Scaramuzza, "High speed and high dynamic range video with an event camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 6, pp. 1964–1980, 2019.
[51] Y. Zhu, C. Zheng, C. Yuan, X. Huang, and X. Hong, "Camvox: A low-cost and accurate lidar-assisted visual SLAM system," *arXiv preprint arXiv:2011.11357*, 2020.
[52] D. Scaramuzza, A. Harati, and R. Siegwart, "Extrinsic self-calibration of a camera and a 3D laser range finder from natural scenes," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4164–4169, IEEE, 2007.
[53] J. Levinson and S. Thrun, "Automatic online calibration of cameras and lasers," in *Robotics: Science and Systems*, vol. 2, p. 7, Citeseer, 2013.
[54] M. Gehrig, W. Aarents, D. Gehrig, and D. Scaramuzza, "DSEC: A stereo event camera dataset for driving scenarios," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4947–4954, 2021.
[55] C. Zheng, W. Xu, Z. Zou, et al., "Fast-LIVO2: Fast, direct lidar-inertial-visual odometry," *IEEE Transactions on Robotics*, 2024.