Supplementary Material for WACV 2026
UniCalib: Targetless LiDAR-Camera Calibration via Probabilistic Flow on Unified Depth Representations

Shu Han [1] Xubo Zhu [2] Ji Wu [1] Ximeng Cai [1] Wen Yang [2] Huai Yu [2] Gui-Song Xia [3]
[1] School of Computer Science, Wuhan University [2] School of Electronic Information, Wuhan University
[3] School of Artificial Intelligence, Wuhan University
{hanshu, zhuxubo, ji.wu, ximengcai, yangwen, yuhuai, guisong.xia}@whu.edu.cn

Abstract

This document is the supplementary material of our WACV submission. It includes additional visualizations, evaluation metrics, and qualitative results to further demonstrate the effectiveness and generalization of our proposed calibration method UniCalib.

1. Uncertainty and Outlier Probability Maps

To further analyze the behavior of our probabilistic model, we visualize its outputs on a representative frame, as shown in Fig. 1. The projected LiDAR image (a) serves as a reference to illustrate the initial misalignment before calibration. The uncertainty map (b) is the predicted per-pixel variance $\boldsymbol{\sigma^2}$ of the depth flow, derived from our probabilistic formulation. Higher values indicate lower confidence in the prediction. Similarly, (c) visualizes the per-pixel outlier probability map $\boldsymbol{\alpha}$ in our probabilistic framework. Higher values of $\boldsymbol{\alpha}$ indicate a greater likelihood of unreliability, thereby down-weighting the influence of the corresponding pixels during optimization.

Although the two maps are derived from different components of our model, we observe that both exhibit notably higher values in the upper half of the image. This phenomenon arises because these regions are visible in the camera image but not covered by the LiDAR sensor due to its limited vertical field of view, resulting in partial occlusion. Consequently, the depth flow estimation in these areas is inherently unreliable. Our model successfully captures this effect, demonstrating its ability to reason about visibility and cross-modal inconsistency. Notably, the outlier probability map shows overall higher responses in uncertain regions, highlighting its stronger suppression of unreliable gradients during training.

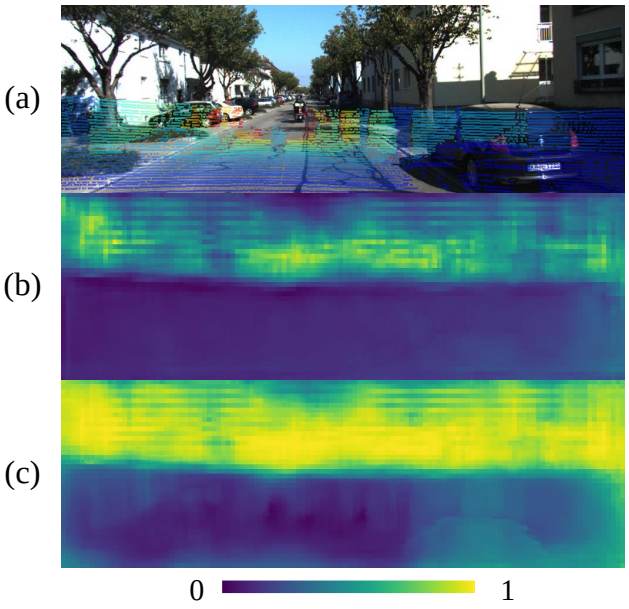


Figure 1. Visualization of model outputs: (a) LiDAR projection using uncalibrated extrinsics, (b) predicted uncertainty map, and (c) learned outlier probability map $\boldsymbol{\alpha}$.

2. Evaluation Metrics

To quantitatively evaluate the LiDAR-camera extrinsic calibration results, we compute the error between the predicted transformation $\mathbf{T}_{\text{est}}$ and the ground truth transformation $\mathbf{T}_{\text{gt}}^{'}$. For each test sample, we first calculate the relative transformation:

$$\Delta\mathbf{T} = \mathbf{T}_{\text{gt}}^{'\,-1} \cdot \mathbf{T}_{\text{est}} = \begin{bmatrix} \Delta\mathbf{R} & \Delta\mathbf{t} \\ \mathbf{0^T} & 1 \end{bmatrix}, \tag{1}$$

where $\Delta\mathbf{R} \in SO(3)$ is the relative rotation matrix and $\Delta\mathbf{t} \in \mathbb{R}^3$ is the relative translation vector.

2.1. Translation Error

The translation error is computed as the element-wise absolute value of the relative translation vector:

$$\text{Err}_{\text{trans},i} = |\Delta t_i|, \quad i \in \{x, y, z\}. \tag{2}$$

2.2. Rotation Error

The rotation error is computed by first converting the relative rotation matrix $\Delta\mathbf{R}$ to Euler angles $(\Delta\phi, \Delta\theta, \Delta\psi)$ representing roll, pitch, and yaw. These are extracted using the standard ZYX (yaw-pitch-roll) decomposition:

$$[\Delta\phi, \Delta\theta, \Delta\psi] = \text{Euler}_{ZYX}(\Delta\mathbf{R}). \tag{3}$$

Each component is then converted from radians to degrees, and the final rotation error is defined as

$$\text{Err}_{\text{rot},i} = |\Delta\theta_i| \cdot \left(\frac{180}{\pi}\right), \quad i \in \{\text{roll}, \text{pitch}, \text{yaw}\}. \tag{4}$$

3. Error Distributions

To provide a more comprehensive analysis of the calibration accuracy of our method, we visualize the translation error distributions per axis and rotation in Fig. 2. For each experimental setup (Exp1–Exp5), we separately plot the distributions of translation and rotation errors along the x, y, z axes and roll, pitch, yaw angles. The results demonstrate that most errors are concentrated near zero across all translation and rotation axes, validating the effectiveness and robustness of our approach.

4. Real-Time Qualitative Results

We present real-time qualitative results on the KITTI Odometry [1], KITTI Raw [1], KITTI-360 [2], and Waymo[3] datasets in the accompanying video provided in the supplementary material. To better illustrate the calibration accuracy of our method, we showcase side-by-side comparisons of three configurations: (1) uncalibrated input, (2) calibrated results using our estimated extrinsics, and (3) calibrated results using ground truth extrinsics. These visualizations highlight the effectiveness of our approach in producing accurate and consistent alignment between LiDAR and camera data in diverse real-world scenarios.

5. Qualitative Results

To further assess the effectiveness and generalization ability of our method, we provide qualitative comparisons on three real-world datasets: KITTI Odometry, KITTI Raw, KITTI-360, and Waymo. For each dataset, we visualize the projection of LiDAR points onto the image plane under three settings: (a) uncalibrated extrinsics, (b) calibrated using our estimated extrinsics, and (c) calibrated with ground

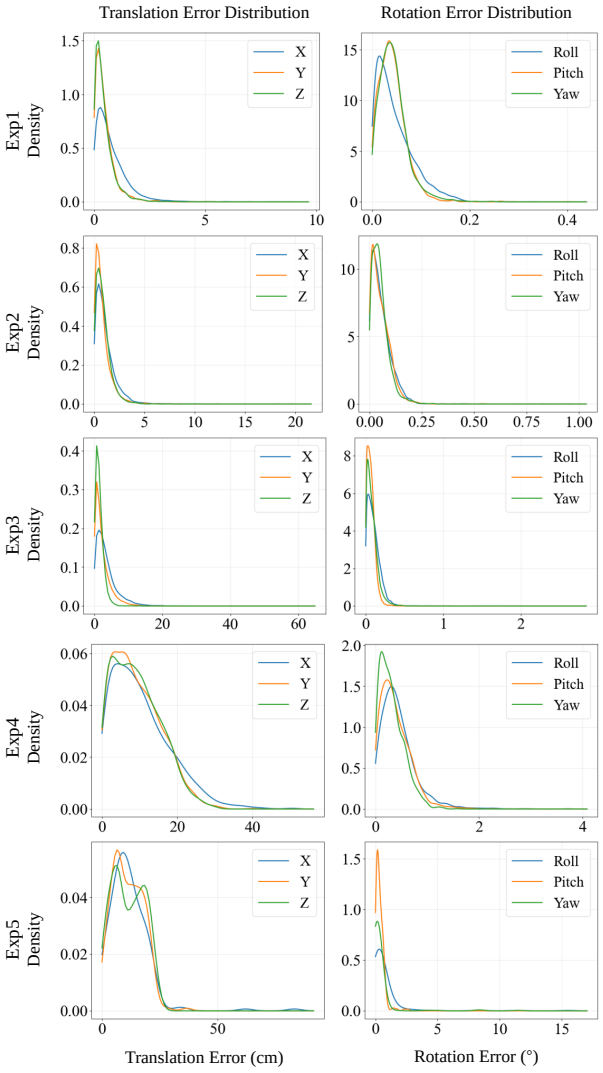


Figure 2. Distributions of translation and rotation errors on each axis under different experimental settings (Exp1–Exp5).

truth extrinsics. As shown in Figs. 3 to 6, our method significantly improves the alignment between the point cloud and the image, achieving results visually comparable to those using ground truth extrinsics. These results further demonstrate the robustness and generalization ability of our method across diverse sensor setups, scene geometries, and acquisition conditions.

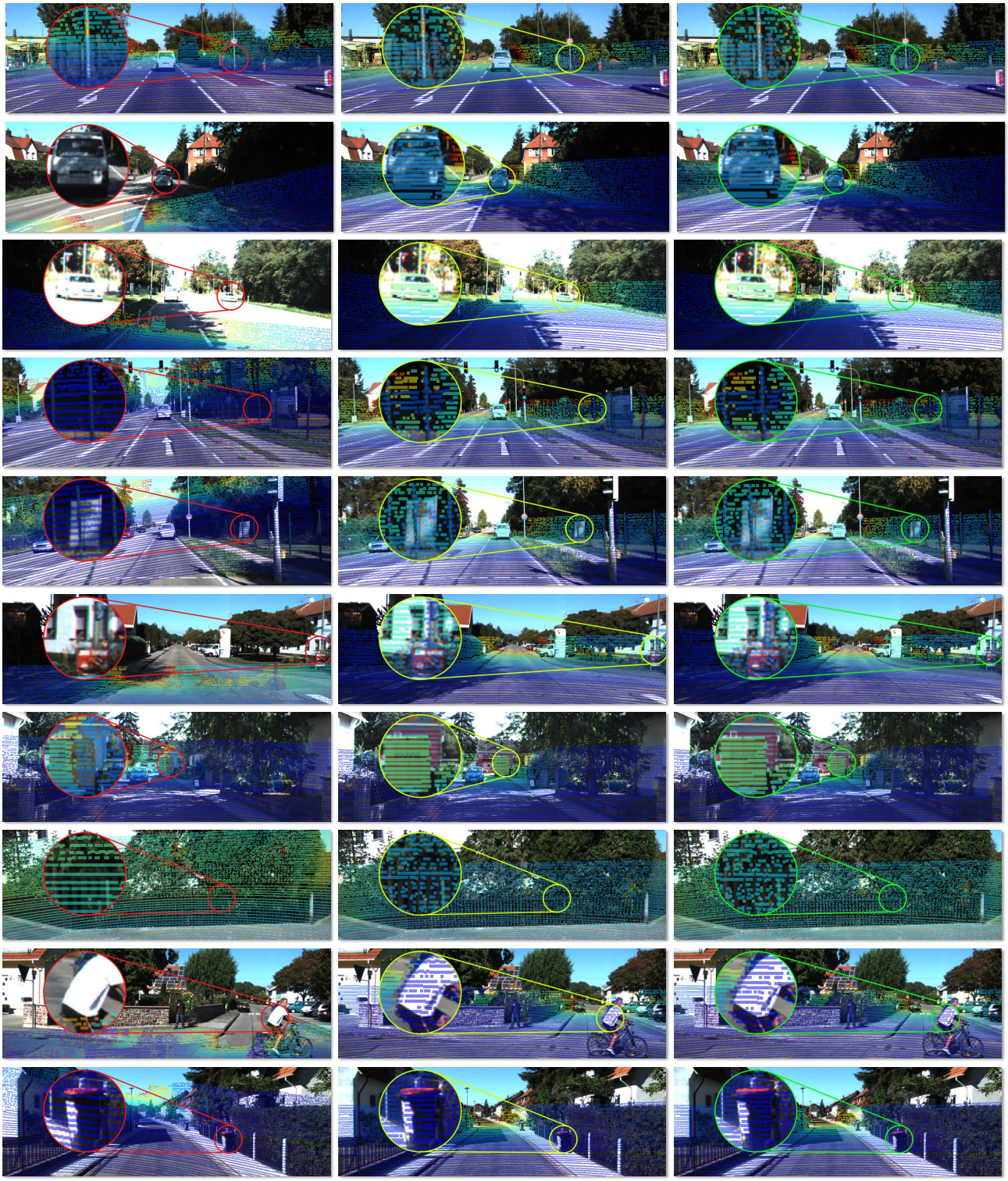
References

[1] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pages 3354–3361, 2012. 2

[2] Yiyi Liao, Jun Xie, and Andreas Geiger. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2D

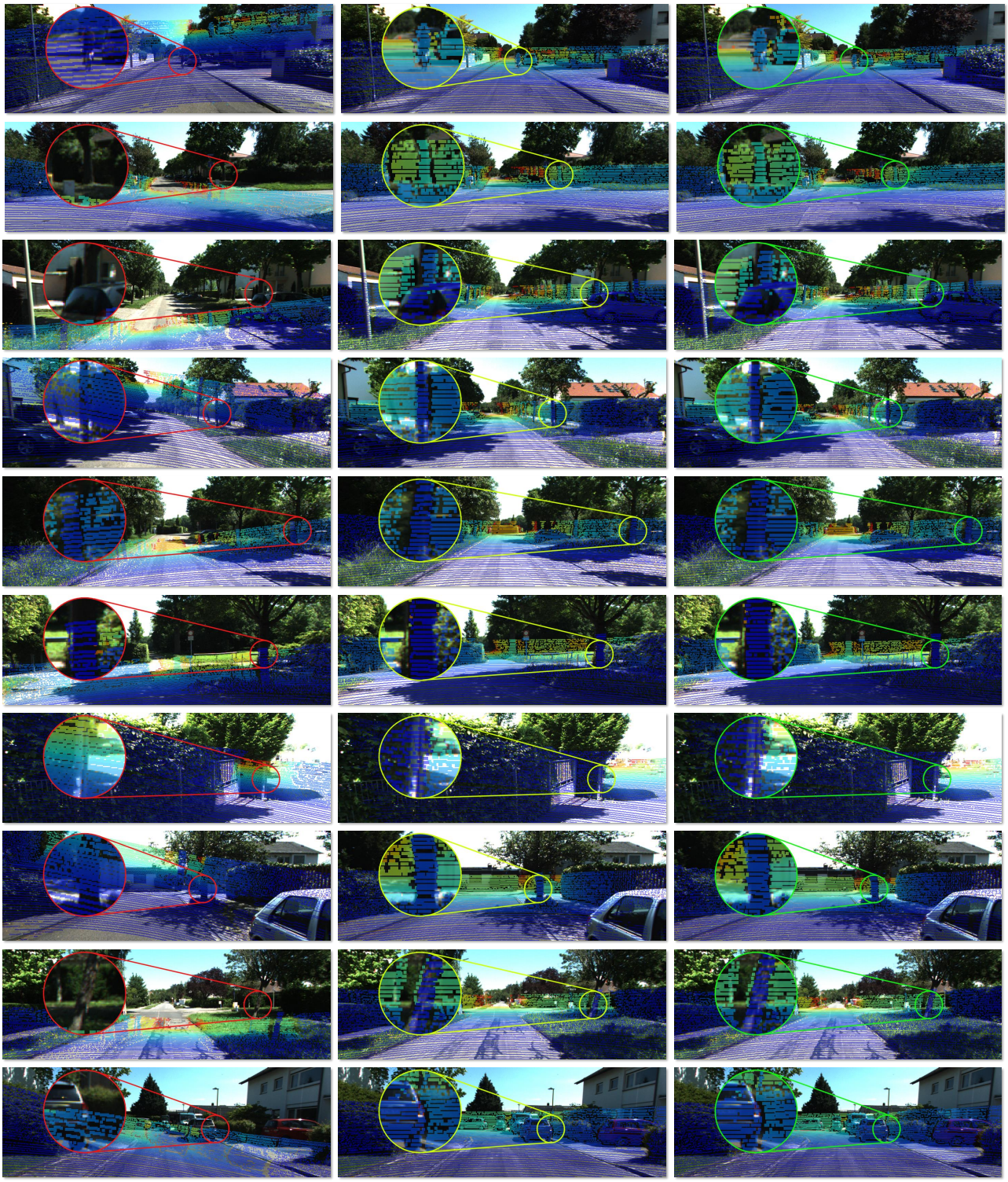
(a) Uncalibrated (b) Calibrated with Estimated Extrinsics (c) Calibrated with GT Extrinsics

Figure 3. Qualitative results on the KITTI Odometry dataset. Each row shows a sample from different scenes under three calibration states: (a) uncalibrated, (b) calibrated with estimated extrinsics, and (c) calibrated with ground truth extrinsics. Our method produces accurate alignment between LiDAR points and image structures.



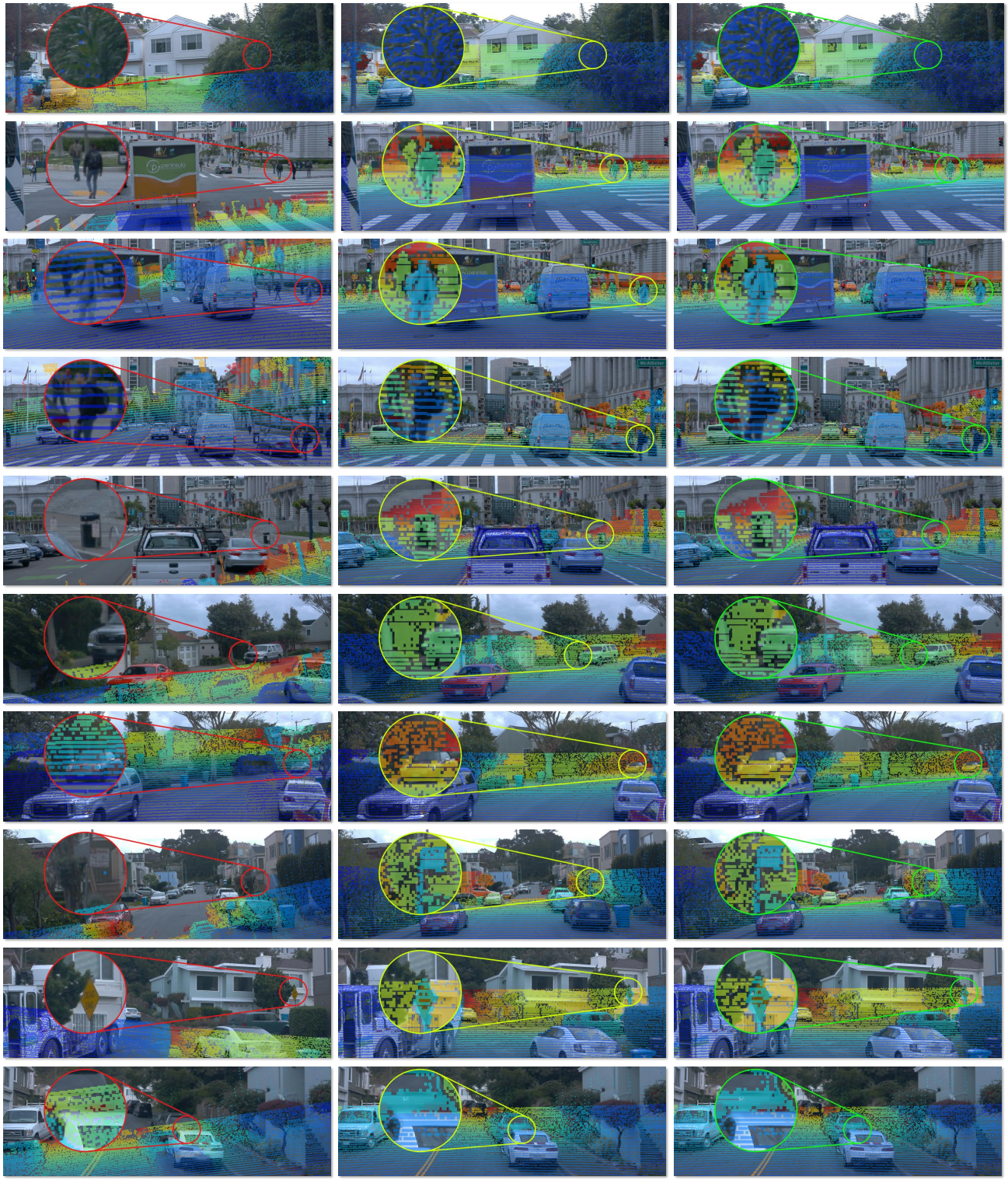
(a) Uncalibrated (b) Calibrated with Estimated Extrinsics (c) Calibrated with GT Extrinsics

Figure 4. Qualitative results on the KITTI Raw dataset. Our method effectively handles diverse urban and residential scenes, producing point-to-image alignment comparable to that of the ground truth extrinsics.



(a) Uncalibrated (b) Calibrated with Estimated Extrinsics (c) Calibrated with GT Extrinsics

Figure 5. Qualitative results on the KITTI-360 dataset. This dataset is not used during training and serves as an evaluation of the generalization ability of our method. Despite the domain gap and more complex urban scenes, our method still achieves accurate and consistent calibration results.



(a) Uncalibrated (b) Calibrated with Estimated Extrinsics (c) Calibrated with GT Extrinsics

Figure 6. Qualitative results on the Waymo dataset. This dataset is not involved in training and is used to evaluate the generalization capability of our approach. Even under the significant domain gap and the diverse, large-scale driving scenarios in Waymo, our method demonstrates robust performance, producing accurate and stable calibration results.

and 3D. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(3):3292–3310, 2022. 2

[3] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2446–2454, 2020. 2