

# LiREC-Net: A Target-Free and Learning-Based Network for LiDAR, RGB, and Event Calibration

## Supplementary Material

Aditya Ranjan Dash<sup>1</sup>   Ramy Batrawy<sup>2</sup>   René Schuster<sup>1,2</sup>   Didier Stricker<sup>1,2</sup>

<sup>1</sup>RPTU – University Kaiserslautern-Landau

<sup>2</sup>DFKI – German Research Center for Artificial Intelligence

firstname.lastname@dfki.de

### A. Overview

Here, we provide the supplementary material to our main paper “*LiREC-Net: A Target-Free and Learning-Based Network for LiDAR, RGB, and Event Calibration*”. We detail the division of data for training and evaluation, list all relevant hyperparameters required for the reproduction of the results, and include more quantitative and qualitative results and comparisons.

### B. Dataset Splits

Tables 1 and 2 summarize the training and evaluation splits used for KITTI [2] and DSEC [1], respectively. Each table lists the sequences included in the training and evaluation sets along with the corresponding number of samples.

### C. Training Hyperparameters

The training schedule for KITTI and DSEC datasets, detailing the number of epochs and learning rate decay milestones per stage, is provided in Tab. 3. Additionally, Tab. 4 lists the hyperparameters used for all experiments.

### D. Comparison with LCCRAFT

In addition to the results presented in the main paper, we further compare LiREC-Net to LCCRAFT [3] using the evaluation metrics defined by LCCRAFT. Specifically, we report the per-axis translation errors ( $e_X, e_Y, e_Z$ ) and the roll, pitch, and yaw rotation errors ( $e_R, e_P, e_Y$ ) in Tab. 5.

Table 1. Training and evaluation splits for the KITTI dataset [2].

Split	Sequences	No. of samples
Training	01–21	39011
Evaluation	00	4541

Table 2. Training and evaluation splits for the DSEC dataset [1].

Split	Sequences	No. of samples
Training	interlaken_00_c–g	52727
	thun_00_a	
	zurich_city_00_a–b	
	zurich_city_01_a–f	
	zurich_city_02_a–e	
	zurich_city_03_a	
	zurich_city_04_a–f	
	zurich_city_05_a–b	
	zurich_city_06_a	
	zurich_city_07_a	
	zurich_city_08_a	
Evaluation	zurich_city_09_a–e	11204
	zurich_city_10_a–b	
	zurich_city_11_a–c	
	interlaken_00_a–b	
	interlaken_01_a	
	thun_01_a–b	
Evaluation	zurich_city_12_a	11204
	zurich_city_13_a–b	
	zurich_city_14_a–c	
	zurich_city_15_a	

Table 3. Training schedule with epochs and learning rate (LR) milestones per stage.

Dataset	Stage	Epochs	Decay Milestones
KITTI [2]	Stage 1	120	20, 50, 70, 100
	Other Stages	50	15, 30, 40
DSEC [1]	Stage 1	150	100, 110, 120, 130
	Other Stages	70	45, 50, 55, 60

Table 4. Hyperparamters and corresponding values of LiREC-Net.

Parameter(s)	Description	Value(s)
$H \times W$	Image size (KITTI)	$376 \times 1240$
	Image size (DSEC)	$1080 \times 1440$
$H' \times W'$	Model input size	$256 \times 512$
$H'' \times W''$	Feature size	$16 \times 32$
$d$	Correlation radius	4
$\mu_{RGB}$	Mean value (KITTI)	[0.485, 0.456, 0.406]
	Mean value (DSEC)	[0.265, 0.283, 0.300]
$\sigma_{RGB}$	Standard deviation (KITTI)	[0.229, 0.224, 0.225]
	Standard deviation (DSEC)	[0.245, 0.270, 0.301]
$z_{max}$	Maximum depth value	80m

Table 5. Comparison between LCCRAFT [3] and our LiREC-Net on KITTI. Errors are reported per translation axis ( $e_X$ ,  $e_Y$ ,  $e_Z$  in cm) and rotation ( $e_R$ ,  $e_P$ ,  $e_Y$  in  $^\circ$ ).

Network	Translation				Rotations			
	$e_X$ [cm]	$e_Y$ [cm]	$e_Z$ [cm]	$\frac{e_X + e_Y + e_Z}{3}$ [cm]	$e_R$ $^\circ$	$e_P$ $^\circ$	$e_Y$ $^\circ$	$\frac{e_R + e_P + e_Y}{3}$ $^\circ$
LCCRAFT [3]	<b>0.821</b>	0.771	1.631	1.074	<b>0.029</b>	0.117	0.041	0.062
LiREC-Net	1.348	<b>0.523</b>	<b>0.747</b>	<b>0.873</b>	0.043	<b>0.092</b>	<b>0.030</b>	<b>0.055</b>

## E. Detailed Results

The detailed per-stage results for all modality–dataset combinations are provided in Tab. 6a–Tab. 6d.

All results are obtained under a maximum perturbation of  $\pm\{20^\circ/150\text{ cm}\}$  and reflect the 5-stage iterative calibration setup. For each stage, we report the per-axis translation errors ( $e_X, e_Y, e_Z$ ) and the roll, pitch, and yaw error components ( $e_R, e_P, e_Y$ ) (cf. Sec. D) together with the mean translation error  $e_t$  and the mean rotation error  $e_r$ .

The stage-wise results reveal a consistent improvement in both translation and rotation accuracy, highlighting the effectiveness of calibration using LiREC-Net for LiDAR-RGB and LiDAR-Event on both KITTI and DSEC.

## F. Additional Qualitative Results

To further illustrate the behavior of LiREC-Net across different stages of refinement, we provide additional per-stage visualizations for both KITTI [2] and DSEC [1], covering LiDAR-RGB and LiDAR-Event pairs (see Figs. 1 to 4). For each sample, we apply the maximum perturbations used during evaluation to produce challenging miscalibration scenarios. As shown, the miscalibrated input (Stage 0) can lead to severe misalignment, where, for KITTI (Figs. 1 and 2) in particular, no LiDAR points fall inside the camera frustum due to the large initial translation and rotation errors.

Despite these extreme perturbations, LiREC-Net consistently recovers accurate calibration. Across Stage 1 to Stage 5, the network progressively realigns the LiDAR point cloud to the corresponding RGB or event frame. After Stage 3, the projections are visually indistinguishable from the ground-truth calibration, demonstrating the robustness of our LiREC-Net. These examples highlight that LiREC-Net can handle even difficult and highly perturbed inputs, producing stable and accurate alignment for both LiDAR-RGB and LiDAR-Event across both the datasets.

## References

- [1] Mathias Gehrig, Willem Aarents, Daniel Gehrig, and Davide Scaramuzza. DSEC: A Stereo Event Camera Dataset for Driving Scenarios. *IEEE Robotics and Automation Letters (RA-L)*, 2021. 1, 2
- [2] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012. 1, 2
- [3] Yu-Chen Lee and Kuan-Wen Chen. LCCRAFT: LiDAR and Camera Calibration Using Recurrent All-Pairs Field Transforms Without Precise Initial Guess. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024. 1, 2

Table 6. Detailed per-stage results of LIREC-Net.

(a) LiDAR-RGB on DSEC.

Stage	Translation				Rotation			
	$e_X$ [cm]	$e_Y$ [cm]	$e_Z$ [cm]	$e_t$ [cm]	$e_R$ [°]	$e_P$ [°]	$e_Y$ [°]	$e_r$ [°]
1	9.54	2.33	4.39	11.82	0.24	0.48	0.31	0.72
2	2.93	0.93	2.22	4.23	0.10	0.46	0.13	0.54
3	1.86	0.63	1.37	2.66	0.07	0.37	0.09	0.42
4	1.17	0.57	1.57	2.25	0.07	0.17	0.08	0.24
5	1.05	0.70	1.95	2.51	0.06	0.07	0.07	0.14

(b) LiDAR-Event on DSEC.

Stage	Translation				Rotation			
	$e_X$ [cm]	$e_Y$ [cm]	$e_Z$ [cm]	$e_t$ [cm]	$e_R$ [°]	$e_P$ [°]	$e_Y$ [°]	$e_r$ [°]
1	9.29	2.37	4.50	11.66	0.22	0.47	0.29	0.69
2	2.80	0.83	2.50	4.33	0.07	0.42	0.08	0.46
3	1.61	0.53	1.38	2.44	0.05	0.33	0.06	0.36
4	1.15	0.92	2.63	3.24	0.05	0.15	0.05	0.19
5	0.91	0.27	0.49	1.18	0.02	0.04	0.04	0.07

(c) LiDAR-RGB on KITTI.

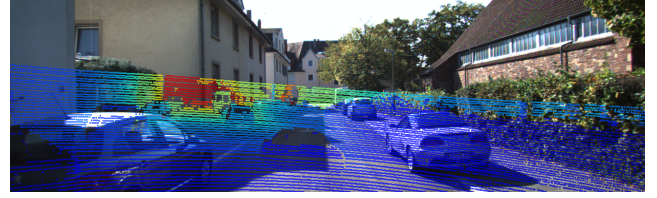
Stage	Translation				Rotation			
	$e_X$ [cm]	$e_Y$ [cm]	$e_Z$ [cm]	$e_t$ [cm]	$e_R$ [°]	$e_P$ [°]	$e_Y$ [°]	$e_r$ [°]
1	5.34	3.99	5.74	10.04	0.22	0.39	0.32	0.63
2	2.89	2.01	2.32	4.76	0.10	0.17	0.11	0.26
3	1.89	1.31	1.82	3.25	0.07	0.11	0.06	0.16
4	1.48	0.61	1.28	2.26	0.04	0.09	0.04	0.12
5	1.35	0.52	0.75	1.80	0.04	0.09	0.03	0.11

(d) LiDAR-Event on KITTI.

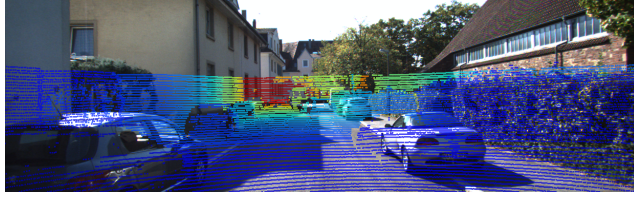
Stage	Translation				Rotation			
	$e_X$ [cm]	$e_Y$ [cm]	$e_Z$ [cm]	$e_t$ [cm]	$e_R$ [°]	$e_P$ [°]	$e_Y$ [°]	$e_r$ [°]
1	4.89	3.25	3.64	7.91	0.20	0.36	0.31	0.60
2	2.75	1.79	2.88	4.91	0.07	0.17	0.07	0.22
3	1.92	1.60	1.91	3.46	0.06	0.11	0.08	0.16
4	1.44	0.58	1.31	2.24	0.03	0.09	0.05	0.12
5	1.30	0.51	0.88	1.82	0.04	0.09	0.04	0.12



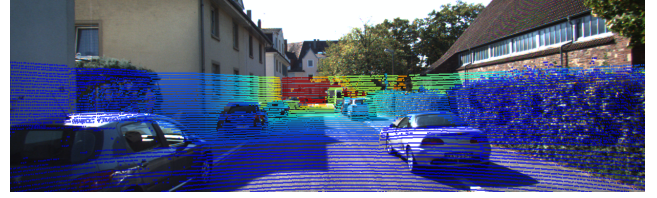
(a) Miscalibrated input [ $e_t = 173.88\text{cm}$ ,  $e_r = 27.10^\circ$ ]



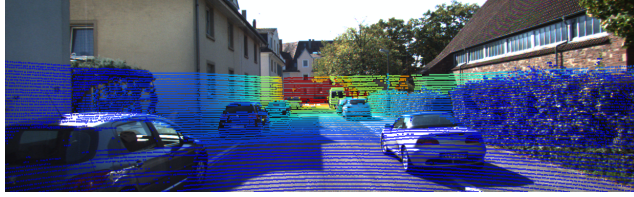
(b) Stage 1 prediction [ $e_t = 106.92\text{cm}$ ,  $e_r = 18.97^\circ$ ]



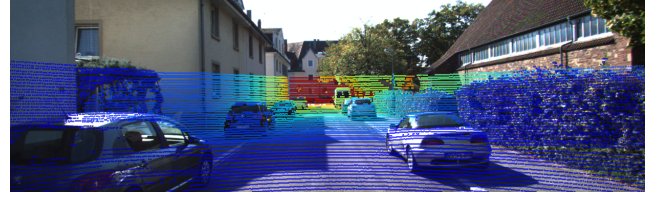
(c) Stage 2 prediction [ $e_t = 61.72\text{cm}$ ,  $e_r = 6.69^\circ$ ]



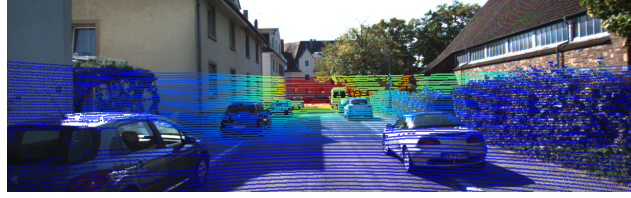
(d) Stage 3 prediction [ $e_t = 6.07\text{cm}$ ,  $e_r = 1.24^\circ$ ]



(e) Stage 4 prediction [ $e_t = 3.60\text{cm}$ ,  $e_r = 0.19^\circ$ ]



(f) Stage 5 prediction [ $e_t = 2.81\text{cm}$ ,  $e_r = 0.10^\circ$ ]



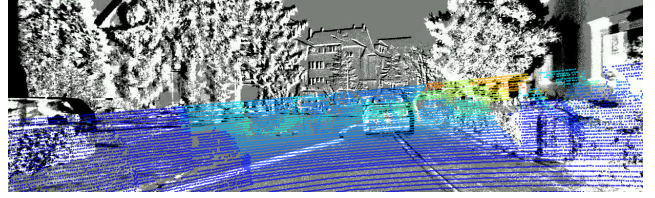
(g) Ground truth

Figure 1. Per-stage visualization for LiDAR-RGB calibration on KITTI. The exact translation and rotation error is provided for the perturbed input and each stage. For miscalibrated input, there are no LiDAR points inside the RGB image boundary because of the large perturbation.

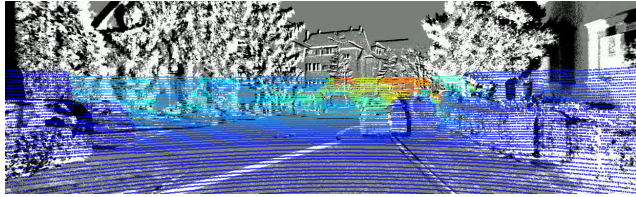




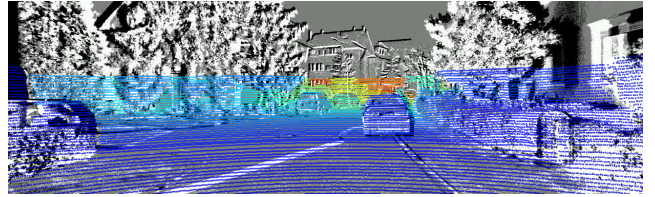
(a) Miscalibrated input [ $e_t = 177.95\text{cm}$ ,  $e_r = 26.54^\circ$ ]



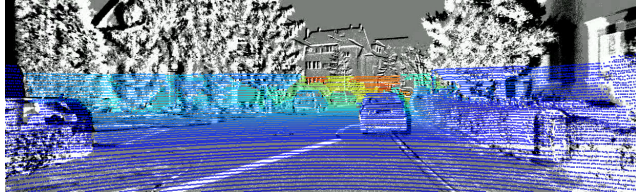
(b) Stage 1 prediction [ $e_t = 129.84\text{cm}$ ,  $e_r = 16.64^\circ$ ]



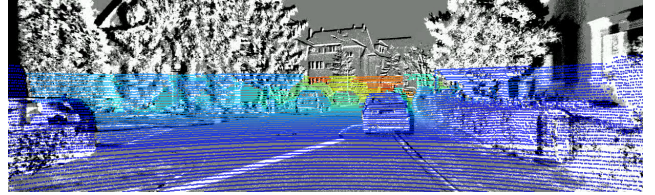
(c) Stage 2 prediction [ $e_t = 19.93\text{cm}$ ,  $e_r = 3.93^\circ$ ]



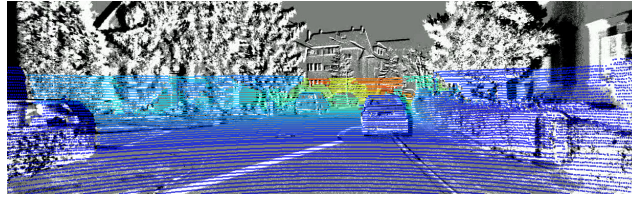
(d) Stage 3 prediction [ $e_t = 2.62\text{cm}$ ,  $e_r = 0.11^\circ$ ]



(e) Stage 4 prediction [ $e_t = 1.90\text{cm}$ ,  $e_r = 0.16^\circ$ ]

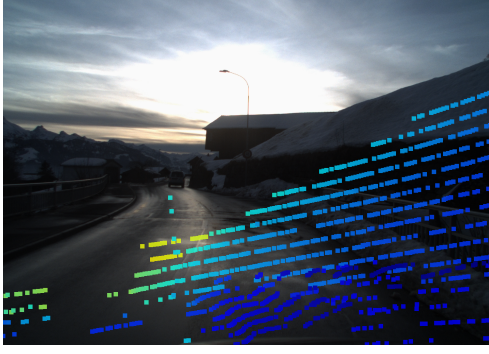


(f) Stage 5 prediction [ $e_t = 1.62\text{cm}$ ,  $e_r = 0.19^\circ$ ]



(g) Ground truth

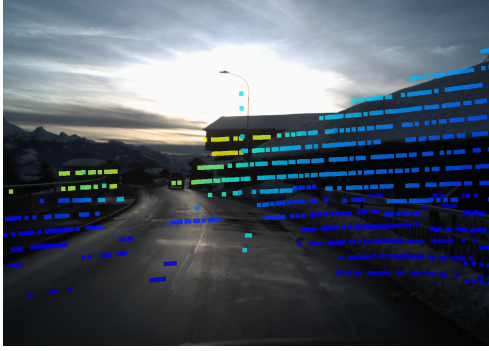
Figure 2. Per-stage visualization for LiDAR-Event calibration on KITTI. The exact translation and rotation error is provided for the perturbed input and each stage. For miscalibrated input, there are no LiDAR points inside the RGB image boundary because of the large perturbation.



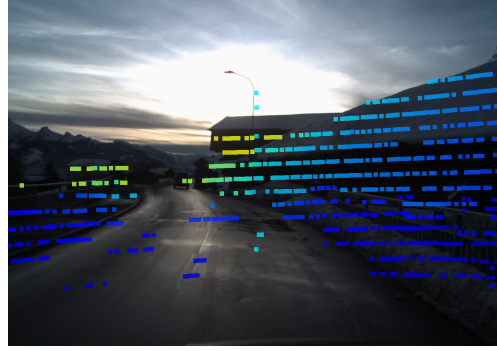
(a) Miscalibrated input [ $e_t = 207.90\text{cm}, e_r = 18.27^\circ$ ]



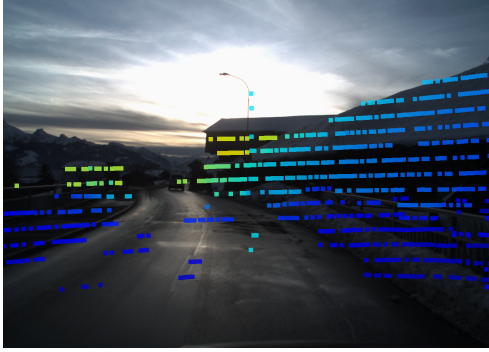
(b) Stage 1 prediction [ $e_t = 189.45\text{cm}, e_r = 3.39^\circ$ ]



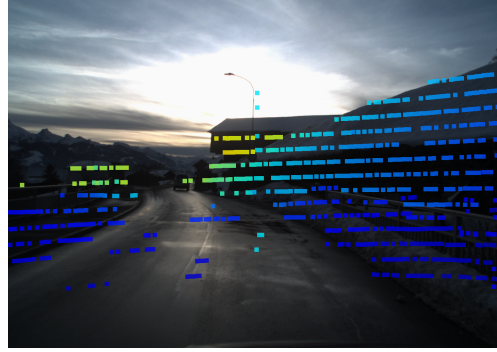
(c) Stage 2 prediction [ $e_t = 44.22\text{cm}, e_r = 1.46^\circ$ ]



(d) Stage 3 prediction [ $e_t = 2.64\text{cm}, e_r = 0.32^\circ$ ]



(e) Stage 4 prediction [ $e_t = 1.76\text{cm}, e_r = 0.13^\circ$ ]



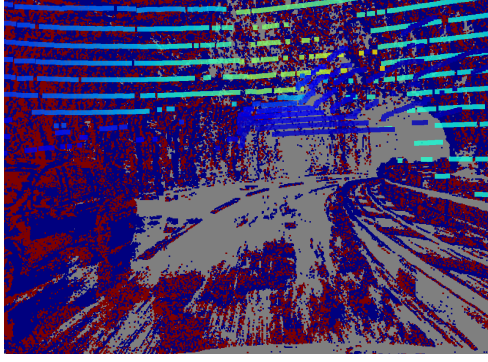
(f) Stage 5 prediction [ $e_t = 4.35\text{cm}, e_r = 0.13^\circ$ ]



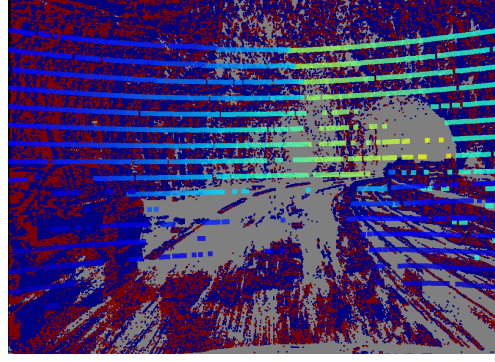
(g) Ground truth

Figure 3. Per-stage visualization for LiDAR-RGB calibration on DSEC. The exact translation and rotation error is provided for the perturbed input and each stage.

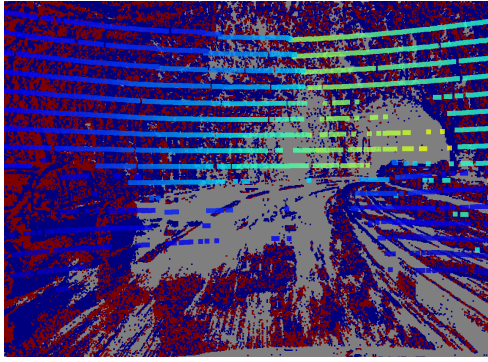




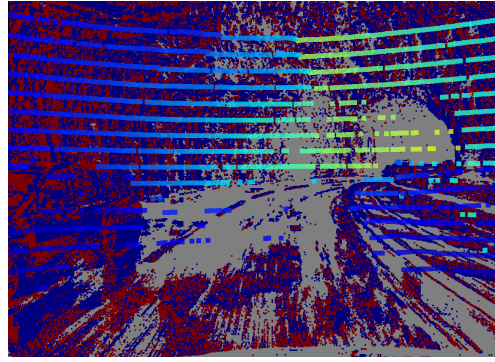
(a) Miscalibrated input [ $e_t = 161.03\text{cm}$ ,  $e_r = 10.53^\circ$ ]



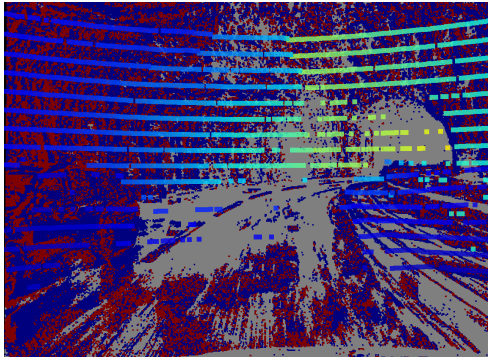
(b) Stage 1 prediction [ $e_t = 29.22\text{cm}$ ,  $e_r = 1.64^\circ$ ]



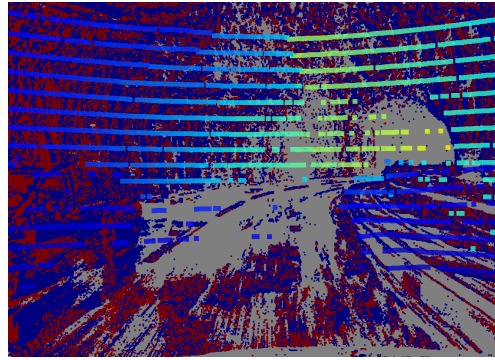
(c) Stage 2 prediction [ $e_t = 40.36\text{cm}$ ,  $e_r = 0.36^\circ$ ]



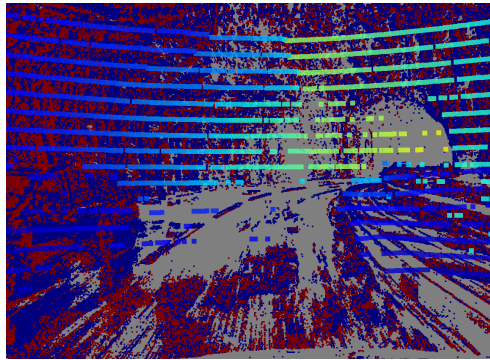
(d) Stage 3 prediction [ $e_t = 4.25\text{cm}$ ,  $e_r = 1.15^\circ$ ]



(e) Stage 4 prediction [ $e_t = 7.63\text{cm}$ ,  $e_r = 0.21^\circ$ ]



(f) Stage 5 prediction [ $e_t = 2.27\text{cm}$ ,  $e_r = 0.21^\circ$ ]



(g) Ground truth

Figure 4. Per-stage visualization for LiDAR-Event calibration on DSEC. The exact translation and rotation error is provided for the perturbed input and each stage.