A. Proof of Lemma 1

Let $\boldsymbol{\gamma}(t) = \boldsymbol{T}\exp(t\boldsymbol{\xi}^\wedge)$. We write the objective function f in (12) as a function of this perturbation in t as

$$f(\boldsymbol{\gamma}(t)) = \sum_{ij} c_{ij} k(\boldsymbol{x}_i, \exp(-t\boldsymbol{\xi}^\wedge)\boldsymbol{T}^{-1}\boldsymbol{z}_j),$$

where the sum is over $i \in 1, ..., |X|$ and $j = 1, ..., |Z|$, respectively. Taking the directional derivative, we obtain

$$\begin{aligned}
&\mathrm{D}f(\boldsymbol{\gamma}(t))[\boldsymbol{\xi}^\wedge] = \sum_{ij} c_{ij} \mathrm{D}k(\boldsymbol{T})[\boldsymbol{\xi}^\wedge] \qquad (23)\\
&= -\frac{1}{2}\sum_{ij} c_{ij} k(\boldsymbol{x}_i, \boldsymbol{T}^{-1}\boldsymbol{z}_j) \lim_{t\to 0}\frac{1}{t}\{\\
&\quad \langle \boldsymbol{x}_i - (\boldsymbol{I} - t\boldsymbol{\xi}^\wedge + o(t^2))\boldsymbol{T}^{-1}\boldsymbol{z}_j,\\
&\quad \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - (\boldsymbol{I} - t\boldsymbol{\xi}^\wedge + o(t^2))\boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle\\
&\quad - \langle \boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle\}\\
&= -\sum_{ij} c_{ij} k(\boldsymbol{x}_i, \boldsymbol{T}^{-1}\boldsymbol{z}_j)\langle \boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j, \boldsymbol{\Sigma}_{ij}^{-1}\boldsymbol{\xi}^\wedge(\boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle.
\end{aligned}$$

Using (3) and identifying the gradients for translation and rotational components from (23) yields the stated result.

B. Line search

For the Taylor expansion in the line search, we obtain expressions similar to [16, Sec. 7] except that we use anisotropic kernels, with coefficients

$$g_{ij}^1 = \beta_{ij}, \qquad (24a)$$

$$g_{ij}^2 = \gamma_{ij} + \frac{1}{2}\beta_{ij}^2, \qquad (24b)$$

$$g_{ij}^3 = \delta_{ij} + \beta_{ij}\gamma_{ij} + \frac{1}{6}\beta_{ij}^3, \qquad (24c)$$

$$g_{ij}^4 = \varepsilon_{ij} + \beta_{ij}\delta_{ij} + \frac{1}{2}\beta_{ij}^2\gamma_{ij} + \frac{1}{2}\gamma_{ij}^2 + \frac{1}{24}\beta_{ij}^4. \qquad (24d)$$

However, our expressions differ in the following parameters

$$\beta_{ij} = \langle \boldsymbol{\omega}^\wedge \boldsymbol{z}_j + \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle, \qquad (24e)$$

$$\begin{aligned}
\gamma_{ij} = {}& (-0.5)\langle \boldsymbol{\omega}^\wedge \boldsymbol{z}_j + \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{\omega}^\wedge \boldsymbol{z}_j + \boldsymbol{v})\rangle\\
& \langle (\boldsymbol{\omega}^\wedge)^2 \boldsymbol{z}_j + \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle, \qquad (24f)
\end{aligned}$$

$$\begin{aligned}
\delta_{ij} = {}& \Big(\langle -\boldsymbol{\omega}^\wedge \boldsymbol{z}_j - \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{\omega}^\wedge)^2 \boldsymbol{z}_j + \boldsymbol{\omega}^\wedge \boldsymbol{v})\rangle +\\
& \langle -(\boldsymbol{\omega}^\wedge)^3 \boldsymbol{z}_j - (\boldsymbol{\omega}^\wedge)^2 \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle\Big), \qquad (24g)
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{ij} = {}& -0.5\langle (\boldsymbol{\omega}^\wedge)^2 z_j + \boldsymbol{\omega}^\wedge \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}\Big((\boldsymbol{\omega}^\wedge)^2 z_j + \boldsymbol{\omega}^\wedge \boldsymbol{v}\Big)\rangle +\\
& 2\langle \boldsymbol{\omega}^\wedge \boldsymbol{z}_j + \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}\Big((\boldsymbol{\omega}^\wedge)^3 \boldsymbol{z}_j + (\boldsymbol{\omega}^\wedge)^2 \boldsymbol{v}\Big)\rangle +\\
& 2\langle (\boldsymbol{\omega}^\wedge)^4 \boldsymbol{z}_j + (\boldsymbol{\omega}^\wedge)^3 \boldsymbol{v}, \boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{x}_i - \boldsymbol{T}^{-1}\boldsymbol{z}_j)\rangle. \qquad (24h)
\end{aligned}$$

B.1. Loss Slices

First, we give intuition on the behavior of the RKHS loss under different perturbations of the ground truth pose. To this end, we sample 8 point clouds from our self-collected LiDAR datasets, later described in Sec. 4.2. Fig. 9 shows the value of the objective function when translating the target point cloud with a pose $\boldsymbol{T} = \exp(\boldsymbol{\xi}^\wedge)$, where $\boldsymbol{\xi} = (0, y, z, \phi, 0, 0)^\top$ and x, y, ϕ are varied independently over $[-3, 3]$ for translations and $[-0.5, 0.5]$ for rotations. Near the ground truth pose of each pair of point clouds, there is a single maximum of the objective function (12). Besides, the curvature of the loss is greater on the z axis with the kernel choice in (13), resulting in tighter fits along surfaces.

C. Derivation of the explicit form of correction term $\Gamma(\cdot, \cdot)$

Let

$$\boldsymbol{\xi} = \begin{bmatrix}\boldsymbol{\omega}\\ \boldsymbol{v}\end{bmatrix}, \qquad \boldsymbol{g} = \mathrm{grad} f(\boldsymbol{T})^\vee = \begin{bmatrix} g_\omega \\ g_v\end{bmatrix}. \qquad (25)$$

From (6), the vee representation of the connection term is

$$\Gamma^\vee(\boldsymbol{\xi}^\wedge, \mathrm{grad} f(\boldsymbol{T})) = \frac{1}{2}\left(\mathrm{ad}_{\boldsymbol{\xi}}\boldsymbol{g} - \mathrm{ad}_{\boldsymbol{\xi}}^*\boldsymbol{g} - \mathrm{ad}_{\boldsymbol{g}}^*\boldsymbol{\xi}\right). \qquad (26)$$

Using (6),

$$\mathrm{ad}_{\boldsymbol{\xi}} = \begin{bmatrix}\boldsymbol{\omega}^\wedge & 0\\ \boldsymbol{v}^\wedge & \boldsymbol{\omega}^\wedge\end{bmatrix}, \qquad \mathrm{ad}_{\boldsymbol{\xi}}^* = \begin{bmatrix}-\boldsymbol{\omega}^\wedge & -\boldsymbol{v}^\wedge\\ 0 & -\boldsymbol{\omega}^\wedge\end{bmatrix}. \qquad (27)$$

Therefore,

$$\mathrm{ad}_{\boldsymbol{\xi}}\boldsymbol{g} = \begin{bmatrix}\boldsymbol{\omega}^\wedge g_\omega\\ \boldsymbol{v}^\wedge g_\omega + \boldsymbol{\omega}^\wedge g_v\end{bmatrix}, \qquad (28)$$

$$\mathrm{ad}_{\boldsymbol{\xi}}^*\boldsymbol{g} = \begin{bmatrix}-\boldsymbol{\omega}^\wedge g_\omega - \boldsymbol{v}^\wedge g_v\\ -\boldsymbol{\omega}^\wedge g_v\end{bmatrix}, \qquad \mathrm{ad}_{\boldsymbol{g}}^*\boldsymbol{\xi} = \begin{bmatrix}-g_\omega^\wedge \boldsymbol{\omega} - g_v^\wedge \boldsymbol{v}\\ -g_\omega^\wedge \boldsymbol{v}\end{bmatrix}. \qquad (29)$$

Substituting into (26), we obtain

$$\Gamma^\vee(\boldsymbol{\xi}^\wedge, \mathrm{grad} f(\boldsymbol{T})) = \frac{1}{2}\begin{bmatrix}\boldsymbol{\omega}^\wedge g_\omega + \boldsymbol{\omega}^\wedge g_\omega + \boldsymbol{v}^\wedge g_v + g_\omega^\wedge \boldsymbol{\omega} + g_v^\wedge \boldsymbol{v}\\ \boldsymbol{v}^\wedge g_\omega + 2\boldsymbol{\omega}^\wedge g_v + g_\omega^\wedge \boldsymbol{v}\end{bmatrix}. \qquad (30)$$

Using $a^\wedge b = -b^\wedge a$, we have

$$\boldsymbol{v}^\wedge g_v + g_v^\wedge \boldsymbol{v} = 0, \ \boldsymbol{\omega}^\wedge g_\omega + g_\omega^\wedge \boldsymbol{\omega} = 0, \ \boldsymbol{v}^\wedge g_\omega + g_\omega^\wedge \boldsymbol{v} = 0. \qquad (31)$$

Hence,

$$\Gamma^\vee(\boldsymbol{\xi}^\wedge, \mathrm{grad} f(\boldsymbol{T})) = \begin{bmatrix}-\frac{1}{2}g_\omega^\wedge \boldsymbol{\omega}\\ -g_v^\wedge \boldsymbol{\omega}\end{bmatrix} = \begin{bmatrix}-\frac{1}{2}g_\omega^\wedge & 0\\ -g_v^\wedge & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\omega}\\ \boldsymbol{v}\end{bmatrix}. \qquad (32)$$

Therefore,

$$\Gamma(\boldsymbol{\xi}^\wedge, \mathrm{grad} f(\boldsymbol{T})) = \left(\begin{bmatrix}-\frac{1}{2}g_\omega^\wedge & 0\\ -g_v^\wedge & 0\end{bmatrix}\boldsymbol{\xi}\right)^\wedge. \qquad (33)$$

Table 5. *Impact of uniform vs non-uniform downsampling*

KITTI	G-CVO-2 w. voxel grid		G-CVO-2 w. random	
seq. id	trans.	rot.	trans.	rot.
03	1.5743	0.0087	4.0074	0.0172
05	1.0557	0.0060	1.6303	0.0084
06	1.1641	0.0070	1.7082	0.0065
07	1.1106	0.0065	1.0130	0.0082

D. Examples of the self-collected datasets

Fig. 10 shows the scenes and the example point cloud observations in Sec. 4.2. The first row is an example frame from the `Dirt Track` sequence, where there is dust floating behind the vehicle, illustrated in red in the LiDAR visualization. The second row is an example frame from the `Race Track` sequence, where there are fewer features to track, and the car travels at greater speeds. The third row shows the `Skid Pad` dataset, a large paved surface.

E. Qualitative tracking results

We provide more qualitative results for frame-to-frame tracking on the KITTI LiDAR dataset and ETH3D RGB-D dataset in Fig. 11 and 12, respectively. We stacked the point clouds based on the poses estimated from the frame-to-frame tracking results in KITTI `00` from Sec. 4.1 and the first 400 frames of ETH3D `table_3` from Sec. 4.3.

F. Ablation on point cloud density

As LiDAR inputs are generally not uniformly dense, we evaluate two types of down-sampling. We consider a voxel grid to achieve a more uniform density, and also evaluate random downsampling, which maintains non-uniformity. Tab. 5 shows that G-CVO achieves higher accuracy with voxel-grid downsampling than with random downsampling, likely because the estimated covariance may be noisy with uneven point densities. Nonetheless, G-CVO remains competitive even when point clouds are non-uniform.

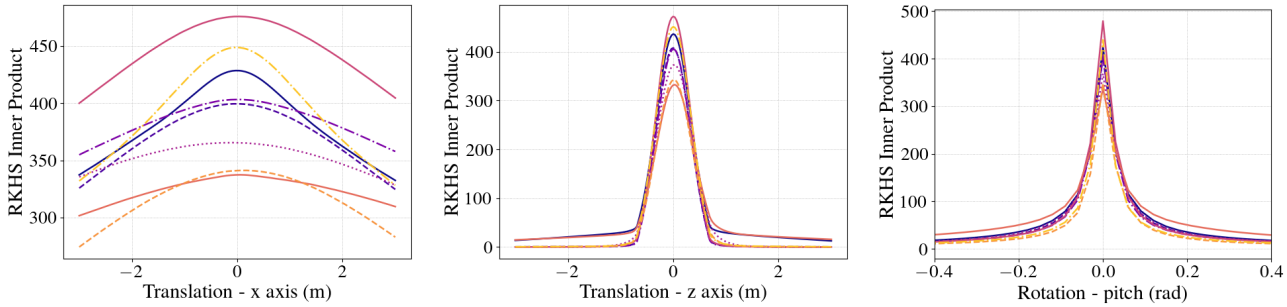


Figure 9. The RKHS inner product with translational and rotational perturbations. Each line corresponds to a different input pair.

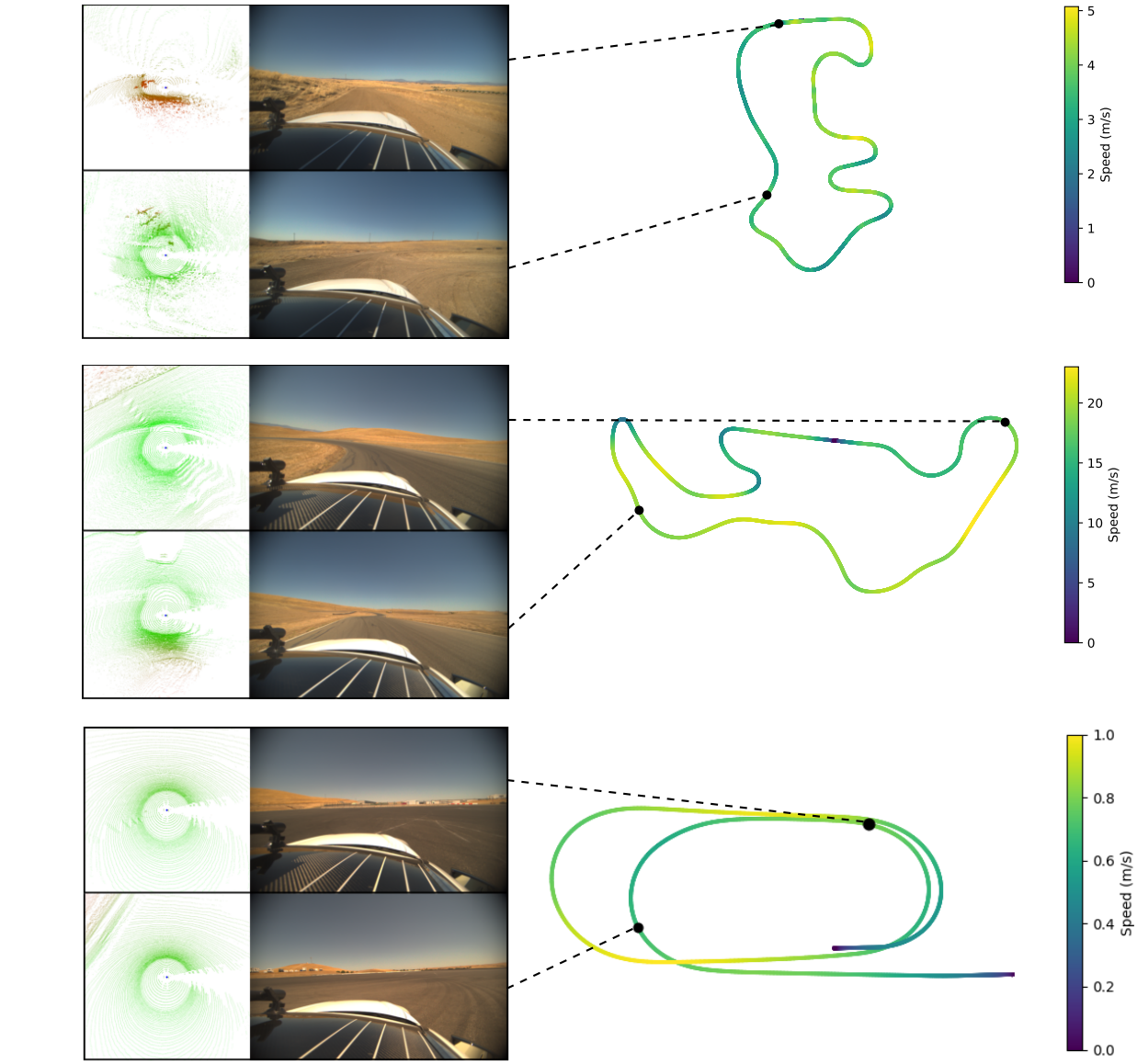
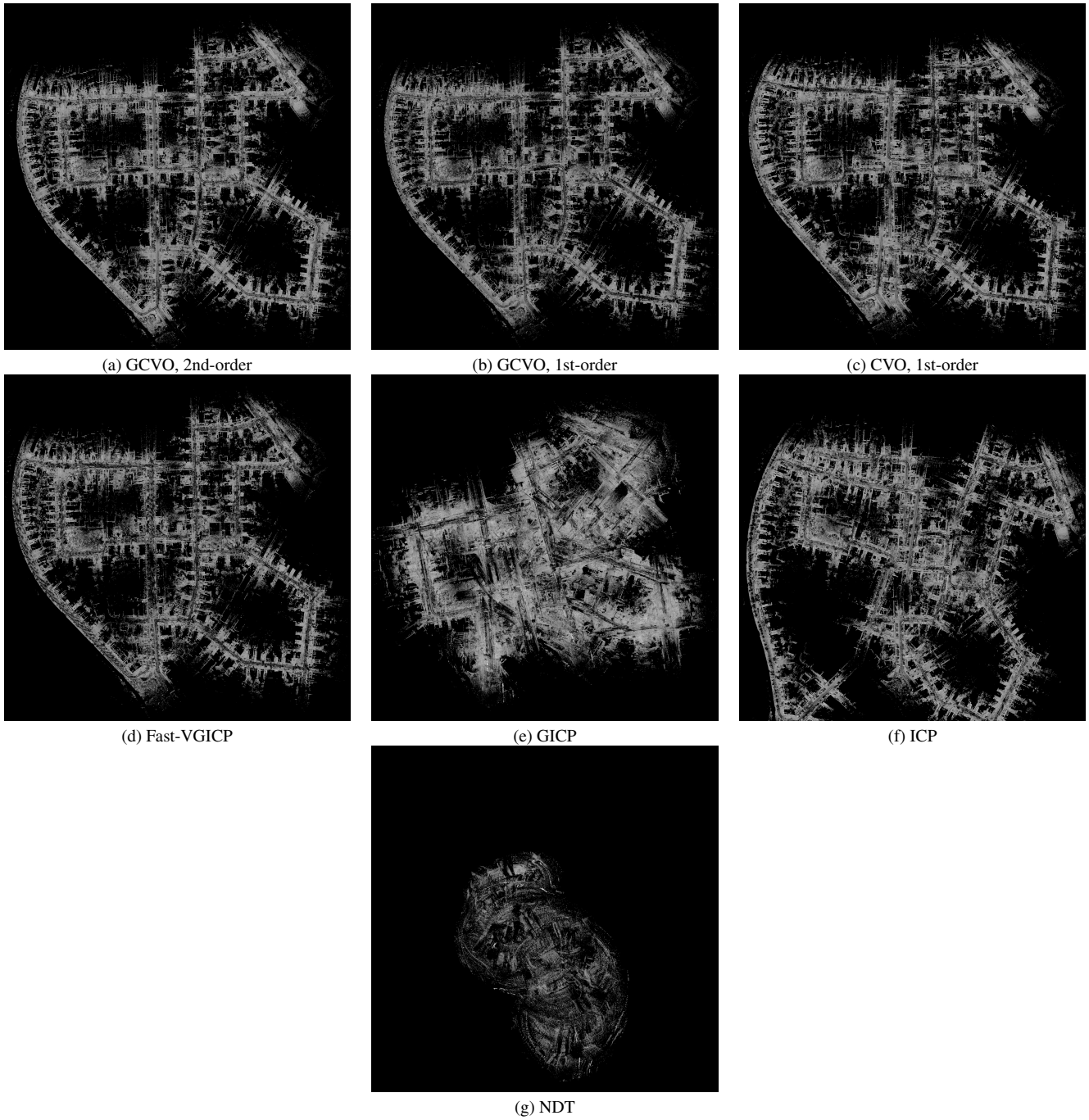


Figure 10. Illustration of the `Dirt Track` dataset (top), `Race Track` dataset (middle), and `Skid Pad` dataset (bottom), with lidar point clouds and images at two points shown at two points along the track.



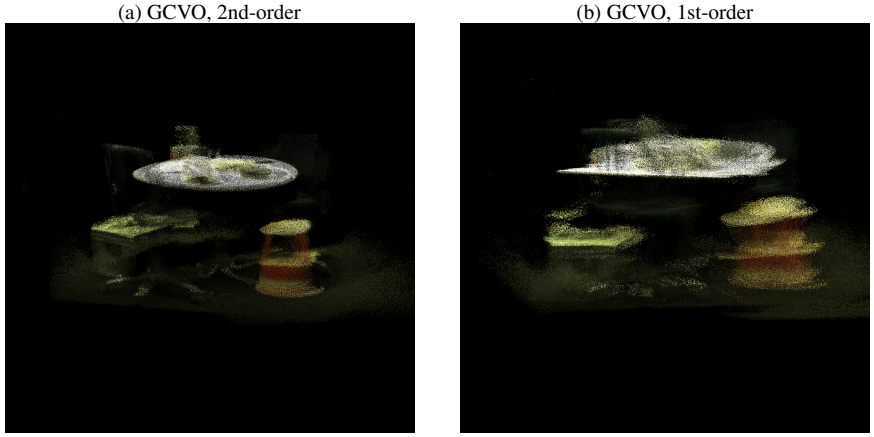
(a) GCVO, 2nd-order (b) GCVO, 1st-order (c) CVO, 1st-order

(d) Fast-VGICP (e) GICP (f) ICP

(g) NDT

Figure 11. The stacked point clouds using the frame-to-frame tracking poses from KITTI LiDAR sequence `00`.

(a) GCVO, 2nd-order (b) GCVO, 1st-order (c) Fast-VGICP



(d) GICP (e) ICP

Figure 12. The stacked point clouds using the frame-to-frame tracking poses from ETH3D RGB-D sequence `table_3`.