HyperGCT: A Dynamic Hyper-GNN-Learned Geometric Constraint for 3D Registration

Supplementary Material

6. More Method Details

GF-NMS. First, we transform the incidence matrix $\mathbf{H}$ output by HyperGCT into an adjacent matrix $\mathbf{A}$ as follows:

$$\mathbf{A} = \left[\mathbf{H} + \mathbf{H}^{\top} - 1\right]_{+}. \tag{17}$$

The motivation is to check the consistency between vertices. If e_i contains v_j and e_j contains v_i (*i.e.*, $\mathbf{H}(v_i, e_j) = \mathbf{H}(v_j, e_i) = 1$), then v_i and v_j are considered consistent, with $\mathbf{A}(v_i, v_j) = \mathbf{A}(v_j, v_i) = 1$. Second, we apply a graph filter based on the degree signal of $\mathbf{A}$ to compute scores for each correspondence:

$$\mathbf{s}^{\mathrm{GF}} = \mathrm{MinMax}((\mathbf{D}_{\mathbf{A}} - \mathbf{A})\mathbf{D}(\mathbf{A})), \tag{18}$$

where $\mathbf{D}_{\mathbf{A}}$ is the diagonal degree matrix of $\mathbf{A}$ and $\mathbf{D}(\mathbf{A})$ is the degree vector of $\mathbf{A}$. MinMax is the min-max normalization operation. Higher values of $\mathbf{s}_i^{\mathrm{GF}}$ indicate that v_i has strong local connectivity, reflecting its structural importance within the graph. Third, we employ standard NMS to select correspondences with confidence scores $\hat{\mathbf{s}}$ that are local maxima, resulting in N_1 seeds ($N_1 \ll N_s$). The remaining correspondences are then ranked by $\mathbf{s}^{\mathrm{GF}}$ in descending order, and the top $N_s - N_1$ ones are selected as the other part of the seed set. As shown in Table 13, GF-NMS generates twice as many inliers as NMS, indicating its effectiveness.

Table 13. The average number of inliers among seeds.

		NMS	GF-NMS
3DMatch	FPFH	64	**152.65**
	FCGF	232.57	**425.96**
3DLoMatch	FPFH	11.44	**28.27**
	FCGF	49.69	**96.40**

7. More Analysis.

Using Different Order Graphs. We replace the second-order graph (SOG) with the first-order graph (FOG) to construct the initial hypergraph in Sect. 3.2 and conduct experiments on 3DMatch/3DLoMatch, KITTI-10m, and KITTI-LC. The results from Tables 14, 15 and 16 show minimal performance differences between different order graphs, indicating that HyperGCT is robust and not sensitive to graph orders.

Combined with Transformer-based Methods. We conduct experiments on 3DMatch and 3DLoMatch using the correspondence generated by CoFiNet [48] and GeoTrans [29]. We apply the same RE/TE criteria and comparison methods in Sect. 4.2. Following [29], point cloud pairs from adjacent frames are excluded from the evaluation. As shown in Table 17, on the 3DMatch dataset, the quality of correspondences generated by the transformer-based methods is already quite high. Hence, the performance improvements brought by all compared methods are relatively limited. However, HyperGCT has achieved the largest improvements of all. Specifically, HyperGCT boosts CoFiNet by 1.56%, reaching a recall of 93.04%, and improves GeoTrans by 0.54%, achieving a recall of 94.99%. As indicated in Table 18, on the 3DLoMatch dataset, due to the limited quality of the matches produced by the transformer-based methods, all the compared approaches show significant performance gains. HyperGCT improves CoFiNet by 6.96%, reaching a recall of 70.63%, and increases GeoTrans by 1.50%, achieving a recall of 78.33%.

Table 14. Registration results on 3DMatch/3DLoMatch.

3DMatch / 3DLoMatch		**RR** (%)	RE (°)	TE (cm)
FCGF	w. SOG	94.45 / 63.73	2.04 / 4.18	6.34 / 10.17
	w. FOG	94.39 / 63.50	2.04 / 4.05	6.33 / 10.08
FPFH	w. SOG	85.89 / 42.28	2.26 / 4.81	6.74 / 10.80
	w. FOG	85.77 / 42.22	2.25 / 4.59	6.75 / 10.81

Table 15. Regstration results on KITTI-10m.

KITTI-10m		**RR** (%)	RE (°)	TE (cm)
FCGF	w. SOG	98.92	0.32	19.82
	w. FOG	98.92	0.32	19.83
FPFH	w. SOG	99.10	0.34	7.74
	w. FOG	99.10	0.34	7.72

Table 16. Registration results on KITTI-LC.

KITTI-LC		**RR** (%)	RE (°)	TE (cm)
0-10m	w. SOG	97.05	0.27	7.47
	w. FOG	97.05	0.27	7.45
10-20m	w. SOG	72.46	0.55	15.24
	w. FOG	72.37	0.55	15.40
20-30m	w. SOG	25.95	0.76	22.84
	w. FOG	26.43	0.75	21.70

Model Complexity. In terms of practical benefits, our method achieves consistent registration, robustness, and generalization advantages. Furthermore, as shown in Tables 1 and 2, our runtime (∼0.2s) is comparable to VBReg (CVPR 2023), and our model has fewer parameters than both PointDSC (CVPR 2021) and VBReg (Table 19). Therefore, the complexity introduced is well justified by the gain in performance and flexibility.

Inductive Biases. Our entire architecture is designed to incorporate inductive biases beyond the non-local compo-

Table 17. Registration results on 3DMatch.

	CoFiNet [48]			GeoTrans [29]		
	RR (%)	RE (°)	TE (cm)	**RR** (%)	RE (°)	TE (cm)
Origin	91.48	2.59	8.19	94.45	1.85	6.11
SC^2-PCR [5]	92.89	2.16	6.93	94.61	1.84	6.14
MAC [46]	92.73	2.22	**6.55**	94.53	1.92	**5.74**
PointDSC [2]	92.65	2.15	6.94	94.53	1.84	6.14
PG-Net [38]	92.65	2.15	6.94	94.61	1.84	6.15
VBReg [19]	92.89	2.16	6.94	94.68	1.84	6.15
Hunter [47]	91.63	2.15	6.94	94.68	1.85	6.19
3DPCP [37]	92.18	**2.09**	6.82	94.61	**1.83**	6.25
HyperGCT	**93.04**	2.15	6.94	**94.99**	1.84	6.16

Table 18. Registration results on 3DLoMatch.

	CoFiNet [48]			GeoTrans [29]		
	RR (%)	RE (°)	TE (cm)	**RR** (%)	RE (°)	TE (cm)
Origin	63.67	4.20	11.47	76.83	2.84	8.69
SC^2-PCR [5]	69.00	3.40	9.73	77.58	2.86	8.78
MAC [46]	70.51	3.54	9.82	**78.33**	3.01	8.82
PointDSC [2]	68.37	3.38	9.75	77.11	2.83	8.73
PG-Net [38]	68.95	3.39	9.74	77.52	2.86	8.80
VBReg [19]	70.34	3.40	9.79	78.04	2.83	8.77
Hunter [47]	65.82	3.36	9.70	77.35	2.87	8.90
3DPCP [37]	67.56	**3.21**	**9.43**	76.53	**2.75**	**8.61**
HyperGCT	**70.63**	3.39	9.84	**78.33**	2.82	8.74

Table 19. Parameter scale of compared learning-based methods.

Method	PointDSC	PG-Net	VBReg	Hunter	3DPCP	**Ours**
# params (M)	1.05	0.96	3.25	0.01	0.27	0.98

nent, in order to effectively learn high-order geometric constraints. Specifically: 1) We leverage SOG, constructed using dynamic compatibility thresholds, as a strong geometric prior. This helps the model adapt to varying input distributions by injecting structured information directly into the learning process. 2) The backbone of HyperGCT is a graph neural network, which inherently encodes a relational inductive bias. Its shared update functions over vertices and hyperedges allow it to generalize across graphs of different sizes and connectivity patterns. 3) Our multi-layer architecture gradually reduces the ratio of node–hyperedge associations across layers. This enables broad information propagation in shallow layers, while enforcing more reliable, selective message passing in deeper layers. We believe these combined inductive biases contribute significantly to the robustness and generalization capability of our method.

Generalization. Our method's improved generalization comes from two aspects. First, unlike conventional correspondence-based approaches that depend heavily on dataset-specific discriminative features, our method treats relationships across different data modalities as a generalized, dataset-agnostic hypergraph structure. This helps the front-end network mitigate biases tied to specific data modalities, enhancing generalization and robustness. Second, unlike traditional outlier rejection strategies that rely solely on learned discriminative features, our back-end solver leverages learned geometric constraints to explore the solution space in a more nuanced manner systematically. This approach avoids the limitations of feature-centric outlier filtering and instead captures a broader range of potential solutions, significantly enhancing hypothesis generation and verification processes.

8. Dataset Statistics

We present the details for all datasets in Table 20, including the average number and inlier ratios for FPFH matches.

Table 20. Dataset information.

Dataset	Type	Modality	# Pairs	# Matches	Inlier ratio
3DMatch [49]	Indoor	RGB-D	1623	4710	6.84%
3DLoMatch [18]	Indoor	RGB-D	1781	4653	1.68%
KITTI-10m [28]	Outdoor	LiDAR	554	8000	4.52%
KITTI-LC [28]	Outdoor	LiDAR	914 1151 1260	8000	6.25% 1.94% 0.81%
ETH [26]	Outdoor	LiDAR	713	5000	0.84%

9. Visualizations

We provide visualizations of registration results on the indoor scene 3DMatch (Fig. 4), outdoor scene KITTI (Fig. 5), and outdoor scene ETH (Fig. 6). Our approach accurately registers the data where PointDSC, Hunter, and MAC fail. This demonstrates that HyperGCT offers stronger constraint capability for generating correct transformation hypotheses.

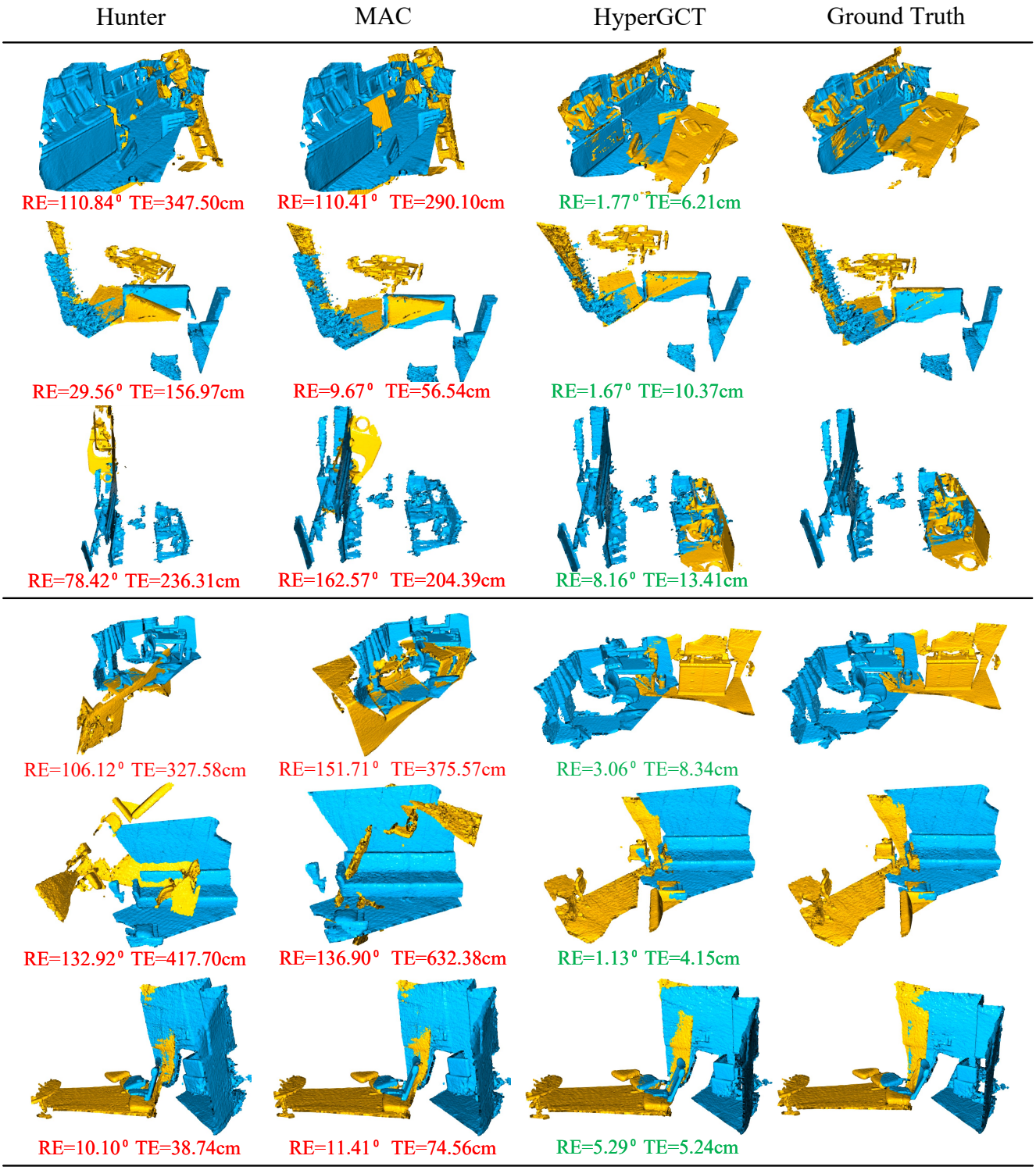

Figure 4. Visualizations of registration results. Rows 1-3 are from 3DMatch, Rows 4-6 are from 3DLoMatch.

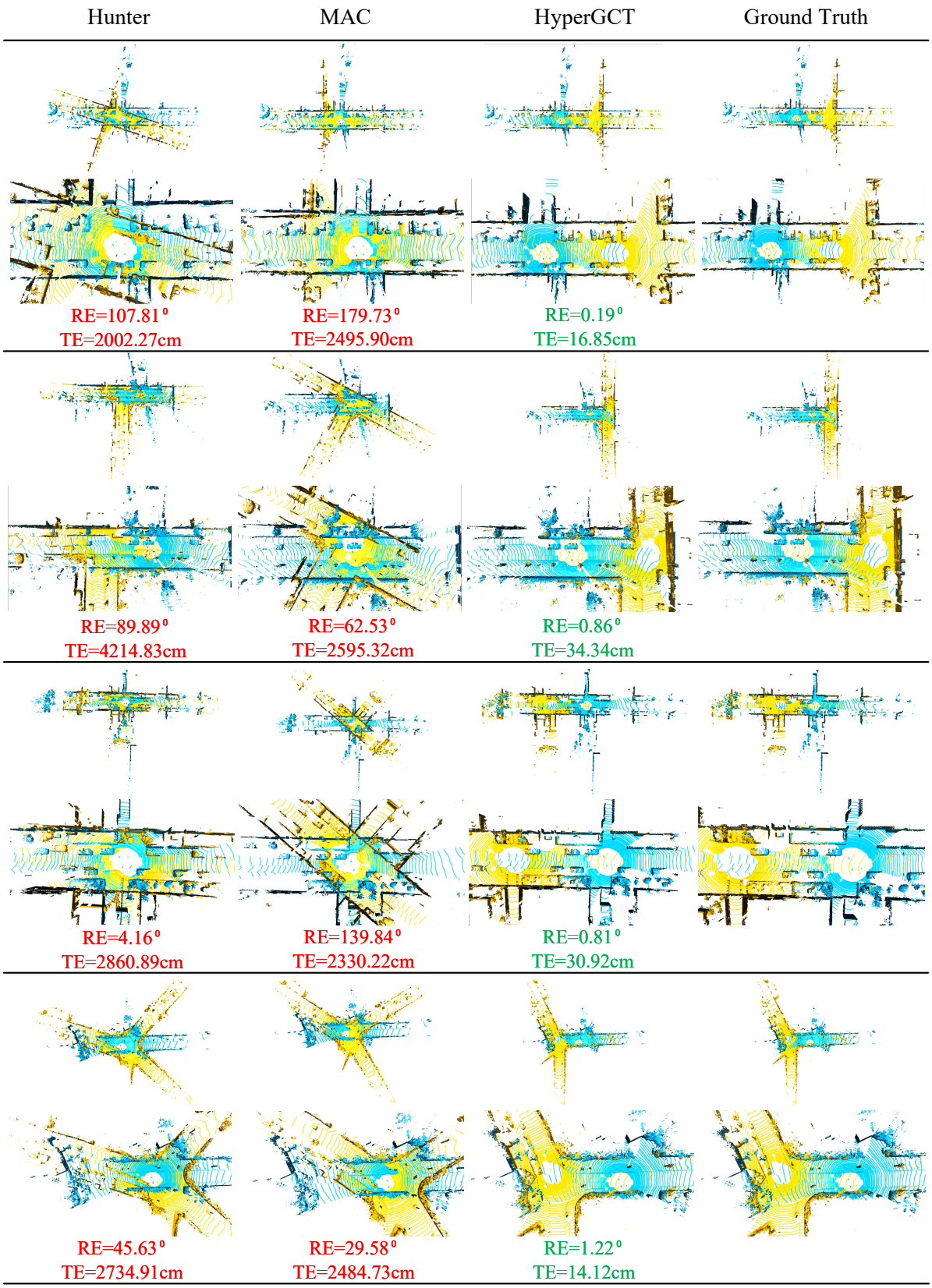

Figure 5. Visualizations of registration results on KITTI-LC. Each visualization provides both a global view (upper row) and a detailed local perspective (lower row).

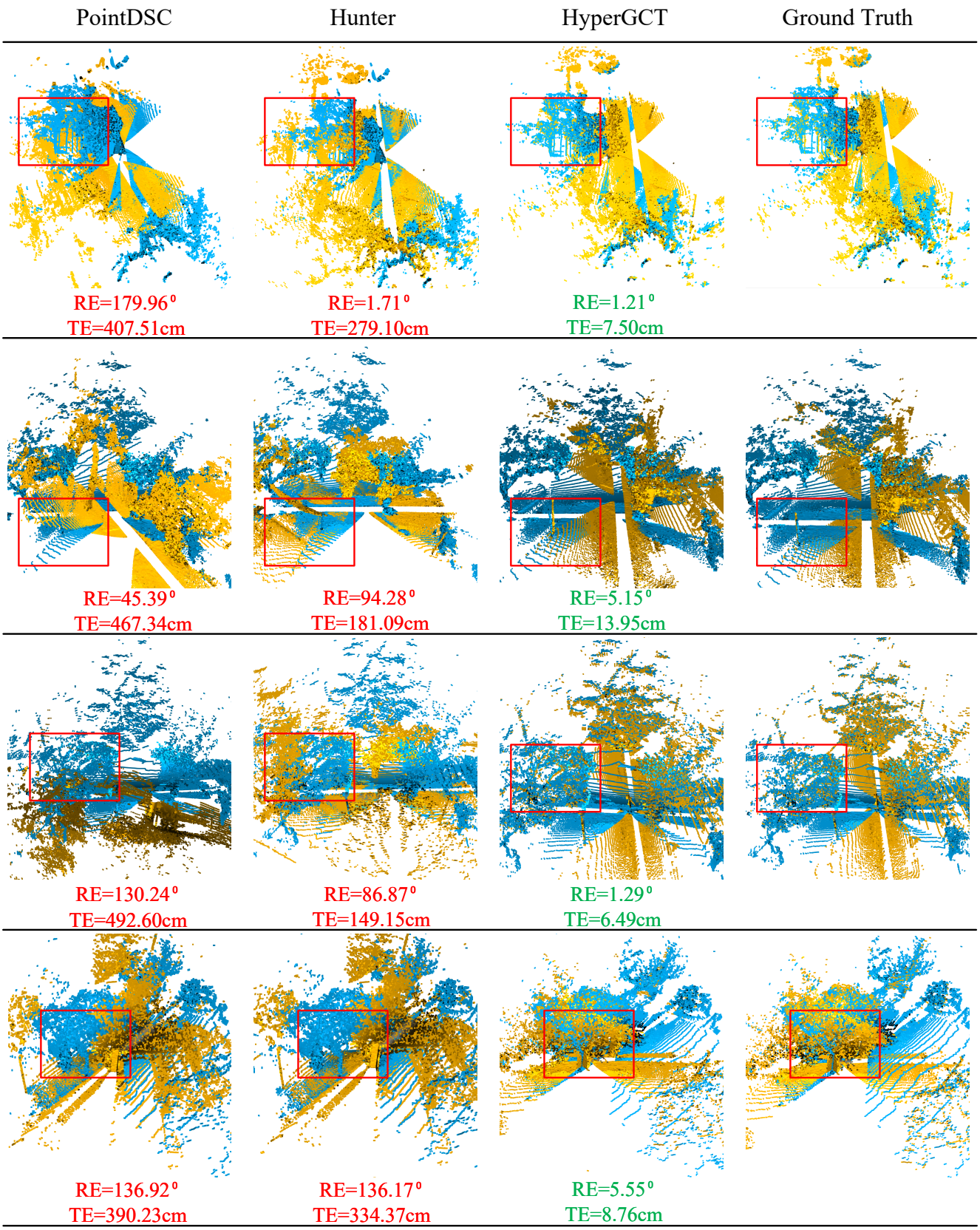

Figure 6. Visualizations of generalization results on ETH. For each set of results, we highlight the same point cloud region with a red box to compare differences across methods.