7 Supplementary Material

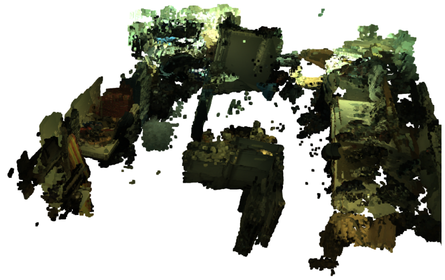
7.1 Small Motion Registration on Real-world ETH3D Odometry

This section presents additional qualitative and quantitative results using the ETH3D RGB-D dataset. Specifically, odometry tests are conducted to perform frame-to-frame registration of sequential 3D point cloud pairs across the four testing sequences, `sfm_lab_room_1`, `plant_1`, `sfm_bench`, and `table_3`. Unlike Section 5.2, there isn't a random 10° perturbation applied on each ground truth rotation, leading to an average of 1.271° initial rotation error. The registration errors are calculated by averaging the rotation and translation errors across all point cloud pairs within each sequence. For a fair comparison, color information is excluded in Geometric-CVO and *EquivAlign* by setting the label function $l_X(x) = 1$ in Eq. (2), because the baselines similarly abstain from using color.

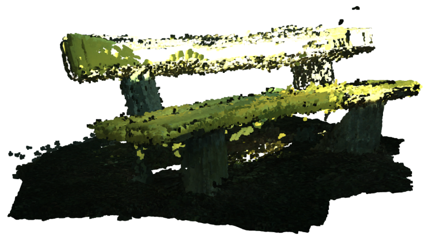
Fig. 5 shows the point cloud reconstruction results based on *EquivAlign*'s frame-to-frame registration poses for four testing sequences. Table 4 shows the sequence-wise quantitative registration errors of each SE(3) registration method. Note that one of the baselines, E2PN, is SE(3) equivariant, but the official implementation does not include translation predictions for this odometry experiment, and thus is not included in this table. Among the competing baselines, *EquivAlign* achieves lower rotation and translation errors in two of the four test sequences, while GICP has lower errors in one of the four sequences. This suggests that the proposed method registers small motions with accuracy comparable to classical fine registration techniques on noisy real-world data. Conversely, while methods like FGR and RANSAC can handle larger motions, they require additional fine registration steps for optimal resolution. Under smaller angles, classical Geometric-CVO's results closely match those of *EquivAlign*, attributed to the pivotal role of equivariant features learned during curriculum training in aligning point clouds with significant angular discrepancies, as explained in Section 5.1. Conversely, when the rotational differences are minimal, it's the coordinate aspect of the representation that primarily facilitates the final adjustments needed for the alignment process. As a result, the proposed method excels at both large and small motion registrations without the need for a coarse-to-fine two-step strategy.

Type	Method	`sfm_lab_room_1`		`plant_1`		`table_3`		`sfm_bench`	
		Rot. Error (°)	Trans. Error (m)	Rot. Error (°)	Trans. Error (m)	Rot. Error (°)	Trans. Error (m)	Rot. Error (°)	Trans. Error (m)
Non-Learning	ICP	0.519	0.031	0.519	0.011	0.451	0.011	1.851	0.031
	GICP	**0.222**	0.031	0.471	0.010	0.394	0.010	**0.300**	**0.007**
	Geometric-CVO	0.243	**0.006**	**0.372**	**0.007**	0.325	0.008	0.643	0.010
InvariantFeatures	FPFH + RANSAC	2.026	0.099	0.752	0.088	0.482	0.074	1.851	0.076
	FPFH + FGR	1.693	0.037	1.989	0.034	1.010	0.024	1.610	0.036
Equivariant Features	*EquivAlign*	0.247	**0.006**	**0.372**	**0.007**	**0.319**	**0.007**	0.581	0.011

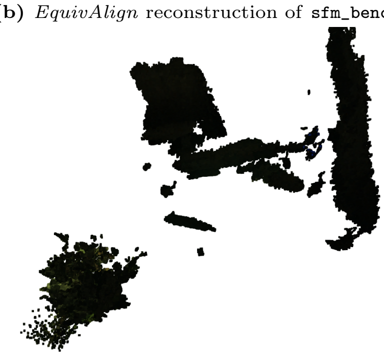
Table 4: ETH3D SE(3) Odometry Test Results per Sequence: This table shows the pairwise registration errors per ETH3D sequence. Among competing baselines, *EquivAlign* achieves lower rotation and translation errors in two of the four test sequences.



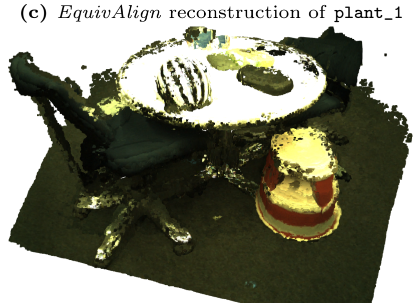
(a) *EquivAlign* reconstruction of `sfm_lab_room_1`



(b) *EquivAlign* reconstruction of `sfm_bench`



(c) *EquivAlign* reconstruction of `plant_1`



(d) *EquivAlign* reconstruction of `sfm_table_3`

Fig. 5: Reconstruction Results on the ETH3D Dataset: Scene reconstruction results achieved by applying the frame-to-frame transformations estimated by *EquivAlign*. Each reconstruction uses the first 150 frames of the corresponding sequence.