

n_imgs	ATE↓				RTA@3↑				AUC-R&T @ 3↑				RTA@1↑				AUC-R&T @ 1↑				
	MBA (Ours)	FastMap	GLOMAP	COLMAP	MBA (Ours)	FastMap	GLOMAP	COLMAP	MBA (Ours)	FastMap	GLOMAP	COLMAP	MBA (Ours)	FastMap	GLOMAP	COLMAP	MBA (Ours)	FastMap	GLOMAP	COLMAP	
mipnerf360 (9)	215.6	5.0e-4	4.2e-4	3.3e-5	5.8e-5	99.6	99.9	100.0	100.0	85.5	97.4	98.2	97.2	87.3	99.8	100.0	99.7	66.1	92.4	94.6	91.9
tnt_advanced (6)	337.8	1.7e-2	6.4e-3	1.2e-2	1.2e-3	59.9	71.4	79.1	98.5	34.7	42.6	75.3	94.8	25.2	42.3	77.5	97.0	12.5	16.7	69.8	90.0
tnt_intermediate (8)	268.6	5.7e-3	7.8e-5	1.9e-5	2.6e-4	89.8	99.9	100.0	99.8	61.7	94.1	99.0	98.9	62.8	99.3	99.9	99.5	35.4	83.1	96.9	97.3
tnt_training (7)	470.1	3.8e-3	3.0e-3	1.1e-2	3.0e-4	88.8	87.8	88.7	99.9	63.2	77.2	87.9	99.5	63.6	82.1	88.6	99.9	31.9	60.5	86.3	98.7
nerf_osr (8)	402.8	1.4e-3	1.6e-3	1.1e-3	1.3e-3	89.8	91.7	92.0	92.1	69.3	70.9	71.9	71.7	54.4	71.1	71.9	71.7	35.0	43.2	45.2	44.7
drone_deploy (9)	524.7	2.4e-3	4.9e-3	4.3e-3	2.0e-3	89.6	97.9	98.2	91.3	71.8	79.2	81.1	65.2	72.4	89.6	91.5	73.5	46.6	50.4	53.5	40.2
urban_scene (3)	3824	8.1e-5	1.7e-5	1.4e-5	1.4e-5	99.0	99.9	99.9	100.0	85.9	95.3	97.0	97.0	94.1	99.5	99.6	99.6	63.2	86.3	91.2	91.3
mill19_building	1920	5.1e-5	3.0e-4	1.3e-2	1.9e-5	99.6	99.9	0.1	99.9	93.5	95.5	0.0	95.6	95.9	99.3	0.0	99.3	81.8	87.0	0.0	87.4
mill19_rubble	1657	4.3e-5	3.6e-5	6.4e-5	3.4e-5	99.9	99.9	99.8	99.9	95.8	93.6	94.5	94.6	98.4	98.6	98.6	98.7	87.8	81.6	84.7	84.8
eyeful_apartment	3804	3.6e-3	2.8e-3	9.4e-3	2.2e-3	57.8	86.8	75.0	90.2	34.5	45.5	50.5	62.0	21.2	51.1	61.3	71.7	8.1	6.4	18.2	21.9
eyeful_kitchen	6042	9.7e-4	3.1e-3	7.4e-3	-	72.3	85.0	59.9	-	46.0	38.1	41.2	-	33.7	46.7	51.7	-	13.4	4.6	14.4	-

Table Caption: Pose accuracy comparison on MipNeRF360, Tanks and Temples, NeRF-OSR, DroneDeploy, Urbanscene3D, Mill-19, and Eyeful Tower. We follow the evaluation protocol established by FastMap for reporting pose accuracy metrics on these datasets. For datasets with multiple scenes, we denote the average number of images as dataset-name (#scenes). Results are listed separately for each scene in Mill-19 and Eyeful Tower. Metrics are color-coded with dark green for best performance and light green for competitive performance. Red denotes complete failures and gray indicates timeout (did not finish in one week).