# Block-Sparse Global Attention for Efficient Multi-View Geometry Transformers

## Supplementary Material

## A. Theoretical FLOPs Reduction

We derive the number of floating point operations (FLOPs) that is required for a forward pass of a model with block-sparse attention using a sparsity ratio of $\rho < 1$ and the orginal VGGT model with dense attention ($\rho = 1$). The ratio between these two numbers serves as an upper bound of the achievable speed-up for a given sparsity ratio. We take only the encoder and the aggregator into account for these calculations. Following common practice, we only consider FLOPs of matrix multiplications in the FFN and Attention layers, counting a multiply-add as a single instruction. We ignore operations related to patchification, embedding layers, register tokens, and task-specific output heads.

We base our calculations on a model using a per-frame transformer encoder $E$ with model dimension $d_E$ and $L_E$ transformer blocks. The aggregator is a transformer with model dimension $d_A$, employing $2L_A$ transformer blocks alternating between frame-wise and global attention. The number of FLOPs needed to apply an MLP layer to a single tokens is

$$\text{FLOP}_{\text{MLP}}(d) = d \cdot d_{\text{FF}} + d_{\text{FF}} \cdot d = 8d^2, \tag{3}$$

where $d_{\text{FF}}$ is the intermediate dimension of the MLP; for the models discussed in this paper, $d_{\text{FF}} = 4d$. The number of FLOPs for a self-attention layer consists of the FLOPs necessary for the linear projections and the computation of the query-key-value products. Block-sparse attention reduces the FLOPs required for computing the query-key-value interactions by the sparsity factor $\rho$. For $t$ tokens and model dimension $d$, this gives

$$\text{FLOP}_{\text{Attn}}(t, d, \rho) = 4td^2 + (1 - \rho)2t^2d. \tag{4}$$

The flops required for computing a single transformer block (consisting of either frame-wise or global attention and an MLP layer) are then

$$\text{FLOP}_{\text{Block}}(t, d) = \text{FLOP}_{\text{Attn}}(t, d, 0) + T \cdot \text{FLOP}_{\text{MLP}}(d). \tag{5}$$

A VGGT-like model contains $L_F$ frame-wise transformer blocks (including the encoder layers) and $L_G$ global attention blocks, such that we can compute the total number of FLOPs as

$$\text{FLOP}(N, T, \rho) = L_F N \cdot \text{FLOP}_{\text{B}}(T, d) + \text{FLOP}_{\text{B}}(NT, d, \rho). \tag{6}$$

Based on these notations, we compute the theoretical speedup as

$$\text{speedup} = \frac{\text{FLOP}(N, T, 0)}{\text{FLOP}(N, T, \rho)}. \tag{7}$$

We show the theoretical speed-up for several configurations in Tab. A-1.

| $N$ | Sparsity ratio $\rho$ 0.25 | 0.50 | 0.75 | 0.90 |
|---|---|---|---|---|
| 100 | 1.2 | 1.7 | 2.6 | 4.0 |
| 300 | 1.3 | 1.8 | 3.3 | 6.4 |
| 500 | 1.3 | 1.9 | 3.5 | 7.4 |
| 1000 | 1.3 | 1.9 | 3.7 | 8.5 |

Table A-1. **Theoretical end-to-end speed-up** of VGGT using block-sparse global attention *vs*. dense global attention, at different number of frames $N$ and sparsity ratios $\rho$. Assuming a resolution of $392 \times 518$, corresponding to $T = 28 \times 37 = 1036$ tokens/frame.

## B. Long-Sequence Experiments

To validate that our method scales to longer input sequences, we evaluate the performance of block-sparse global attention models on complete sequences from the ScanNet [7] test set. In order to keep compute demands reasonable, we evaluate on every second test set sequence, totalling 50 sequences. Instead of taking the first $N$ frames with a fixed temporal stride (as in the main text), here we sample $N$ frames evenly spaced over the whole trajectory. This sampling scheme increases the task difficulty considerably: not only is the number of input frames much larger, the trajectories are also significantly longer and more complex. We show results for 100, 300, 500, and 1000 input frames in Fig. A-1. For 100 input frames, VGGT and MapAnything reach an ATE of 0.18 and 0.24, respectively. With increasing input sequence length, their task performance drops slightly. $\pi^3$ shows the best results, achieving an ATE of 0.15 regardless of the number of input frames.

Our sparse variants keep similar task performance up to 60% effective sparsity, resulting in a 1.5× end-to-end speed-up at 100 input frames and 2× at more than 300 frames. At around 75% effective sparsity rate, the task performance starts to degrade, but is still comparable to or better than other state-of-the-art models.

## C. Ablations

We present the results for two ablations of our method. In the first ablation, we evaluate whether it is necessary to distinguish between patch tokens and special tokens. In the second ablation, we investigate whether training linear projections on top of the pooled queries and keys improves the robustness of the method to sparsity.

### C.1. Treating the special tokens special

In the main text, we distinguish between the special tokens, camera embedding and register tokens, and the patch tokens. We apply the block-sparse attention only on the patch-to-patch attention, while we compute the special-to-patch, patch-to-special, and the special-to-special attention as usual, *i.e*. dense. In this experiment, we compare this approach with an implementation that does not distinguish between the camera embeddings, register tokens, and patch tokens. The results are shown in Fig. A-2. At low sparsity levels, the chance of skipping computations concerning special tokens is small, since there are far more patch tokens than special tokens, and both methods perform similarly. At high sparsity levels, however, our strategy of always keeping all interactions con-

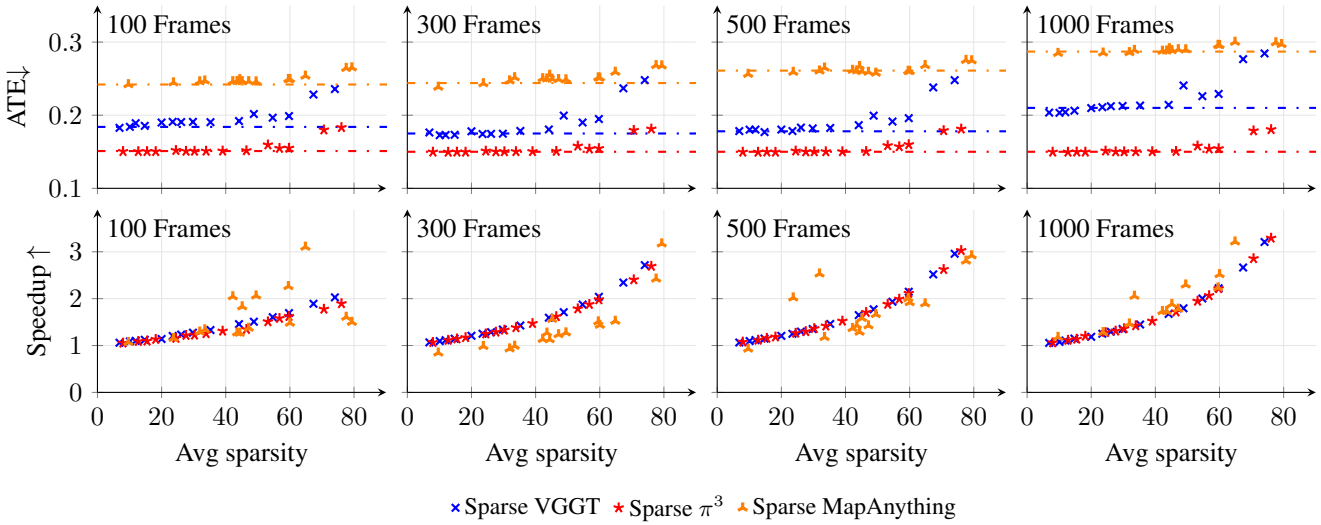


Figure A-1. **Pose estimation and speed-up for long sequences.** Evaluated on a 50-sequence subset of the ScanNet [7] validation set; frames are evenly spaced along the entire sequence. This benchmark is much harder than the experiments in the main text, since the trajectories are longer and more complex. Our method achieves a 2× speed-up at 60% effective sparsity with little to no drop in task performance compared to the original model. On longer sequences, end-to-end inference time is reduced up to 3×. Inaccurate time measurements for MapAnything are due to extensive pre- and post-processing applied by the model.

cerning the camera and register tokens significantly reduces the performance degradation compared to the naive strategy.

## C.2. Learning additional linear projections

We train additional linear projection layers on top of the pooled key and query representations following SeerAttention [15]. At training time, the linear projections are optimized to predict the entries of the downsampled attention map; at inference time, the predicted low-res attention map is converted into a binary block mask, as in the main text. With this experiment, we investigate whether the linear projections improve the robustness of the model for higher sparsity ratios compared to the default key and query representations.

Our training setup follows SeerAttention [15], that is, one linear projection per attention head. The training objective is a self-supervised KL loss between the downsampled ground-truth attention matrix and the logits of the predicted low-res mask. We train on BlendedMVS [47] because it contains real indoor and outdoor scenes and is similar to VGGT's training data. Since no gradients need to be passed through the actual model, training is both fast and lightweight. We train the projection layers for 3k steps with a batch size of 16, and a sequence length of 8 frames at resolutions sampled from $518^2$ and $518 \times 378$, using the AdamW optimizer with a learning rate of $10^{-3}$ and weight decay of 0.01. The block size is set to 128 for the queries and 64 for the keys. Training finishes in around three hours on a single H100.

The results shown in Fig. A-3 show little to no improvements over the training-free baseline.

## C.3. Full Layer Skip Ablations

We further investigate the impact of skipping global attention layers in the style of Fig. 5 in the main text. Results for VGGT are shown in Fig. A-4, and for MapAnything in Fig. A-5. We do omit $\pi^3$ because it behaves very similar to VGGT.

For VGGT, the experiments indicate a strong sensitivity of the model to perturbations in the middle layers, regardless of the task. On Co3Dv2, for example, removing just four (out of 24) global attention layers in the middle of the aggregator reduces VGGT's task performance from state-of-the-art levels to zero AUC@30. Removing four of the first or last global attention layers, in contrast, retains most of the task performance on ScanNet, ETH3D, and NRGBD. On Co3Dv2, we observe a moderate performance drop when removing layers towards the model output, but a much smaller drop when removing layers early in the aggregator.

In contrast to VGGT, MapAnything shows much more robust performance when we remove layers in the middle of the aggregator, and task performance seems to rely much more on the last few global attention layers. Removing the first nine (out of twelve) global attention layers retains the same task performance as removing the last layer. We hypothesize that this behaviour stems from the MapAnything's architecture, which uses the image patch tokens to predict camera poses instead of special camera tokens like VGGT.

For both models, the performance drops due to layer skipping are much more severe than the drop that we observe when applying our block-sparse attention scheme.

# D. Additional Visualizations

## D.1. Additional Qualitative Results

We provide further visualizations of reconstructed point clouds and estimated trajectories for all three tested models in Fig. A-6.

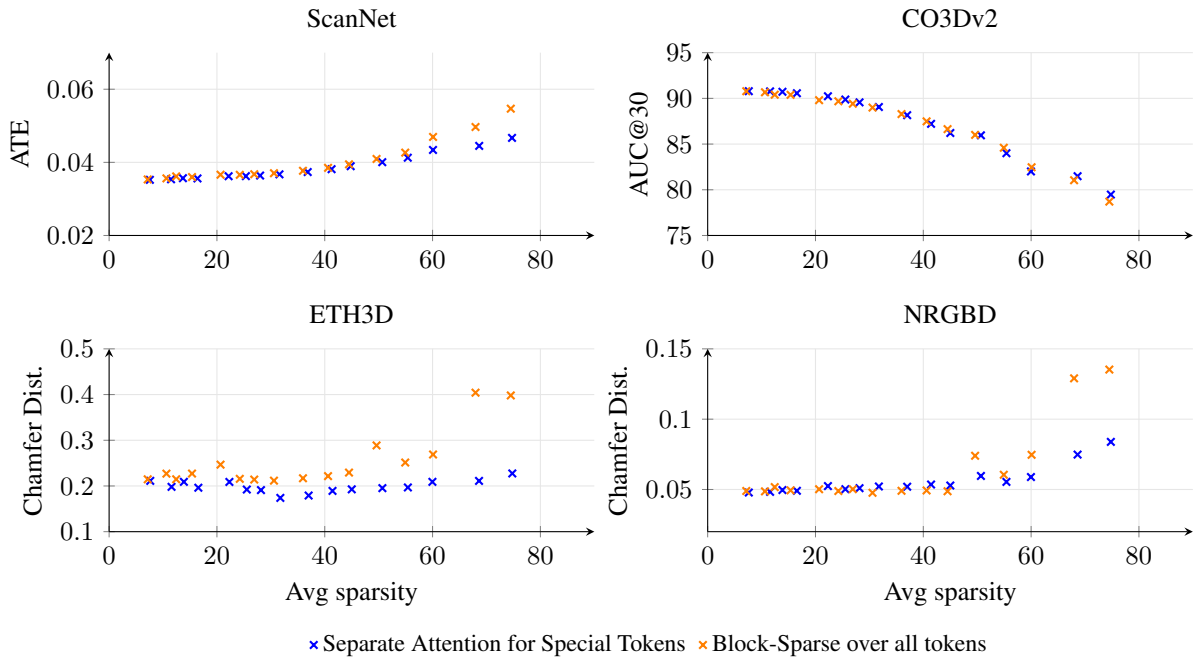


Figure A-2. **Ablation on special treatment of special tokens.** We compare the downstream performance of simple block-sparse attention over all tokens against our approach of separating patch tokens from camera and register tokens. For higher sparsity ratios, the chance of dropping special tokens rises, and at the same time, the task performance drops significantly more sharply for the simplified variant.

### D.2. Attention map statistics

We plot the average and maximum value of the global attention matrix for layers in the aggregator of VGGT, $\pi^3$, and MapAnything in Fig. A-8, in the same style as Fig. 4 in the main text. While $\pi^3$ uses less global attention layers in the aggregator than VGGT, the statistics of the remaining layers closely resemble those of the corresponding layers in VGGT. We hypothesize that the reason for this remarkable similarity is the fact that $\pi^3$ was finetuned from a VGGT checkpoint [44]. MapAnything shows consistently high maximum activations in the patch-to-patch attention in all layers. Compared to VGGT and $\pi^3$, the peak maximum activation is lower. The average special-to-special attention score is much higher than for the other models, and we hypothesize that this is due to the fact that MapAnything only employs a single special token (the metric scale token).

### D.3. High-res Attention Maps

We provide further visualizations of VGGT's attention maps in Fig. A-9. Note that all attention map visualizations in this paper are done with inputs rescaled to at most 224px for better visibility of token activations.

### D.4. Further Correspondence Visualizations

We provide additional qualitative results for correspondence estimation in Fig. A-10, demonstrating how VGGT and $\pi^3$ establish matches even in challenging scenarios with repeated structures and significant viewpoint changes.

## E. Full Results Tables

For completeness, we provide the full results for all models in Fig. 7 and Fig. 9 in the tables on the following pages.

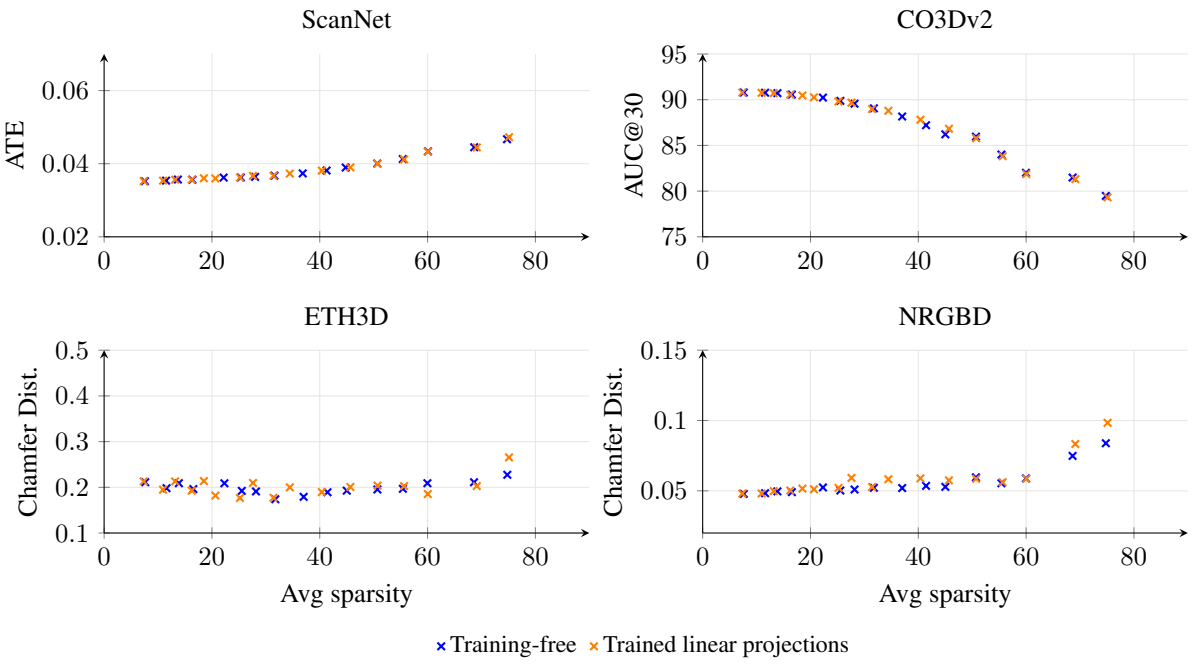


Figure A-3. **Ablation on learned linear projections.** We compare the downstream performance of training-free block-sparse mask prediction with an approach following SeerAttention [15], that is, training linear projections on top of pooled query and key features. The model with additional trained linear layers does not show significantly improved performance or robustness against increased sparsity compared to the training-free variant.

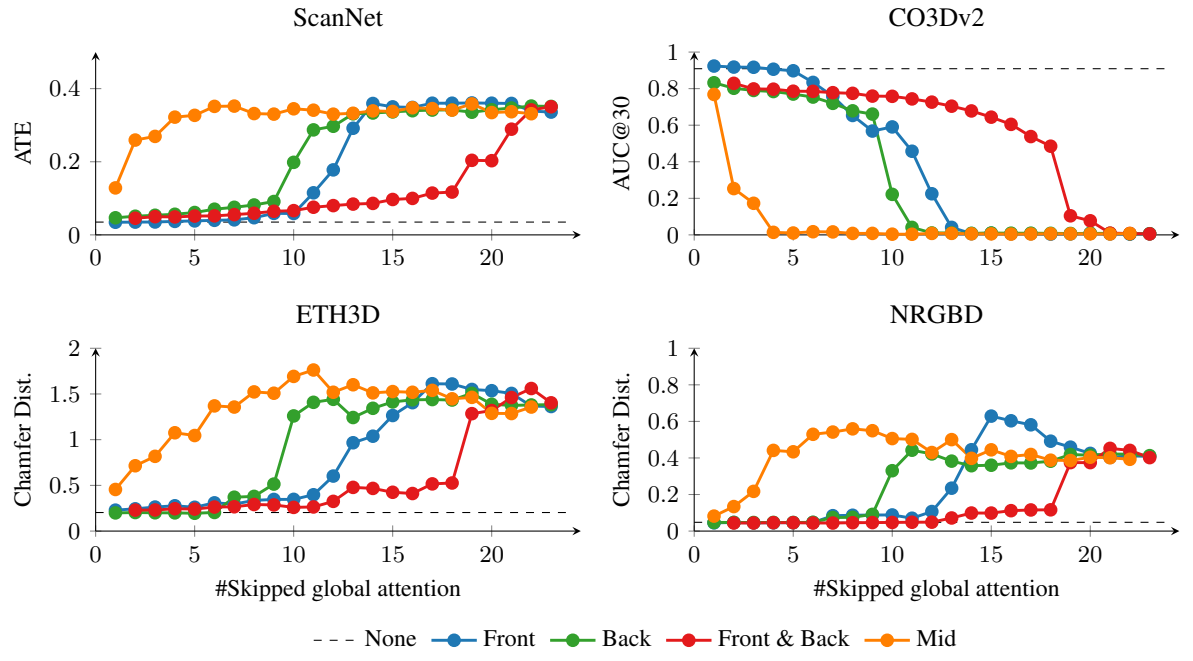


Figure A-4. **VGGT is sensitive to pruning of the middle aggregator layers.** We skip the computation of different global attention layers in the aggregator starting with the earliest (Front), last (Back), alternating (Front & Back), or from the middle layers (Middle), and evaluate the performance drop on different tasks. The x-axis denotes the total number of skipped layers. The experiment shows that VGGT is especially sensitive to pruning of the center layers, and robust against pruning the early and late layers. MapAnything, in contrast, is much more sensitive to alterations in the last layers.

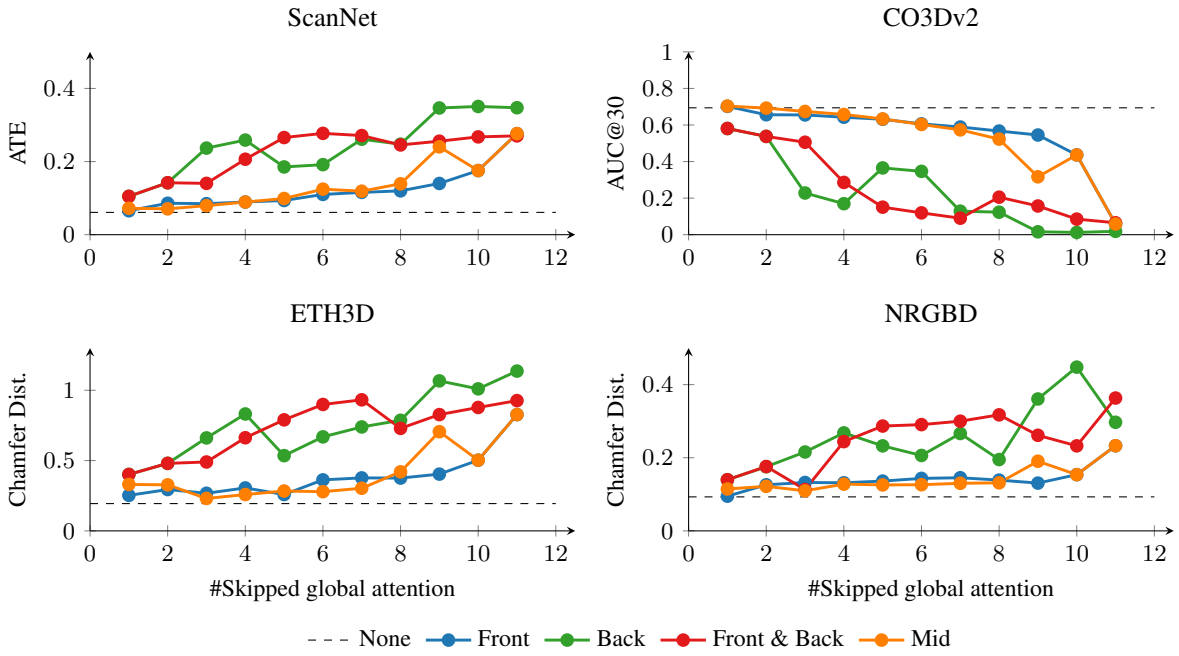


Figure A-5. **MapAnything is sensitive to pruning of the later aggregator layers.** We skip the computation of different global attention layers in the aggregator starting with the earliest (Front), last (Back), alternating (Front & Back), or from the middle layers (Middle), and evaluate the performance drop on different tasks. The x-axis denotes the total number of skipped layers. MapAnything, in contrast to VGGT, is sensitive to alterations in the last aggregator layers.


Figure A-6. **Qualitative point map estimation results.** We show the reconstructed pointcloud of the *terrace* scene of the ETH3D [33] dataset. Increasing sparsity leads to small perturbations in the reconstructed pointclouds, but the overall structure of the reconstructed scene stays consistent.

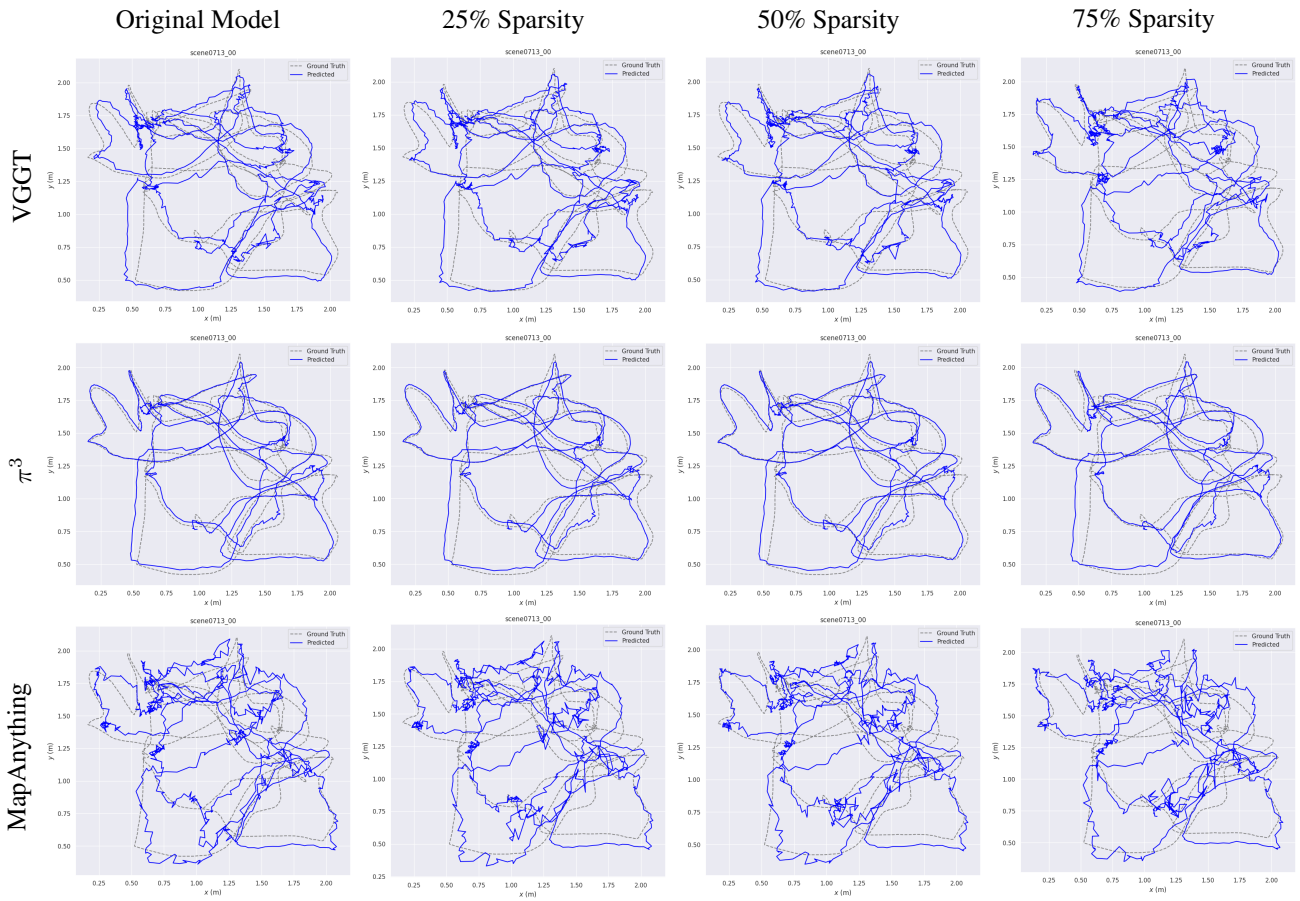


Figure A-7. **Qualitative long-sequence pose estimation results.** We show the estimated camera trajectory of a 1000-frame sequence of the ScanNet [7] validation set. Even with high sparsity, the estimated camera poses are similar to the poses estimated by the original model.

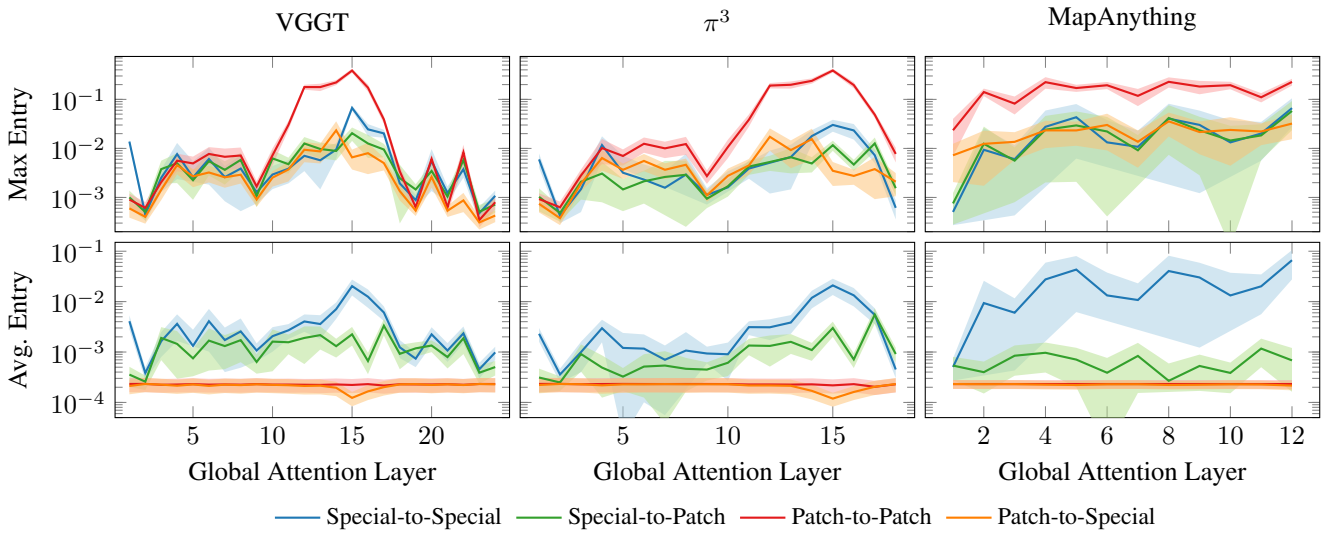


Figure A-8. **Average & maximum attention scores in global attention maps.** Note that the models use different numbers of global attention blocks in the aggregator. Since $\pi^3$ started training from a VGGT checkpoint, the statistics are highly similar to VGGT. MapAnything uses only a single special token, which explains the comparatively high average special-to-special attention score.

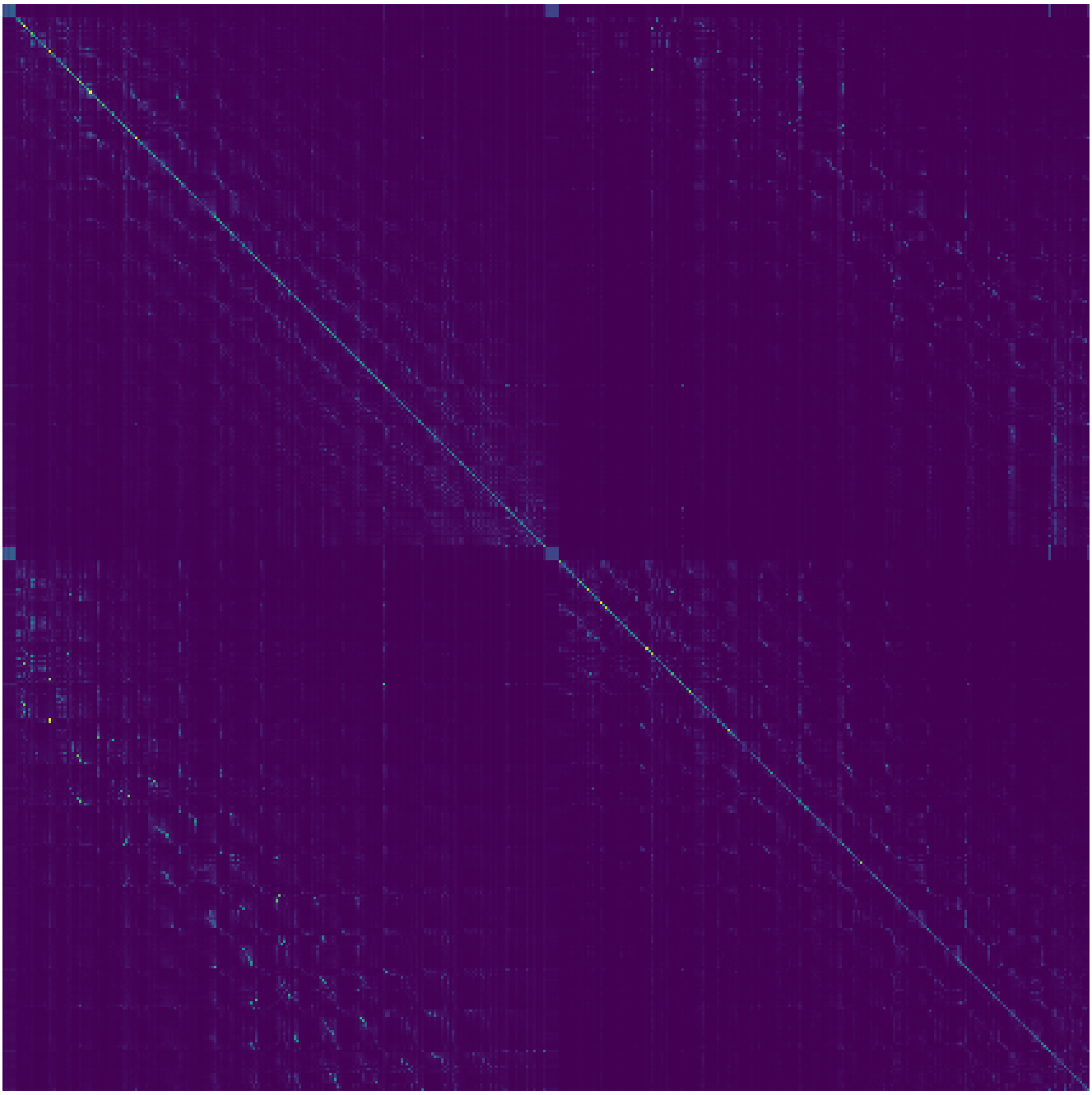

Figure A-9. Larger visualization of the global attention matrix of aggregator layer 15 of VGGT. We show the average over all heads after the softmax.

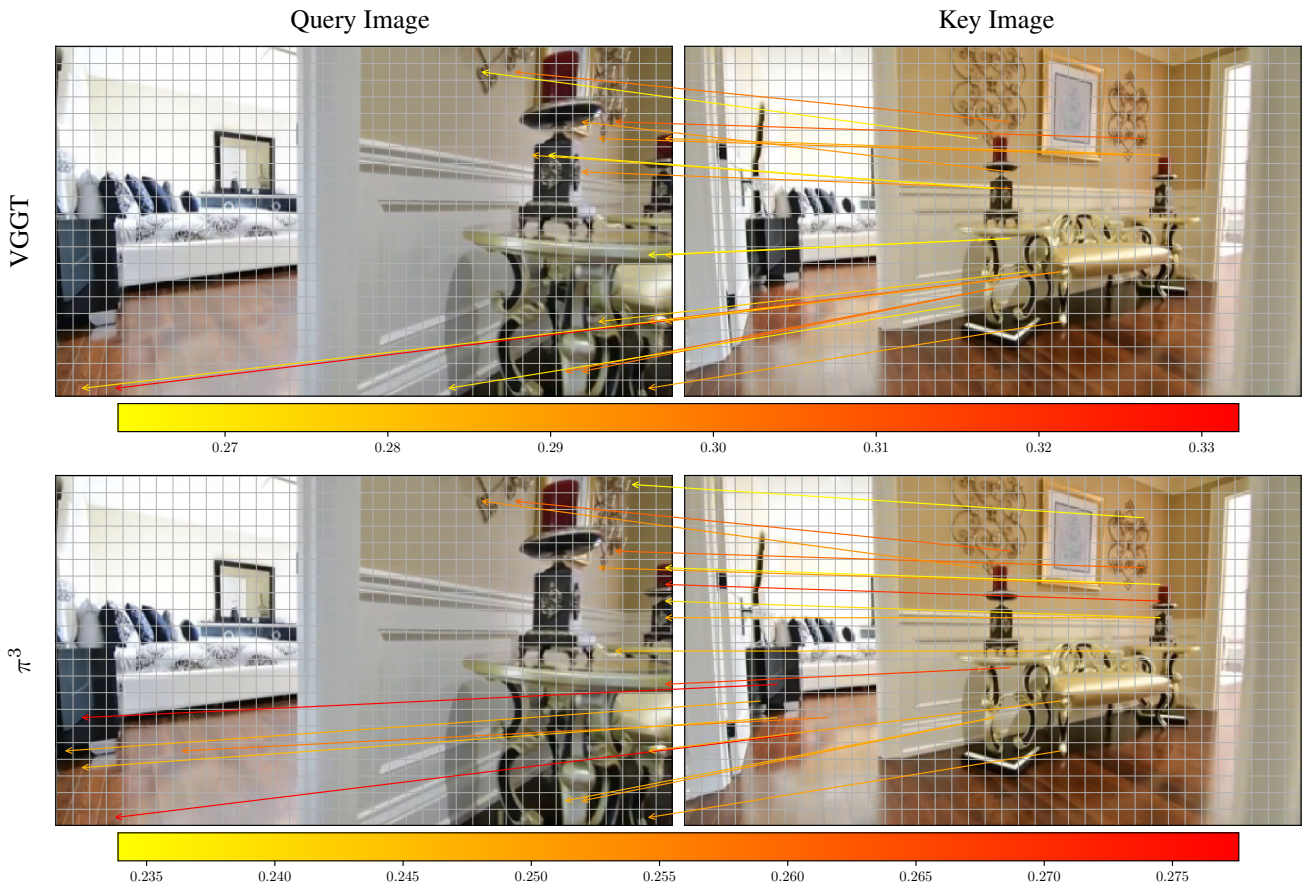


Figure A-10. Visualization of layer-15 correspondences for VGGT (top) and $\pi^3$ (bottom). Each row depicts a query image (left) and a key image (right), with arrows indicating matched points across views, where warmer colors indicates a higher attention value for the matches. Two visually identical Candles are present in both images. In both methods, the correspondences associated with one Candle in the key image are consistently mapped to the same Candle in the query image, rather than being confused with the second instance. This demonstrates the ability of the models to resolve ambiguities in scenes containing repeated objects and to maintain consist correspondences across viewpoint changes.

## References


[1] Dejan Azinović, Ricardo Martin-Brualla, Dan B Goldman, Matthias Nießner, and Justus Thies. Neural rgb-d surface reconstruction. In *CVPR*, 2022. 6, 7, 15

[2] Iz Beltagy, Matthew E Peters, and Arman Cohan. Longformer: The long-document transformer. *arXiv preprint arXiv:2004.05150*, 2020. 5

[3] Yohann Cabon, Lucas Stoffl, Leonid Antsfeld, Gabriela Csurka, Boris Chidlovskii, Jerome Revaud, and Vincent Leroy. Must3r: Multi-view network for stereo 3d reconstruction. In *CVPR*, 2025. 2

[4] Beidi Chen, Tri Dao, Eric Winsor, Zhao Song, Atri Rudra, and Christopher Ré. Scatterbrain: Unifying sparse and low-rank attention approximation. *NeurIPS*, 34:17413–17426, 2021. 3

[5] Beidi Chen, Tri Dao, Kaizhao Liang, Jiaming Yang, Zhao Song, Atri Rudra, and Christopher Re. Pixelated butterfly: Simple and efficient sparse training for neural network models. In *ICLR*, 2022. 3

[6] Rewon Child, Scott Gray, Alec Radford, and Ilya Sutskever. Generating long sequences with sparse transformers. *arXiv preprint arXiv:1904.10509*, 2019. 5

[7] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5828–5839, 2017. 6, 7, 8, 1, 2, 17, 22, 23, 24, 25

[8] Tri Dao. Flashattention-2: Faster attention with better parallelism and work partitioning. *arXiv preprint arXiv:2307.08691*, 2023. 1, 4, 8

[9] Tri Dao, Dan Fu, Stefano Ermon, Atri Rudra, and Christopher Ré. Flashattention: Fast and memory-efficient exact attention with io-awareness. *NeurIPS*, 2022. 3, 4, 5, 8

[10] Timothée Darcet, Maxime Oquab, Julien Mairal, and Piotr Bojanowski. Vision transformers need registers, 2023. 3, 6

[11] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, G Heigold, S Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2020. 3

[12] Bardienus Duisterhof, Lojze Zust, Philippe Weinzaepfel, Vincent Leroy, Yohann Cabon, and Jerome Revaud. Mast3r-sfm: a fully-integrated solution for unconstrained structure-from-motion. *arXiv preprint arXiv:2409.19152*, 2024. 2

[13] Sven Elflein, Qunjie Zhou, and Laura Leal-Taixé. Light3r-sfm: Towards feed-forward structure-from-motion. In *CVPR*, 2025. 2

[14] Elias Frantar and Dan Alistarh. Sparsegpt: Massive language models can be accurately pruned in one-shot. In *ICML*, pages 10323–10337. PMLR, 2023. 3

[15] Yizhao Gao, Zhichen Zeng, Dayou Du, Shijie Cao, Peiyuan Zhou, Jiaxing Qi, Junjie Lai, Hayden Kwok-Hay So, Ting Cao, Fan Yang, et al. Seerattention: Learning intrinsic sparse attention in your llms. *arXiv preprint arXiv:2410.13276*, 2024. 2, 3, 8, 4

[16] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 2

[17] Ali Hassani, Steven Walton, Jiachen Li, Shen Li, and Humphrey Shi. Neighborhood attention transformer. In *CVPR*, pages 6185–6194, 2023. 3

[18] Wonbong Jang, Philippe Weinzaepfel, Vincent Leroy, Lourdes Agapito, and Jerome Revaud. Pow3r: Empowering unconstrained 3d reconstruction with camera and scene priors. In *CVPR*, 2025. 2

[19] Rasmus Jensen, Anders Dahl, George Vogiatzis, Engil Tola, and Henrik Aanæs. Large scale multi-view stereopsis evaluation. In *CVPR*, 2014. 6, 7, 13

[20] Huiqiang Jiang, Yucheng Li, Chengruidong Zhang, Qianhui Wu, Xufang Luo, Surin Ahn, Zhenhua Han, Amir H Abdi, Dongsheng Li, Chin-Yew Lin, et al. Minference 1.0: Accelerating pre-filling for long-context llms via dynamic sparse attention. *NeurIPS*, 37:52481–52515, 2024. 3

[21] Nikhil Keetha, Norman Müller, Johannes Schönberger, Lorenzo Porzi, Yuchen Zhang, Tobias Fischer, Arno Knapitsch, Duncan Zauss, Ethan Weber, Nelson Antunes, et al. Mapanything: Universal feed-forward metric 3d reconstruction. *arXiv preprint arXiv:2509.13414*, 2025. 2, 3, 6, 7

[22] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM TOG*, 2017. 6, 8, 19, 20, 21

[23] Xunhao Lai, Jianqiao Lu, Yao Luo, Yiyuan Ma, and Xun Zhou. Flexprefill: A context-aware sparse attention mechanism for efficient long-sequence inference. *arXiv preprint arXiv:2502.20766*, 2025. 3

[24] Vincent Leroy, Yohann Cabon, and Jérôme Revaud. Grounding image matching in 3d with mast3r. In *ECCV*, 2024. 1, 2

[25] David G Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004. 3

[26] Enzhe Lu, Zhejun Jiang, Jingyuan Liu, Yulun Du, Tao Jiang, Chao Hong, Shaowei Liu, Weiran He, Enming Yuan, Yuzhi Wang, et al. Moba: Mixture of block attention for long-context llms. *arXiv preprint arXiv:2502.13189*, 2025. 3

[27] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023. 3

[28] Linfei Pan, Dániel Baráth, Marc Pollefeys, and Johannes L Schönberger. Global structure-from-motion revisited. In *ECCV*, 2024. 1, 2

[29] René Ranftl, Alexey Bochkovskiy, and Vladlen Koltun. Vision transformers for dense prediction. In *CVPR*, 2021. 2, 3

[30] Jeremy Reizenstein, Roman Shapovalov, Philipp Henzler, Luca Sbordone, Patrick Labatut, and David Novotny. Common objects in 3d: Large-scale learning and evaluation of real-life 3d category reconstruction. In *ICCV*, 2021. 6, 7, 12

[31] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *CVPR*, 2016. 1, 2, 3

[32] Johannes Lutz Schönberger, Enliang Zheng, Marc Pollefeys, and Jan-Michael Frahm. Pixelwise view selection for unstructured multi-view stereo. In *ECCV*, 2016. 2, 3

[33] Thomas Schöps, Johannes L. Schönberger, Silvano Galliani, Torsten Sattler, Konrad Schindler, Marc Pollefeys, and Andreas Geiger. A multi-view stereo benchmark with high-resolution images and multi-camera videos. In *CVPR*, 2017. 6, 7, 5, 14

[34] Jay Shah, Ganesh Bikshandi, Ying Zhang, Vijay Thakkar, Pradeep Ramani, and Tri Dao. Flashattention-3: Fast and accurate attention with asynchrony and low-precision. *NeurIPS*, 2024. 8

[35] Jamie Shotton, Ben Glocker, Christopher Zach, Shahram Izadi, Antonio Criminisi, and Andrew Fitzgibbon. Scene coordinate regression forests for camera relocalization in rgb-d images. In *CVPR*, 2013. 6, 7, 11

[36] Jürgen Sturm, Nikolas Engelhard, Felix Endres, Wolfram Burgard, and Daniel Cremers. A benchmark for the evaluation of rgb-d slam systems. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pages 573–580. IEEE, 2012. 6, 7, 18

[37] Mingjie Sun, Zhuang Liu, Anna Bair, and J Zico Kolter. A simple and effective pruning approach for large language models. *arXiv preprint arXiv:2306.11695*, 2023. 3

[38] Zhenggang Tang, Yuchen Fan, Dilin Wang, Hongyu Xu, Rakesh Ranjan, Alexander Schwing, and Zhicheng Yan. Mv-dust3r+: Single-stage scene reconstruction from sparse views in 2 seconds. In *CVPR*, 2025. 1

[39] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *NeurIPS*, 2017. 4

[40] Hengyi Wang and Lourdes Agapito. 3d reconstruction with spatial memory. In *3DV*, 2024. 2

[41] Jianyuan Wang, Minghao Chen, Nikita Karaev, Andrea Vedaldi, Christian Rupprecht, and David Novotny. Vggt: Visual geometry grounded transformer. In *CVPR*, 2025. 1, 2, 3, 6, 7

[42] Qianqian Wang, Yifei Zhang, Aleksander Holynski, Alexei A Efros, and Angjoo Kanazawa. Continuous 3d perception model with persistent state. In *CVPR*, 2025. 2, 7

[43] Shuzhe Wang, Vincent Leroy, Yohann Cabon, Boris Chidlovskii, and Jerome Revaud. Dust3r: Geometric 3d vision made easy. In *CVPR*, 2024. 1, 2

[44] Yifan Wang, Jianjun Zhou, Haoyi Zhu, Wenzheng Chang, Yang Zhou, Zizun Li, Junyi Chen, Jiangmiao Pang, Chunhua Shen, and Tong He. $\pi^3$: Scalable permutation-equivariant visual geometry learning. *arXiv preprint arXiv:2507.13347*, 2025. 2, 3, 4, 6, 7

[45] Jianing Yang, Alexander Sax, Kevin J Liang, Mikael Henaff, Hao Tang, Ang Cao, Joyce Chai, Franziska Meier, and Matt Feiszli. Fast3r: Towards 3d reconstruction of 1000+ images in one forward pass. In *CVPR*, 2025. 2, 7

[46] Yifan Yang, Kai Zhen, Bhavana Ganesh, Aram Galstyan, Goeric Huybrechts, Markus Müller, Jonas M Kübler, Rupak Vignesh Swaminathan, Athanasios Mouchtaris, Sravan Babu Bodapati, et al. Wanda++: Pruning large language models via regional gradients. *arXiv preprint arXiv:2503.04992*, 2025. 3

[47] Yao Yao, Zixin Luo, Shiwei Li, Jingyang Zhang, Yufan Ren, Lei Zhou, Tian Fang, and Long Quan. Blendedmvs: A large-scale dataset for generalized multi-view stereo networks. In *CVPR*, 2020. 2

[48] Jingyang Yuan, Huazuo Gao, Damai Dai, Junyu Luo, Liang Zhao, Zhengyan Zhang, Zhenda Xie, Yuxing Wei, Lean Wang, Zhiping Xiao, et al. Native sparse attention: Hardware-aligned and natively trainable sparse attention. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 23078–23097, 2025. 3

[49] Junyi Zhang, Charles Herrmann, Junhwa Hur, Varun Jampani, Trevor Darrell, Forrester Cole, Deqing Sun, and Ming-Hsuan Yang. Monst3r: A simple approach for estimating geometry in the presence of motion. In *ICLR*, 2025. 7

[50] Jintao Zhang, Chendong Xiang, Haofeng Huang, Haocheng Xi, Jun Zhu, Jianfei Chen, et al. Spargeattention: Accurate and training-free sparse attention accelerating any model inference. In *ICML*, 2025. 3, 4, 6, 8

[51] Peiyuan Zhang, Yongqi Chen, Haofeng Huang, Will Lin, Zhengzhong Liu, Ion Stoica, Eric P Xing, and Hao Zhang. Faster video diffusion with trainable sparse attention. In *NeurIPS*, 2025. 3

[52] Shangzhan Zhang, Jianyuan Wang, Yinghao Xu, Nan Xue, Christian Rupprecht, Xiaowei Zhou, Yujun Shen, and Gordon Wetzstein. Flare: Feed-forward geometry, appearance and camera estimation from uncalibrated sparse views. In *CVPR*, 2025. 4, 7

[53] Tinghui Zhou, Richard Tucker, John Flynn, Graham Fyffe, and Noah Snavely. Stereo magnification: learning view synthesis using multiplane images. *ACM TOG*, 2018. 6, 7, 16

| Model | Acc | Comp | N.C. | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.050 | 0.062 | 0.750 | 10 | 0.970 | 7.630 |
| VGGT | 0.053 | 0.062 | 0.746 | 30 | 0.970 | 11.580 |
| VGGT | 0.052 | 0.061 | 0.746 | 50 | 0.970 | 13.870 |
| VGGT | 0.053 | 0.062 | 0.749 | 40 | 0.950 | 16.550 |
| VGGT | 0.057 | 0.063 | 0.747 | 60 | 0.930 | 22.310 |
| VGGT | 0.054 | 0.061 | 0.747 | 50 | 0.900 | 25.540 |
| VGGT | 0.052 | 0.059 | 0.748 | 50 | 0.880 | 28.170 |
| VGGT | 0.050 | 0.055 | 0.749 | 50 | 0.850 | 31.760 |
| VGGT | 0.050 | 0.055 | 0.749 | 50 | 0.800 | 37.010 |
| VGGT | 0.053 | 0.058 | 0.747 | 50 | 0.750 | 41.440 |
| VGGT | 0.053 | 0.057 | 0.747 | 50 | 0.700 | 45.010 |
| VGGT | 0.050 | 0.059 | 0.747 | 70 | 0.700 | 50.690 |
| VGGT | 0.046 | 0.055 | 0.749 | 60 | 0.600 | 55.420 |
| VGGT | 0.050 | 0.059 | 0.746 | 60 | 0.400 | 59.970 |
| VGGT | 0.053 | 0.073 | 0.740 | 80 | 0.500 | 68.620 |
| VGGT | 0.064 | 0.084 | 0.729 | 80 | 0.400 | 74.790 |
| $\pi^3$ | 0.055 | 0.073 | 0.743 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.056 | 0.074 | 0.740 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.051 | 0.066 | 0.746 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.052 | 0.067 | 0.745 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.058 | 0.076 | 0.741 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.052 | 0.067 | 0.743 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.052 | 0.066 | 0.743 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.053 | 0.066 | 0.742 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.055 | 0.066 | 0.741 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.056 | 0.064 | 0.740 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.058 | 0.064 | 0.736 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.077 | 0.086 | 0.720 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.070 | 0.074 | 0.723 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.074 | 0.075 | 0.716 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.083 | 0.084 | 0.709 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.090 | 0.085 | 0.701 | 80 | 0.400 | 77.030 |
| MapAnything | 0.084 | 0.083 | 0.728 | 10 | 0.970 | 10.077 |
| MapAnything | 0.080 | 0.082 | 0.728 | 30 | 0.970 | 22.514 |
| MapAnything | 0.098 | 0.094 | 0.717 | 50 | 0.700 | 49.348 |
| MapAnything | 0.095 | 0.094 | 0.718 | 50 | 0.750 | 48.348 |
| MapAnything | 0.096 | 0.097 | 0.720 | 50 | 0.800 | 46.706 |
| MapAnything | 0.087 | 0.088 | 0.724 | 50 | 0.850 | 44.231 |
| MapAnything | 0.081 | 0.084 | 0.726 | 50 | 0.900 | 40.464 |
| MapAnything | 0.082 | 0.084 | 0.726 | 50 | 0.970 | 30.522 |
| MapAnything | 0.092 | 0.094 | 0.719 | 60 | 0.400 | 60.604 |
| MapAnything | 0.093 | 0.092 | 0.719 | 60 | 0.600 | 59.944 |
| MapAnything | 0.083 | 0.074 | 0.726 | 70 | 0.400 | 70.418 |
| MapAnything | 0.096 | 0.091 | 0.723 | 70 | 0.700 | 63.918 |
| MapAnything | 0.126 | 0.108 | 0.711 | 80 | 0.400 | 79.392 |
| MapAnything | 0.127 | 0.108 | 0.715 | 80 | 0.500 | 77.154 |

Table A-2. Full results for Seven Scenes [35]

| Model | RRA@30 | RTA@30 | AUC@30 | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.975 | 0.916 | 0.908 | 10 | 0.970 | 7.630 |
| VGGT | 0.975 | 0.915 | 0.908 | 30 | 0.970 | 11.580 |
| VGGT | 0.975 | 0.915 | 0.907 | 50 | 0.970 | 13.870 |
| VGGT | 0.973 | 0.914 | 0.906 | 40 | 0.950 | 16.550 |
| VGGT | 0.971 | 0.912 | 0.902 | 60 | 0.930 | 22.310 |
| VGGT | 0.967 | 0.910 | 0.899 | 50 | 0.900 | 25.540 |
| VGGT | 0.964 | 0.909 | 0.896 | 50 | 0.880 | 28.170 |
| VGGT | 0.959 | 0.906 | 0.891 | 50 | 0.850 | 31.760 |
| VGGT | 0.950 | 0.900 | 0.882 | 50 | 0.800 | 37.010 |
| VGGT | 0.941 | 0.895 | 0.872 | 50 | 0.750 | 41.440 |
| VGGT | 0.932 | 0.888 | 0.862 | 50 | 0.700 | 45.010 |
| VGGT | 0.932 | 0.886 | 0.860 | 70 | 0.700 | 50.690 |
| VGGT | 0.914 | 0.874 | 0.840 | 60 | 0.600 | 55.420 |
| VGGT | 0.896 | 0.862 | 0.820 | 60 | 0.400 | 59.970 |
| VGGT | 0.894 | 0.857 | 0.815 | 80 | 0.500 | 68.620 |
| VGGT | 0.876 | 0.844 | 0.795 | 80 | 0.400 | 74.790 |
| $\pi^3$ | 0.976 | 0.917 | 0.910 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.976 | 0.917 | 0.910 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.976 | 0.916 | 0.909 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.975 | 0.915 | 0.908 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.974 | 0.913 | 0.905 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.971 | 0.911 | 0.902 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.969 | 0.909 | 0.899 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.965 | 0.907 | 0.895 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.959 | 0.902 | 0.888 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.953 | 0.899 | 0.881 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.946 | 0.895 | 0.875 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.943 | 0.888 | 0.867 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.928 | 0.883 | 0.855 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.915 | 0.876 | 0.842 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.906 | 0.861 | 0.825 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.887 | 0.849 | 0.805 | 80 | 0.400 | 77.030 |
| MapAnything | 0.810 | 0.745 | 0.694 | 10 | 0.970 | 10.077 |
| MapAnything | 0.809 | 0.744 | 0.693 | 30 | 0.970 | 22.514 |
| MapAnything | 0.803 | 0.736 | 0.685 | 50 | 0.700 | 49.348 |
| MapAnything | 0.803 | 0.736 | 0.685 | 50 | 0.750 | 48.348 |
| MapAnything | 0.803 | 0.737 | 0.685 | 50 | 0.800 | 46.706 |
| MapAnything | 0.804 | 0.737 | 0.686 | 50 | 0.850 | 44.231 |
| MapAnything | 0.804 | 0.738 | 0.687 | 50 | 0.900 | 40.464 |
| MapAnything | 0.807 | 0.741 | 0.690 | 50 | 0.970 | 30.522 |
| MapAnything | 0.798 | 0.730 | 0.678 | 60 | 0.400 | 60.604 |
| MapAnything | 0.798 | 0.731 | 0.678 | 60 | 0.600 | 59.944 |
| MapAnything | 0.788 | 0.720 | 0.667 | 70 | 0.400 | 70.418 |
| MapAnything | 0.790 | 0.722 | 0.669 | 70 | 0.700 | 63.918 |
| MapAnything | 0.765 | 0.698 | 0.641 | 80 | 0.400 | 79.392 |
| MapAnything | 0.765 | 0.698 | 0.642 | 80 | 0.500 | 77.154 |

Table A-3. Full results for Co3Dv2 [30]

| Model | Acc | Comp | N.C. | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 1.075 | 1.331 | 0.613 | 10 | 0.970 | 7.774 |
| VGGT | 1.068 | 1.330 | 0.612 | 30 | 0.970 | 12.173 |
| VGGT | 1.072 | 1.335 | 0.613 | 40 | 0.950 | 17.249 |
| VGGT | 1.172 | 1.286 | 0.609 | 50 | 0.750 | 41.589 |
| VGGT | 1.215 | 1.281 | 0.609 | 50 | 0.700 | 45.062 |
| VGGT | 1.114 | 1.316 | 0.611 | 50 | 0.850 | 32.250 |
| VGGT | 1.100 | 1.328 | 0.610 | 50 | 0.880 | 28.764 |
| VGGT | 1.145 | 1.303 | 0.611 | 50 | 0.800 | 37.314 |
| VGGT | 1.071 | 1.336 | 0.611 | 50 | 0.970 | 14.699 |
| VGGT | 1.089 | 1.330 | 0.611 | 50 | 0.900 | 26.206 |
| VGGT | 1.613 | 1.287 | 0.614 | 60 | 0.400 | 60.123 |
| VGGT | 1.344 | 1.269 | 0.612 | 60 | 0.600 | 55.626 |
| VGGT | 1.081 | 1.333 | 0.611 | 60 | 0.930 | 23.219 |
| VGGT | 1.222 | 1.275 | 0.609 | 70 | 0.700 | 51.120 |
| VGGT | 1.952 | 1.342 | 0.617 | 80 | 0.400 | 74.832 |
| VGGT | 1.627 | 1.291 | 0.614 | 80 | 0.500 | 68.760 |
| $\pi^3$ | 1.564 | 1.356 | 0.609 | 10 | 0.970 | 8.787 |
| $\pi^3$ | 1.565 | 1.362 | 0.609 | 30 | 0.970 | 13.939 |
| $\pi^3$ | 1.855 | 1.470 | 0.613 | 50 | 0.700 | 47.267 |
| $\pi^3$ | 1.742 | 1.441 | 0.612 | 50 | 0.750 | 44.617 |
| $\pi^3$ | 1.667 | 1.414 | 0.610 | 50 | 0.800 | 40.792 |
| $\pi^3$ | 1.614 | 1.392 | 0.609 | 50 | 0.850 | 35.800 |
| $\pi^3$ | 1.594 | 1.382 | 0.610 | 50 | 0.880 | 32.175 |
| $\pi^3$ | 1.582 | 1.373 | 0.609 | 50 | 0.900 | 29.450 |
| $\pi^3$ | 1.567 | 1.365 | 0.609 | 50 | 0.970 | 16.621 |
| $\pi^3$ | 2.747 | 1.714 | 0.619 | 60 | 0.400 | 60.123 |
| $\pi^3$ | 2.196 | 1.556 | 0.616 | 60 | 0.600 | 57.786 |
| $\pi^3$ | 1.576 | 1.370 | 0.610 | 60 | 0.930 | 26.157 |
| $\pi^3$ | 1.893 | 1.443 | 0.613 | 70 | 0.700 | 55.174 |
| $\pi^3$ | 3.172 | 1.771 | 0.620 | 80 | 0.400 | 76.939 |
| $\pi^3$ | 2.767 | 1.689 | 0.619 | 80 | 0.500 | 72.056 |
| MapAnything | 2.409 | 2.496 | 0.621 | 10 | 0.970 | 10.077 |
| MapAnything | 2.419 | 2.484 | 0.621 | 30 | 0.970 | 22.514 |
| MapAnything | 2.366 | 2.452 | 0.622 | 50 | 0.700 | 49.348 |
| MapAnything | 2.369 | 2.453 | 0.622 | 50 | 0.750 | 48.348 |
| MapAnything | 2.363 | 2.447 | 0.622 | 50 | 0.800 | 46.706 |
| MapAnything | 2.361 | 2.450 | 0.622 | 50 | 0.850 | 44.231 |
| MapAnything | 2.348 | 2.458 | 0.621 | 50 | 0.900 | 40.464 |
| MapAnything | 2.350 | 2.468 | 0.621 | 50 | 0.970 | 30.522 |
| MapAnything | 2.427 | 2.446 | 0.622 | 60 | 0.400 | 60.604 |
| MapAnything | 2.438 | 2.447 | 0.622 | 60 | 0.600 | 59.944 |
| MapAnything | 2.520 | 2.435 | 0.623 | 70 | 0.400 | 70.418 |
| MapAnything | 2.502 | 2.455 | 0.623 | 70 | 0.700 | 63.918 |
| MapAnything | 2.680 | 2.440 | 0.624 | 80 | 0.400 | 79.392 |
| MapAnything | 2.673 | 2.440 | 0.624 | 80 | 0.500 | 77.154 |

Table A-4. Full results for DTU [19]

| Model | Acc | Comp | N.C. | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.232 | 0.190 | 0.886 | 10 | 0.970 | 7.630 |
| VGGT | 0.220 | 0.176 | 0.888 | 30 | 0.970 | 11.580 |
| VGGT | 0.238 | 0.180 | 0.884 | 50 | 0.970 | 13.870 |
| VGGT | 0.218 | 0.174 | 0.887 | 40 | 0.950 | 16.550 |
| VGGT | 0.231 | 0.187 | 0.874 | 60 | 0.930 | 22.310 |
| VGGT | 0.212 | 0.172 | 0.885 | 50 | 0.900 | 25.540 |
| VGGT | 0.213 | 0.169 | 0.884 | 50 | 0.880 | 28.170 |
| VGGT | 0.195 | 0.153 | 0.886 | 50 | 0.850 | 31.760 |
| VGGT | 0.201 | 0.158 | 0.880 | 50 | 0.800 | 37.010 |
| VGGT | 0.208 | 0.170 | 0.877 | 50 | 0.750 | 41.440 |
| VGGT | 0.212 | 0.173 | 0.869 | 50 | 0.700 | 45.010 |
| VGGT | 0.207 | 0.183 | 0.863 | 70 | 0.700 | 50.690 |
| VGGT | 0.217 | 0.177 | 0.863 | 60 | 0.600 | 55.420 |
| VGGT | 0.233 | 0.185 | 0.856 | 60 | 0.400 | 59.970 |
| VGGT | 0.234 | 0.188 | 0.842 | 80 | 0.500 | 68.620 |
| VGGT | 0.250 | 0.205 | 0.837 | 80 | 0.400 | 74.790 |
| $\pi^3$ | 0.107 | 0.085 | 0.898 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.113 | 0.089 | 0.897 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.115 | 0.091 | 0.898 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.115 | 0.089 | 0.892 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.135 | 0.104 | 0.882 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.135 | 0.101 | 0.880 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.138 | 0.103 | 0.878 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.146 | 0.108 | 0.876 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.159 | 0.114 | 0.866 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.170 | 0.117 | 0.860 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.189 | 0.126 | 0.856 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.209 | 0.146 | 0.850 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.230 | 0.145 | 0.841 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.274 | 0.170 | 0.832 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.320 | 0.194 | 0.823 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.386 | 0.222 | 0.800 | 80 | 0.400 | 77.030 |
| MapAnything | 0.221 | 0.164 | 0.843 | 10 | 0.970 | 10.077 |
| MapAnything | 0.246 | 0.197 | 0.843 | 30 | 0.970 | 22.514 |
| MapAnything | 0.231 | 0.178 | 0.843 | 50 | 0.700 | 49.348 |
| MapAnything | 0.229 | 0.180 | 0.843 | 50 | 0.750 | 48.348 |
| MapAnything | 0.227 | 0.178 | 0.843 | 50 | 0.800 | 46.706 |
| MapAnything | 0.225 | 0.176 | 0.843 | 50 | 0.850 | 44.231 |
| MapAnything | 0.230 | 0.172 | 0.842 | 50 | 0.900 | 40.464 |
| MapAnything | 0.229 | 0.174 | 0.842 | 50 | 0.970 | 30.522 |
| MapAnything | 0.269 | 0.207 | 0.838 | 60 | 0.400 | 60.604 |
| MapAnything | 0.265 | 0.206 | 0.839 | 60 | 0.600 | 59.944 |
| MapAnything | 0.252 | 0.200 | 0.834 | 70 | 0.400 | 70.418 |
| MapAnything | 0.247 | 0.195 | 0.838 | 70 | 0.700 | 63.918 |
| MapAnything | 0.467 | 0.304 | 0.812 | 80 | 0.400 | 79.392 |
| MapAnything | 0.443 | 0.282 | 0.808 | 80 | 0.500 | 77.154 |

Table A-5. Full results for ETH3D [33]

| Model | Acc | Comp | N.C. | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.047 | 0.049 | 0.897 | 10 | 0.970 | 7.630 |
| VGGT | 0.047 | 0.049 | 0.899 | 30 | 0.970 | 11.580 |
| VGGT | 0.049 | 0.051 | 0.899 | 50 | 0.970 | 13.870 |
| VGGT | 0.048 | 0.050 | 0.898 | 40 | 0.950 | 16.550 |
| VGGT | 0.051 | 0.054 | 0.893 | 60 | 0.930 | 22.310 |
| VGGT | 0.049 | 0.051 | 0.897 | 50 | 0.900 | 25.540 |
| VGGT | 0.050 | 0.052 | 0.894 | 50 | 0.880 | 28.170 |
| VGGT | 0.051 | 0.053 | 0.892 | 50 | 0.850 | 31.760 |
| VGGT | 0.051 | 0.053 | 0.890 | 50 | 0.800 | 37.010 |
| VGGT | 0.053 | 0.054 | 0.885 | 50 | 0.750 | 41.440 |
| VGGT | 0.053 | 0.053 | 0.881 | 50 | 0.700 | 45.010 |
| VGGT | 0.058 | 0.061 | 0.875 | 70 | 0.700 | 50.690 |
| VGGT | 0.054 | 0.056 | 0.885 | 60 | 0.600 | 55.420 |
| VGGT | 0.057 | 0.060 | 0.875 | 60 | 0.400 | 59.970 |
| VGGT | 0.072 | 0.078 | 0.860 | 80 | 0.500 | 68.620 |
| VGGT | 0.078 | 0.090 | 0.849 | 80 | 0.400 | 74.790 |
| $\pi^3$ | 0.028 | 0.028 | 0.907 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.029 | 0.029 | 0.904 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.030 | 0.030 | 0.902 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.032 | 0.031 | 0.897 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.033 | 0.032 | 0.898 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.032 | 0.032 | 0.897 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.033 | 0.032 | 0.896 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.035 | 0.033 | 0.892 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.039 | 0.036 | 0.884 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.043 | 0.040 | 0.877 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.048 | 0.043 | 0.868 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.051 | 0.047 | 0.878 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.058 | 0.049 | 0.862 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.069 | 0.055 | 0.842 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.074 | 0.063 | 0.835 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.083 | 0.067 | 0.815 | 80 | 0.400 | 77.030 |
| MapAnything | 0.093 | 0.093 | 0.853 | 10 | 0.970 | 10.077 |
| MapAnything | 0.092 | 0.092 | 0.856 | 30 | 0.970 | 22.514 |
| MapAnything | 0.100 | 0.090 | 0.856 | 50 | 0.700 | 49.348 |
| MapAnything | 0.100 | 0.091 | 0.856 | 50 | 0.750 | 48.348 |
| MapAnything | 0.100 | 0.091 | 0.856 | 50 | 0.800 | 46.706 |
| MapAnything | 0.099 | 0.091 | 0.856 | 50 | 0.850 | 44.231 |
| MapAnything | 0.098 | 0.091 | 0.855 | 50 | 0.900 | 40.464 |
| MapAnything | 0.097 | 0.092 | 0.855 | 50 | 0.970 | 30.522 |
| MapAnything | 0.114 | 0.118 | 0.841 | 60 | 0.400 | 60.604 |
| MapAnything | 0.114 | 0.116 | 0.843 | 60 | 0.600 | 59.944 |
| MapAnything | 0.138 | 0.128 | 0.828 | 70 | 0.400 | 70.418 |
| MapAnything | 0.136 | 0.129 | 0.828 | 70 | 0.700 | 63.918 |
| MapAnything | 0.174 | 0.149 | 0.820 | 80 | 0.400 | 79.392 |
| MapAnything | 0.201 | 0.174 | 0.815 | 80 | 0.500 | 77.154 |

Table A-6. Full results for NRGBD [1]

| Model | RRA@30 | RTA@30 | AUC@30 | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.977 | 0.763 | 0.760 | 10 | 0.970 | 7.630 |
| VGGT | 0.977 | 0.762 | 0.759 | 30 | 0.970 | 11.580 |
| VGGT | 0.976 | 0.762 | 0.759 | 50 | 0.970 | 13.870 |
| VGGT | 0.976 | 0.759 | 0.756 | 40 | 0.950 | 16.550 |
| VGGT | 0.975 | 0.754 | 0.751 | 60 | 0.930 | 22.310 |
| VGGT | 0.976 | 0.752 | 0.749 | 50 | 0.900 | 25.540 |
| VGGT | 0.975 | 0.750 | 0.746 | 50 | 0.880 | 28.170 |
| VGGT | 0.975 | 0.745 | 0.742 | 50 | 0.850 | 31.760 |
| VGGT | 0.975 | 0.738 | 0.735 | 50 | 0.800 | 37.010 |
| VGGT | 0.974 | 0.732 | 0.729 | 50 | 0.750 | 41.440 |
| VGGT | 0.974 | 0.726 | 0.723 | 50 | 0.700 | 45.010 |
| VGGT | 0.972 | 0.721 | 0.718 | 70 | 0.700 | 50.690 |
| VGGT | 0.972 | 0.712 | 0.709 | 60 | 0.600 | 55.420 |
| VGGT | 0.971 | 0.699 | 0.696 | 60 | 0.400 | 59.970 |
| VGGT | 0.967 | 0.687 | 0.683 | 80 | 0.500 | 68.620 |
| VGGT | 0.965 | 0.675 | 0.671 | 80 | 0.400 | 74.790 |
| $\pi^3$ | 0.992 | 0.871 | 0.870 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.992 | 0.868 | 0.867 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.992 | 0.866 | 0.865 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.992 | 0.861 | 0.860 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.992 | 0.854 | 0.853 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.992 | 0.847 | 0.846 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.992 | 0.841 | 0.840 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.991 | 0.833 | 0.832 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.991 | 0.821 | 0.820 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.991 | 0.812 | 0.810 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.990 | 0.805 | 0.804 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.990 | 0.798 | 0.796 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.989 | 0.784 | 0.783 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.988 | 0.776 | 0.774 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.986 | 0.762 | 0.760 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.984 | 0.747 | 0.745 | 80 | 0.400 | 77.030 |
| MapAnything | 0.982 | 0.776 | 0.773 | 10 | 0.970 | 10.077 |
| MapAnything | 0.983 | 0.773 | 0.771 | 30 | 0.970 | 22.514 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.700 | 49.348 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.750 | 48.348 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.800 | 46.706 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.850 | 44.231 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.900 | 40.464 |
| MapAnything | 0.983 | 0.768 | 0.766 | 50 | 0.970 | 30.522 |
| MapAnything | 0.983 | 0.759 | 0.757 | 60 | 0.400 | 60.604 |
| MapAnything | 0.983 | 0.759 | 0.757 | 60 | 0.600 | 59.944 |
| MapAnything | 0.982 | 0.741 | 0.739 | 70 | 0.400 | 70.418 |
| MapAnything | 0.982 | 0.744 | 0.743 | 70 | 0.700 | 63.918 |
| MapAnything | 0.977 | 0.692 | 0.690 | 80 | 0.400 | 79.392 |
| MapAnything | 0.977 | 0.695 | 0.693 | 80 | 0.500 | 77.154 |

Table A-7. Full results for Real Estate 10K [53]

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.035 | 0.015 | 0.382 | 10 | 0.970 | 7.560 |
| VGGT | 0.035 | 0.015 | 0.382 | 30 | 0.970 | 11.490 |
| VGGT | 0.036 | 0.016 | 0.393 | 50 | 0.970 | 13.690 |
| VGGT | 0.036 | 0.015 | 0.384 | 40 | 0.950 | 16.390 |
| VGGT | 0.036 | 0.016 | 0.387 | 60 | 0.930 | 22.170 |
| VGGT | 0.036 | 0.016 | 0.389 | 50 | 0.900 | 25.370 |
| VGGT | 0.036 | 0.016 | 0.392 | 50 | 0.880 | 28.010 |
| VGGT | 0.037 | 0.016 | 0.395 | 50 | 0.850 | 31.600 |
| VGGT | 0.037 | 0.016 | 0.402 | 50 | 0.800 | 36.840 |
| VGGT | 0.038 | 0.017 | 0.410 | 50 | 0.750 | 41.270 |
| VGGT | 0.039 | 0.017 | 0.421 | 50 | 0.700 | 44.810 |
| VGGT | 0.040 | 0.017 | 0.428 | 70 | 0.700 | 50.670 |
| VGGT | 0.041 | 0.018 | 0.451 | 60 | 0.600 | 55.400 |
| VGGT | 0.043 | 0.019 | 0.481 | 60 | 0.400 | 60.080 |
| VGGT | 0.044 | 0.019 | 0.484 | 80 | 0.500 | 68.650 |
| VGGT | 0.047 | 0.020 | 0.527 | 80 | 0.400 | 74.740 |
| $\pi^3$ | 0.031 | 0.013 | 0.345 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.031 | 0.013 | 0.345 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.030 | 0.013 | 0.346 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.031 | 0.013 | 0.346 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.031 | 0.013 | 0.348 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.031 | 0.013 | 0.349 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.031 | 0.013 | 0.351 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.031 | 0.013 | 0.353 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.032 | 0.013 | 0.356 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.032 | 0.013 | 0.359 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.032 | 0.014 | 0.363 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.033 | 0.014 | 0.369 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.033 | 0.014 | 0.375 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.034 | 0.014 | 0.382 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.035 | 0.015 | 0.397 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.037 | 0.015 | 0.418 | 80 | 0.400 | 77.030 |
| MapAnything | 0.061 | 0.027 | 0.987 | 10 | 0.970 | 10.077 |
| MapAnything | 0.061 | 0.027 | 0.991 | 30 | 0.970 | 22.514 |
| MapAnything | 0.061 | 0.029 | 1.007 | 50 | 0.700 | 49.348 |
| MapAnything | 0.061 | 0.029 | 1.005 | 50 | 0.750 | 48.348 |
| MapAnything | 0.061 | 0.029 | 1.015 | 50 | 0.800 | 46.706 |
| MapAnything | 0.061 | 0.028 | 1.023 | 50 | 0.850 | 44.231 |
| MapAnything | 0.061 | 0.028 | 1.093 | 50 | 0.900 | 40.464 |
| MapAnything | 0.061 | 0.028 | 0.998 | 50 | 0.970 | 30.522 |
| MapAnything | 0.062 | 0.028 | 1.104 | 60 | 0.400 | 60.604 |
| MapAnything | 0.062 | 0.028 | 1.093 | 60 | 0.600 | 59.944 |
| MapAnything | 0.064 | 0.029 | 1.114 | 70 | 0.400 | 70.418 |
| MapAnything | 0.064 | 0.028 | 1.108 | 70 | 0.700 | 63.918 |
| MapAnything | 0.067 | 0.030 | 1.044 | 80 | 0.400 | 79.392 |
| MapAnything | 0.067 | 0.029 | 1.025 | 80 | 0.500 | 77.154 |

Table A-8. Full results for ScanNet [7] (first 90 frames with temporal stride of 3).

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.012 | 0.010 | 0.352 | 10 | 0.970 | 7.560 |
| VGGT | 0.012 | 0.010 | 0.352 | 30 | 0.970 | 11.490 |
| VGGT | 0.012 | 0.010 | 0.353 | 50 | 0.970 | 13.690 |
| VGGT | 0.012 | 0.010 | 0.354 | 40 | 0.950 | 16.390 |
| VGGT | 0.013 | 0.011 | 0.356 | 60 | 0.930 | 22.170 |
| VGGT | 0.013 | 0.011 | 0.358 | 50 | 0.900 | 25.370 |
| VGGT | 0.013 | 0.011 | 0.359 | 50 | 0.880 | 28.010 |
| VGGT | 0.014 | 0.011 | 0.361 | 50 | 0.850 | 31.600 |
| VGGT | 0.014 | 0.012 | 0.366 | 50 | 0.800 | 36.840 |
| VGGT | 0.015 | 0.012 | 0.370 | 50 | 0.750 | 41.270 |
| VGGT | 0.016 | 0.013 | 0.376 | 50 | 0.700 | 44.810 |
| VGGT | 0.017 | 0.013 | 0.379 | 70 | 0.700 | 50.670 |
| VGGT | 0.018 | 0.014 | 0.390 | 60 | 0.600 | 55.400 |
| VGGT | 0.020 | 0.015 | 0.414 | 60 | 0.400 | 60.080 |
| VGGT | 0.020 | 0.015 | 0.410 | 80 | 0.500 | 68.650 |
| VGGT | 0.024 | 0.016 | 0.453 | 80 | 0.400 | 74.740 |
| $\pi^3$ | 0.014 | 0.009 | 0.349 | 10 | 0.970 | 8.650 |
| $\pi^3$ | 0.014 | 0.009 | 0.350 | 30 | 0.970 | 13.510 |
| $\pi^3$ | 0.014 | 0.009 | 0.350 | 50 | 0.970 | 16.040 |
| $\pi^3$ | 0.014 | 0.009 | 0.351 | 40 | 0.950 | 19.130 |
| $\pi^3$ | 0.014 | 0.009 | 0.353 | 60 | 0.930 | 25.500 |
| $\pi^3$ | 0.015 | 0.010 | 0.354 | 50 | 0.900 | 29.030 |
| $\pi^3$ | 0.015 | 0.010 | 0.355 | 50 | 0.880 | 31.830 |
| $\pi^3$ | 0.015 | 0.010 | 0.357 | 50 | 0.850 | 35.570 |
| $\pi^3$ | 0.016 | 0.011 | 0.360 | 50 | 0.800 | 40.750 |
| $\pi^3$ | 0.017 | 0.011 | 0.361 | 50 | 0.750 | 44.680 |
| $\pi^3$ | 0.017 | 0.011 | 0.363 | 50 | 0.700 | 47.350 |
| $\pi^3$ | 0.019 | 0.012 | 0.371 | 70 | 0.700 | 55.120 |
| $\pi^3$ | 0.019 | 0.012 | 0.373 | 60 | 0.600 | 57.650 |
| $\pi^3$ | 0.020 | 0.013 | 0.375 | 60 | 0.400 | 59.950 |
| $\pi^3$ | 0.022 | 0.014 | 0.387 | 80 | 0.500 | 72.190 |
| $\pi^3$ | 0.023 | 0.014 | 0.401 | 80 | 0.400 | 77.030 |
| MapAnything | 0.026 | 0.017 | 0.419 | 10 | 0.970 | 10.077 |
| MapAnything | 0.026 | 0.017 | 0.418 | 30 | 0.970 | 22.514 |
| MapAnything | 0.027 | 0.018 | 0.419 | 50 | 0.700 | 49.348 |
| MapAnything | 0.027 | 0.018 | 0.419 | 50 | 0.750 | 48.348 |
| MapAnything | 0.027 | 0.018 | 0.420 | 50 | 0.800 | 46.706 |
| MapAnything | 0.027 | 0.018 | 0.420 | 50 | 0.850 | 44.231 |
| MapAnything | 0.027 | 0.018 | 0.420 | 50 | 0.900 | 40.464 |
| MapAnything | 0.027 | 0.018 | 0.419 | 50 | 0.970 | 30.522 |
| MapAnything | 0.028 | 0.018 | 0.421 | 60 | 0.400 | 60.604 |
| MapAnything | 0.028 | 0.018 | 0.422 | 60 | 0.600 | 59.944 |
| MapAnything | 0.029 | 0.018 | 0.424 | 70 | 0.400 | 70.418 |
| MapAnything | 0.029 | 0.018 | 0.423 | 70 | 0.700 | 63.918 |
| MapAnything | 0.032 | 0.018 | 0.424 | 80 | 0.400 | 79.392 |
| MapAnything | 0.032 | 0.018 | 0.424 | 80 | 0.500 | 77.154 |

Table A-9. Full results for TUM [36].

| Model | RRA@5 | RRA@15 | RRA@30 | RTA@5 | RTA@15 | RTA@30 | AUC@5 | AUC@15 | AUC@30 | ATE | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VGGT | 0.840 | 0.895 | 0.913 | 0.817 | 0.885 | 0.911 | 0.768 | 0.852 | 0.880 | 0.024 | 10 | 0.970 | 7.273 |
| VGGT | 0.836 | 0.894 | 0.912 | 0.815 | 0.884 | 0.910 | 0.764 | 0.851 | 0.879 | 0.025 | 30 | 0.970 | 11.093 |
| VGGT | 0.828 | 0.890 | 0.910 | 0.811 | 0.882 | 0.909 | 0.757 | 0.847 | 0.877 | 0.025 | 40 | 0.950 | 15.976 |
| VGGT | 0.741 | 0.859 | 0.894 | 0.753 | 0.860 | 0.898 | 0.660 | 0.812 | 0.858 | 0.024 | 50 | 0.700 | 44.533 |
| VGGT | 0.779 | 0.873 | 0.900 | 0.778 | 0.870 | 0.902 | 0.701 | 0.827 | 0.866 | 0.024 | 50 | 0.800 | 36.316 |
| VGGT | 0.798 | 0.880 | 0.905 | 0.791 | 0.875 | 0.905 | 0.722 | 0.835 | 0.871 | 0.024 | 50 | 0.850 | 31.045 |
| VGGT | 0.808 | 0.883 | 0.906 | 0.797 | 0.877 | 0.907 | 0.732 | 0.838 | 0.873 | 0.024 | 50 | 0.880 | 27.458 |
| VGGT | 0.814 | 0.885 | 0.907 | 0.801 | 0.879 | 0.907 | 0.739 | 0.841 | 0.874 | 0.024 | 50 | 0.900 | 24.842 |
| VGGT | 0.834 | 0.892 | 0.911 | 0.814 | 0.883 | 0.910 | 0.762 | 0.849 | 0.878 | 0.024 | 50 | 0.970 | 13.124 |
| VGGT | 0.633 | 0.822 | 0.875 | 0.677 | 0.832 | 0.884 | 0.546 | 0.769 | 0.837 | 0.024 | 60 | 0.400 | 59.683 |
| VGGT | 0.691 | 0.842 | 0.885 | 0.718 | 0.847 | 0.891 | 0.607 | 0.792 | 0.848 | 0.024 | 60 | 0.600 | 54.867 |
| VGGT | 0.821 | 0.887 | 0.908 | 0.806 | 0.881 | 0.909 | 0.747 | 0.844 | 0.875 | 0.024 | 60 | 0.930 | 21.490 |
| VGGT | 0.606 | 0.811 | 0.871 | 0.655 | 0.824 | 0.880 | 0.516 | 0.757 | 0.831 | 0.025 | 70 | 0.400 | 68.666 |
| VGGT | 0.726 | 0.853 | 0.891 | 0.742 | 0.857 | 0.897 | 0.644 | 0.805 | 0.855 | 0.024 | 70 | 0.700 | 49.820 |
| VGGT | 0.596 | 0.803 | 0.868 | 0.635 | 0.812 | 0.873 | 0.500 | 0.745 | 0.826 | 0.025 | 80 | 0.400 | 74.143 |
| VGGT | 0.631 | 0.818 | 0.875 | 0.664 | 0.825 | 0.880 | 0.539 | 0.762 | 0.835 | 0.024 | 80 | 0.500 | 67.759 |
| $\pi^3$ | 0.848 | 0.905 | 0.926 | 0.844 | 0.905 | 0.928 | 0.796 | 0.875 | 0.902 | 0.015 | 10 | 0.970 | 8.240 |
| $\pi^3$ | 0.846 | 0.904 | 0.924 | 0.842 | 0.904 | 0.928 | 0.794 | 0.874 | 0.901 | 0.015 | 30 | 0.970 | 13.130 |
| $\pi^3$ | 0.840 | 0.901 | 0.923 | 0.838 | 0.903 | 0.928 | 0.787 | 0.871 | 0.900 | 0.015 | 40 | 0.950 | 18.652 |
| $\pi^3$ | 0.783 | 0.881 | 0.914 | 0.810 | 0.893 | 0.924 | 0.723 | 0.849 | 0.891 | 0.015 | 50 | 0.700 | 46.717 |
| $\pi^3$ | 0.808 | 0.890 | 0.918 | 0.822 | 0.897 | 0.925 | 0.750 | 0.859 | 0.894 | 0.015 | 50 | 0.800 | 39.745 |
| $\pi^3$ | 0.820 | 0.894 | 0.919 | 0.828 | 0.899 | 0.926 | 0.763 | 0.863 | 0.896 | 0.015 | 50 | 0.850 | 34.638 |
| $\pi^3$ | 0.826 | 0.896 | 0.921 | 0.831 | 0.900 | 0.927 | 0.771 | 0.866 | 0.898 | 0.015 | 50 | 0.880 | 31.002 |
| $\pi^3$ | 0.830 | 0.898 | 0.921 | 0.833 | 0.901 | 0.927 | 0.775 | 0.867 | 0.898 | 0.015 | 50 | 0.900 | 28.288 |
| $\pi^3$ | 0.843 | 0.902 | 0.922 | 0.840 | 0.903 | 0.928 | 0.790 | 0.872 | 0.900 | 0.015 | 50 | 0.970 | 15.558 |
| $\pi^3$ | 0.724 | 0.860 | 0.904 | 0.775 | 0.882 | 0.918 | 0.660 | 0.827 | 0.880 | 0.015 | 60 | 0.400 | 59.683 |
| $\pi^3$ | 0.742 | 0.866 | 0.907 | 0.786 | 0.886 | 0.920 | 0.679 | 0.834 | 0.884 | 0.015 | 60 | 0.600 | 56.939 |
| $\pi^3$ | 0.834 | 0.899 | 0.922 | 0.835 | 0.902 | 0.928 | 0.780 | 0.869 | 0.900 | 0.015 | 60 | 0.930 | 24.825 |
| $\pi^3$ | 0.680 | 0.834 | 0.887 | 0.741 | 0.865 | 0.908 | 0.613 | 0.801 | 0.864 | 0.017 | 70 | 0.400 | 69.252 |
| $\pi^3$ | 0.766 | 0.872 | 0.910 | 0.794 | 0.888 | 0.921 | 0.703 | 0.841 | 0.887 | 0.015 | 70 | 0.700 | 53.773 |
| $\pi^3$ | 0.651 | 0.817 | 0.875 | 0.714 | 0.850 | 0.900 | 0.580 | 0.781 | 0.851 | 0.017 | 80 | 0.400 | 76.202 |
| $\pi^3$ | 0.684 | 0.833 | 0.886 | 0.736 | 0.862 | 0.907 | 0.614 | 0.797 | 0.862 | 0.016 | 80 | 0.500 | 70.919 |
| MapAnything | 0.687 | 0.836 | 0.881 | 0.687 | 0.825 | 0.878 | 0.586 | 0.776 | 0.839 | 0.023 | 10 | 0.970 | 9.785 |
| MapAnything | 0.686 | 0.836 | 0.882 | 0.687 | 0.827 | 0.880 | 0.586 | 0.776 | 0.840 | 0.022 | 30 | 0.970 | 24.555 |
| MapAnything | 0.686 | 0.834 | 0.880 | 0.686 | 0.827 | 0.880 | 0.584 | 0.775 | 0.839 | 0.022 | 40 | 0.950 | 32.998 |
| MapAnything | 0.684 | 0.836 | 0.882 | 0.684 | 0.827 | 0.880 | 0.582 | 0.777 | 0.840 | 0.022 | 50 | 0.700 | 49.660 |
| MapAnything | 0.683 | 0.837 | 0.883 | 0.685 | 0.828 | 0.881 | 0.581 | 0.777 | 0.841 | 0.021 | 50 | 0.800 | 47.981 |
| MapAnything | 0.683 | 0.837 | 0.882 | 0.685 | 0.827 | 0.881 | 0.581 | 0.777 | 0.841 | 0.021 | 50 | 0.850 | 46.257 |
| MapAnything | 0.683 | 0.838 | 0.883 | 0.685 | 0.827 | 0.881 | 0.581 | 0.778 | 0.841 | 0.021 | 50 | 0.880 | 44.772 |
| MapAnything | 0.683 | 0.837 | 0.881 | 0.685 | 0.827 | 0.881 | 0.581 | 0.777 | 0.840 | 0.021 | 50 | 0.900 | 43.486 |
| MapAnything | 0.684 | 0.836 | 0.882 | 0.688 | 0.828 | 0.881 | 0.584 | 0.777 | 0.841 | 0.022 | 50 | 0.970 | 34.473 |
| MapAnything | 0.676 | 0.834 | 0.881 | 0.679 | 0.824 | 0.879 | 0.573 | 0.773 | 0.839 | 0.022 | 60 | 0.400 | 60.019 |
| MapAnything | 0.676 | 0.833 | 0.880 | 0.678 | 0.824 | 0.878 | 0.572 | 0.772 | 0.838 | 0.022 | 60 | 0.600 | 59.707 |
| MapAnything | 0.681 | 0.837 | 0.883 | 0.683 | 0.826 | 0.881 | 0.580 | 0.776 | 0.841 | 0.022 | 60 | 0.930 | 45.707 |
| MapAnything | 0.664 | 0.830 | 0.879 | 0.669 | 0.822 | 0.878 | 0.560 | 0.768 | 0.837 | 0.022 | 70 | 0.400 | 70.026 |
| MapAnything | 0.667 | 0.829 | 0.879 | 0.669 | 0.821 | 0.877 | 0.562 | 0.768 | 0.836 | 0.022 | 70 | 0.700 | 65.915 |
| MapAnything | 0.643 | 0.814 | 0.866 | 0.651 | 0.809 | 0.868 | 0.538 | 0.752 | 0.823 | 0.026 | 80 | 0.400 | 79.583 |
| MapAnything | 0.646 | 0.814 | 0.866 | 0.651 | 0.808 | 0.868 | 0.540 | 0.752 | 0.823 | 0.026 | 80 | 0.500 | 78.170 |

Table A-10. Full results for Tanks & Temples [22] (50 frames).

| Model | RRA@5 | RRA@15 | RRA@30 | RTA@5 | RTA@15 | RTA@30 | AUC@5 | AUC@15 | AUC@30 | ATE | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VGGT | 0.837 | 0.893 | 0.912 | 0.809 | 0.879 | 0.908 | 0.761 | 0.847 | 0.877 | 0.016 | 10 | 0.970 | 7.273 |
| VGGT | 0.836 | 0.893 | 0.912 | 0.808 | 0.879 | 0.908 | 0.759 | 0.846 | 0.877 | 0.016 | 30 | 0.970 | 11.093 |
| VGGT | 0.832 | 0.891 | 0.910 | 0.806 | 0.879 | 0.908 | 0.753 | 0.845 | 0.876 | 0.017 | 40 | 0.950 | 15.976 |
| VGGT | 0.740 | 0.856 | 0.891 | 0.746 | 0.854 | 0.894 | 0.653 | 0.806 | 0.855 | 0.017 | 50 | 0.700 | 44.533 |
| VGGT | 0.780 | 0.872 | 0.899 | 0.772 | 0.865 | 0.900 | 0.696 | 0.823 | 0.864 | 0.016 | 50 | 0.800 | 36.316 |
| VGGT | 0.798 | 0.877 | 0.902 | 0.784 | 0.869 | 0.903 | 0.716 | 0.830 | 0.867 | 0.017 | 50 | 0.850 | 31.045 |
| VGGT | 0.808 | 0.882 | 0.905 | 0.790 | 0.872 | 0.905 | 0.727 | 0.834 | 0.870 | 0.017 | 50 | 0.880 | 27.458 |
| VGGT | 0.816 | 0.885 | 0.907 | 0.794 | 0.874 | 0.906 | 0.734 | 0.837 | 0.872 | 0.017 | 50 | 0.900 | 24.842 |
| VGGT | 0.836 | 0.892 | 0.911 | 0.806 | 0.879 | 0.908 | 0.757 | 0.846 | 0.877 | 0.016 | 50 | 0.970 | 13.124 |
| VGGT | 0.639 | 0.820 | 0.872 | 0.674 | 0.826 | 0.879 | 0.546 | 0.765 | 0.833 | 0.017 | 60 | 0.400 | 59.683 |
| VGGT | 0.695 | 0.840 | 0.883 | 0.712 | 0.841 | 0.888 | 0.604 | 0.787 | 0.845 | 0.017 | 60 | 0.600 | 54.867 |
| VGGT | 0.825 | 0.889 | 0.909 | 0.800 | 0.877 | 0.908 | 0.744 | 0.842 | 0.875 | 0.016 | 60 | 0.930 | 21.490 |
| VGGT | 0.611 | 0.810 | 0.867 | 0.653 | 0.818 | 0.876 | 0.517 | 0.754 | 0.827 | 0.017 | 70 | 0.400 | 68.666 |
| VGGT | 0.732 | 0.855 | 0.891 | 0.739 | 0.853 | 0.895 | 0.644 | 0.804 | 0.855 | 0.017 | 70 | 0.700 | 49.820 |
| VGGT | 0.603 | 0.806 | 0.865 | 0.640 | 0.812 | 0.872 | 0.505 | 0.748 | 0.825 | 0.018 | 80 | 0.400 | 74.143 |
| VGGT | 0.645 | 0.823 | 0.875 | 0.671 | 0.825 | 0.880 | 0.549 | 0.766 | 0.835 | 0.017 | 80 | 0.500 | 67.759 |
| $\pi^3$ | 0.854 | 0.909 | 0.926 | 0.843 | 0.904 | 0.928 | 0.795 | 0.875 | 0.902 | 0.012 | 10 | 0.970 | 8.240 |
| $\pi^3$ | 0.851 | 0.908 | 0.926 | 0.841 | 0.903 | 0.928 | 0.791 | 0.873 | 0.902 | 0.012 | 30 | 0.970 | 13.130 |
| $\pi^3$ | 0.847 | 0.906 | 0.925 | 0.838 | 0.902 | 0.927 | 0.786 | 0.872 | 0.901 | 0.012 | 40 | 0.950 | 18.652 |
| $\pi^3$ | 0.789 | 0.886 | 0.915 | 0.806 | 0.891 | 0.921 | 0.721 | 0.849 | 0.889 | 0.012 | 50 | 0.700 | 46.717 |
| $\pi^3$ | 0.810 | 0.894 | 0.919 | 0.816 | 0.894 | 0.923 | 0.743 | 0.856 | 0.893 | 0.012 | 50 | 0.800 | 39.745 |
| $\pi^3$ | 0.821 | 0.898 | 0.921 | 0.822 | 0.896 | 0.924 | 0.755 | 0.861 | 0.895 | 0.012 | 50 | 0.850 | 34.638 |
| $\pi^3$ | 0.829 | 0.900 | 0.922 | 0.826 | 0.898 | 0.925 | 0.764 | 0.864 | 0.897 | 0.012 | 50 | 0.880 | 31.002 |
| $\pi^3$ | 0.834 | 0.902 | 0.923 | 0.830 | 0.899 | 0.926 | 0.770 | 0.866 | 0.898 | 0.012 | 50 | 0.900 | 28.288 |
| $\pi^3$ | 0.848 | 0.907 | 0.925 | 0.839 | 0.903 | 0.927 | 0.788 | 0.872 | 0.901 | 0.012 | 50 | 0.970 | 15.558 |
| $\pi^3$ | 0.737 | 0.868 | 0.905 | 0.777 | 0.880 | 0.916 | 0.666 | 0.829 | 0.879 | 0.012 | 60 | 0.400 | 59.683 |
| $\pi^3$ | 0.751 | 0.873 | 0.908 | 0.785 | 0.883 | 0.918 | 0.682 | 0.835 | 0.882 | 0.012 | 60 | 0.600 | 56.939 |
| $\pi^3$ | 0.840 | 0.904 | 0.924 | 0.833 | 0.901 | 0.927 | 0.778 | 0.869 | 0.900 | 0.012 | 60 | 0.930 | 24.825 |
| $\pi^3$ | 0.696 | 0.852 | 0.896 | 0.744 | 0.867 | 0.909 | 0.620 | 0.811 | 0.869 | 0.012 | 70 | 0.400 | 69.252 |
| $\pi^3$ | 0.773 | 0.878 | 0.910 | 0.792 | 0.885 | 0.919 | 0.704 | 0.841 | 0.885 | 0.012 | 70 | 0.700 | 53.773 |
| $\pi^3$ | 0.670 | 0.838 | 0.889 | 0.717 | 0.856 | 0.903 | 0.589 | 0.795 | 0.861 | 0.012 | 80 | 0.400 | 76.202 |
| $\pi^3$ | 0.701 | 0.851 | 0.897 | 0.740 | 0.865 | 0.909 | 0.622 | 0.808 | 0.869 | 0.012 | 80 | 0.500 | 70.919 |
| MapAnything | 0.690 | 0.834 | 0.877 | 0.683 | 0.824 | 0.879 | 0.582 | 0.772 | 0.836 | 0.016 | 10 | 0.970 | 9.785 |
| MapAnything | 0.687 | 0.832 | 0.876 | 0.682 | 0.822 | 0.877 | 0.580 | 0.771 | 0.835 | 0.016 | 30 | 0.970 | 24.555 |
| MapAnything | 0.685 | 0.832 | 0.876 | 0.680 | 0.821 | 0.876 | 0.578 | 0.770 | 0.834 | 0.016 | 40 | 0.950 | 32.998 |
| MapAnything | 0.683 | 0.832 | 0.877 | 0.678 | 0.821 | 0.877 | 0.574 | 0.769 | 0.835 | 0.016 | 50 | 0.700 | 49.660 |
| MapAnything | 0.682 | 0.831 | 0.876 | 0.677 | 0.820 | 0.876 | 0.573 | 0.768 | 0.834 | 0.016 | 50 | 0.800 | 47.981 |
| MapAnything | 0.681 | 0.830 | 0.876 | 0.677 | 0.820 | 0.876 | 0.573 | 0.768 | 0.834 | 0.016 | 50 | 0.850 | 46.257 |
| MapAnything | 0.681 | 0.830 | 0.876 | 0.677 | 0.820 | 0.876 | 0.574 | 0.767 | 0.833 | 0.016 | 50 | 0.880 | 44.772 |
| MapAnything | 0.681 | 0.830 | 0.876 | 0.677 | 0.820 | 0.876 | 0.574 | 0.767 | 0.833 | 0.016 | 50 | 0.900 | 43.486 |
| MapAnything | 0.686 | 0.830 | 0.875 | 0.680 | 0.820 | 0.875 | 0.579 | 0.768 | 0.833 | 0.016 | 50 | 0.970 | 34.473 |
| MapAnything | 0.677 | 0.830 | 0.877 | 0.673 | 0.820 | 0.876 | 0.568 | 0.767 | 0.835 | 0.016 | 60 | 0.400 | 60.019 |
| MapAnything | 0.678 | 0.831 | 0.878 | 0.674 | 0.820 | 0.877 | 0.568 | 0.767 | 0.835 | 0.016 | 60 | 0.600 | 59.707 |
| MapAnything | 0.677 | 0.828 | 0.875 | 0.674 | 0.819 | 0.875 | 0.570 | 0.765 | 0.833 | 0.016 | 60 | 0.930 | 45.707 |
| MapAnything | 0.667 | 0.826 | 0.874 | 0.666 | 0.817 | 0.875 | 0.558 | 0.763 | 0.832 | 0.016 | 70 | 0.400 | 70.026 |
| MapAnything | 0.666 | 0.826 | 0.874 | 0.665 | 0.817 | 0.875 | 0.557 | 0.762 | 0.831 | 0.016 | 70 | 0.700 | 65.915 |
| MapAnything | 0.651 | 0.819 | 0.870 | 0.653 | 0.810 | 0.869 | 0.542 | 0.754 | 0.825 | 0.017 | 80 | 0.400 | 79.583 |
| MapAnything | 0.652 | 0.819 | 0.869 | 0.653 | 0.810 | 0.870 | 0.542 | 0.753 | 0.825 | 0.017 | 80 | 0.500 | 78.170 |

Table A-11. Full results for Tanks & Temples [22] (100 frames).

| Model | RRA@5 | RRA@15 | RRA@30 | RTA@5 | RTA@15 | RTA@30 | AUC@5 | AUC@15 | AUC@30 | ATE | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VGGT | 0.829 | 0.887 | 0.907 | 0.798 | 0.873 | 0.904 | 0.747 | 0.839 | 0.871 | 0.012 | 10 | 0.970 | 7.273 |
| VGGT | 0.826 | 0.886 | 0.906 | 0.797 | 0.873 | 0.903 | 0.745 | 0.838 | 0.870 | 0.012 | 30 | 0.970 | 11.093 |
| VGGT | 0.821 | 0.885 | 0.905 | 0.794 | 0.872 | 0.903 | 0.739 | 0.836 | 0.869 | 0.012 | 40 | 0.950 | 15.976 |
| VGGT | 0.734 | 0.854 | 0.889 | 0.737 | 0.850 | 0.891 | 0.643 | 0.801 | 0.851 | 0.012 | 50 | 0.700 | 44.533 |
| VGGT | 0.773 | 0.868 | 0.896 | 0.763 | 0.860 | 0.896 | 0.685 | 0.817 | 0.859 | 0.012 | 50 | 0.800 | 36.316 |
| VGGT | 0.791 | 0.874 | 0.899 | 0.775 | 0.865 | 0.899 | 0.704 | 0.824 | 0.863 | 0.012 | 50 | 0.850 | 31.045 |
| VGGT | 0.800 | 0.878 | 0.901 | 0.781 | 0.867 | 0.900 | 0.715 | 0.828 | 0.864 | 0.012 | 50 | 0.880 | 27.458 |
| VGGT | 0.806 | 0.880 | 0.902 | 0.785 | 0.869 | 0.901 | 0.721 | 0.830 | 0.866 | 0.012 | 50 | 0.900 | 24.842 |
| VGGT | 0.825 | 0.886 | 0.905 | 0.796 | 0.873 | 0.903 | 0.743 | 0.837 | 0.870 | 0.012 | 50 | 0.970 | 13.124 |
| VGGT | 0.631 | 0.818 | 0.870 | 0.664 | 0.822 | 0.876 | 0.533 | 0.760 | 0.829 | 0.012 | 60 | 0.400 | 59.683 |
| VGGT | 0.689 | 0.838 | 0.880 | 0.704 | 0.838 | 0.885 | 0.594 | 0.782 | 0.841 | 0.012 | 60 | 0.600 | 54.867 |
| VGGT | 0.813 | 0.882 | 0.903 | 0.789 | 0.870 | 0.902 | 0.730 | 0.833 | 0.867 | 0.012 | 60 | 0.930 | 21.490 |
| VGGT | 0.603 | 0.807 | 0.865 | 0.642 | 0.812 | 0.872 | 0.503 | 0.747 | 0.823 | 0.012 | 70 | 0.400 | 68.666 |
| VGGT | 0.726 | 0.850 | 0.886 | 0.729 | 0.847 | 0.890 | 0.634 | 0.796 | 0.848 | 0.012 | 70 | 0.700 | 49.820 |
| VGGT | 0.593 | 0.803 | 0.862 | 0.628 | 0.806 | 0.868 | 0.490 | 0.740 | 0.819 | 0.013 | 80 | 0.400 | 74.143 |
| VGGT | 0.636 | 0.819 | 0.871 | 0.659 | 0.819 | 0.875 | 0.535 | 0.759 | 0.829 | 0.012 | 80 | 0.500 | 67.759 |
| $\pi^3$ | 0.851 | 0.906 | 0.924 | 0.838 | 0.900 | 0.924 | 0.788 | 0.870 | 0.900 | 0.009 | 10 | 0.970 | 8.240 |
| $\pi^3$ | 0.849 | 0.905 | 0.924 | 0.835 | 0.899 | 0.924 | 0.785 | 0.868 | 0.898 | 0.009 | 30 | 0.970 | 13.130 |
| $\pi^3$ | 0.845 | 0.903 | 0.923 | 0.831 | 0.897 | 0.923 | 0.779 | 0.866 | 0.897 | 0.009 | 40 | 0.950 | 18.652 |
| $\pi^3$ | 0.782 | 0.880 | 0.910 | 0.797 | 0.884 | 0.916 | 0.709 | 0.840 | 0.883 | 0.009 | 50 | 0.700 | 46.717 |
| $\pi^3$ | 0.806 | 0.889 | 0.915 | 0.808 | 0.889 | 0.919 | 0.735 | 0.850 | 0.888 | 0.009 | 50 | 0.800 | 39.745 |
| $\pi^3$ | 0.821 | 0.894 | 0.918 | 0.817 | 0.892 | 0.920 | 0.751 | 0.856 | 0.891 | 0.009 | 50 | 0.850 | 34.638 |
| $\pi^3$ | 0.828 | 0.897 | 0.919 | 0.821 | 0.894 | 0.921 | 0.760 | 0.859 | 0.893 | 0.008 | 50 | 0.880 | 31.002 |
| $\pi^3$ | 0.833 | 0.899 | 0.920 | 0.824 | 0.895 | 0.922 | 0.765 | 0.861 | 0.894 | 0.008 | 50 | 0.900 | 28.288 |
| $\pi^3$ | 0.847 | 0.904 | 0.923 | 0.834 | 0.898 | 0.924 | 0.783 | 0.867 | 0.898 | 0.009 | 50 | 0.970 | 15.558 |
| $\pi^3$ | 0.737 | 0.864 | 0.901 | 0.770 | 0.875 | 0.912 | 0.661 | 0.823 | 0.874 | 0.008 | 60 | 0.400 | 59.683 |
| $\pi^3$ | 0.751 | 0.869 | 0.904 | 0.779 | 0.878 | 0.913 | 0.676 | 0.828 | 0.877 | 0.008 | 60 | 0.600 | 56.939 |
| $\pi^3$ | 0.838 | 0.900 | 0.921 | 0.828 | 0.896 | 0.922 | 0.772 | 0.863 | 0.895 | 0.008 | 60 | 0.930 | 24.825 |
| $\pi^3$ | 0.699 | 0.851 | 0.896 | 0.741 | 0.865 | 0.906 | 0.617 | 0.807 | 0.867 | 0.008 | 70 | 0.400 | 69.252 |
| $\pi^3$ | 0.772 | 0.876 | 0.908 | 0.787 | 0.881 | 0.915 | 0.697 | 0.835 | 0.880 | 0.008 | 70 | 0.700 | 53.773 |
| $\pi^3$ | 0.667 | 0.837 | 0.888 | 0.713 | 0.852 | 0.899 | 0.581 | 0.790 | 0.857 | 0.009 | 80 | 0.400 | 76.202 |
| $\pi^3$ | 0.698 | 0.849 | 0.894 | 0.736 | 0.861 | 0.904 | 0.614 | 0.804 | 0.864 | 0.008 | 80 | 0.500 | 70.919 |
| MapAnything | 0.679 | 0.824 | 0.870 | 0.667 | 0.812 | 0.870 | 0.567 | 0.759 | 0.826 | 0.011 | 10 | 0.970 | 9.785 |
| MapAnything | 0.679 | 0.824 | 0.870 | 0.667 | 0.812 | 0.870 | 0.566 | 0.759 | 0.827 | 0.011 | 30 | 0.970 | 24.555 |
| MapAnything | 0.677 | 0.823 | 0.869 | 0.666 | 0.812 | 0.870 | 0.564 | 0.758 | 0.826 | 0.011 | 40 | 0.950 | 32.998 |
| MapAnything | 0.673 | 0.821 | 0.868 | 0.664 | 0.811 | 0.870 | 0.560 | 0.756 | 0.825 | 0.011 | 50 | 0.700 | 49.660 |
| MapAnything | 0.673 | 0.821 | 0.868 | 0.664 | 0.811 | 0.869 | 0.560 | 0.756 | 0.825 | 0.011 | 50 | 0.800 | 47.981 |
| MapAnything | 0.673 | 0.821 | 0.867 | 0.663 | 0.811 | 0.869 | 0.560 | 0.756 | 0.824 | 0.011 | 50 | 0.850 | 46.257 |
| MapAnything | 0.673 | 0.821 | 0.868 | 0.663 | 0.811 | 0.869 | 0.560 | 0.756 | 0.824 | 0.011 | 50 | 0.880 | 44.772 |
| MapAnything | 0.673 | 0.821 | 0.868 | 0.663 | 0.811 | 0.869 | 0.560 | 0.756 | 0.824 | 0.011 | 50 | 0.900 | 43.486 |
| MapAnything | 0.676 | 0.822 | 0.868 | 0.665 | 0.811 | 0.869 | 0.564 | 0.757 | 0.825 | 0.011 | 50 | 0.970 | 34.473 |
| MapAnything | 0.666 | 0.818 | 0.866 | 0.660 | 0.809 | 0.868 | 0.555 | 0.754 | 0.823 | 0.011 | 60 | 0.400 | 60.019 |
| MapAnything | 0.666 | 0.818 | 0.866 | 0.660 | 0.809 | 0.868 | 0.555 | 0.753 | 0.823 | 0.011 | 60 | 0.600 | 59.707 |
| MapAnything | 0.669 | 0.819 | 0.866 | 0.661 | 0.809 | 0.868 | 0.558 | 0.754 | 0.822 | 0.011 | 60 | 0.930 | 45.707 |
| MapAnything | 0.660 | 0.816 | 0.863 | 0.653 | 0.806 | 0.865 | 0.547 | 0.749 | 0.819 | 0.012 | 70 | 0.400 | 70.026 |
| MapAnything | 0.658 | 0.814 | 0.863 | 0.652 | 0.804 | 0.864 | 0.546 | 0.748 | 0.818 | 0.012 | 70 | 0.700 | 65.915 |
| MapAnything | 0.644 | 0.804 | 0.854 | 0.638 | 0.795 | 0.856 | 0.532 | 0.738 | 0.809 | 0.013 | 80 | 0.400 | 79.583 |
| MapAnything | 0.643 | 0.803 | 0.853 | 0.638 | 0.794 | 0.856 | 0.532 | 0.737 | 0.809 | 0.014 | 80 | 0.500 | 78.170 |

Table A-12. Full results for Tanks & Temples [22] (200 frames).

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.183 | 0.117 | 4.632 | 10 | 0.970 | 6.879 |
| VGGT | 0.184 | 0.118 | 4.659 | 30 | 0.970 | 10.103 |
| VGGT | 0.185 | 0.118 | 4.709 | 40 | 0.950 | 14.749 |
| VGGT | 0.192 | 0.116 | 4.851 | 50 | 0.700 | 44.131 |
| VGGT | 0.190 | 0.119 | 4.739 | 50 | 0.800 | 35.240 |
| VGGT | 0.191 | 0.117 | 4.801 | 50 | 0.850 | 29.774 |
| VGGT | 0.191 | 0.116 | 4.711 | 50 | 0.880 | 26.108 |
| VGGT | 0.191 | 0.117 | 4.745 | 50 | 0.900 | 23.455 |
| VGGT | 0.189 | 0.116 | 4.796 | 50 | 0.970 | 11.921 |
| VGGT | 0.199 | 0.120 | 5.231 | 60 | 0.400 | 59.727 |
| VGGT | 0.197 | 0.122 | 5.158 | 60 | 0.600 | 54.648 |
| VGGT | 0.190 | 0.120 | 4.697 | 60 | 0.930 | 19.995 |
| VGGT | 0.202 | 0.113 | 5.071 | 70 | 0.700 | 48.770 |
| VGGT | 0.236 | 0.135 | 5.970 | 80 | 0.400 | 74.039 |
| VGGT | 0.228 | 0.124 | 5.707 | 80 | 0.500 | 67.357 |
| $\pi^3$ | 0.150 | 0.081 | 3.999 | 10 | 0.970 | 7.972 |
| $\pi^3$ | 0.150 | 0.081 | 4.000 | 30 | 0.970 | 12.738 |
| $\pi^3$ | 0.150 | 0.081 | 4.001 | 40 | 0.950 | 18.160 |
| $\pi^3$ | 0.152 | 0.082 | 4.033 | 50 | 0.700 | 46.446 |
| $\pi^3$ | 0.151 | 0.081 | 4.019 | 50 | 0.800 | 39.029 |
| $\pi^3$ | 0.151 | 0.081 | 4.014 | 50 | 0.850 | 33.866 |
| $\pi^3$ | 0.151 | 0.081 | 4.009 | 50 | 0.880 | 30.267 |
| $\pi^3$ | 0.151 | 0.081 | 4.007 | 50 | 0.900 | 27.614 |
| $\pi^3$ | 0.150 | 0.081 | 4.002 | 50 | 0.970 | 15.467 |
| $\pi^3$ | 0.155 | 0.083 | 4.071 | 60 | 0.400 | 59.727 |
| $\pi^3$ | 0.155 | 0.083 | 4.061 | 60 | 0.600 | 56.759 |
| $\pi^3$ | 0.152 | 0.081 | 4.012 | 60 | 0.930 | 24.466 |
| $\pi^3$ | 0.159 | 0.083 | 4.076 | 70 | 0.700 | 53.153 |
| $\pi^3$ | 0.183 | 0.091 | 4.238 | 80 | 0.400 | 76.082 |
| $\pi^3$ | 0.180 | 0.090 | 4.227 | 80 | 0.500 | 70.609 |
| MapAnything | 0.242 | 0.139 | 6.021 | 10 | 0.970 | 9.648 |
| MapAnything | 0.245 | 0.140 | 6.145 | 30 | 0.970 | 23.704 |
| MapAnything | 0.246 | 0.140 | 6.091 | 40 | 0.950 | 31.897 |
| MapAnything | 0.246 | 0.140 | 6.261 | 50 | 0.700 | 49.465 |
| MapAnything | 0.246 | 0.140 | 6.127 | 50 | 0.800 | 47.159 |
| MapAnything | 0.246 | 0.140 | 6.234 | 50 | 0.850 | 45.141 |
| MapAnything | 0.246 | 0.140 | 6.256 | 50 | 0.880 | 43.532 |
| MapAnything | 0.246 | 0.140 | 6.217 | 50 | 0.900 | 42.204 |
| MapAnything | 0.247 | 0.140 | 6.111 | 50 | 0.970 | 33.467 |
| MapAnything | 0.249 | 0.141 | 6.106 | 60 | 0.400 | 60.016 |
| MapAnything | 0.249 | 0.141 | 6.269 | 60 | 0.600 | 59.550 |
| MapAnything | 0.249 | 0.141 | 6.315 | 60 | 0.930 | 44.385 |
| MapAnything | 0.254 | 0.142 | 6.095 | 70 | 0.700 | 64.838 |
| MapAnything | 0.265 | 0.145 | 6.538 | 80 | 0.400 | 79.337 |
| MapAnything | 0.264 | 0.145 | 6.562 | 80 | 0.500 | 77.571 |

Table A-13. Full results for ScanNet [7] (100 frames evenly sampled from full sequence).

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.176 | 0.069 | 2.485 | 10 | 0.970 | 6.879 |
| VGGT | 0.173 | 0.061 | 2.383 | 30 | 0.970 | 10.103 |
| VGGT | 0.173 | 0.060 | 2.386 | 40 | 0.950 | 14.749 |
| VGGT | 0.181 | 0.076 | 2.576 | 50 | 0.700 | 44.131 |
| VGGT | 0.178 | 0.069 | 2.522 | 50 | 0.800 | 35.240 |
| VGGT | 0.175 | 0.063 | 2.424 | 50 | 0.850 | 29.774 |
| VGGT | 0.175 | 0.064 | 2.417 | 50 | 0.880 | 26.108 |
| VGGT | 0.174 | 0.064 | 2.428 | 50 | 0.900 | 23.455 |
| VGGT | 0.173 | 0.060 | 2.396 | 50 | 0.970 | 11.921 |
| VGGT | 0.195 | 0.083 | 2.830 | 60 | 0.400 | 59.727 |
| VGGT | 0.190 | 0.077 | 2.649 | 60 | 0.600 | 54.648 |
| VGGT | 0.178 | 0.070 | 2.504 | 60 | 0.930 | 19.995 |
| VGGT | 0.200 | 0.078 | 2.834 | 70 | 0.700 | 48.770 |
| VGGT | 0.248 | 0.085 | 3.601 | 80 | 0.400 | 74.039 |
| VGGT | 0.237 | 0.078 | 3.258 | 80 | 0.500 | 67.357 |
| $\pi^3$ | 0.150 | 0.041 | 1.994 | 10 | 0.970 | 7.972 |
| $\pi^3$ | 0.150 | 0.041 | 1.994 | 30 | 0.970 | 12.738 |
| $\pi^3$ | 0.150 | 0.041 | 1.994 | 40 | 0.950 | 18.160 |
| $\pi^3$ | 0.151 | 0.042 | 2.009 | 50 | 0.700 | 46.446 |
| $\pi^3$ | 0.150 | 0.041 | 2.003 | 50 | 0.800 | 39.029 |
| $\pi^3$ | 0.150 | 0.041 | 2.000 | 50 | 0.850 | 33.866 |
| $\pi^3$ | 0.150 | 0.041 | 1.999 | 50 | 0.880 | 30.267 |
| $\pi^3$ | 0.150 | 0.041 | 1.998 | 50 | 0.900 | 27.614 |
| $\pi^3$ | 0.150 | 0.041 | 1.994 | 50 | 0.970 | 15.467 |
| $\pi^3$ | 0.155 | 0.042 | 2.034 | 60 | 0.400 | 59.727 |
| $\pi^3$ | 0.154 | 0.042 | 2.030 | 60 | 0.600 | 56.759 |
| $\pi^3$ | 0.151 | 0.041 | 2.000 | 60 | 0.930 | 24.466 |
| $\pi^3$ | 0.158 | 0.043 | 2.033 | 70 | 0.700 | 53.153 |
| $\pi^3$ | 0.181 | 0.046 | 2.097 | 80 | 0.400 | 76.082 |
| $\pi^3$ | 0.179 | 0.045 | 2.088 | 80 | 0.500 | 70.609 |
| MapAnything | 0.239 | 0.088 | 3.596 | 10 | 0.970 | 9.648 |
| MapAnything | 0.243 | 0.089 | 3.742 | 30 | 0.970 | 23.704 |
| MapAnything | 0.247 | 0.089 | 3.591 | 40 | 0.950 | 31.897 |
| MapAnything | 0.249 | 0.086 | 3.508 | 50 | 0.700 | 49.465 |
| MapAnything | 0.249 | 0.086 | 3.543 | 50 | 0.800 | 47.159 |
| MapAnything | 0.250 | 0.086 | 3.440 | 50 | 0.850 | 45.141 |
| MapAnything | 0.250 | 0.086 | 3.430 | 50 | 0.880 | 43.532 |
| MapAnything | 0.250 | 0.086 | 3.556 | 50 | 0.900 | 42.204 |
| MapAnything | 0.252 | 0.088 | 3.511 | 50 | 0.970 | 33.467 |
| MapAnything | 0.251 | 0.085 | 3.514 | 60 | 0.400 | 60.016 |
| MapAnything | 0.251 | 0.085 | 3.506 | 60 | 0.600 | 59.550 |
| MapAnything | 0.255 | 0.087 | 3.492 | 60 | 0.930 | 44.385 |
| MapAnything | 0.259 | 0.084 | 3.654 | 70 | 0.700 | 64.838 |
| MapAnything | 0.268 | 0.087 | 3.832 | 80 | 0.400 | 79.337 |
| MapAnything | 0.268 | 0.087 | 3.949 | 80 | 0.500 | 77.571 |

Table A-14. Full results for ScanNet [7] (300 frames evenly sampled from full sequence).

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.178 | 0.056 | 1.814 | 10 | 0.970 | 6.879 |
| VGGT | 0.180 | 0.060 | 1.868 | 30 | 0.970 | 10.103 |
| VGGT | 0.177 | 0.053 | 1.805 | 40 | 0.950 | 14.749 |
| VGGT | 0.186 | 0.063 | 2.274 | 50 | 0.700 | 44.131 |
| VGGT | 0.183 | 0.061 | 2.010 | 50 | 0.800 | 35.240 |
| VGGT | 0.182 | 0.061 | 1.923 | 50 | 0.850 | 29.774 |
| VGGT | 0.183 | 0.066 | 2.098 | 50 | 0.880 | 26.108 |
| VGGT | 0.178 | 0.056 | 1.993 | 50 | 0.900 | 23.455 |
| VGGT | 0.180 | 0.060 | 1.900 | 50 | 0.970 | 11.921 |
| VGGT | 0.196 | 0.068 | 2.237 | 60 | 0.400 | 59.727 |
| VGGT | 0.191 | 0.063 | 2.078 | 60 | 0.600 | 54.648 |
| VGGT | 0.180 | 0.059 | 1.905 | 60 | 0.930 | 19.995 |
| VGGT | 0.199 | 0.068 | 2.260 | 70 | 0.700 | 48.770 |
| VGGT | 0.248 | 0.069 | 2.793 | 80 | 0.400 | 74.039 |
| VGGT | 0.238 | 0.070 | 2.594 | 80 | 0.500 | 67.357 |
| $\pi^3$ | 0.150 | 0.030 | 1.406 | 10 | 0.970 | 7.972 |
| $\pi^3$ | 0.149 | 0.030 | 1.405 | 30 | 0.970 | 12.738 |
| $\pi^3$ | 0.150 | 0.030 | 1.405 | 40 | 0.950 | 18.160 |
| $\pi^3$ | 0.151 | 0.031 | 1.416 | 50 | 0.700 | 46.446 |
| $\pi^3$ | 0.150 | 0.031 | 1.411 | 50 | 0.800 | 39.029 |
| $\pi^3$ | 0.150 | 0.031 | 1.409 | 50 | 0.850 | 33.866 |
| $\pi^3$ | 0.150 | 0.030 | 1.408 | 50 | 0.880 | 30.267 |
| $\pi^3$ | 0.150 | 0.030 | 1.407 | 50 | 0.900 | 27.614 |
| $\pi^3$ | 0.150 | 0.030 | 1.405 | 50 | 0.970 | 15.467 |
| $\pi^3$ | 0.160 | 0.031 | 1.436 | 60 | 0.400 | 59.727 |
| $\pi^3$ | 0.157 | 0.031 | 1.430 | 60 | 0.600 | 56.759 |
| $\pi^3$ | 0.151 | 0.030 | 1.408 | 60 | 0.930 | 24.466 |
| $\pi^3$ | 0.158 | 0.032 | 1.433 | 70 | 0.700 | 53.153 |
| $\pi^3$ | 0.181 | 0.034 | 1.483 | 80 | 0.400 | 76.082 |
| $\pi^3$ | 0.179 | 0.034 | 1.474 | 80 | 0.500 | 70.609 |
| MapAnything | 0.256 | 0.079 | 3.128 | 10 | 0.970 | 9.648 |
| MapAnything | 0.259 | 0.079 | 3.115 | 30 | 0.970 | 23.704 |
| MapAnything | 0.261 | 0.078 | 3.092 | 40 | 0.950 | 31.897 |
| MapAnything | 0.258 | 0.074 | 3.211 | 50 | 0.700 | 49.465 |
| MapAnything | 0.259 | 0.074 | 3.238 | 50 | 0.800 | 47.159 |
| MapAnything | 0.260 | 0.075 | 3.108 | 50 | 0.850 | 45.141 |
| MapAnything | 0.261 | 0.075 | 3.256 | 50 | 0.880 | 43.532 |
| MapAnything | 0.262 | 0.076 | 3.095 | 50 | 0.900 | 42.204 |
| MapAnything | 0.265 | 0.078 | 3.199 | 50 | 0.970 | 33.467 |
| MapAnything | 0.260 | 0.073 | 3.030 | 60 | 0.400 | 60.016 |
| MapAnything | 0.260 | 0.073 | 3.033 | 60 | 0.600 | 59.550 |
| MapAnything | 0.267 | 0.076 | 3.146 | 60 | 0.930 | 44.385 |
| MapAnything | 0.268 | 0.073 | 3.246 | 70 | 0.700 | 64.838 |
| MapAnything | 0.275 | 0.074 | 3.241 | 80 | 0.400 | 79.337 |
| MapAnything | 0.275 | 0.075 | 3.303 | 80 | 0.500 | 77.571 |

Table A-15. Full results for ScanNet [7] (500 frames evenly sampled from full sequence).

| Model | ATE | RPEtrans | RPErot | $\rho$ (%) | CDF | Sparsity |
|---|---|---|---|---|---|---|
| VGGT | 0.204 | 0.056 | 1.587 | 10 | 0.970 | 6.879 |
| VGGT | 0.203 | 0.058 | 1.608 | 30 | 0.970 | 10.103 |
| VGGT | 0.206 | 0.061 | 1.678 | 40 | 0.950 | 14.749 |
| VGGT | 0.214 | 0.060 | 1.892 | 50 | 0.700 | 44.131 |
| VGGT | 0.213 | 0.059 | 1.770 | 50 | 0.800 | 35.240 |
| VGGT | 0.213 | 0.061 | 1.876 | 50 | 0.850 | 29.774 |
| VGGT | 0.212 | 0.062 | 1.899 | 50 | 0.880 | 26.108 |
| VGGT | 0.211 | 0.060 | 1.822 | 50 | 0.900 | 23.455 |
| VGGT | 0.204 | 0.059 | 1.742 | 50 | 0.970 | 11.921 |
| VGGT | 0.229 | 0.068 | 2.277 | 60 | 0.400 | 59.727 |
| VGGT | 0.226 | 0.067 | 2.218 | 60 | 0.600 | 54.648 |
| VGGT | 0.210 | 0.058 | 1.660 | 60 | 0.930 | 19.995 |
| VGGT | 0.241 | 0.065 | 2.337 | 70 | 0.700 | 48.770 |
| VGGT | 0.284 | 0.067 | 3.361 | 80 | 0.400 | 74.039 |
| VGGT | 0.277 | 0.069 | 2.658 | 80 | 0.500 | 67.357 |
| $\pi^3$ | 0.150 | 0.021 | 0.895 | 10 | 0.970 | 7.972 |
| $\pi^3$ | 0.150 | 0.021 | 0.896 | 30 | 0.970 | 12.738 |
| $\pi^3$ | 0.150 | 0.021 | 0.896 | 40 | 0.950 | 18.160 |
| $\pi^3$ | 0.151 | 0.022 | 0.908 | 50 | 0.700 | 46.446 |
| $\pi^3$ | 0.150 | 0.022 | 0.902 | 50 | 0.800 | 39.029 |
| $\pi^3$ | 0.150 | 0.021 | 0.900 | 50 | 0.850 | 33.866 |
| $\pi^3$ | 0.150 | 0.021 | 0.899 | 50 | 0.880 | 30.267 |
| $\pi^3$ | 0.150 | 0.021 | 0.898 | 50 | 0.900 | 27.614 |
| $\pi^3$ | 0.150 | 0.021 | 0.896 | 50 | 0.970 | 15.467 |
| $\pi^3$ | 0.154 | 0.022 | 0.925 | 60 | 0.400 | 59.727 |
| $\pi^3$ | 0.154 | 0.022 | 0.924 | 60 | 0.600 | 56.759 |
| $\pi^3$ | 0.151 | 0.021 | 0.898 | 60 | 0.930 | 24.466 |
| $\pi^3$ | 0.158 | 0.022 | 0.922 | 70 | 0.700 | 53.153 |
| $\pi^3$ | 0.180 | 0.025 | 0.968 | 80 | 0.400 | 76.082 |
| $\pi^3$ | 0.179 | 0.024 | 0.959 | 80 | 0.500 | 70.609 |
| MapAnything | 0.285 | 0.075 | 3.040 | 10 | 0.970 | 9.648 |
| MapAnything | 0.285 | 0.074 | 2.945 | 30 | 0.970 | 23.704 |
| MapAnything | 0.286 | 0.074 | 3.053 | 40 | 0.950 | 31.897 |
| MapAnything | 0.290 | 0.069 | 2.895 | 50 | 0.700 | 49.465 |
| MapAnything | 0.290 | 0.070 | 2.831 | 50 | 0.800 | 47.159 |
| MapAnything | 0.289 | 0.071 | 3.019 | 50 | 0.850 | 45.141 |
| MapAnything | 0.288 | 0.072 | 2.928 | 50 | 0.880 | 43.532 |
| MapAnything | 0.288 | 0.072 | 2.926 | 50 | 0.900 | 42.204 |
| MapAnything | 0.289 | 0.074 | 3.301 | 50 | 0.970 | 33.467 |
| MapAnything | 0.296 | 0.069 | 2.928 | 60 | 0.400 | 60.016 |
| MapAnything | 0.296 | 0.069 | 2.854 | 60 | 0.600 | 59.550 |
| MapAnything | 0.291 | 0.073 | 2.972 | 60 | 0.930 | 44.385 |
| MapAnything | 0.300 | 0.073 | 3.181 | 70 | 0.700 | 64.838 |
| MapAnything | 0.297 | 0.071 | 3.147 | 80 | 0.400 | 79.337 |
| MapAnything | 0.299 | 0.072 | 3.170 | 80 | 0.500 | 77.571 |

Table A-16. Full results for ScanNet [7] (1000 frames evenly sampled from full sequence).