— Supplementary Material —
Global-Aware Edge Prioritization for Pose Graph Initialization

Tong Wei[1], Giorgos Tolias[1], Jiří Matas[1], Daniel Barath[2,3]

[1] Visual Recognition Group, FEE, Czech Technical University in Prague

[2] Computer Vision and Geometry Group, ETH Zurich, [3] HUN-REN SZTAKI

{weitong, toliageo, matas}@fel.cvut.cz, {danielbela.barath}@inf.ethz.ch

1. Node Prioritization for VGGT

Our proposed global edge prioritization ranks all potential image pairs and utilize most promising edges to improve COLMAP reconstruction. To further extend its applicability, we aggregate edge-level scores into node-level scores, which can be used to filter out outlier viewpoints for VGGT. We determine the outlier image set in two steps based on the predicted scores. First, we take the top-30 images per row and count the frequency of all images. Images that appear fewer than once are selected as candidates. Second, all the rows on the selected images from the first step are averaged and used as a metric to finalize the outlier set. If the mean score of an image exceeds a fixed threshold, ie, 0.2, it is excluded from the outlier set. The remaining outliers are filtered out and passed to VGGT for reconstruction.

In the table below, AUC scores and the median of relative pose, rotation, and translation errors on all/filtered pairs by ours are shown, on 10 scenes from IMC23 (each containing fewer than 400 images due to CUDA memory constraints). The evaluation is done on the filtered image sets (first two rows) and the full (last two rows). Our method consistently improves the accuracy of the reconstructed cameras.

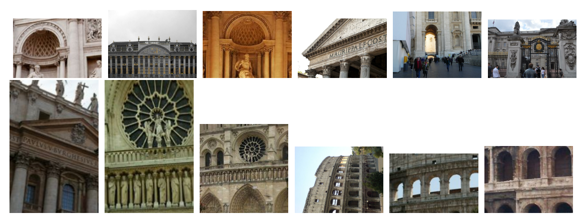
	AUC@2.5° ↑	AUC@5° ↑	Pose err.↓	R err.↓	t err.↓
Evaluation on Filtered IMC23 (10 scenes)					
VGGT	34.5	52.0	1.94	0.65	1.69
VGGT + Ours	**36.6**	**53.9**	**1.78**	**0.59**	**1.55**
Evaluation on IMC23 (10 scenes)					
VGGT	34.3	51.7	1.95	0.66	1.71
VGGT + Ours	**35.8**	**52.8**	**1.83**	**0.60**	**1.60**

2. Graph Clustering

Graph clustering is applied to large-scale image collections based on the intermediate tokens output from our trained encoder. After it is applied, the matchability scores are predicted by GNN and averaged on the overlapping edges among different subgraphs. In our setup, the clustering is only applied on scenes with more than 500 images. We tried applying the clustering approach applied to all scenes (including small ones), leading to competitive AUCs.

3. Failure Cases

In the figure below, failure examples of MegaLoc (unregistered cameras), registered *successfully* by our model, are shown in the first row. In the second row, there are failure cases of both MegaLoc and ours (low resolution, small landmark part shown).



4. KNN-based Image Pair Selection

Except for running COLMAP on MST pairs, we measure AUC@5° on IMC23 for MegaLoc and our method using kNN-based image-pair selection. As shown in the table below, our method outperforms the pre-trained model by a large margin on most values of k, with an exception at $k = 2$.

Method / kNNs	k=1	k=2	k=3	k=5
MegaLoc	1.2	**63.9**	64.0	66.5
Ours	**5.3**	61.2	**71.1**	**73.1**

5. Full Run-time Comparison

The table below reports the timing (*in seconds*) for all steps of Ours and of MegaLoc, averaged on IMC23. Image pairs selected by two MSTs are fed in SfM. Edge prediction for ours is slower than MegaLoc due to GNN, but this step takes negligible time compared to COLMAP. Note that COLMAP typically runs faster for ours due to better pair selection.

Method	Encoder	Predictor	MST	COLMAP
MegaLoc	2.91	**0.01**	**0.25**	2.3k
Ours		0.08	0.30	**2.1k**

6. Backbone Discussion

In the main paper, we have shown the performance of the model trained with MegaLoc and SALAD. To show how the model benefits from GNN predictor, we performed an experiment using DINOv2 (finetuning last 4 layers) and our GNN edge matchability predictor on top. We observe that the GNN improves AUC@5° on IMC23 compared to the pretrained DINOv2 model.

Method / k-MSTs	k=1	k=2	k=3	k=5
pretrained DINOv2	56.0	67.4	70.1	72.4
DINOv2 + GNN	**60.4**	**68.8**	**72.5**	**73.5**