Leveraging Camera Triplets for Efficient and Accurate Structure-from-Motion Supplementary Material

Lalit Manam and Venu Madhav Govindu
Indian Institute of Science
Bengaluru - 560012, INDIA
{lalitmanam,venug}@iisc.ac.in

A. Proof for Theorem

In Sec. 3 of the main paper, we stated Thm. 1. We first rewrite the optimization problem and the thresholding scheme here for reference.

$$\max_{\substack{s_{ij}\in\{0,1\},\\(i,j)\in\mathcal{E}}} \frac{\sum_{(i,j)\in\mathcal{E}} s_{ij}q_{ij}}{\sum_{(i,j)\in\mathcal{E}} s_{ij}} - \lambda \frac{\sum_{(i,j)\in\mathcal{E}}(1-s_{ij})q_{ij}}{\sum_{(i,j)\in\mathcal{E}}(1-s_{ij})}, \quad \text{(S1)}$$

$$s_{ij} = \begin{cases} 1 & \text{if } q_{ij} \geq \tau, \\ 0 & \text{otherwise,} \end{cases} \quad \text{(S2)}$$

where $\lambda \geq 0$ is a regularization parameter and τ denotes a threshold. Now, we provide the proof for Thm. 1.

Theorem 1. *For a given λ in Eqn. S1, there exists a threshold τ such that the values of s_{ij} obtained by solving the problems given by Eqn. S1 and Eqn. S2 are the same.*

Proof. In Eqn. S1, the first term $\frac{\sum_{(i,j)\in\mathcal{E}} s_{ij}q_{ij}}{\sum_{(i,j)\in\mathcal{E}} s_{ij}}$, is the average score of the retained edges and the second term, $\frac{\sum_{(i,j)\in\mathcal{E}}(1-s_{ij})q_{ij}}{\sum_{(i,j)\in\mathcal{E}}(1-s_{ij})}$, is the average score of the removed edges. We use the fact that given a set of numbers, removing numbers from the set less than their average or adding numbers to the set greater than their average increases the average of the final set.

The problem in Eqn. S1 is to maximize the function by improving the average score of the retained edges (first term) with the low average score of the removed edges (second term). This is achieved when the edges with top k scores are retained while others are removed. This is because it leads to the maximum possible increase in the first term and the minimum possible increase in the second term. Here, k is dependent on the value of λ. This is the same as thresholding the edges based on scores with $sq_{k+1} < \tau \leq sq_k$, where $sq_z, z = \{1, 2, \cdots, |\mathcal{E}|\}$ are the sorted scores q_{ij} in descending order. ∎

B. Additional Results

In Sec. 4 of the main paper, we provided reconstruction statistics and visual results on generic and ambiguous datasets. Here, we provide details of the datasets (for the largest connected component) along with additional results on them. We restate notations for the graphs below and introduce new notations which will be used further.

- $\mathcal{G}$: Original viewgraph.
- $\mathcal{G}_{LCT}$: Graph containing edges contributing to the largest connected component of the triplet graph $\mathcal{G}_T$.
- $\mathcal{G}_{Dopp}$: Graph obtained after applying Doppelgangers [3] on the original graph $\mathcal{G}$.
- $\mathcal{G}_F(m)$: Graph obtained after applying our method (Algo. 1) with minimum edge score m.
- $|\mathcal{V}_{sub}|$: Number of nodes in the graph $\mathcal{G}_{sub}$, where sub is one of the subscripts of the graphs used above.
- $|\mathcal{E}_{sub}|$: Number of edges in the graph $\mathcal{G}_{sub}$, where sub is one of the subscripts of the graphs used above.

B.1. Generic Datasets

In Sec. 4.1 of the main paper, we presented reconstruction results on generic datasets. Here, we provide details on different aspects in the subsequent paragraphs.

Dataset details: In Table S1, we show the number of nodes and edges present in the datasets after applying our method with different values of minimum edge score $m = \{0.6, 0.7, 0.8, 0.9\}$. It can be seen that, after sparsifying the graphs with our method, the number of edges is 10% to 30% of the original graphs ($\mathcal{G}$). The reduction of edges in the sparsified graphs is similar when compared to the graphs consisting of triplets ($\mathcal{G}_{LCT}$). Moreover, there is a steady decrease in the number of edges with an increase in m, but the number of nodes abruptly reduces for $m = \{0.8, 0.9\}$. This reveals that the original graphs are well connected, due to which removing edges with smaller values of m did not lead to a drastic reduction in the number of nodes. For higher values of m, the graphs become sparse, leading to a substantial reduction in the number of nodes.

Reconstruction details: In Table S2, we show the reconstruction statistics of the graphs obtained with different values of m using COLMAP [36]. It can be seen that most of the cameras and 3D points are reconstructed with the filtered graphs ($\mathcal{G}_F(m)$) compared to the original graph ($\mathcal{G}$). Moreover, the number of 3D points reconstructed does not decrease drastically even if the number of nodes reconstructed reduces considerably for $m = \{0.8, 0.9\}$. This indicates that most parts of the reconstructions are still recovered even after the removal of many nodes. This shows the advantage of obtaining edge scores based on the connectivity of 3D points. We also observe that applying our method leads to better-quality reconstructions by avoiding ghost artifacts.

In Table S3, we provide the mean reprojection error for the reconstructed 3D points. It can be seen that reprojection errors reduce with an increase in m, revealing that the camera parameters (intrinsics and motion parameters) and 3D points are more consistent to the epipolar inliers in the sparsified graphs. We also compare the camera motions of different graphs, keeping the motions obtained from the original graphs as a reference. We only compare those datasets where ghost artifacts are not found in the original graphs. This ensures that the reference camera motions are reliable. It can be seen that the mean rotation difference is less than $1°$ for almost all the datasets. Also, the camera translations obtained from different graphs are close to the original graphs. This shows that the accuracy of the reconstructions is maintained after the sparsification of graphs.

Time taken: In Table S4, we show the reconstruction time taken by COLMAP [36] on sparsified graphs. It can be seen for all the datasets, reconstruction times reduce by a margin of 50% to 80% for the sparsified graphs, even for lower values of m, when compared to the original graphs.

In Table S5, we show the time taken for our method. We refer to preprocessing time as time taken for Steps 1-3 and filter time for Steps 4-11 in Algo. 1. It can be seen that for most of the datasets, our method takes less than 1% of the time taken for reconstruction (Table S4). The Rome dataset is very well connected, due to which the triplets are large in number, resulting in a longer computation time for our method.

Visual results: In Figs. S1, S2, S3, and S4, we provide visual results on generic datasets with varying values of the minimum edge score m. From Figs. S1, S2, and S3, it can be seen that reconstructions obtained after sparsifying the graphs are visually similar to that of the original graphs even after losing some nodes with faster reconstruction times. Moreover, the main structure in the scenes is recovered for reconstructions of sparsified graphs even for a high value of $m = 0.9$ for most datasets. This is a consequence of scoring edges based on 3D point connectivity. For Quad, the cameras are not well connected, due to which only a small part of the reconstruction of the original graph is recovered in the sparsified graph with $m = 0.9$ ($\mathcal{G}_F(0.9)$). In Fig. S4, we see that our method leads to improved reconstruction quality by avoiding ghost artifacts.

Choosing minimum edge score: Given the results shown above, we recommend setting the value of minimum edge score m between 0.6 and 0.7 for graph sparsification on generic datasets, keeping a trade-off between reconstruction quality and time.

B.2. Ambiguous Datasets

In Sec. 4.2 of the main paper, we presented reconstruction results on ambiguous datasets. Here, we provide the results after applying our method for different values of minimum edge score m. For large-scale datasets, we show for $m = \{0.6, 0.7, 0.8, 0.9\}$ and for medium and small-scale datasets, we show for $m = \{0.3, 0.4, 0.5, 0.6\}$. We also provide results from Doppelgangers [3], where we use probability threshold $p = 0.8$ for all datasets except for Louvre [46], where we use $p = 0.9$. Doppelgangers [3] did not disambiguate some datasets for any probability threshold we checked and thus are marked as not disambiguated.

Dataset details: In Table S6, we show the number of nodes and edges present in the datasets after applying our method with different values of minimum edge score m and after Doppelgangers [3] on the original graph. In these datasets, our method not only removes false edges but also sparsifies the graphs. Similar to the observation for generic datasets, here as well, there is a steady decrease in the number of edges with increasing m, but the number of nodes abruptly decreases for $m = \{0.8, 0.9\}$. This shows that original graphs are well connected, due to which smaller values of m lead to the removal of a few nodes. For larger values of m, the graph becomes sparser, leading to more nodes being lost. We also observe that the number of edges after applying our method is lesser compared to Doppelgangers [3] for most datasets due to the sparsification behaviour of our method. The number of nodes retained by our method is larger than that of Doppelgangers [3] for large-scale datasets [46, 47], similar for medium-scale [17], and lesser for small-scale datasets [49]. For small-scale datasets [49], the redundancy of edges in the graphs is low, which causes many nodes to be lost.

Performance comparison: In Table S7, we provide a comparison of our method with other disambiguation methods [17, 46, 6, 49, 3] by following the experimental procedure in [3]. It can be seen that COLMAP [36] fails to disambiguate on all datasets except Louvre [46] and Temple of Heaven [49]. Other methods are able to disambiguate many datasets with our method performing the best. Our method disambiguates all datasets except Cup [49]. For small-scale datasets [49], our method tends to over-split the reconstructions. As discussed in Sec. 4.2 of the main paper, the datasets from [49] have low edge redundancy, due to which the sparsification behaviour of our method leads to reconstruction over-splits.

Reconstruction details: In Table S8, we provide the reconstruction statistics on ambiguous datasets. It can be seen that large-scale datasets [46, 47] are disambiguated using $m = 0.6$ (which is also recommended for graph sparsification), except for highly ambiguous datasets, Louvre [46] and Sacre Coeur [46], where higher values of m are required. Usage of high values of m is not recommended in general since it removes most of the edges, thus failing to reconstruct, as seen for Seville [46] and Yorkminster [47] for $m = 0.9$. For medium [17] and small-scale [49] datasets, $m = 0.3$ disambiguates all datasets except Cup, which is disambiguated with $m = \{0.5, 0.6\}$. It can also be seen that the number of cameras and 3D points recovered are similar for our method and Doppelgangers [3] for large [46, 47] and medium-scale [17] datasets, even though our method sparsifies the graphs.

In Table S9, we provide the mean reprojection error for the reconstructed 3D points, which reduces with increase in m. This shows that, for ambiguous datasets as well, the camera and 3D points reconstructed are more consistent with that of the epipolar inliers with an increase in m. We also compare the camera motions of different graphs to that of the Doppelgangers [3] graph. We only compare camera motions for the datasets which have been disambiguated by Doppelgangers [3] to ensure only correctly estimated cameras are used as a reference for comparison. It can be seen that for the cases when our method disambiguates repeated structures, camera motions are close to the one obtained from Doppelgangers [3]. For Seville [46] and Yorkminster [47], no reconstruction was obtained from COLMAP [36] for $\mathcal{G}_F(0.9)$ since most of the cameras and edges were lost in the largest connected component. For Street [49] and Radcliffe Camera [17], some reconstructions did not have common cameras with Doppelgangers [3] reconstructions since different facades of the buildings were reconstructed.

Time taken: In Table S10, we compare the reconstruction time taken by COLMAP [36]. It can be seen that reconstruction time after disambiguation reduces for all datasets for both Doppelgangers [3] and our method with significant reduction for large-scale datasets. Moreover, since our method also sparsifies the graphs, the reconstruction time for our method is less compared to Doppelgangers [3].

In Table S11, we compare the processing time for different disambiguation methods. Similar to the previous subsection, we refer to the preprocessing time for our method as the time taken for Steps 1-3 and filter time for Steps 4-11 in Algo. 1. For Doppelgangers [3], preprocessing involves extracting specific keypoint descriptors and matches based on which their neural network is trained, and filtering refers to inferencing on all the edges using the neural network. For other methods, preprocessing time involves time required for giving specific inputs to their algorithms. It can be seen that our method is the fastest among other methods, taking less than 0.1% of the reconstruction time taken by COLMAP (Table S10) for large [46, 47] and medium-scale [17] datasets and less than 10% for small scale [49] datasets. Moreover, our method takes significantly less time to filter due to vectorized and parallelized operations. Our overall time taken, including preprocessing time, is significantly less than Doppelgangers [3] since our method scores the edges independent of keypoint descriptors and thus does not require extraction of specific keypoint descriptors.

Visual results: In Figs. S5, S6, S7, S8, S9, and S10, we provide visual results on the ambiguous datasets. Since both original ($\mathcal{G}$) and triplet ($\mathcal{G}_{LCT}$) graphs are not disambiguated, we show results only for $\mathcal{G}_{LCT}$ to fit all comparisons for a given dataset in a single page. It can be seen that for large-scale datasets, our method is able to disambiguate and recover most parts of the reconstructions with faster reconstruction times. Since our method does both disambiguation and graph sparsification, in some cases, removing true edges leads to superimposed reconstructions, as seen in Ellis Island [47] ($\mathcal{G}_F(0.7)$) and Louvre [46] ($\mathcal{G}_F(0.6)$ and $\mathcal{G}_F(0.7)$). However, these are corrected after further increasing the minimum edge score.

Choosing minimum edge score: Given the results on ambiguous datasets, we recommend setting the minimum edge score m between 0.6 and 0.7 for large-scale datasets and between 0.3 and 0.4 for medium and small-scale datasets to keep a trade-off between getting a proper reconstruction and time taken for reconstruction. The minimum edge score should only be increased for highly ambiguous datasets to avoid superimposed reconstructions in such datasets.

Dataset	Number of Nodes						Number of Edges					
	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{V}_{LCT}\rvert$	$\lvert\mathcal{V}_F(0.6)\rvert$	$\lvert\mathcal{V}_F(0.7)\rvert$	$\lvert\mathcal{V}_F(0.8)\rvert$	$\lvert\mathcal{V}_F(0.9)\rvert$	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{E}_{LCT}\rvert$	$\lvert\mathcal{E}_F(0.6)\rvert$	$\lvert\mathcal{E}_F(0.7)\rvert$	$\lvert\mathcal{E}_F(0.8)\rvert$	$\lvert\mathcal{E}_F(0.9)\rvert$
Alamo [47]	1700	1124	777	722	617	386	31986	29964	7966	5750	3758	1563
Gendarmenmarkt [47]	1241	1143	1063	1016	948	753	35819	35510	8885	6649	4469	2161
Madrid Metropolis [47]	1012	759	639	540	438	172	14997	14372	3633	2686	1733	433
Montreal Notre Dame [47]	1815	612	549	521	404	352	46466	25480	5547	4228	2586	1340
NYC Library [47]	1926	869	676	590	545	369	23399	14455	4750	3509	2351	1060
Notre Dame [47]	1430	1430	1328	1270	1134	660	91907	91907	22200	15875	10018	3883
Piccadilly [47]	5087	3670	3431	3308	3106	2633	111238	105868	37754	27477	17568	8347
Roman Forum [47]	1960	1669	1555	1495	1357	1132	36023	34710	13038	9511	6119	3084
Tower of London [47]	1165	900	739	659	500	307	15886	15352	4936	3510	2197	914
Trafalgar [47]	11606	9146	8423	8110	7632	6451	238983	229930	96257	71029	46108	21413
Union Square [47]	3015	1804	1593	1335	1186	880	36718	34037	12537	7366	4761	2307
Vienna Cathedral [47]	4151	2042	1592	1483	1089	892	70674	54435	21473	16013	7921	3913
Dubrovnik [24]	6039	6001	5900	5830	5515	4865	167489	167009	72981	53740	34806	17156
Rome [24]	14537	11333	10688	10515	8904	7849	697172	579978	285592	210828	121487	57543
Quad [5]	6330	5977	5461	5136	4050	1678	105990	104310	35083	25773	16452	3786

Table S1. Dataset details of **generic** datasets. A steady decrease in the number of edges is observed with increasing minimum edge score m but not on the number of nodes.

Dataset	Clean Reconstruction						# Cameras Reconstructed ($\#N_{CR}$)						# 3D Points Reconstructed (in 10^3)					
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Alamo [47]	✓	✓	✓	✓	✓	✓	899	875	652	612	543	341	161	157	128	126	121	78
Gendarmenmarkt [47]	✗	✗	✓	✓	✓	✓	1042	1045	899	857	794	654	207	206	160	151	135	108
Madrid Metropolis [47]	✓	✓	✓	✓	✓	✓	471	453	274	232	229	141	74	72	47	33	34	21
Montreal Notre Dame [47]	✓	✓	✓	✓	✓	✓	575	570	411	398	389	340	145	146	108	103	96	84
Notre Dame [47]	✗	✗	✓	✓	✓	✓	1407	1406	1295	1093	987	659	348	348	332	288	277	227
NYC Library [47]	✓	✓	✓	✓	✓	✓	635	614	509	396	399	299	110	109	93	84	77	57
Piccadilly [47]	✗	✗	✓	✓	✓	✓	3208	3110	2786	2645	2466	2037	374	368	311	288	255	202
Roman Forum [47]	✓	✓	✓	✓	✓	✓	1587	1565	1365	1295	1195	680	324	323	278	263	240	132
Tower of London [47]	✓	✓	✓	✓	✓	✓	735	718	535	495	428	279	165	165	137	129	110	84
Trafalgar [47]	✓	✓	✓	✓	✓	✓	7867	7532	6538	6256	5783	4602	706	695	573	545	491	397
Union Square [47]	✓	✓	✓	✓	✓	✓	1160	1150	891	805	737	366	82	81	62	55	49	24
Vienna Cathedral [47]	✗	✗	✓	✓	✓	✓	1197	1161	1010	965	908	528	304	301	258	248	237	148
Dubrovnik [24]	✓	✓	✓	✓	✓	✓	5883	5850	5576	5464	5209	4777	1252	1250	1127	1073	978	780
Rome [24]	✓	✓	✓	✓	✓	✓	1812	1812	1807	1728	1742	1655	395	395	389	389	389	391
Quad [5]	✓	✓	✓	✓	✓*	✓*	5729	5569	4360	3843	2044	367	1295	1277	1041	882	615	50

Table S2. Reconstruction statistics on **generic** datasets. ✓/✓*/✗: Removed/Removed but over-split/Existing ghost artifacts. Our method sparsifies the graphs, reconstructing most cameras and 3D points avoiding ghost artifacts.

Dataset	Mean Reprojection Errors (px)						Mean Camera Rotation Difference (degrees)					Mean Camera Translation Difference				
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Alamo [47]	0.66	0.66	0.62	0.59	0.56	**0.54**	0.37	0.35	0.34	0.30	0.20	0.15	0.22	0.14	0.11	0.11
Gendarmenmarkt [47]	0.76	0.76	0.61	0.74	0.73	**0.70**	-	-	-	-	-	-	-	-	-	-
Madrid Metropolis [47]	0.62	0.61	0.57	0.58	0.55	**0.52**	0.32	4.33	4.91	4.78	0.19	0.09	0.56	0.51	0.55	0.99
Montreal Notre Dame [47]	0.86	0.86	0.80	0.79	0.77	**0.72**	0.17	0.07	0.06	0.07	0.07	0.03	0.03	0.03	0.04	0.03
Notre Dame [47]	0.76	0.76	0.70	0.68	0.66	**0.62**	-	-	-	-	-	-	-	-	-	-
NYC Library [47]	0.74	0.74	0.71	0.70	0.67	**0.63**	0.22	0.07	0.09	0.07	0.08	0.04	0.07	0.09	0.06	0.02
Piccadilly [47]	0.76	0.76	0.74	0.72	0.70	**0.66**	-	-	-	-	-	-	-	-	-	-
Roman Forum [47]	0.76	0.76	0.73	0.72	0.70	**0.64**	0.10	0.20	0.20	0.21	0.27	0.02	0.02	0.03	0.02	0.04
Tower of London [47]	0.63	0.63	0.62	0.61	0.60	**0.56**	0.09	0.12	0.13	0.17	0.31	0.06	0.03	0.03	0.04	0.03
Trafalgar [47]	0.74	0.74	0.73	0.72	0.69	**0.66**	0.61	2.30	1.14	1.20	3.00	0.93	0.13	0.10	0.11	0.15
Union Square [47]	0.74	0.73	0.69	0.69	0.67	**0.61**	2.20	5.94	2.95	3.27	0.98	0.51	0.41	0.57	0.69	0.27
Vienna Cathedral [47]	0.75	0.75	0.74	0.73	0.71	**0.70**	-	-	-	-	-	-	-	-	-	-
Dubrovnik [24]	0.76	0.76	0.73	0.72	0.70	**0.66**	0.09	0.12	0.13	0.24	0.35	0.01	0.01	0.02	0.02	0.20
Rome [24]	0.90	0.90	0.87	0.85	0.82	**0.77**	0.01	0.03	0.03	0.04	0.06	0.04	0.05	0.03	0.04	0.04
Quad [5]	0.69	0.69	0.67	0.66	0.65	**0.62**	0.34	0.71	1.33	0.72	0.31	0.02	0.05	0.09	0.06	0.04

Table S3. Reprojection errors and camera motion differences (with camera motions from original graphs $\mathcal{G}$ as reference) on **generic** datasets. Camera translation difference is specified in the units obtained from the output of COLMAP [36] on the original graphs $\mathcal{G}$. **Bold** values indicate least reprojection error in each dataset. '-' indicates that reconstruction from the original graph contains ghost artifacts and is not used for comparison of camera motions.

Dataset	Reconstruction Time (mins) (t_R)					
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Alamo [47]	44	44	28	25	28	16
Gendarmenmarkt [47]	372	339	144	103	74	36
Madrid Metropolis [47]	88	90	22	21	16	3
Montreal Notre Dame [47]	114	100	50	52	38	25
Notre Dame [47]	2054	2131	1323	758	514	206
NYC Library [47]	105	101	40	40	34	15
Piccadilly [47]	1969	1409	1060	1000	1035	452
Roman Forum [47]	784	810	477	453	269	60
Tower of London [47]	95	86	44	42	36	16
Trafalgar [47]	6875	5601	3852	3698	2795	1962
Union Square [47]	220	185	80	86	60	51
Vienna Cathedral [47]	793	703	398	376	225	107
Dubrovnik [24]	582	556	521	428	367	276
Rome [24]	284	280	279	265	227	139
Quad [5]	417	418	318	249	133	23

Table S4. Reconstruction time using COLMAP [36] on **generic** datasets. Applying our method for graph sparsification recovers most part of the reconstruction of the original graph with reduced reconstruction time.

Dataset	Time Taken for Filter (sec)					Time Taken for Preprocessing + Filter (sec)				
	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Alamo [47]	13	10	11	15	11	13	23	25	28	24
Gendarmenmarkt [47]	12	12	14	18	14	12	25	27	31	27
Madrid Metropolis [47]	2	2	2	2	2	2	4	4	5	4
Montreal Notre Dame [47]	23	10	11	13	11	23	33	35	36	34
NYC Library [47]	3	1	2	1	1	3	4	5	4	4
Notre Dame [47]	113	125	169	182	143	113	238	282	295	257
Piccadilly [47]	83	87	97	113	101	83	170	181	197	184
Roman Forum [47]	7	7	8	11	8	7	14	15	18	16
Tower of London [47]	2	2	2	3	2	2	4	4	5	4
Trafalgar [47]	405	438	503	526	512	405	844	909	932	918
Union Square [47]	12	11	12	17	13	12	23	24	29	24
Vienna Cathedral [47]	53	42	53	56	47	53	94	105	109	99
Dubrovnik [24]	182	223	233	246	446	182	405	414	428	628
Rome [24]	5038	4440	5305	4798	4792	5038	9478	10343	9836	9830
Quad [5]	73	83	99	92	93	73	152	173	166	167

Table S5. Time taken by our method (Algo. 1) on **generic** datasets. Preprocessing time: Steps 1-3 of Algo. 1. Filter time: Steps 4-11 of Algo. 1. Our method takes $< 1\%$ of the reconstruction time for most datasets.

Dataset	Number of Nodes							Number of Edges						
	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{V}_{LCT}\rvert$	$\lvert\mathcal{V}_{Dopp}\rvert$	$\lvert\mathcal{V}_F(0.6)\rvert$	$\lvert\mathcal{V}_F(0.7)\rvert$	$\lvert\mathcal{V}_F(0.8)\rvert$	$\lvert\mathcal{V}_F(0.9)\rvert$	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{E}_{LCT}\rvert$	$\lvert\mathcal{E}_{Dopp}\rvert$	$\lvert\mathcal{E}_F(0.6)\rvert$	$\lvert\mathcal{E}_F(0.7)\rvert$	$\lvert\mathcal{E}_F(0.8)\rvert$	$\lvert\mathcal{E}_F(0.9)\rvert$
Louvre [46]	3443	1349	557	1232	1144	948	763	47063	31206	3766	9518	6963	4328	2245
Notre Dame [46]	12720	11836	8199	11190	10686	10011	8298	447199	441860	271483	210531	156583	103759	50583
Sacre Coeur [46]	5228	4927	4701	4576	4420	4226	3647	218469	217083	161611	76229	55757	36101	17606
Seville [46]	2262	1586	1474	1496	1295	1194	-	42390	33467	21543	11974	7893	5303	-
Ellis Island [47]	2075	1677	840	1541	1490	1373	1068	40266	39236	10871	17622	13168	8769	4445
Piazza del Popolo [47]	1666	1148	1069	991	946	862	713	29392	26341	16643	8153	6030	3925	1903
Yorkminster [47]	2534	1935	1104	1694	1550	1249	-	40483	36416	16825	12590	9120	5826	-
	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{V}_{LCT}\rvert$	$\lvert\mathcal{V}_{Dopp}\rvert$	$\lvert\mathcal{V}_F(0.3)\rvert$	$\lvert\mathcal{V}_F(0.4)\rvert$	$\lvert\mathcal{V}_F(0.5)\rvert$	$\lvert\mathcal{V}_F(0.6)\rvert$	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{E}_{LCT}\rvert$	$\lvert\mathcal{E}_{Dopp}\rvert$	$\lvert\mathcal{E}_F(0.3)\rvert$	$\lvert\mathcal{E}_F(0.4)\rvert$	$\lvert\mathcal{E}_F(0.5)\rvert$	$\lvert\mathcal{E}_F(0.6)\rvert$
Alexander Nevsky Cathedral [17]	448	448	447	431	427	426	420	29111	29111	12580	2773	2419	2073	1705
Arc de Triomphe [17]	434	434	415	416	412	406	400	14565	14556	9597	4303	3669	3072	2461
Berliner Dom [17]	1618	1618	1615	1613	1603	1591	1580	92255	92224	77981	46634	39835	32866	26099
Big Ben [17]	402	402	397	391	388	383	373	19598	19598	10685	3547	3083	2637	2181
Brandenburg Gate [17]	175	175	153	133	123	112	101	9300	9300	7151	401	340	273	206
Church on Spilled Blood [17]	277	277	265	142	141	139	133	14765	14765	7387	644	579	503	415
Indoor [17]	152	152	152	42	42	33	28	8787	8787	2382	41	41	32	27
Radcliffe Camera [17]	282	282	281	177	121	115	114	13900	13900	6456	929	591	500	422
Books [49]	21	21	21	9	7	7	7	208	208	92	11	9	9	9
Cereal [49]	25	25	25	7	7	7	7	294	294	215	9	6	6	6
Cup [49]	64	64	63	64	64	40	40	2016	2016	125	81	72	44	42
Desk [49]	31	31	31	12	11	11	11	446	446	188	12	11	11	11
Oats [49]	23	23	23	9	9	9	9	252	252	140	11	11	10	10
Street [49]	19	19	10	19	19	10	9	171	171	13	19	19	9	8
Temple of Heaven [49]	338	338	338	338	338	338	338	20116	20116	1782	11234	8854	6385	4142

Table S6. Dataset details of **ambiguous** datasets. '-' represents that the largest connected component does not have enough nodes and edges to obtain a reconstruction. A steady decrease in the number of edges is observed with increasing minimum edge score m but not on the number of nodes.

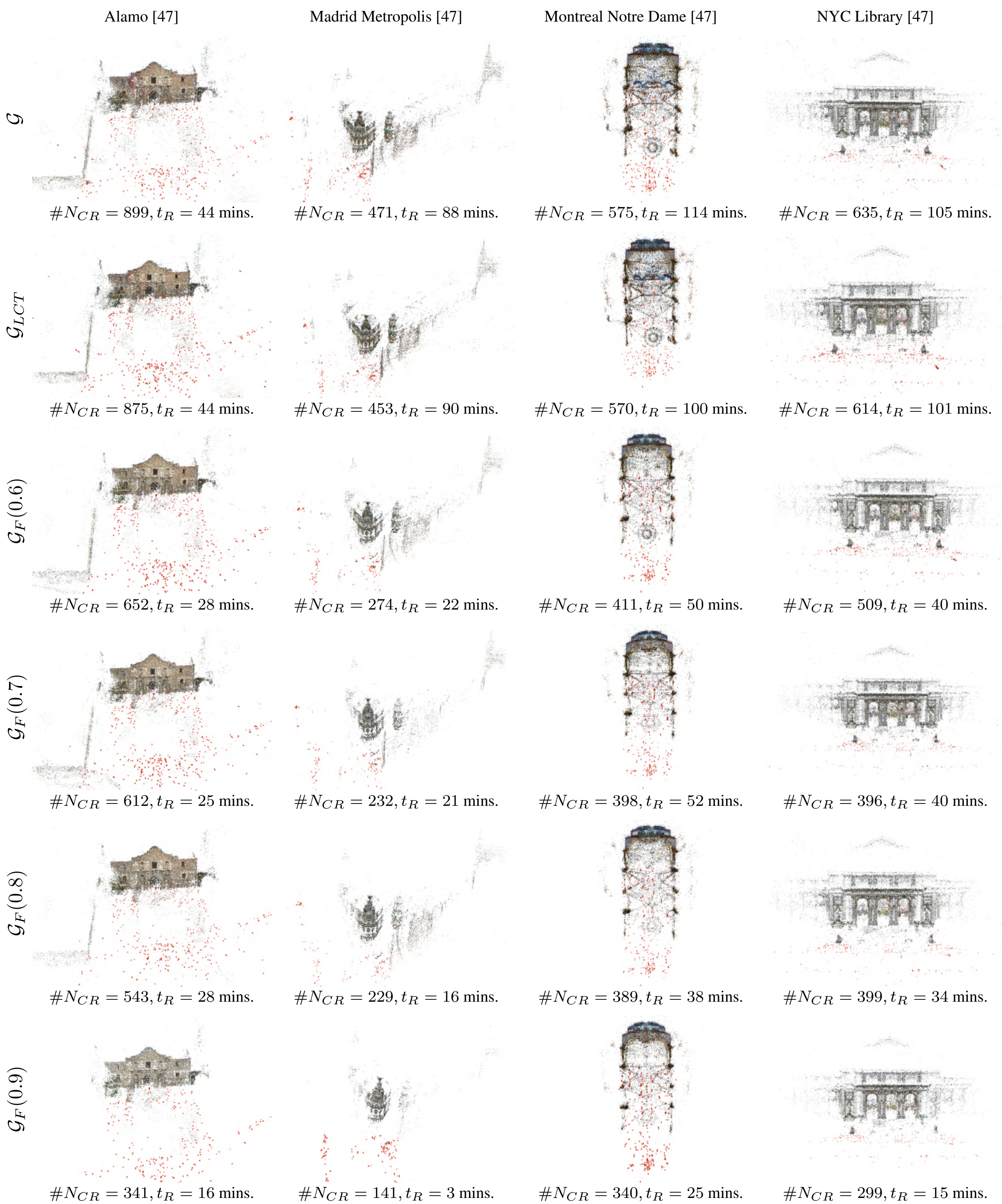

Figure S1. Reconstructions obtained with different graphs on **generic** datasets from [47]. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m): 0.6 to 0.7.

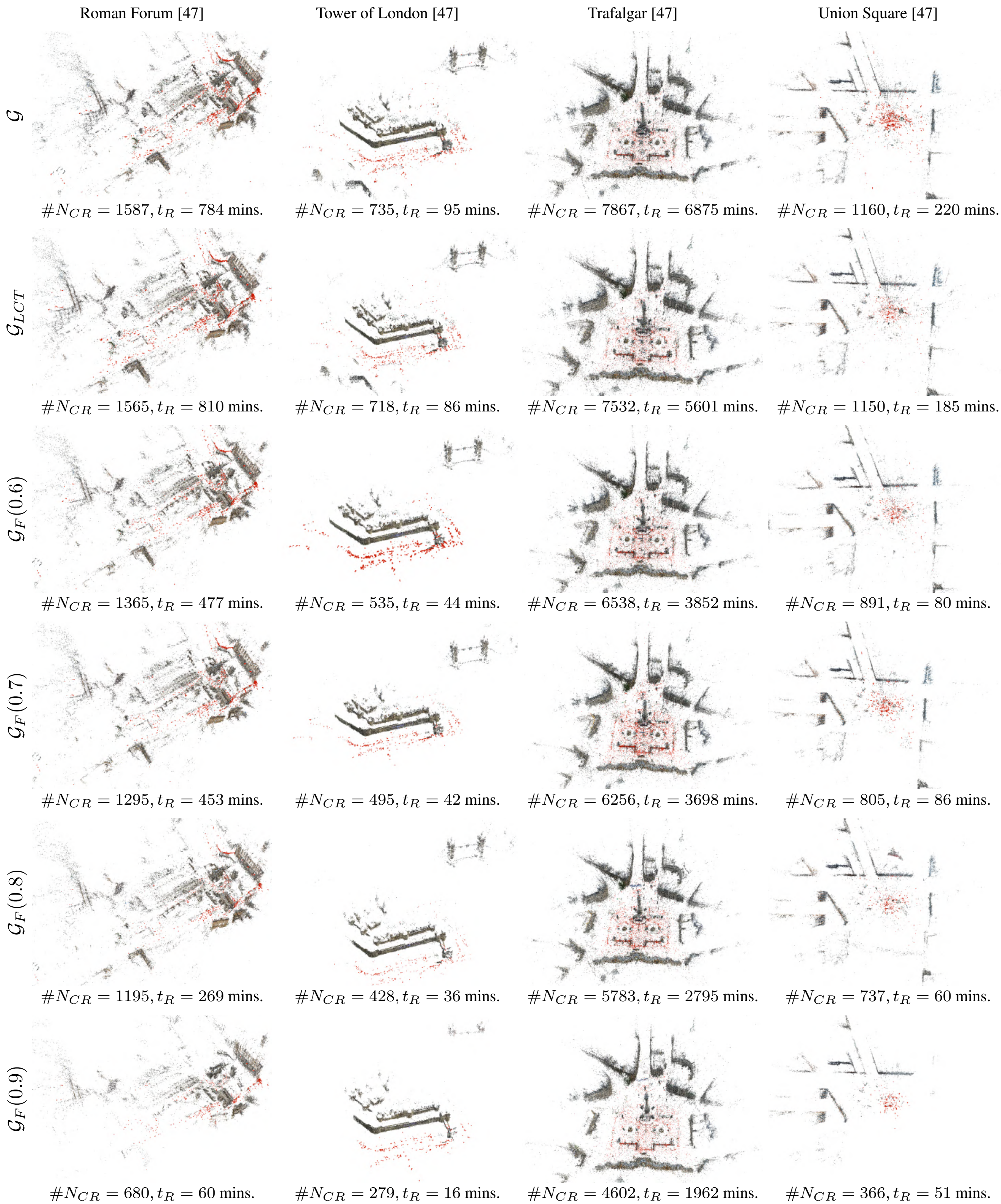

Figure S2. Reconstructions obtained with different graphs on **generic** datasets from [47]. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m): 0.6 to 0.7.

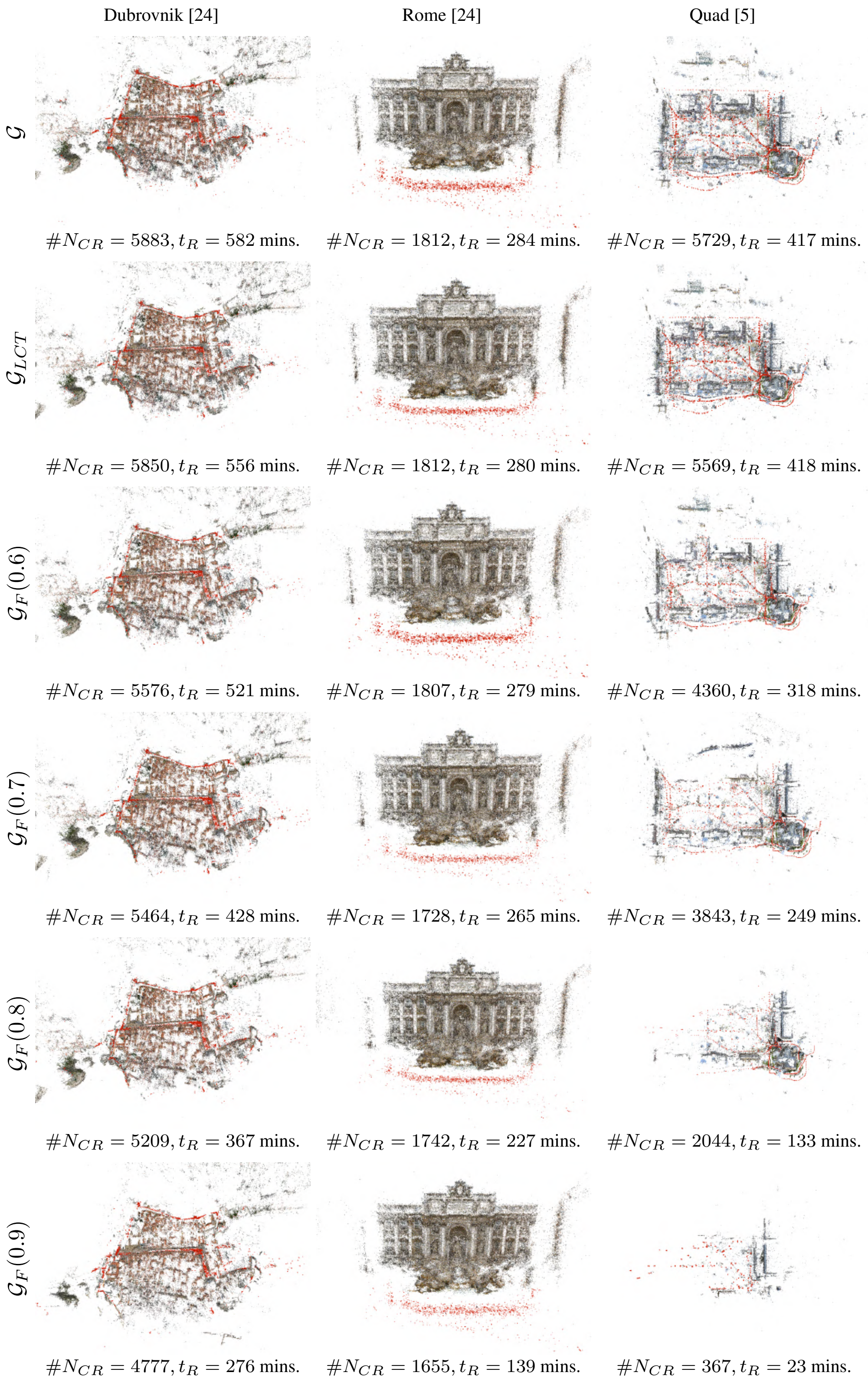

Figure S3. Reconstructions obtained with different graphs on **generic** datasets from [24] and [5]. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m): 0.6 to 0.7.

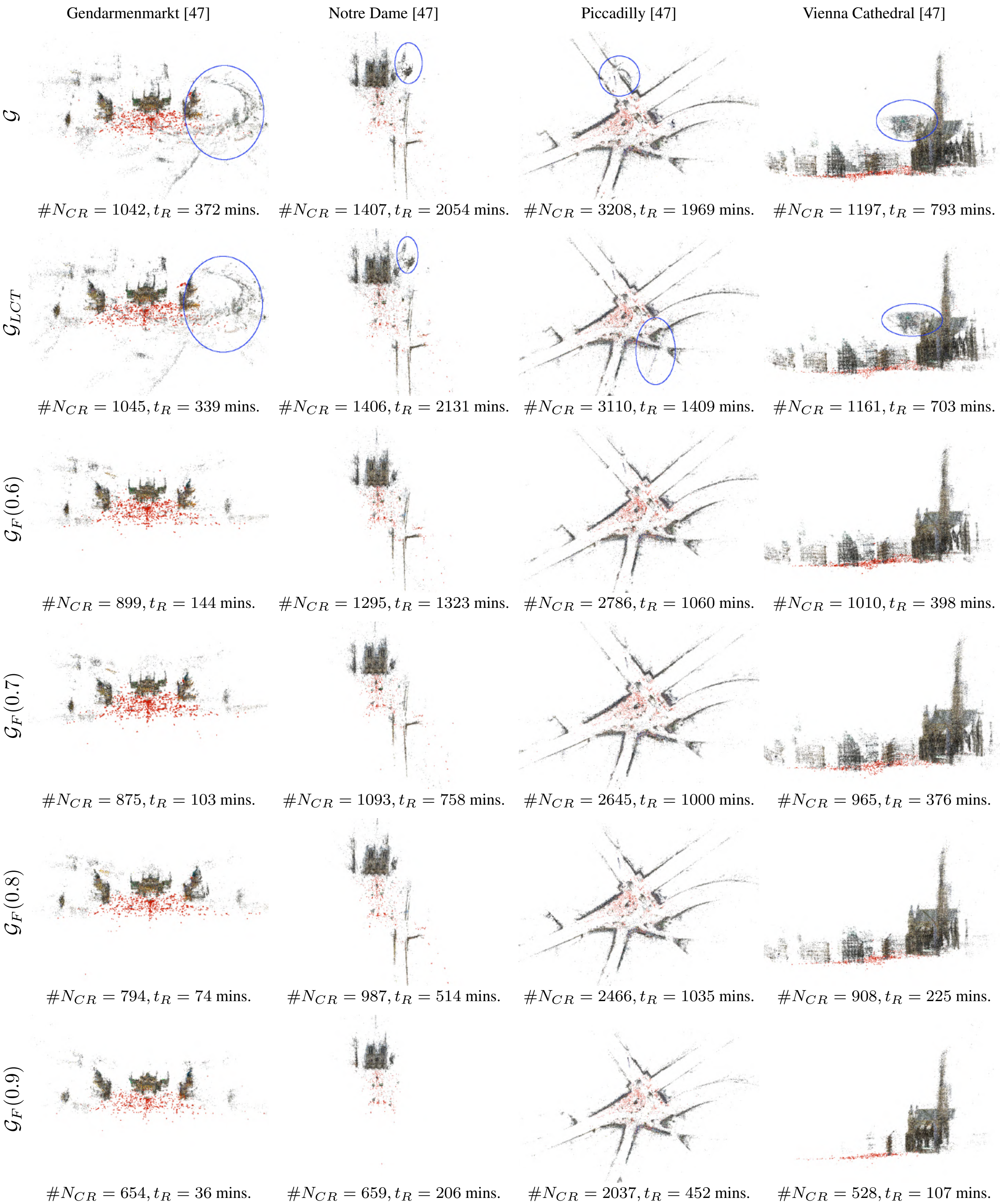

Figure S4. Reconstructions obtained with different graphs on **generic** datasets from [47] where ghost artifacts are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m): 0.6 to 0.7.

Dataset	COLMAP [36]	Heinly [17]	Wilson [46]	Cui [6]	Yan [49]	Cai [3]	Ours
Louvre [46]	✓	-	✗	✓	✓	✓	✓
Notre Dame [46]	✗	-	✗	✓	-*	✓	✓
Sacre Coeur [46]	✗	✓	✓	✓	✓	✓	✓
Seville [46]	✗	✗	✓*	✓*	✓*	✓	✓
Ellis Island [47]	✗	-	✓	✓*	✓	✓	✓
Piazza del Popolo [47]	✗	-	✗	✓*	✓*	✓	✓
Yorkminster [47]	✗	-	✓	✓	✓	✓	✓
Alexander Nevsky Cathedral [17]	✗	✓	✗	✓	✓	✓	✓
Arc de Triomphe [17]	✗	✓	✗	✗	✓	✓	✓
Berliner Dom [17]	✗	✓	✓*	✓	✓	✓	✓
Big Ben [17]	✗	✓	✓	✗	✗	✓	✓
Brandenburg Gate [17]	✗	✓	✗	✓*	✓*	✓*	✓*
Church on Spilled Blood [17]	✗	✓	✗	✓	✗	✗	✓
Indoor [17]	✗	✓	✗	✓	✓	✓	✓
Radcliffe Camera [17]	✗	✓	✓*	✓*	✓*	✓*	✓*
Books [49]	✗	-	✗	✗	✗	✗	✓*
Cereal [49]	✗	✓	✗	✗	✓	✗	✓*
Cup [49]	✗	✗	✗	✗	✗	✓	✗
Desk [49]	✗	-	✗	✓	✓	✓	✓*
Oats [49]	✗	-	✗	✗	✗	✗	✓*
Street [49]	✗	✗	✗	✓	✗	✓*	✓
Temple of Heaven [49]	✓	-	✓	✓	✓*	✓	✓

Table S7. Comparison of our proposed method with other methods on **ambiguous** datasets. ✓/✓*/✗: Disambiguated/Disambiguated but oversplit/Non-disambiguated reconstructions. Results of [17] are shown as reported in [3] ('-' means results not reported in [3]). '-*' means code failed to execute. Our method disambiguates all datasets except Cup.

Dataset	Disambiguated							# Cameras Reconstructed ($\#N_{CR}$)							# 3D Points Reconstructed (in 10^3)						
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Louvre [46]	✓	✓	✓	✗	✗	✓	✓	367	359	320	320	317	213	186	128	129	108	113	110	77	69
Notre Dame [46]	✗	✗	✓	✓	✓	✓	✓	7952	7778	5481	5420	5277	4921	3977	1785	1775	1314	1306	1280	1226	1053
Sacre Coeur [46]	✗	✗	✓	✗	✗	✗	✓	4492	4443	3796	2740	2645	2498	1984	711	706	548	475	450	449	347
Seville [46]	✗	✗	✓	✓	✓	✓	✗*	1510	1498	447	475	253	231	-	353	353	109	117	74	71	-
Ellis Island [47]	✗	✗	✓	✓	✗	✓	✓	880	910	314	333	614	319	288	164	171	84	89	122	87	75
Piazza del Popolo [47]	✗	✗	✓	✓	✓	✓*	✓*	1023	1012	922	865	818	420	299	138	136	123	110	101	52	40
Yorkminster [47]	✗	✗	✓	✓	✓	✓	✗*	1065	1027	585	520	457	388	-	284	280	173	162	141	126	-
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$
Alexander Nevsky Cathedral [17]	✗	✗	✓	✓	✓	✓	✓	447	447	445	429	426	425	419	100	100	90	83	82	81	79
Arc de Triomphe [17]	✗	✗	✓	✓	✓	✓	✓	427	425	392	394	387	324	317	81	81	69	70	69	53	52
Berliner Dom [17]	✗	✗	✓	✓	✓	✓	✓	1603	1603	1600	1588	1585	1575	1560	242	242	238	234	233	232	231
Big Ben [17]	✗	✗	✓	✓	✓	✓	✓	397	398	394	379	375	365	357	74	74	73	67	66	65	63
Brandenburg Gate [17]	✗	✗	✓*	✓*	✓*	✓*	✓*	172	172	151	129	121	110	100	24	24	21	14	14	12	12
Church on Spilled Blood [17]	✗	✗	✗	✓	✓	✓	✓	273	274	258	136	135	131	127	69	69	64	30	30	29	28
Indoor [17]	✗	✗	✓	✓	✓	✓*	✓*	152	152	152	42	42	33	28	73	73	58	9	9	9	8
Radcliffe Camera [17]	✗	✗	✓*	✓*	✓*	✓*	✓*	279	280	94	177	120	115	114	76	76	28	40	28	27	26
Books [49]	✗	✗	✗	✓*	✓*	✓*	✓*	21	21	21	9	7	7	7	8	8	7	2	2	2	2
Cereal [49]	✗	✗	✗	✓*	✓*	✓*	✓*	25	25	25	7	7	7	7	12	12	12	3	3	3	3
Cup [49]	✗	✗	✓	✗	✗	✓*	✓*	64	64	63	64	64	40	40	6	6	6	5	5	3	3
Desk [49]	✓	✓	✓	✓*	✓*	✓*	✓*	31	31	31	12	11	11	11	14	14	12	3	3	3	3
Oats [49]	✗	✗	✗	✓*	✓*	✓*	✓*	23	23	23	9	9	9	9	8	8	7	3	3	3	3
Street [49]	✗	✗	✓*	✓	✓	✓*	✓*	19	19	10	19	19	10	9	4	4	1	2	2	1	1
Temple of Heaven [49]	✓	✓	✓	✓	✓	✓	✓	338	338	338	338	338	338	338	192	192	185	196	199	200	198

Table S8. Reconstruction statistics on **ambiguous** datasets. ✓/✓*/✗/✗*: Disambiguated/Disambiguated but oversplit/Non-disambiguated reconstructions/No reconstruction obtained. Our method removes false edges and sparsifies the viewgraphs recovering most of the cameras and 3D points when compared to Doppelgangaers [3].

Dataset	Mean Reprojection Errors (px)							Mean Camera Rotation Difference (degrees)						Mean Camera Translation Difference					
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Louvre [46]	0.60	0.60	0.60	0.60	0.58	**0.58**	0.60	10.39	11.00	26.44	25.75	0.26	0.26	0.55	0.60	1.21	1.19	0.28	0.14
Notre Dame [46]	0.73	0.74	0.74	0.72	0.71	0.69	**0.67**	0.10	0.10	0.21	0.27	0.18	0.51	0.01	0.02	0.02	0.03	0.02	0.07
Sacre Coeur [46]	0.70	0.71	0.71	0.68	0.68	0.67	**0.61**	0.92	2.54	2.49	1.08	0.91	1.79	0.32	0.31	0.16	0.20	0.15	0.32
Seville [46]	0.68	0.68	0.64	0.62	0.65	**0.63**	-	0.17	0.22	0.11	0.11	0.15	-	0.02	0.05	0.07	0.11	0.15	-
Ellis Island [47]	0.76	0.75	0.80	0.79	0.74	0.76	**0.74**	1.23	1.24	2.40	2.44	3.64	1.39	0.16	0.16	0.13	0.15	0.21	0.08
Piazza del Popolo [47]	0.69	0.69	0.68	0.67	0.66	0.68	**0.65**	87.09	87.21	1.76	1.75	0.42	0.08	3.71	3.70	0.23	0.23	0.08	0.03
Yorkminster [47]	0.75	0.75	0.77	0.76	0.76	**0.74**	-	0.39	0.34	0.13	0.15	0.28	-	0.06	0.06	0.03	0.08	0.19	-
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$
Alexander Nevsky Cathedral [17]	0.67	0.67	0.68	0.62	0.61	0.60	**0.58**	68.33	67.97	0.06	0.08	0.08	0.59	3.40	3.40	0.01	0.02	0.01	0.04
Arc de Triomphe [17]	0.66	0.66	0.66	0.64	0.64	0.63	**0.62**	68.06	67.84	2.98	1.23	1.21	0.30	2.78	2.72	0.29	0.24	0.18	0.15
Berliner Dom [17]	0.70	0.70	0.70	0.68	0.67	0.66	**0.65**	45.54	45.70	0.09	0.09	0.10	0.09	2.49	2.50	0.03	0.03	0.03	0.03
Big Ben [17]	0.64	0.64	0.64	0.61	0.60	0.59	**0.59**	50.48	50.41	2.44	1.49	1.52	1.03	1.42	1.42	0.04	0.04	0.04	0.03
Brandenburg Gate [17]	0.84	0.84	0.87	0.73	0.72	0.73	**0.71**	0.05	0.05	0.13	0.14	0.14	0.19	0.03	0.03	0.09	0.11	0.12	0.16
Church on Spilled Blood [17]	0.61	0.61	0.61	0.52	0.52	0.51	**0.50**	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*
Indoor [17]	0.52	0.55	0.53	0.40	0.40	0.39	**0.38**	67.49	67.49	0.33	0.33	0.17	0.13	2.74	2.74	0.06	0.06	0.02	0.02
Radcliffe Camera [17]	0.64	0.64	0.60	0.58	0.58	0.57	**0.56**	0.09	0.09	-#	-#	-#	-#	0.02	0.02	-#	-#	-#	-#
Books [49]	0.41	0.41	0.41	**0.31**	**0.31**	**0.31**	**0.31**	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*
Cereal [49]	0.41	0.41	0.41	0.30	**0.27**	**0.27**	**0.27**	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*
Cup [49]	0.61	0.61	0.39	0.42	0.38	0.37	**0.35**	71.42	71.42	66.01	72.54	1.22	1.66	3.51	3.51	3.20	3.59	0.08	0.07
Desk [49]	0.49	0.49	0.48	**0.40**	**0.40**	**0.40**	**0.40**	0.02	0.02	0.18	0.18	0.18	0.18	0.01	0.01	0.01	0.01	0.01	0.01
Oats [49]	0.40	0.39	0.37	0.25	0.25	**0.24**	**0.24**	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*	-*
Street [49]	0.50	0.50	0.37	0.35	0.35	0.33	**0.31**	0.27	0.27	0.27	0.27	0.15	-#	0.04	0.04	0.02	0.02	0.01	-#
Temple of Heaven [49]	0.65	0.65	**0.63**	0.65	0.66	0.67	0.68	1.30	1.27	1.04	0.52	0.62	0.02	0.07	0.07	0.05	0.02	0.03	0.01

Table S9. Reprojection errors and camera motion differences (with camera motions from Doppelgangers [3] graphs $\mathcal{G}_{Dopp}$ as reference) on **ambiguous** datasets. Camera translation difference is specified in the units obtained from the output of COLMAP [36] on the Doppelgangers graphs $\mathcal{G}_{Dopp}$. **Bold** values indicate least reprojection error in each dataset. The best value is checked only across disambiguated/disambiguated but oversplit reconstructions. '-': Reconstruction not obtained. '-*': Doppelgangaers reconstruction is not disambiguated and not used for camera motion comparison. '-#': No common cameras reconstructed between Doppelgangers and our method due to different facades of the building being reconstructed by the graphs.

Dataset	Reconstuction Time (mins) (t_R)						
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
Louvre [46]	95	90	79	42	33	30	52
Notre Dame [46]	1558	1686	1160	1059	1023	736	494
Sacre Coeur [46]	385	405	347	180	159	149	156
Seville [46]	203	113	87	24	13	13	-
Ellis Island [47]	212	195	33	26	80	27	16
Piazza del Popolo [47]	178	184	128	99	92	30	16
Yorkminster [47]	383	412	113	86	82	52	-
	$\mathcal{G}$	$\mathcal{G}_{LCT}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$
Alexander Nevsky Cathedral [17]	32	28	31	13	14	13	13
Arc de Triomphe [17]	22	18	17	15	17	11	9
Berliner Dom [17]	142	118	126	104	95	96	96
Big Ben [17]	40	36	43	33	33	33	30
Brandenburg Gate [17]	8	8	9	2	2	2	2
Church on Spilled Blood [17]	18	18	16	5	4	4	4
Indoor [17]	9	8	8	1	1	1	1
Radcliffe Camera [17]	18	18	6	6	4	3	3
Books [49]	0.3	0.3	0.3	0.1	0.1	0.1	0.1
Cereal [49]	0.5	0.5	0.5	0.1	0.1	0.1	0.1
Cup [49]	0.5	0.6	0.3	0.2	0.2	0.2	0.2
Desk [49]	1.0	1.1	0.8	0.2	0.2	0.2	0.2
Oats [49]	0.4	0.5	0.4	0.1	0.1	0.1	0.1
Street [49]	0.2	0.2	0.1	0.1	0.1	0.1	0.1
Temple of Heaven [49]	78	76	32	54	51	49	46

Table S10. Reconstruction time using COLMAP [36] on **ambiguous** datasets. Applying our method removes false edges and sparsifies the viewgraphs leading to reduced reconstruction time.

Dataset	Time Taken (sec)						
	$\mathcal{G}_{Cui}$	$\mathcal{G}_{Yan}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.6)$	$\mathcal{G}_F(0.7)$	$\mathcal{G}_F(0.8)$	$\mathcal{G}_F(0.9)$
	[6]	[49]	[3]	Ours			
Louvre [46]	1261	259681	16009	48	50	50	50
Notre Dame [46]	10038	-	139065	5314	5781	4923	4930
Sacre Coeur [46]	3154	253093	78242	1269	1382	1141	1142
Seville [46]	1057	42490	13479	22	27	22	22
Ellis Island [47]	27385	1742	12450	34	47	41	38
Piazza del Popolo [47]	356	20370	8925	13	16	16	14
Yorkminster [47]	868	82653	12748	19	26	23	21
	$\mathcal{G}_{Cui}$	$\mathcal{G}_{Yan}$	$\mathcal{G}_{Dopp}$	$\mathcal{G}_F(0.3)$	$\mathcal{G}_F(0.4)$	$\mathcal{G}_F(0.5)$	$\mathcal{G}_F(0.6)$
	[6]	[49]	[3]	Ours			
Alexander Nevsky Cathedral [17]	217	174	11203	96	65	65	65
Arc de Triomphe [17]	141	168	7715	15	11	11	11
Berliner Dom [17]	1095	3161	31719	607	428	427	430
Big Ben [17]	164	158	9011	31	23	23	26
Brandenburg Gate [17]	67	33	3657	11	8	8	8
Church on Spilled Blood [17]	147	98	4360	21	16	16	16
Indoor [17]	78	48	2535	11	8	8	8
Radcliffe Camera [17]	167	106	4381	17	13	13	13
Books [49]	12	6	81	1	1	1	1
Cereal [49]	13	8	104	1	1	1	1
Cup [49]	37	7	706	1	1	1	1
Desk [49]	16	8	167	1	1	1	1
Oats [49]	12	6	100	1	1	1	1
Street [49]	8	6	72	1	1	1	1
Temple of Heaven [49]	1144	211	10790	35	26	26	26

Table S11. Time taken for different disambiguation methods on **ambiguous** datasets. Time taken consists of the time required for specific preprocessing for the algorithms and their filtering times. Our method takes significantly less time compared to other methods.

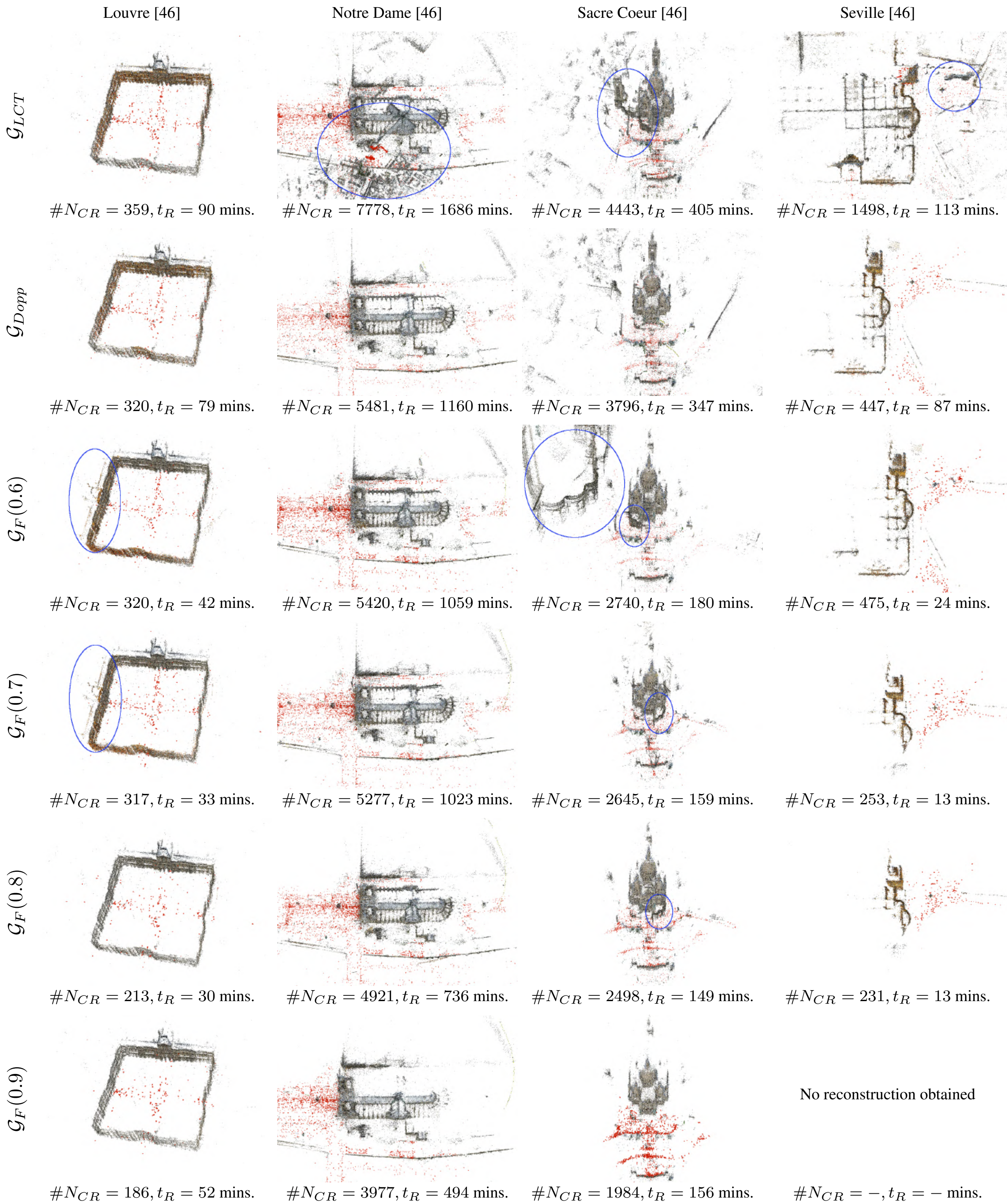

Figure S5. Reconstructions obtained with different graphs on **ambiguous** datasets from [46] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for large-scale datasets: 0.6 to 0.7; for highly ambiguous datasets, recommended $m > 0.8$.

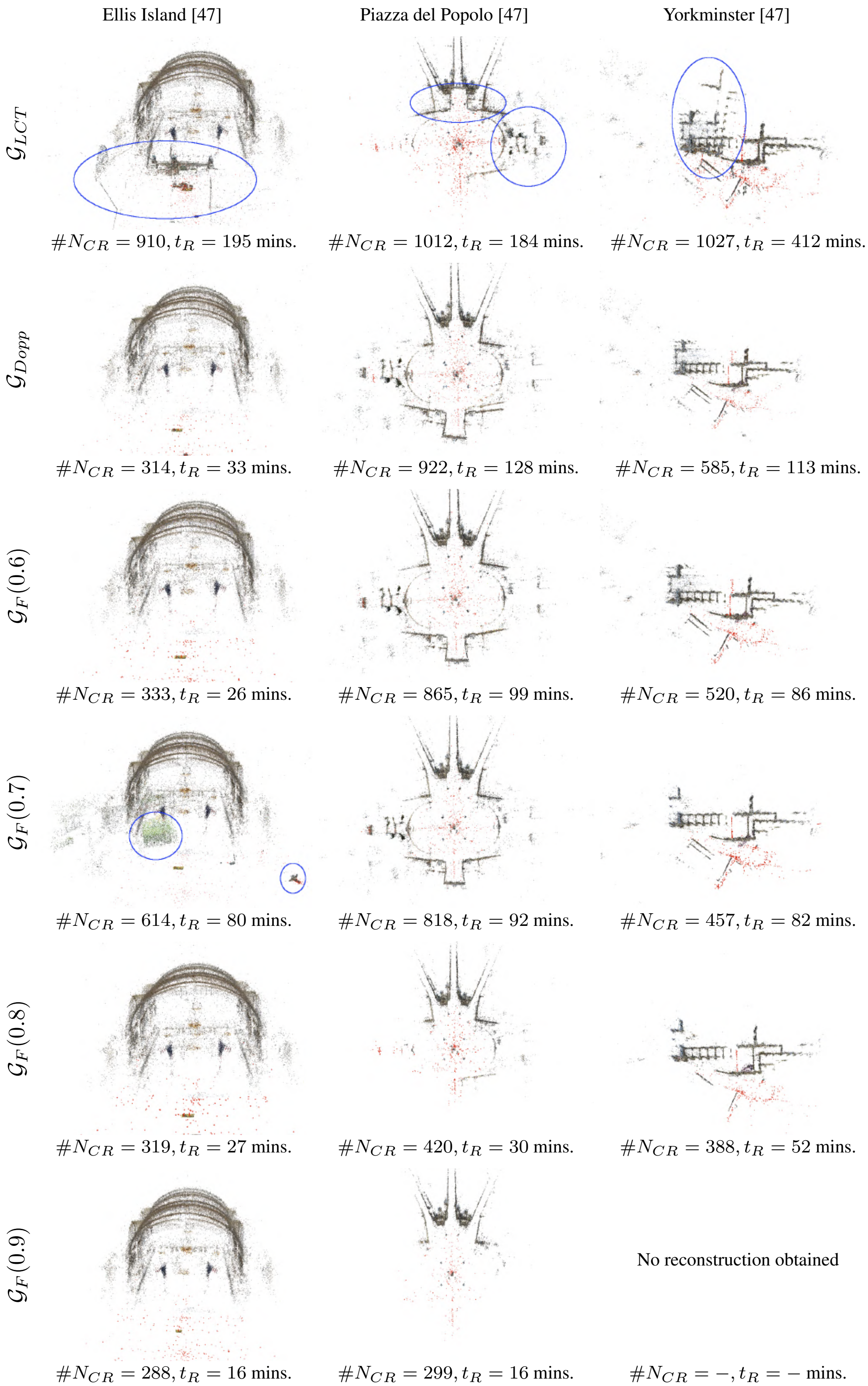

Figure S6. Reconstructions obtained with different graphs on **ambiguous** datasets from [47] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for large-scale datasets: 0.6 to 0.7.

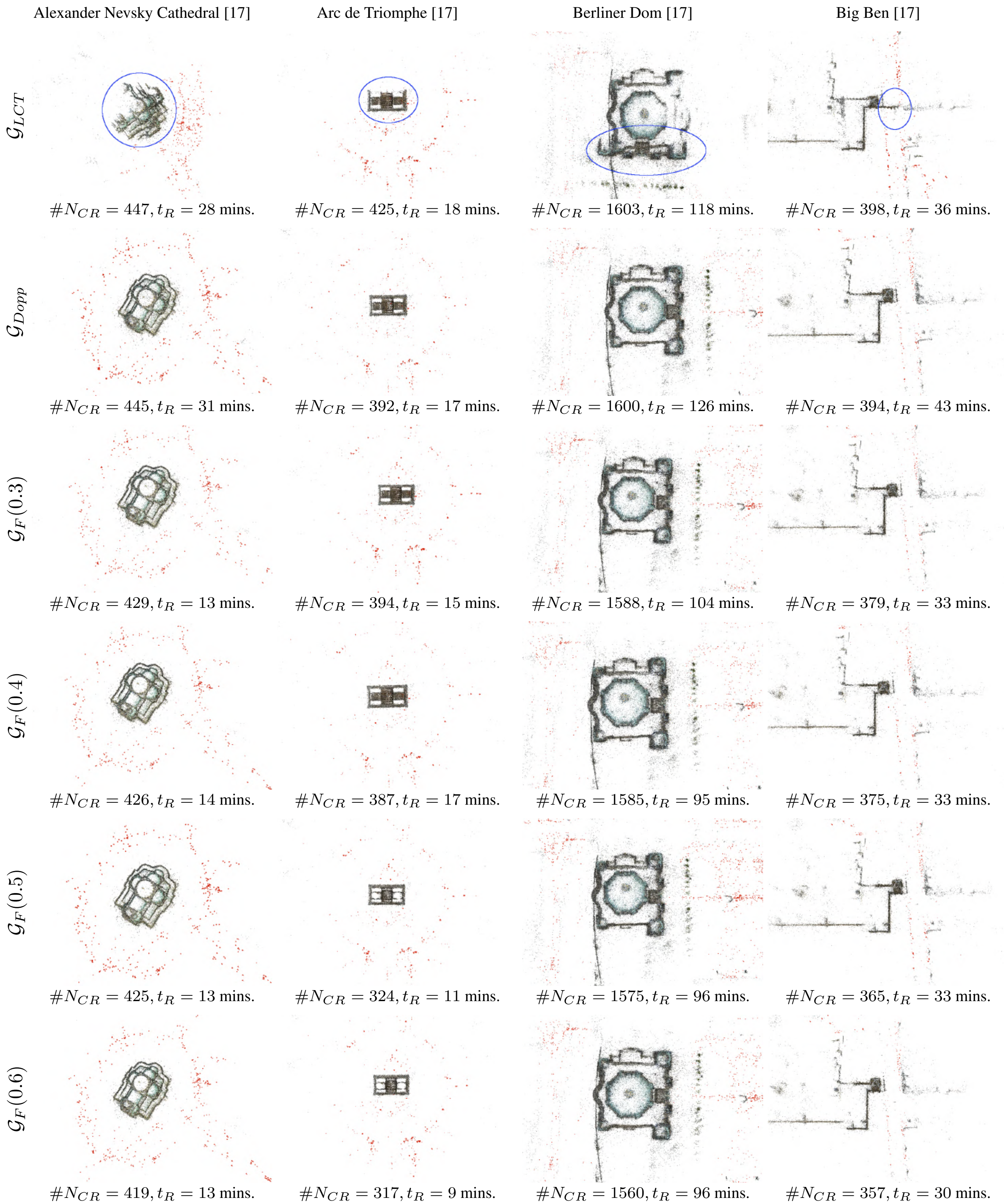

Figure S7. Reconstructions obtained with different graphs on **ambiguous** datasets from [17] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for medium-scale datasets: 0.3 to 0.4.

Figure S8. Reconstructions obtained with different graphs on **ambiguous** datasets from [17] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for medium-scale datasets: 0.3 to 0.4.

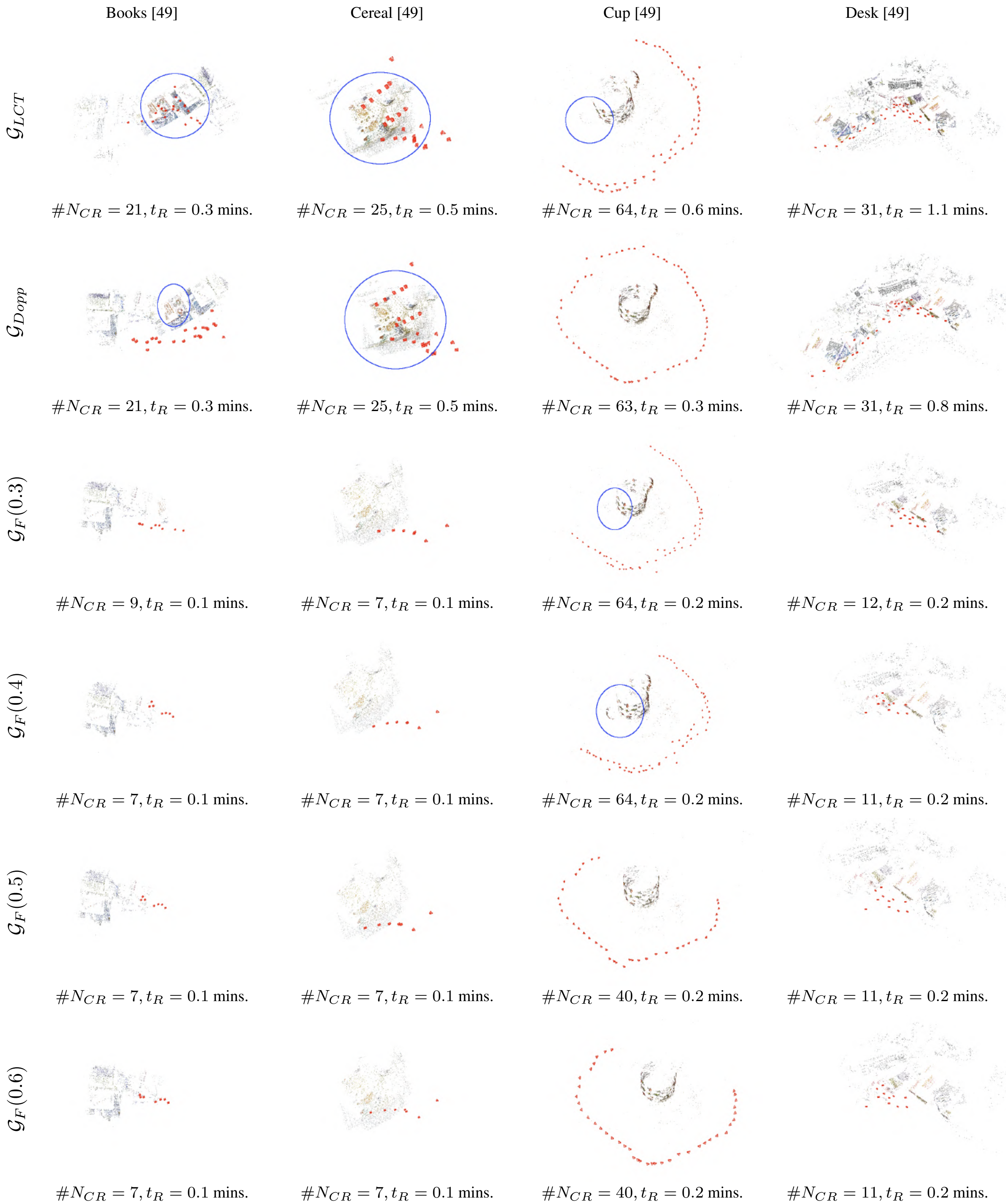

Figure S9. Reconstructions obtained with different graphs on **ambiguous** datasets from [49] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for small-scale datasets: 0.3 to 0.4.

Figure S10. Reconstructions obtained with different graphs on **ambiguous** datasets from [49] where superimposed parts of the reconstructions are marked in blue. $\#N_{CR}$: Number of cameras reconstructed, t_R: Reconstruction time using COLMAP [36]. Recommended minimum edge score (m) for small-scale datasets: 0.3 to 0.4.