CSCI 677 –Advance Computer Vision- Fall 2024 - HW 3

USC ID: 1154164561
NAME: Srija Madarapu
EMAIL: madarapu@usc.edu

Code:

https://colab.research.google.com/drive/1yzrR6MM6MwTKIKma_2t51kOHCygk9cjj?usp=sharing

Images:

https://drive.google.com/drive/folders/1UlEfaZshnCrUVwJw5MVlM-zoAwgF-0DU?usp=sharing

1. SFM: COLMAP

First 10 images:

```
=======================================================================
Feature extraction
=======================================================================
```

images	features	Camera(qw, qx, qy, qz)(tx, ty, tz)
1	12912	(0.999859, -0.00398792, -0.00386372, 0.0158283) (4.16896, 0.567398, 2.4463)
2	8612	(1, 0.000265556, -0.000106344, 0.000286177) (3.36233, 0.347961, 1.93785)
3	8982	(0.99997, -0.00413087, 0.00643097, 0.00109312)(2.62639, 0.262216, 1.35428)
4	9532	(0.998363, 0.0159648, 0.054768, 0.00399297)(1.12452, 0.0466749, 0.210846)
5	10306	(0.997872, 0.0301993, 0.0577848, 0.000817888)(-0.13028, -0.0256912, -0.204667)
6	11756	(0.991168, 0.00915945, 0.132041, -0.00822531)(-1.13319, -0.121328, -0.491623)
7	11849	(0.989593, -0.00530236, 0.141079, -0.0278276) (-2.31573, -0.0945697, -0.527258)
8	11145	(0.983882, -0.0104924, 0.176058, -0.0295064)(-3.89367, -0.0934167, -0.44076)
9	10414	(0.970974, -0.000658693, 0.236845, -0.0333832) (-5.13771, -0.030397, 0.215164)
10	10849	(0.965679, -0.00120932, 0.255726, -0.04545) (-6.77865, 0.119204, 0.509964)

	Camera	Images	Points	Observations	Mean(1)	Mean(2)	Mean(3)
Exhaustive	1	10	2705	10308	3.81072	1030.8	0.567743
Sequential	1	10	2395	9028	3.76952	1003.11	0.533517

```
=======================================================================
Feature matching
=======================================================================
Global bundle adjustment
Registering image #8 (3)
=> Image sees 152 / 1602 points
Registering image #9 (4)
```
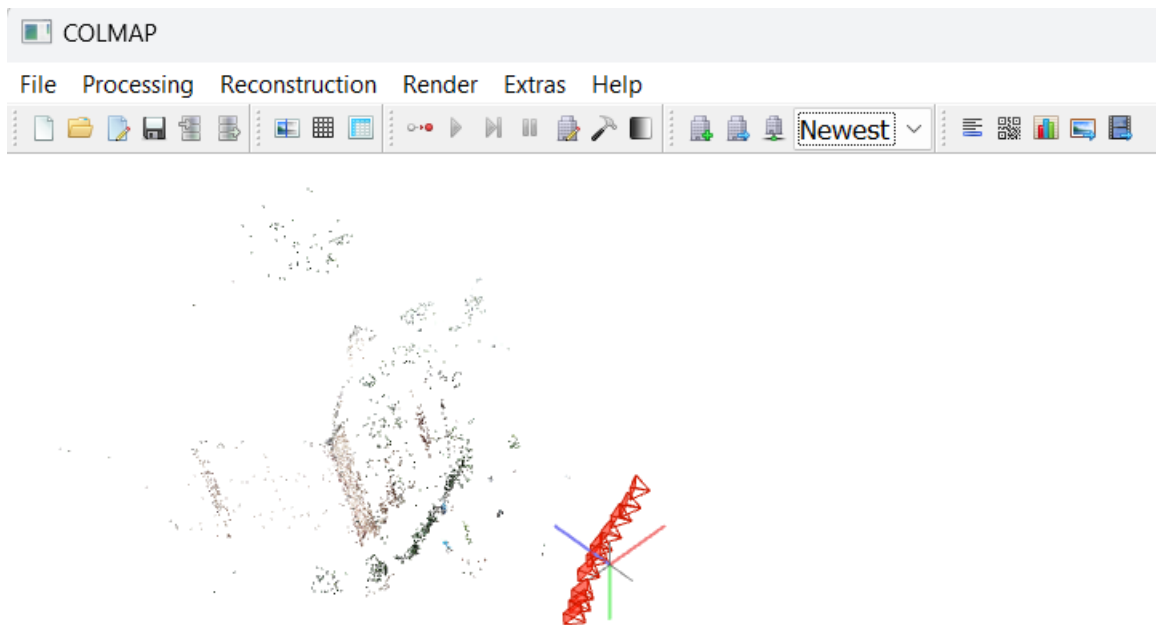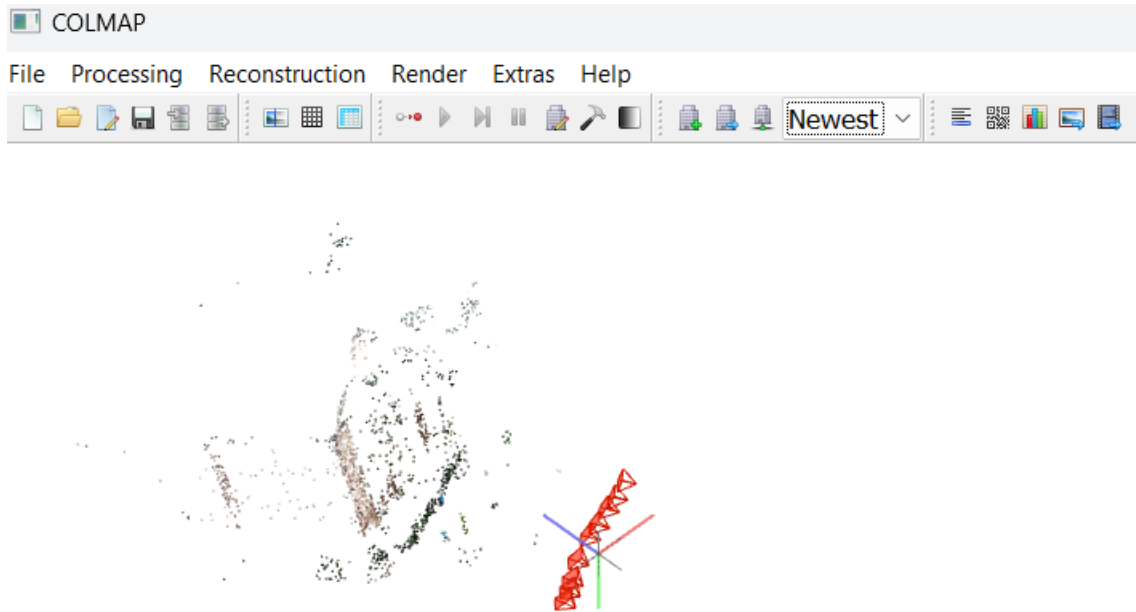
=> Image sees 520 / 1588 points
Registering image #10 (5)
=> Image sees 605 / 1059 points
Registering image #6 (6)
=> Image sees 307 / 1455 points
Registering image #4 (7)
=> Image sees 414 / 1521 points
Registering image #3 (8)
=> Image sees 712 / 2619 points
Registering image #1 (9)
=> Image sees 1294 / 2402 points
Registering image #5 (10)
=> Image sees 541 / 1249 points

Exhaustive Output:

Sequential Output:

All 20 images:

```
================================================================
Feature extraction
================================================================
```

images	features	Camera(qw, qx, qy, qz)(tx, ty, tz)
1	8597	(0.999843, 0.000461982, -0.00956244, 0.0149136) (4.37751, 0.738851, 2.93848)
2	9170	(0.999968, 0.00471436, -0.00636818, -0.000702601) (4.03822, 0.545181, 2.69651)
3	9597	(1, 0.000230292, -0.000259031, 0.000111829) (3.72724, 0.525424, 2.39413)
4	10079	(0.998631, 0.0206517, 0.0479864, 0.00252462) (3.20493, 0.370993, 1.63896)
5	10952	(0.99813, 0.0351094, 0.0500314, -0.000743614) (2.64933, 0.272481, 1.4512)
6	12484	(0.992025, 0.0134502, 0.124914, -0.0101257) (2.3966, 0.252393, 0.876561)
7	12669	(0.990664, -0.00147801, 0.133025, -0.0297722) (1.89297, 0.19704, 0.788813)
8	11905	(0.985386, -0.00701976, 0.167229, -0.031626) (1.24986, 0.195235, 0.614924)
9	11051	(0.973006, 0.00286672, 0.227963, -0.0358315) (0.77321, 0.170386, 0.534664)
10	11358	(0.968209, 0.00222961, 0.245498, -0.0479215) (0.0576351, 0.168477, 0.549135)
11	11784	(0.959136, 0.003897, 0.277791, -0.0536239) (-0.609724, 0.163816, 0.561513)
12	11944	(0.945281, -0.00582422, 0.319013, -0.0681176) (-1.27432, 0.171813, 0.61867)
13	11613	(0.92013, 0.00553692, 0.383427, -0.0794585) (-1.92677, 0.131504, 0.742624)
14	13152	(0.904598, 0.00727188, 0.414731, -0.0982187) (-2.61682, 0.221002, 0.956433)
15	8448	(0.885135, -0.00430989, 0.454764, -0.0985253) (-3.22441, 0.247084, 1.3593)
16	12147	(0.86132, 0.00163896, 0.494891, -0.114925) (-3.66362, 0.331758, 1.58102)
17	11953	(0.8517, 0.00938666, 0.511425, -0.113863) (-4.32834, 0.284482, 1.53246)
18	12186	(0.839172, 0.0196981, 0.529658, -0.121917) (-5.17076, 0.362136, 1.27525)
19	8358	(0.837132, 0.0122385, 0.535816, -0.109373) (-5.67339, 0.10427, 0.893395)
20	12698	(0.835011, 0.0260151, 0.533569, -0.131844) (-5.63974, 0.170753, 0.17961)

	Camera	Images	Points	Observations	Mean(1)	Mean(2)	Mean(3)
Exhaustive	20	20	4746	18750	3.9507	937.5	0.570304
Sequential	20	20	4732	18699	3.95161	934.95	0.569696

===
Feature matching
===
Global bundle adjustment
Registering image #8 (3)
=> Image sees 220 / 1766 points
Registering image #9 (4)
=> Image sees 573 / 1802 points
Registering image #10 (5)
=> Image sees 673 / 1576 points
Registering image #11 (6)
=> Image sees 597 / 1444 points
Registering image #12 (7)
=> Image sees 478 / 1175 points
Registering image #6 (8)
=> Image sees 350 / 1555 points
Registering image #4 (9)
=> Image sees 468 / 1592 points
Registering image #2 (10)
=> Image sees 716 / 2958 points
Registering image #1 (11)
=> Image sees 1169 / 2169 points
Registering image #5 (12)
=> Image sees 596 / 1308 points
Registering image #13 (13)
=> Image sees 297 / 890 points
Registering image #14 (14)
=> Image sees 197 / 771 points
Registering image #16 (15)
=> Image sees 208 / 1302 points
Registering image #15 (16)
=> Image sees 236 / 766 points
Registering image #17 (17)
=> Image sees 347 / 1397 points
Registering image #18 (18)
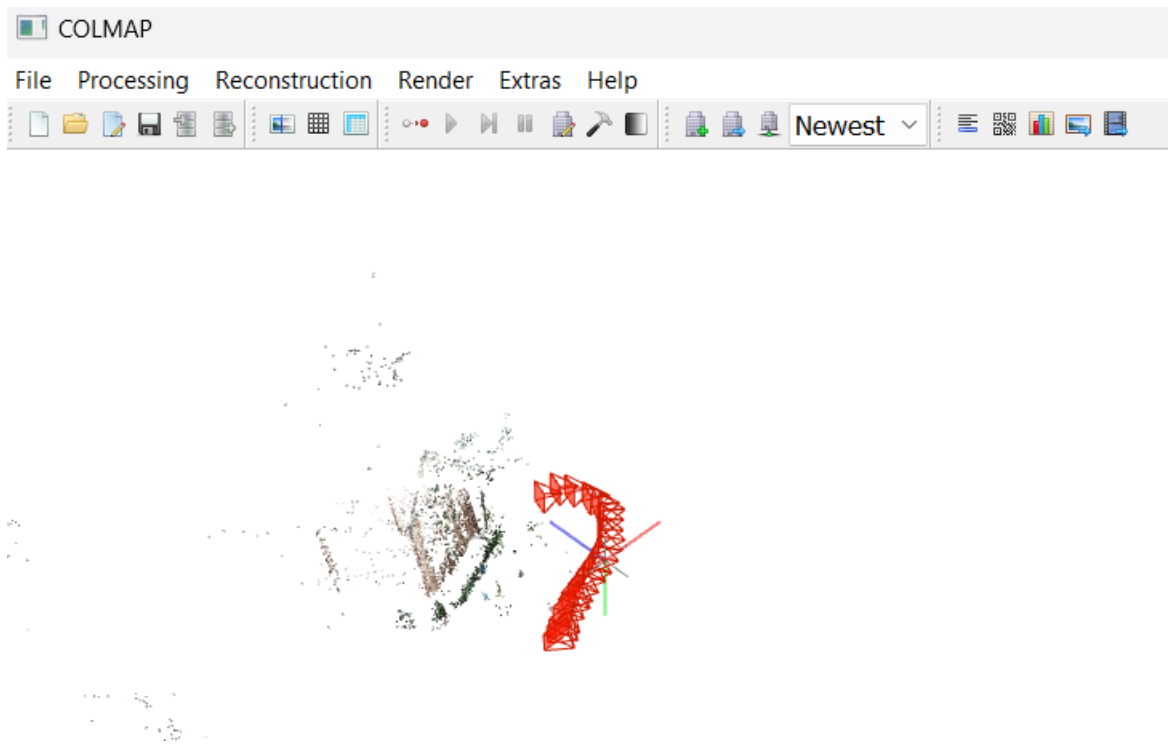=> Image sees 348 / 1957 points
Registering image #20 (19)
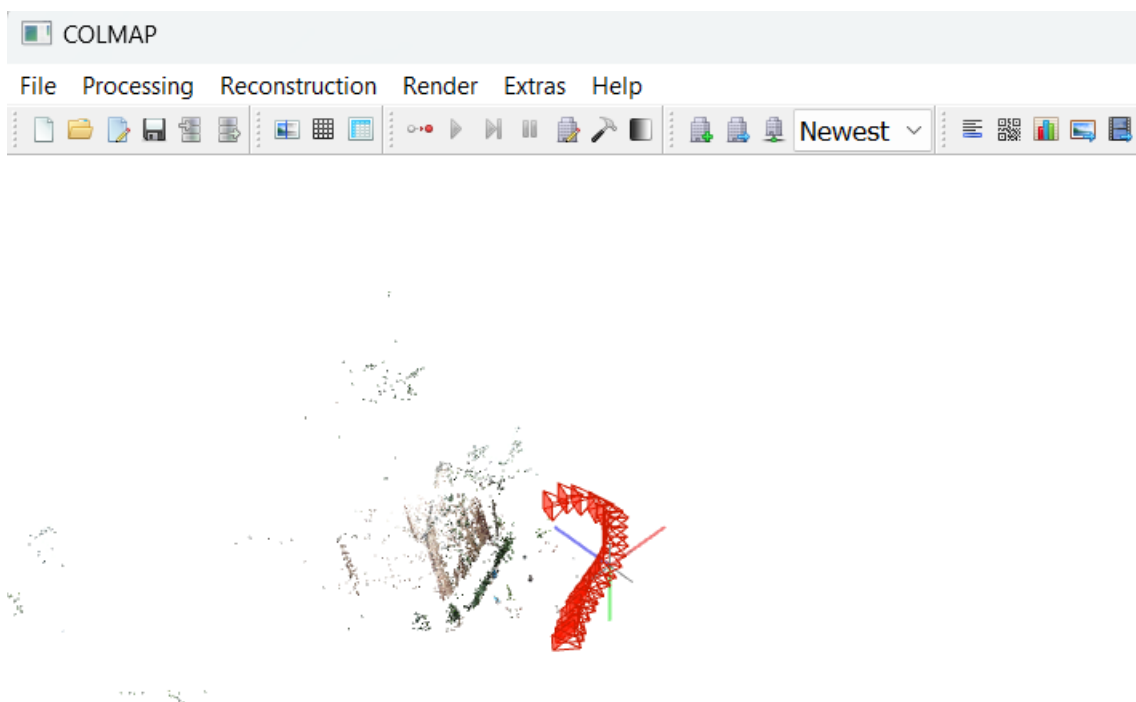=> Image sees 538 / 2098 points
Registering image #19 (20)
=> Image sees 796 / 1661 points

Exhaustive Output:

Sequential Output:

Observestion:

Each entry indicates how many points the registered image can see from the total available points:

- For example, for image #8, it sees 152 points out of a total of 1602 available points, which suggests that only a small fraction of the points in the scene were visible in this image. This can affect the quality of the reconstruction.

- Similarly, image #9 sees 520 points out of 1588, and image #10 sees 605 out of 1059, indicating varying levels of coverage and potentially different challenges in aligning the images.

- The feature extraction process yields a variable number of features depending on the image content and complexity.
- The feature matching process, highlighted by the global bundle adjustment, demonstrates how different images contribute to the overall reconstruction by showing the number of points visible to each image which is expected due to different camera viewpoints and occlusions.
- Images that see more points are likely to produce better results in terms of accuracy and completeness of the final model.

2 Gaussian Splatting

First 10 images excluding 5th image

30000 Iteration:

7000 Iteration:

All 20 images excluding 5th,10th & 15th image

In this experiment, I used the first 20 images while excluding images #5, #10, and #15 from the training set for the Gaussian Splatting model. Rendering the scene based on the camera poses for these excluded images yielded an L1 loss of 0.0915 and a PSNR of 16.41 dB at iteration 7000, indicating moderate quality. The SSIM was 0.5711, and LPIPS was 0.3656, highlighting perceptual differences. In contrast, the training set showed better performance with an L1 loss of 0.0236 and a PSNR of 28.19 dB. Comparing iterations 5000 and 7000 revealed that while training metrics improved, test set performance remained relatively stable.

7000 Iteration:

Images	Iterations	SSIM	PSNR	LPIPS
10	100	0.3571112	11.1152649	0.6627116

10	200	0.4025256	14.0389118	0.6317880
10	30000	0.5012640	14.9089890	0.3759232
20	7000	0.5710956	16.3903465	0.3656400
20	500	0.4229152	15.8951702	0.6097611

Metrics Observation:

- **PSNR**: Measures the peak error between the original and rendered images. Higher values indicate better quality.

- **SSIM**: Measures the similarity between two images. Values range from -1 to 1, with 1 indicating perfect similarity.

- **LPIPS:** captures perceptual similarity using deep learning features, with lower values indicating more similarity.