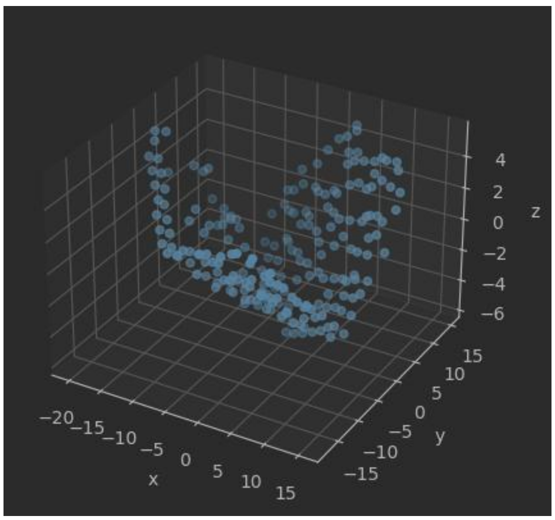
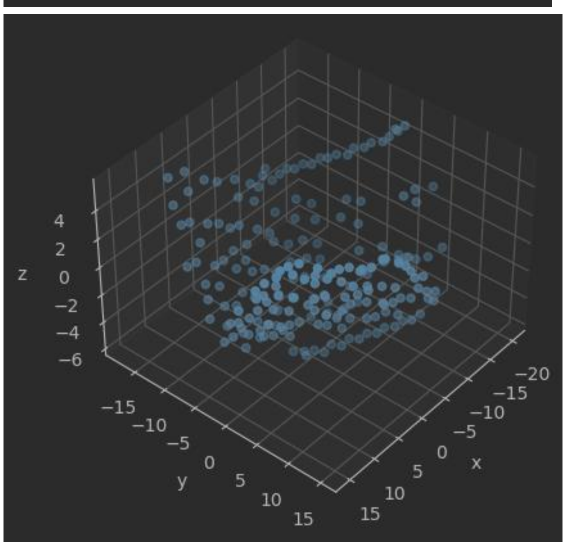
CS543/ECE549 Assignment 5

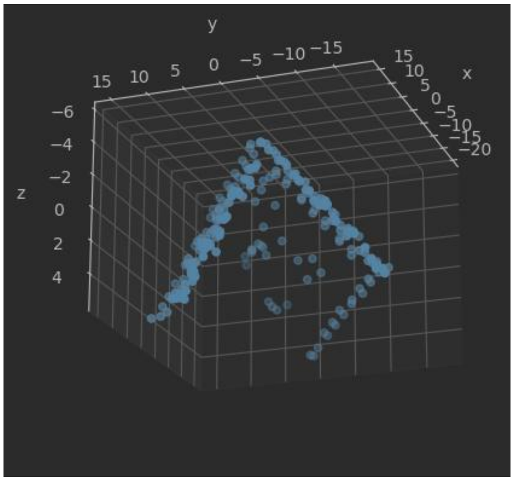
Name: Siyu Ren
NetId: siyuren2

Part 1: Affine factorization

A: Display the 3D structure (you may want to include snapshots from several viewpoints to show the structure clearly). Report the Q matrix you found to eliminate the affine ambiguity. Discuss whether or not the reconstruction has an ambiguity.
3D structure:





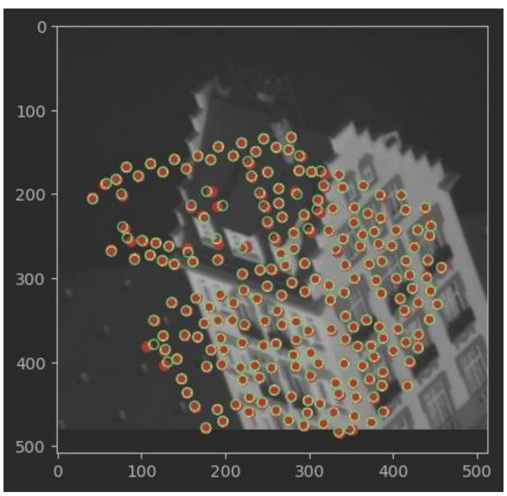


B: Display three frames with both the observed feature points and the estimated projected 3D points overlayed.

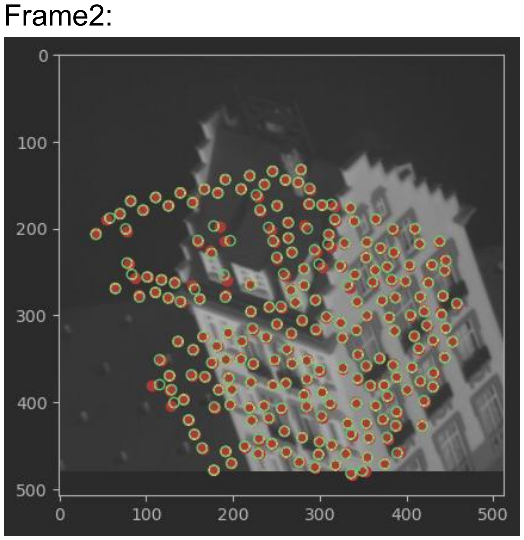
The observed points is the red points.

The estimated points are the green circles.

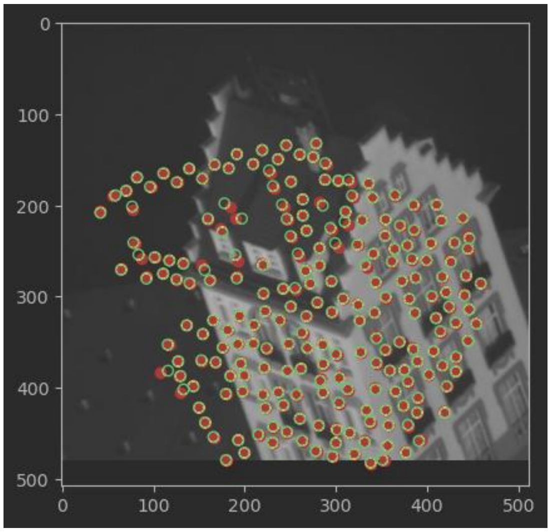
Frame1:



Frame2:



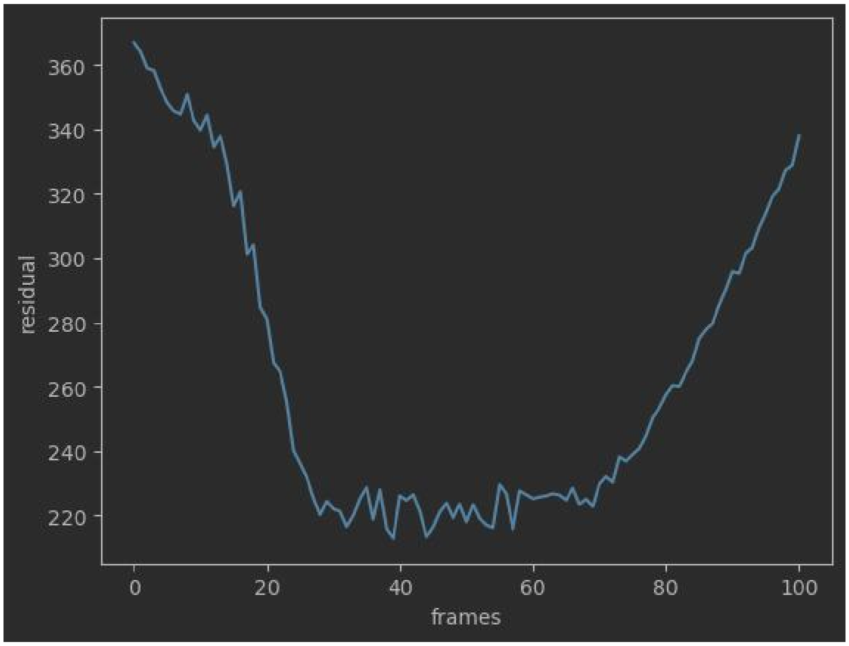
Frame3:



C: Report your total residual (sum of squared Euclidean distances, in pixels, between the observed and the reprojected features) over all the frames, and plot the per-frame residual as a function of the frame number.

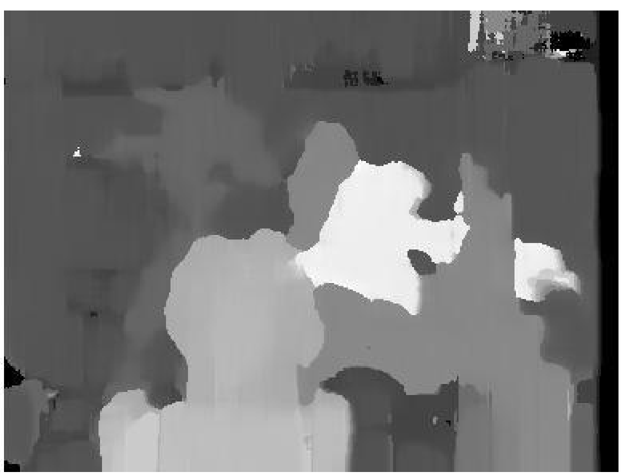
Total residual:

[366.9767987 364.09019959 358.96011927 358.25363934 352.53304969
348.34604267 345.59708926 344.61683825 350.81780577 342.77218058
339.70622487 344.33706819 334.44969983 337.89052734 329.14846611
316.09133198 320.63150003 301.22082641 304.06827237 284.68474511
281.1248818 267.47805293 264.72732726 255.1585075 240.21657594
236.17342945 231.86652927 225.51964492 220.3107584 224.498312
222.09816385 221.34576273 216.62694301 220.08049801 225.40130305
228.81014889 218.88141627 228.04669907 216.02873165 212.91413671
226.19342598 224.79060018 226.51985283 221.31389586 213.40891342
216.48335808 221.10735338 223.91644125 219.34055687 223.64328262
218.10726811 223.4357223 219.11244542 217.1850391 216.24172561
229.62910069 226.81456514 215.86860096 227.71211304 226.51975011
225.2867486 225.80438911 226.15046874 226.82586643 226.38364535
224.8760545 228.53396479 223.55856771 225.13825459 222.8604424
229.78598218 232.11698602 230.36405524 238.24749188 236.96668894
238.9045237 240.79133241 244.81901083 250.39288446 253.44801429
257.46192866 260.36986316 260.09022569 264.48152167 267.97877959
274.96111204 277.72258667 279.68241378 285.2391338 290.13835199
295.75659604 295.20135673 301.43601078 303.20197755 309.35612909
313.7843151 319.10842614 321.35483343 327.29535624 328.87518012
337.96555387]

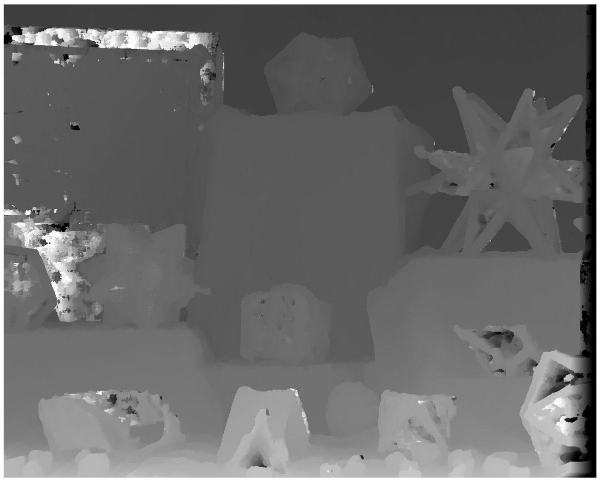


Part 2: Binocular stereo

A: Display best output disparity maps for both pairs.
Tsukuba:

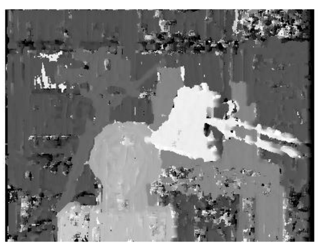
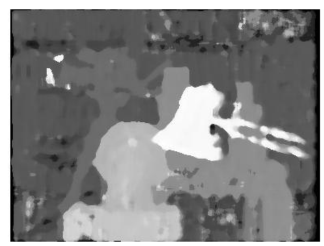


Moebius:



B: Study of implementation parameters:

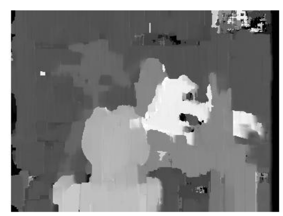
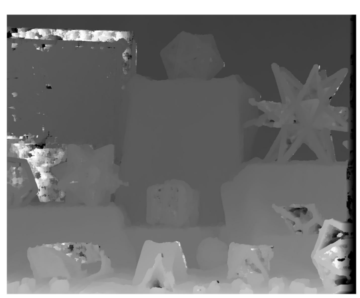
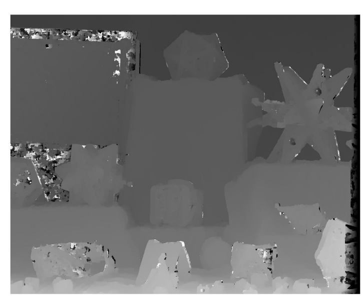
1. **Search window size:** show disparity maps for several window sizes and discuss which window size works the best (or what are the tradeoffs between using different window sizes). How does the running time depend on window size?

Window size	3 * 3	5 * 5
Disparity map	 Time = 1.52s	 Time = 1.66s
Window size	11 * 11	19 * 19
Disparity map	 Time = 2.43s	 Time = 4.15s

There are more noise in the mao when window size is smaller. With larger window size, the noise becomes smaller, but the running time increases and map becomes blurrier. Window size = 15 works the best.

2. **Disparity range:** what is the range of the scanline in the second image that should be traversed in order to find a match for a given location in the first image? Examine the stereo pair to determine what is the maximum disparity value that makes sense, where to start the search on the scanline, and which direction to search in. Report which settings you ended up using.
   For each pixel on the second image, I searched the scanline on the same horizontal line. in range [max(- disparity range, 1 - left boundary of the window)]. It turned out that only searching to the right will still give good results. So I set the range to [0, 1 - left boundary of the window], which means start from the current pixel and searches to right.
   The maximum disparity I used for tsukuba is 15.
   The maximum disparity I used for moebius is 70.

3. **Matching function:** try sum of squared differences (SSD), sum of absolute differences (SAD), and normalized correlation. Discuss whether there is any difference between using these functions, both in terms of quality of the results and in terms of running time.

SSD	SAD	NCC
 Time = 6.22s	 Time = 3.22s	 Time = 2.85s
 Time = 532.27s	 Time = 217.41s	 Time = 197.58s

The quality of results of SSD and SAD is relatively the same. The quality of result of NCC is better than SSD and SAD. you ended up using.
The running time of SSD is much larger than SAD and NCC. The running time of NCC is the smallest.

C: Discuss the shortcomings of your algorithm. Where do the estimated disparity maps look good, and where do they look bad? What would be required to produce better results? Also discuss the running time of your approach and what might be needed to make stereo run faster.

The objects look good in the result image, but the edges of different objects look bad. There is a lot of noise at the edges of objects. We can refine the result image by eliminating the noise using median filter. The median filter takes about 1s.

Part 3: Extra Credit

Post any extra credit for parts 1 or 2 here. Don’t forget to include references, an explanation, and outputs to receive credit. Refer to the assignment for suggested outputs.