CS543/ECE549 Assignment 5

Name: Hanwen Zhang
NetId: hanwen6

<u>Part 1: Affine factorization</u>

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.

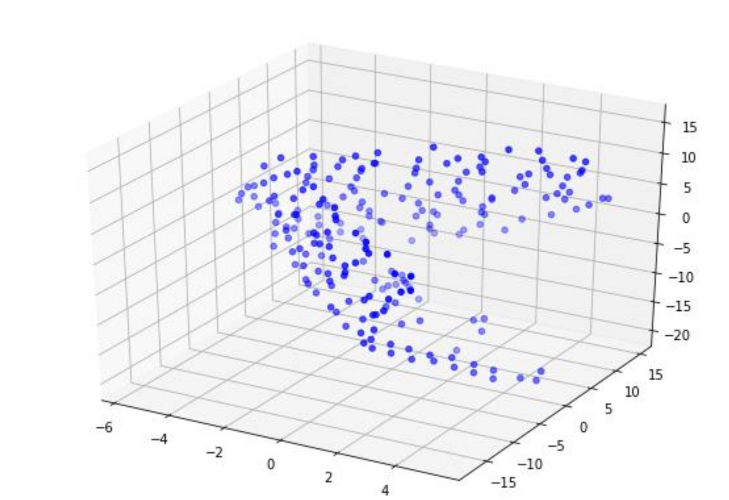


Fig 1.1. 3D structure

Q matrix:

[[0.07969843	0.	0.	]
[0.00070199	0.08500133	0.	]
[0.001331	0.00767057	0.03722294	]]

According to the outputs above, the reconstruction does not have an ambiguity.

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

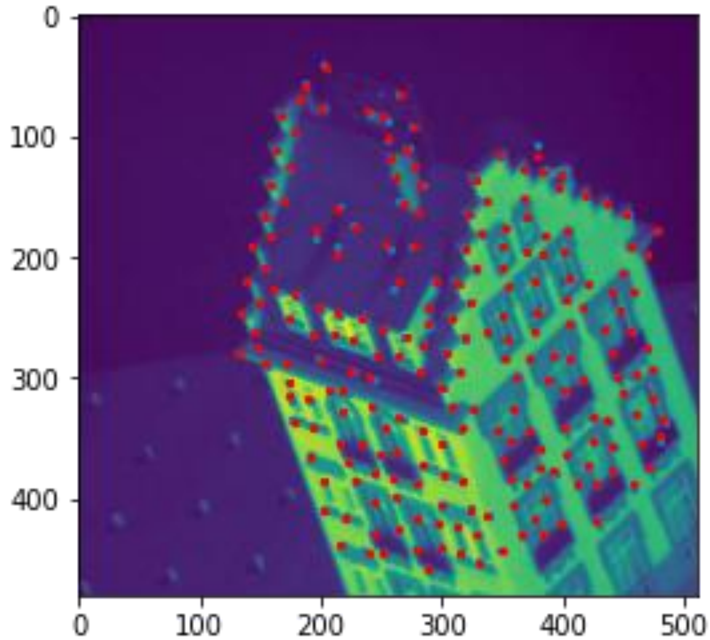


Fig 1.2. Frame 01

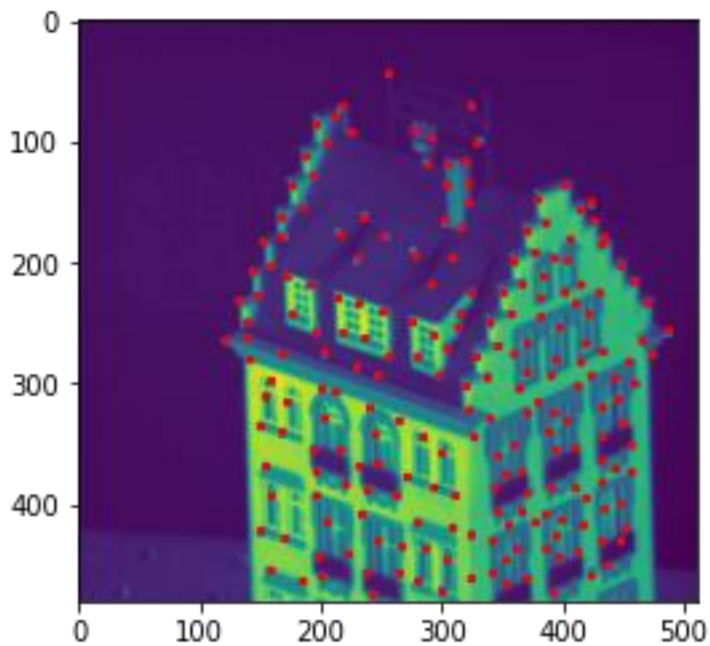


Fig 1.2. Frame 02

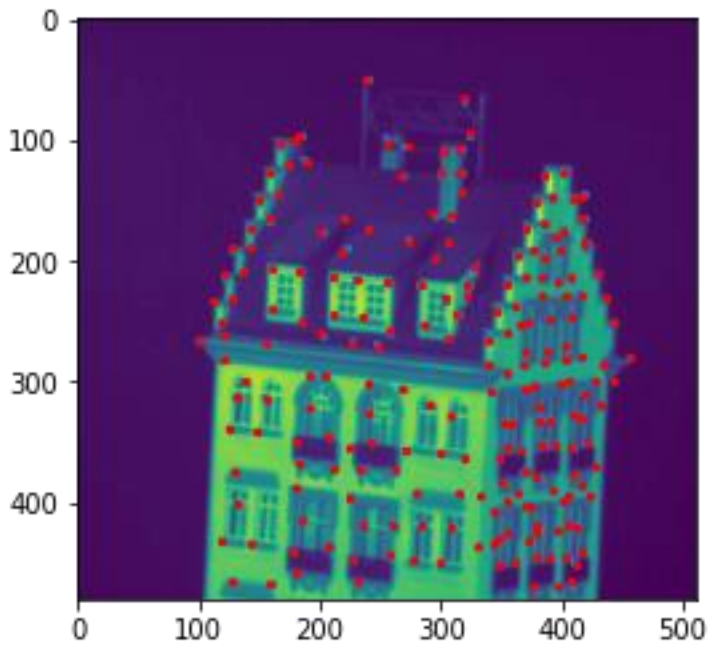


Fig 1.3. Frame 03

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: 16428.33206303

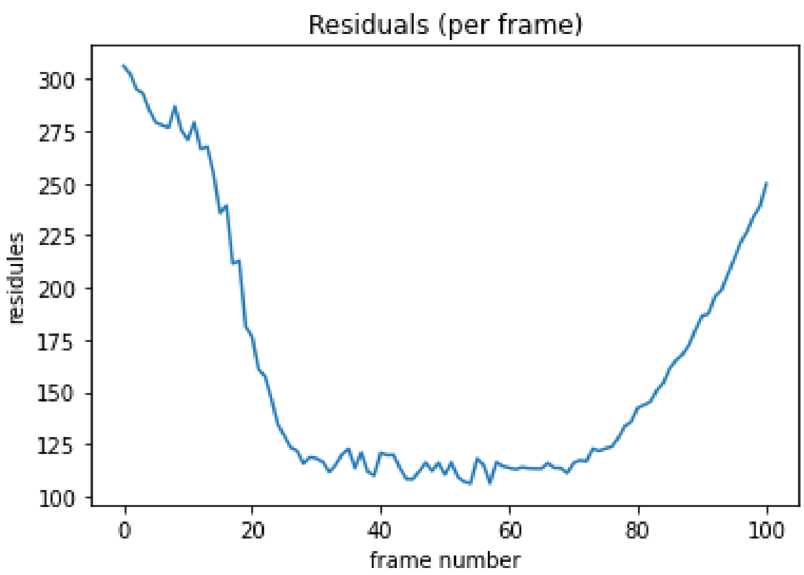


Fig 1.4. Residuals (per frame)

<u>**Part 2: Binocular stereo**</u>

A: Display best output disparity maps for both pairs.

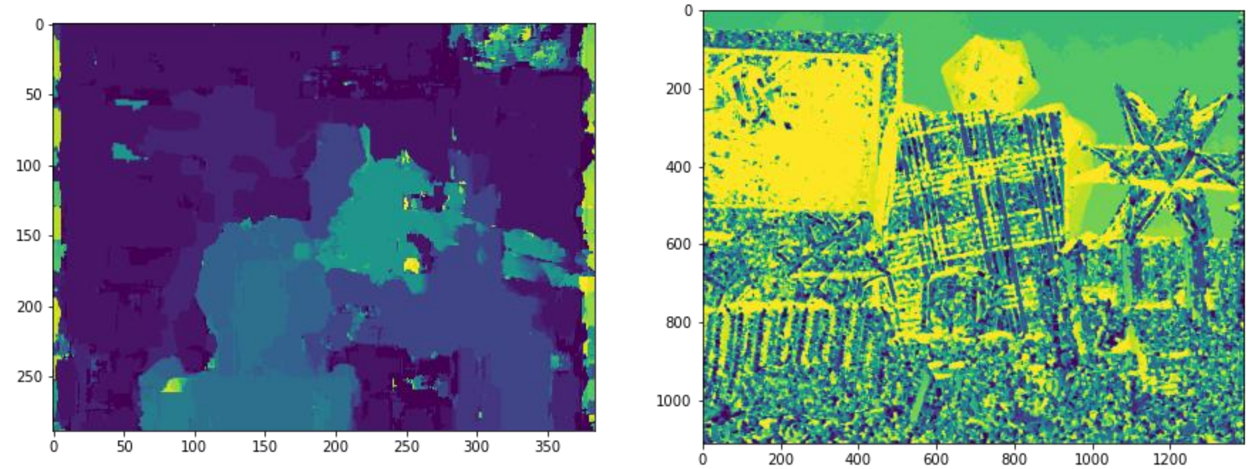


Fig 2.1. Best output disparity maps

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?

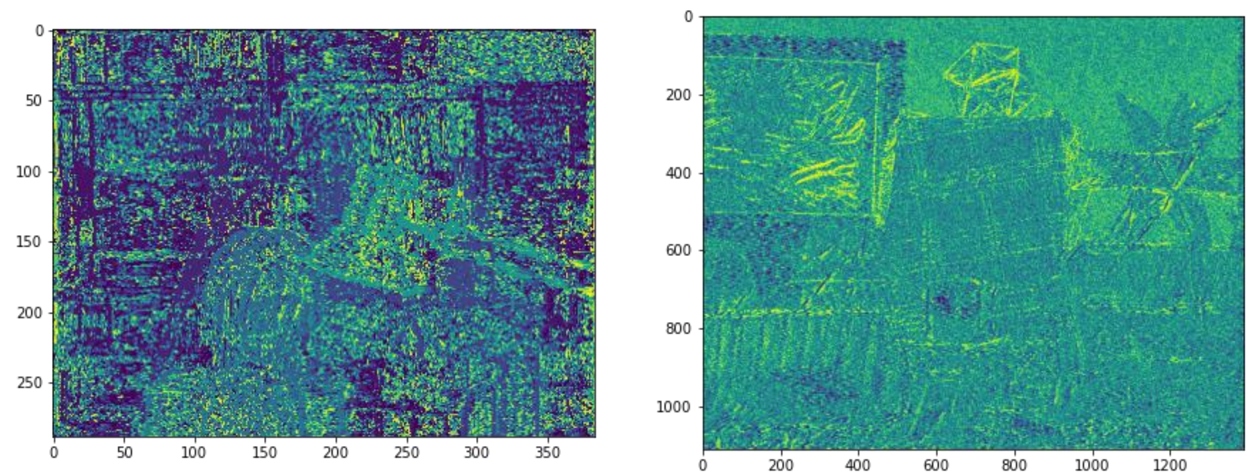


Fig 2.2. Window Size 2

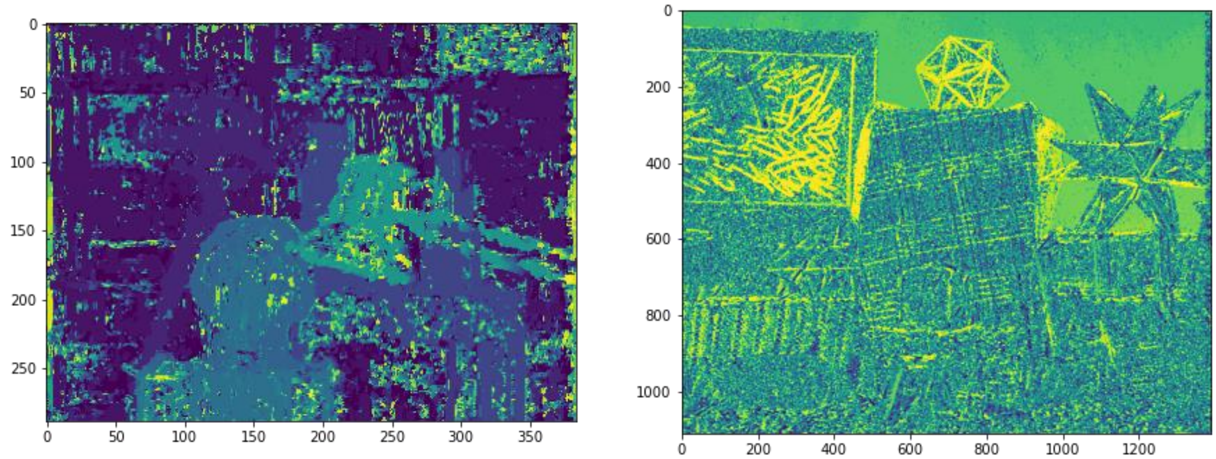


Fig 2.2. Window Size 4

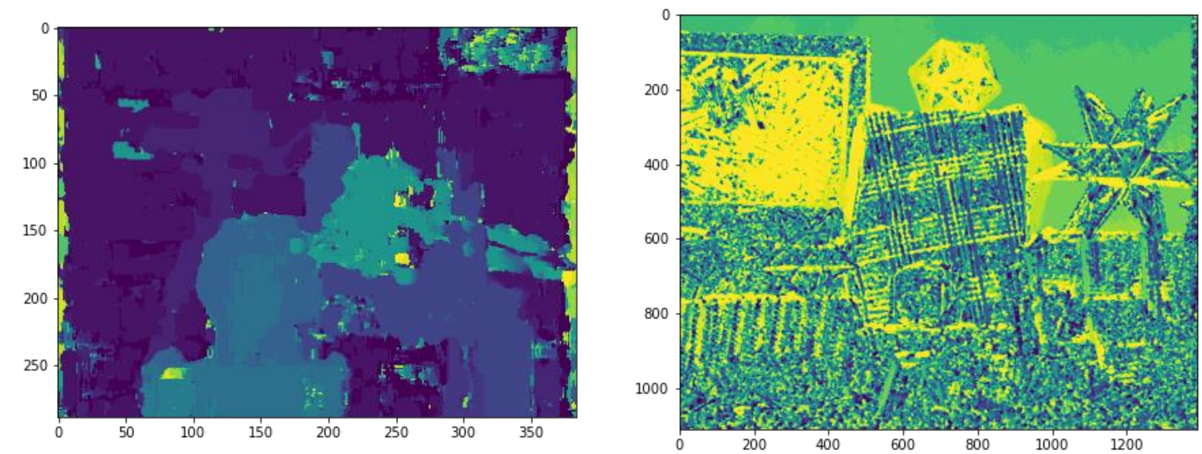


Fig 2.2. Window Size 8

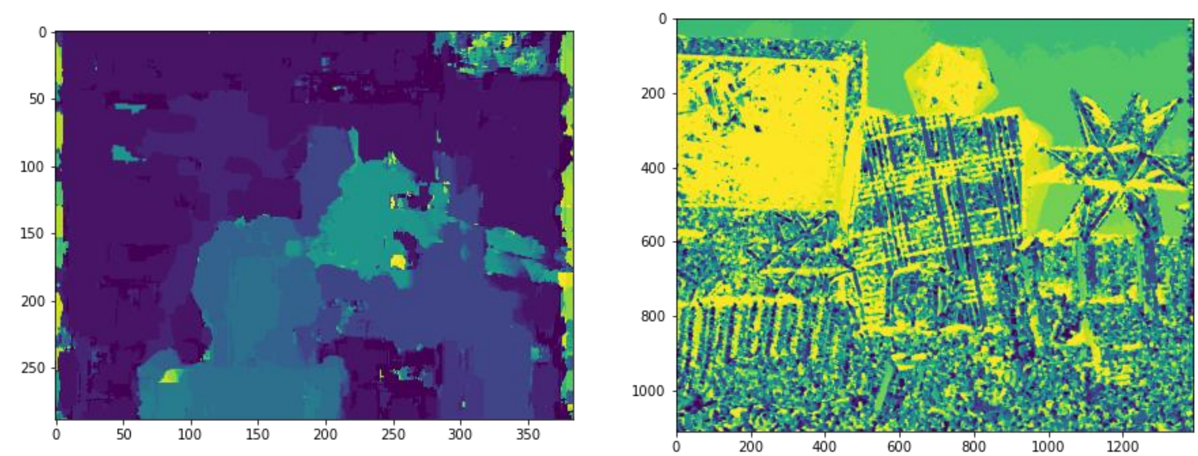


Fig 2.2. Window Size 10

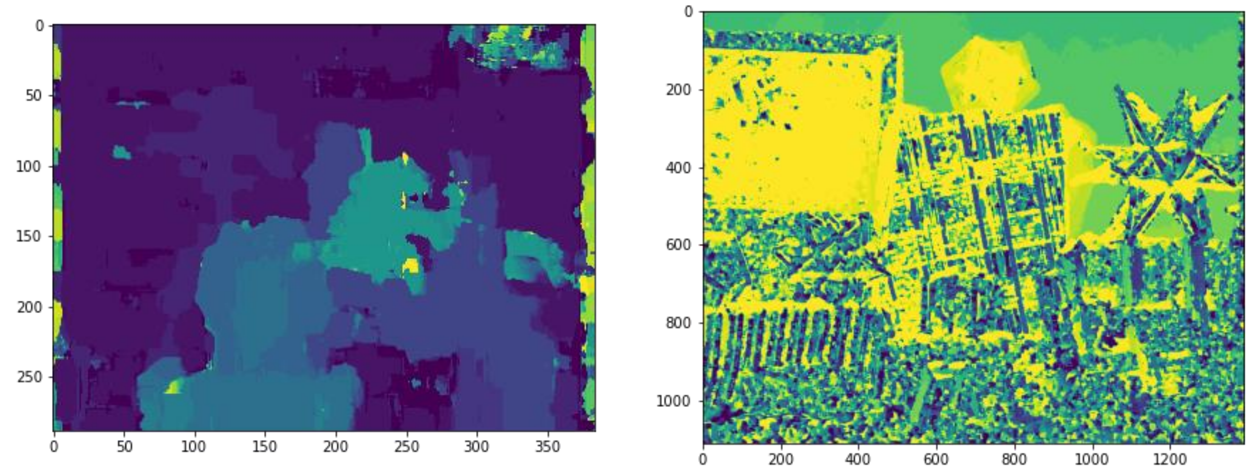


Fig 2.2. Window Size 12

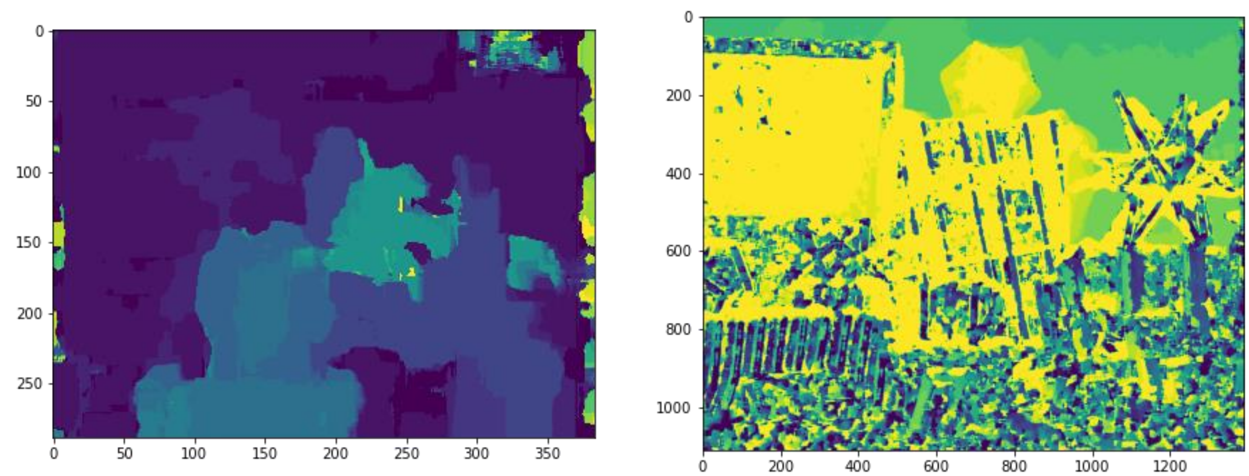


Fig 2.2. Window Size 16

The best window size is: 10

Increasing the window size will ultimately increase the execution time since there is more calculation needs to be done in each pixel. From the different wid values on the figure, in general, when the window value is very small, there are many miscellaneous points. When the window value is too large, the contour features of objects in the image are hidden and cannot be highlighted. From the analysis of the area at the lower right corner of each image, when the wid is small, there are many noise points so that the shape of the object can not be seen clearly. The contour is very thin. When the wid is large, we can see that the contour is very broad, and the original appearance is basically invisible. The parallax map effect is also poor. At the same time, we can find that the image is brighter when it is located in the front, that is, the place near the camera is brighter.

In the experiment, different windows have a certain impact on the matching results. When the window is set large, the image becomes relatively blurred and the details cannot be seen clearly; However, if the window is set too small, many other points will be generated, which will interfere with the representation of objects in the image. Therefore, the appropriate value can be found through multiple window value comparisons.

The accuracy of binocular ranging is difficult to analyze theoretically, and needs to be verified according to the actual situation. Generally speaking, the farther the distance is, the greater the error is. Therefore, binocular ranging is not suitable for measuring targets that are too far away. If you want to get a more reliable depth for a distant target, you need to increase the baseline distance of the camera, but at the same time, the difference between the left and right views will become larger, thus improving the difficulty of stereo matching.

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.

The function searches on the scan line from the original coordinates of the left image, with offsets on either side of the pixels of the right image, to make sure it can find the right window on the right image.

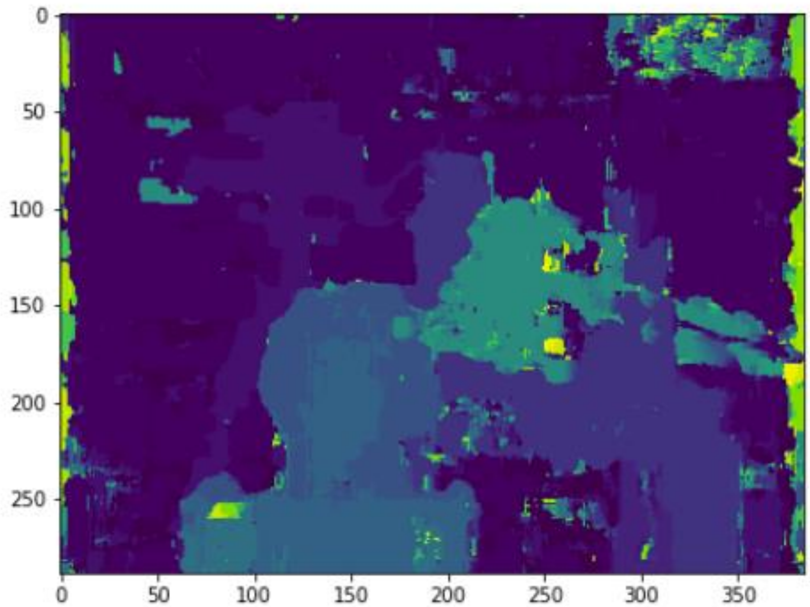
The range of the sacnline is [0, 19]

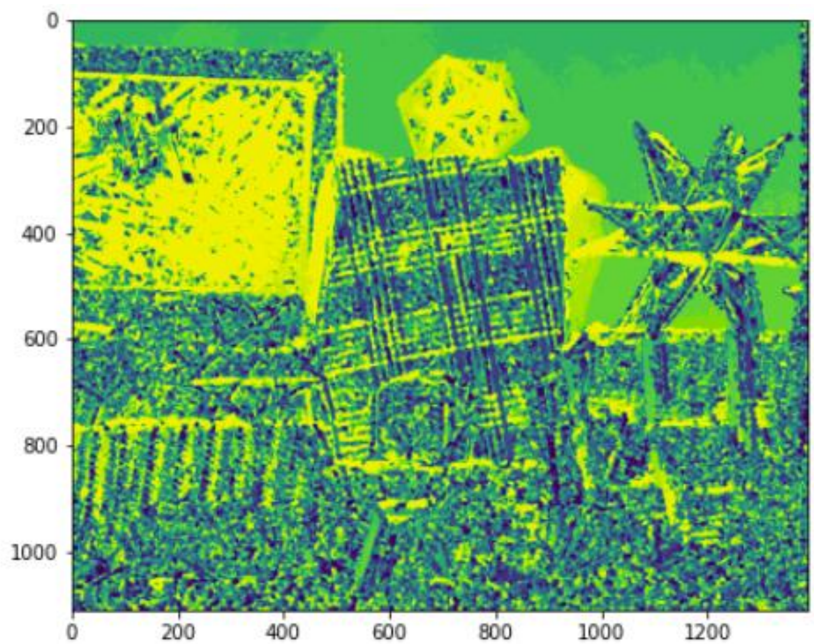
Maximun disparity value that makes sense: 1.0

Start from: 4

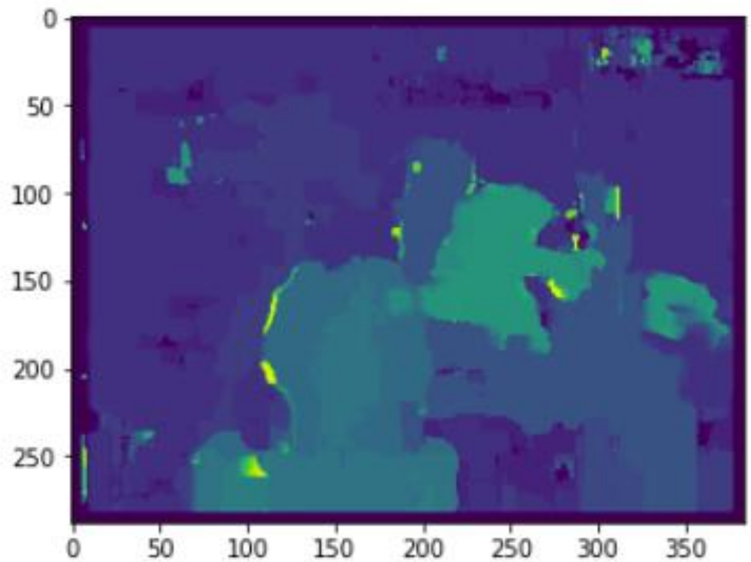
3. **Matching function:** try sum of squared differences (SSD), sum of absolute differences (SAD), and normalized correlation. Show the output disparity maps for each. Discuss whether there is any difference between using these functions, both in terms of quality of the results and in terms of running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

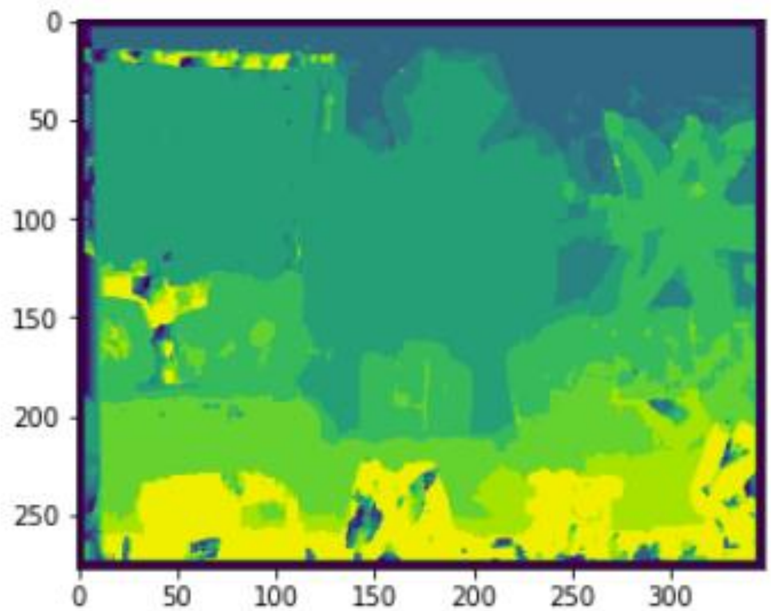
   NCC:



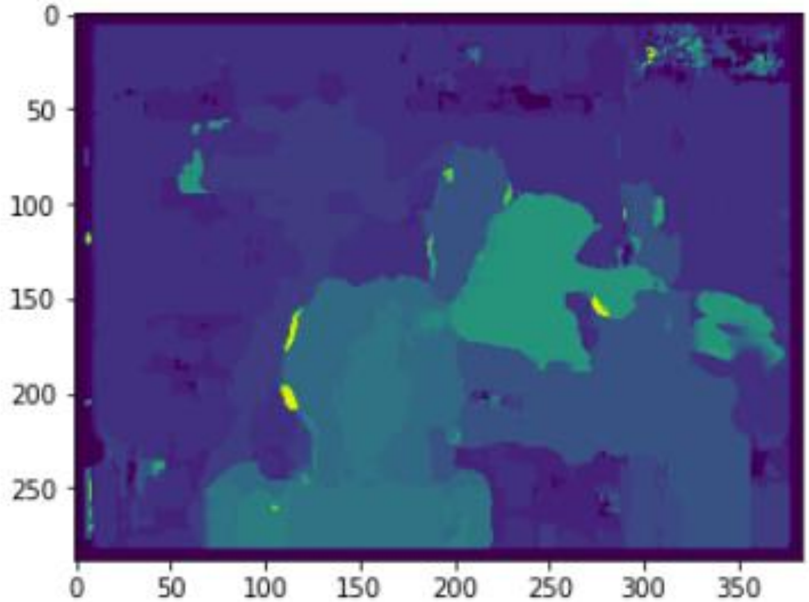


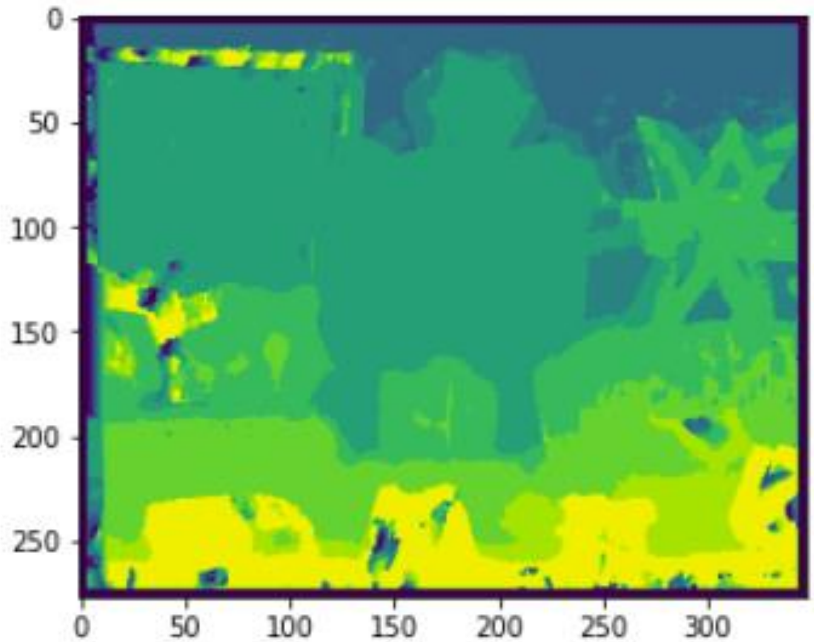
SSD:


0
50
100
150
200
250
0
50
100
150
200
250
300
350


0
50
100
150
200
250
0
50
100
150
200
250
300

SAD:



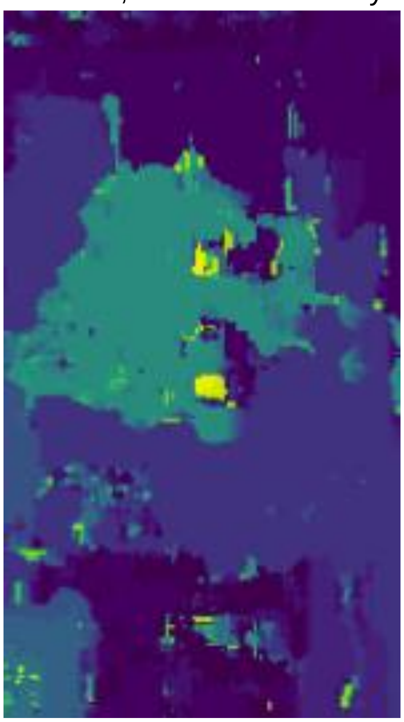


For down sampling, both SSD and SAD's run time are approximately 16s, for not down sampling, both SSD and SAD's run time are approximately 50s. NCC has a significant longer time of 55 seconds due to the amount of computation needed. SAD has more tolerable to noise as it has the clearest edge of all the different methods. NCC has the most detail than any of the methods, but that also means the result will contain more noises than the other two method.

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.

As the figure show below, the camera on the back is with good details and clean outlines.

However, there are still many significant noise on the disparity map as show in figure below.

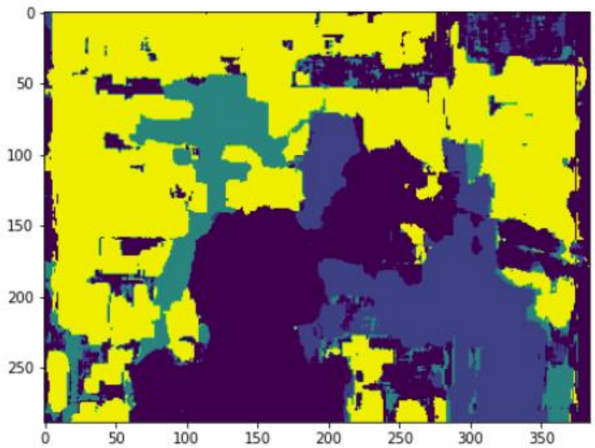


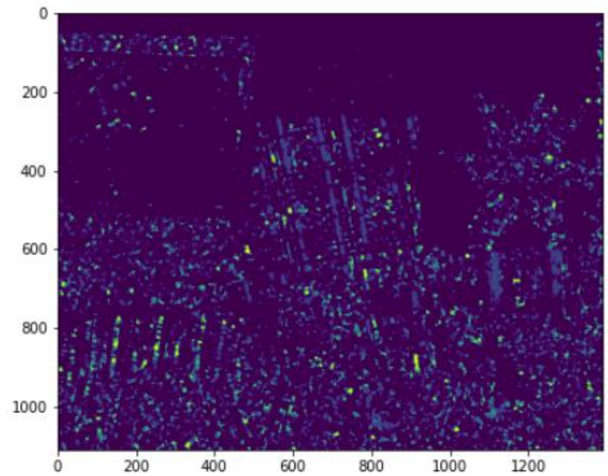
Therefore, it is necessary to do some preprocessing to the input image like use gaussian blur function to lower the noise of the background to produce better result.
For running time, I think down sampling the large input data might make stereo run faster.

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.

Parallax map converted to depth map





Convert depth map to point cloud map