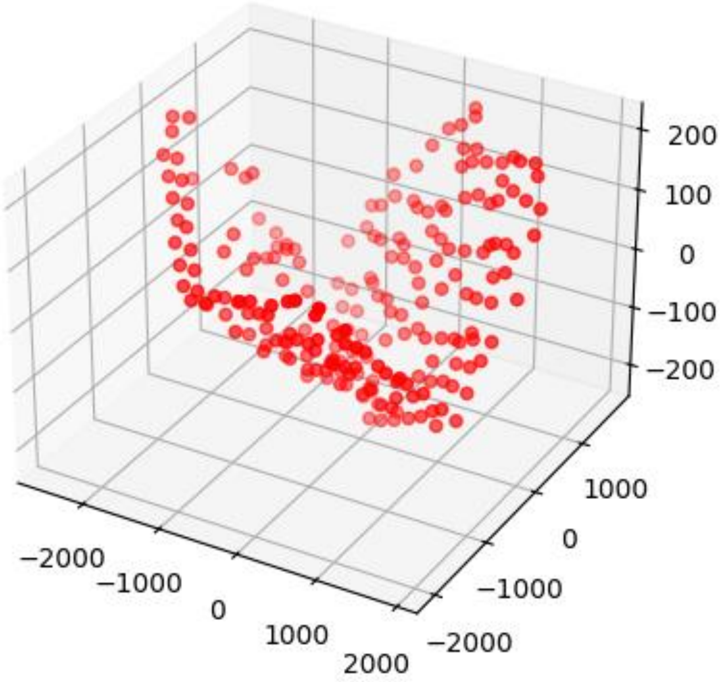
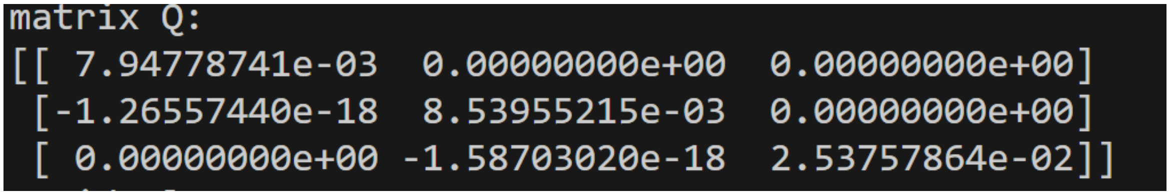
CS543/ECE549 Assignment 5

Name: Hanliang Jiang
NetId: hj33

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.

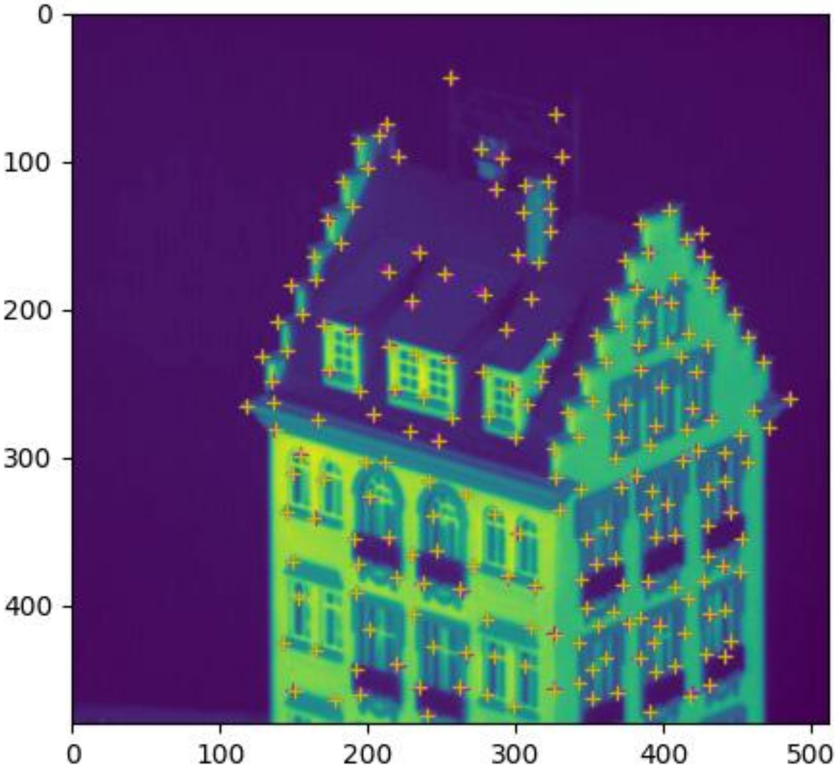


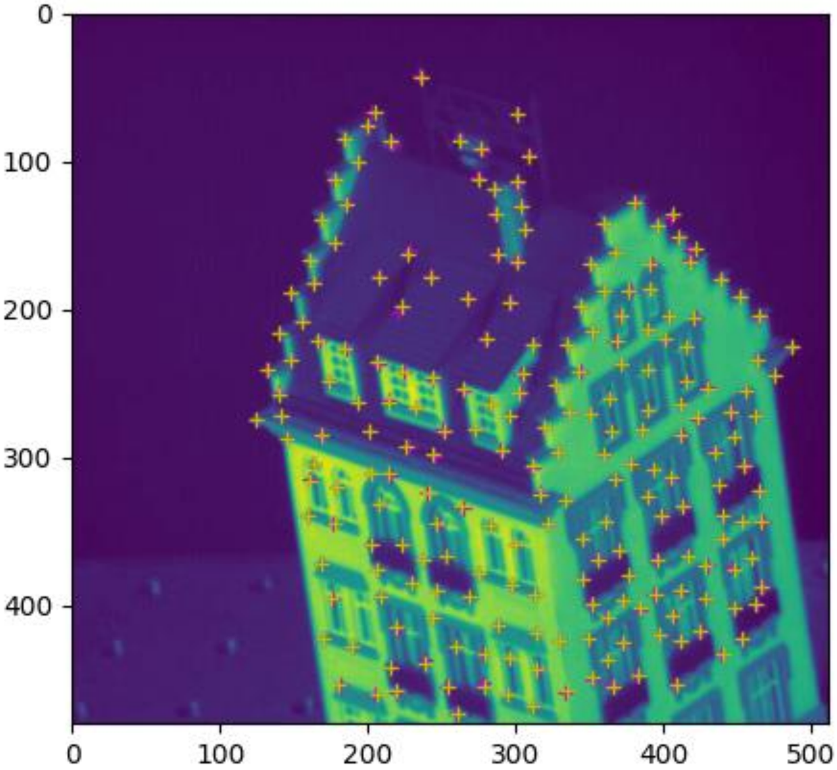


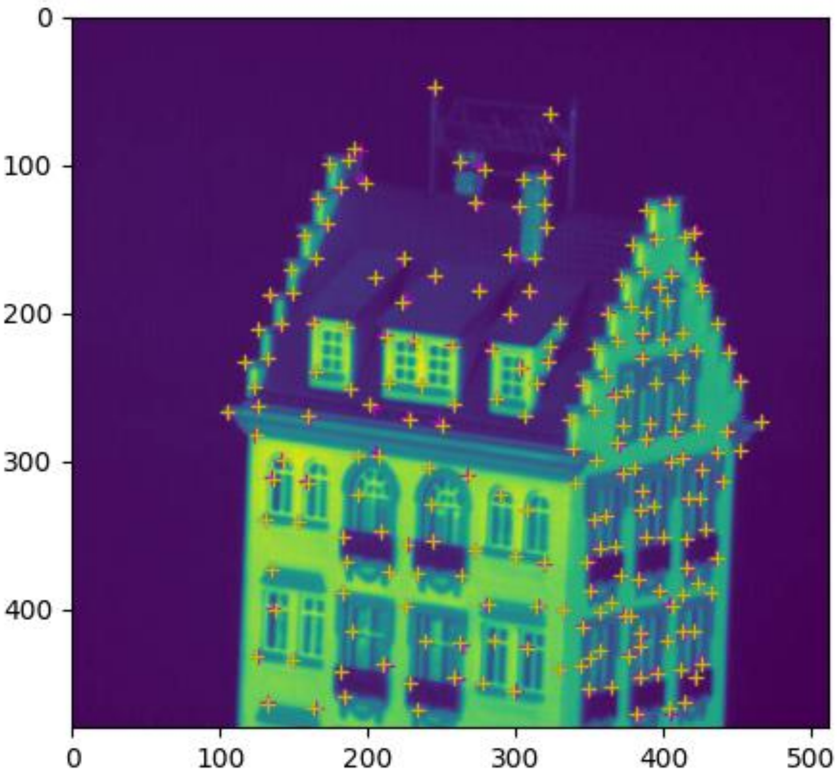
```
matrix Q:
[[ 7.94778741e-03  0.00000000e+00  0.00000000e+00]
 [-1.26557440e-18  8.53955215e-03  0.00000000e+00]
 [ 0.00000000e+00 -1.58703020e-18  2.53757864e-02]]
```

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

With frame = 30, 60, 90:

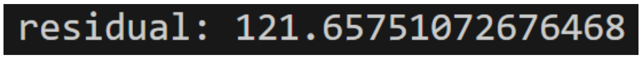

0
100
200
300
400
0
100
200
300
400
500


0
100
200
300
400
0
100
200
300
400
500

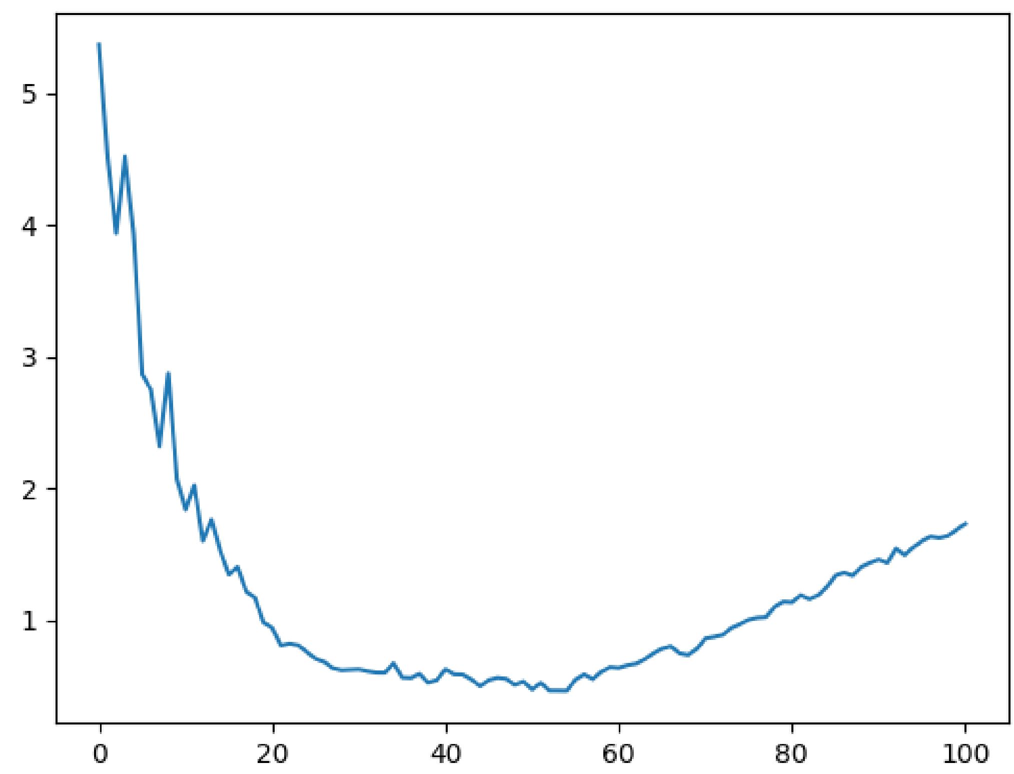


C: Report your total residual (sum of squared Euclidean distances, in pixels, between the observed and the reprojected features) over all the frames, and plot the per-frame residual as a function of the frame number.

Euclidean distances with norm=2

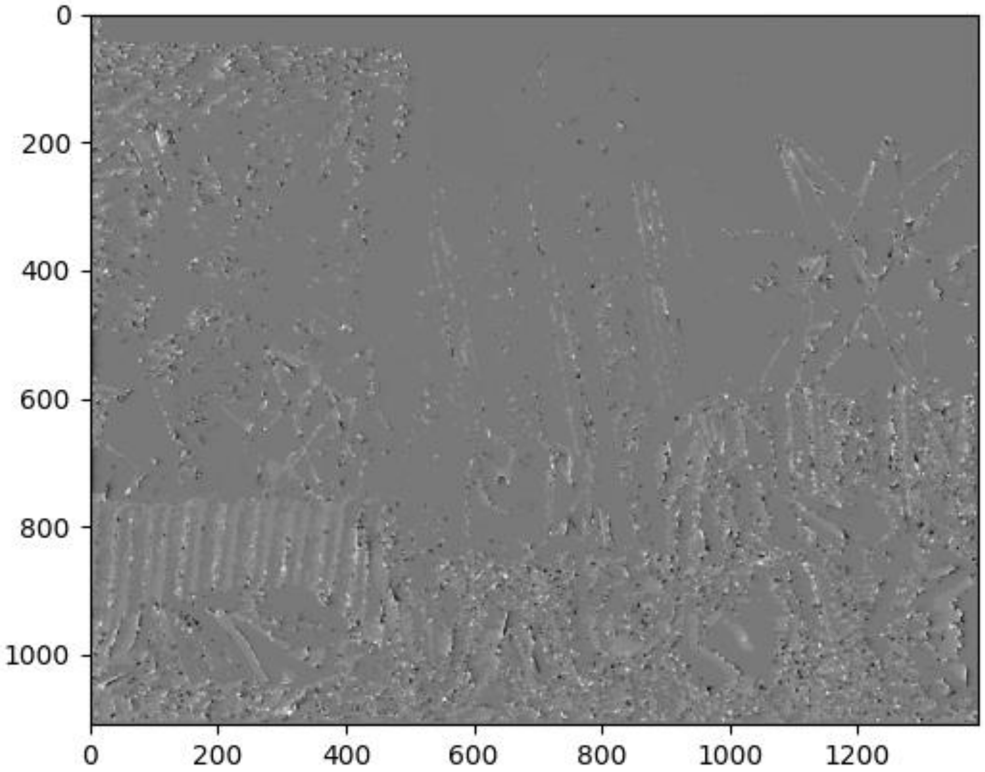


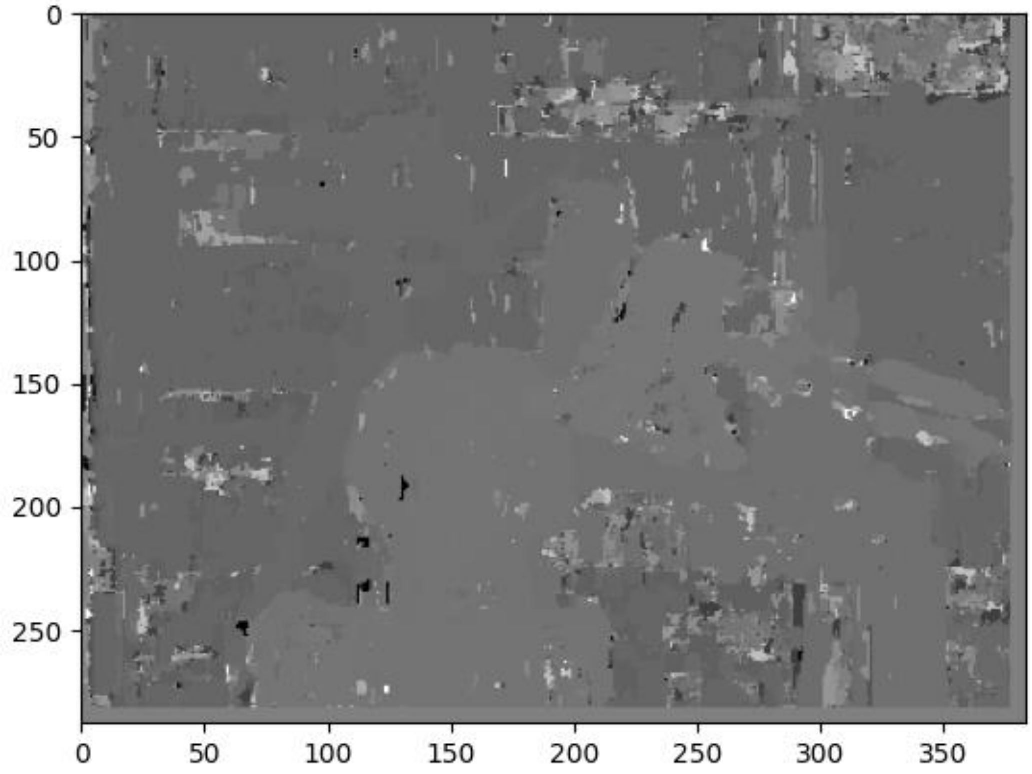
```
residual: 121.65751072676468
```



Part 2: Binocular stereo

A: Display best output disparity maps for both pairs.


0
200
400
600
800
1000
0
200
400
600
800
1000
1200

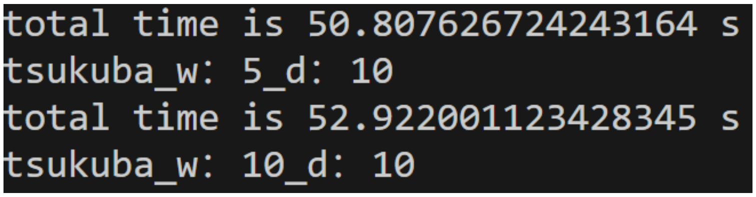


B: Study of implementation parameters:

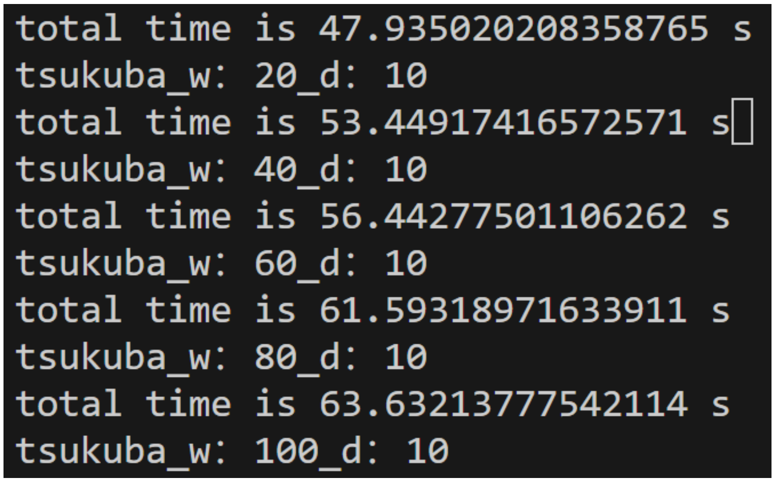
1. **Search window size:** show disparity maps for several window sizes and discuss which window size works the best (or what are the tradeoffs between using different window sizes). How does the running time depend on window size?

   Changing window size in range: [5, 15, 20, 40, 60, 80, 100], while disparity range=10 and method=SSD,

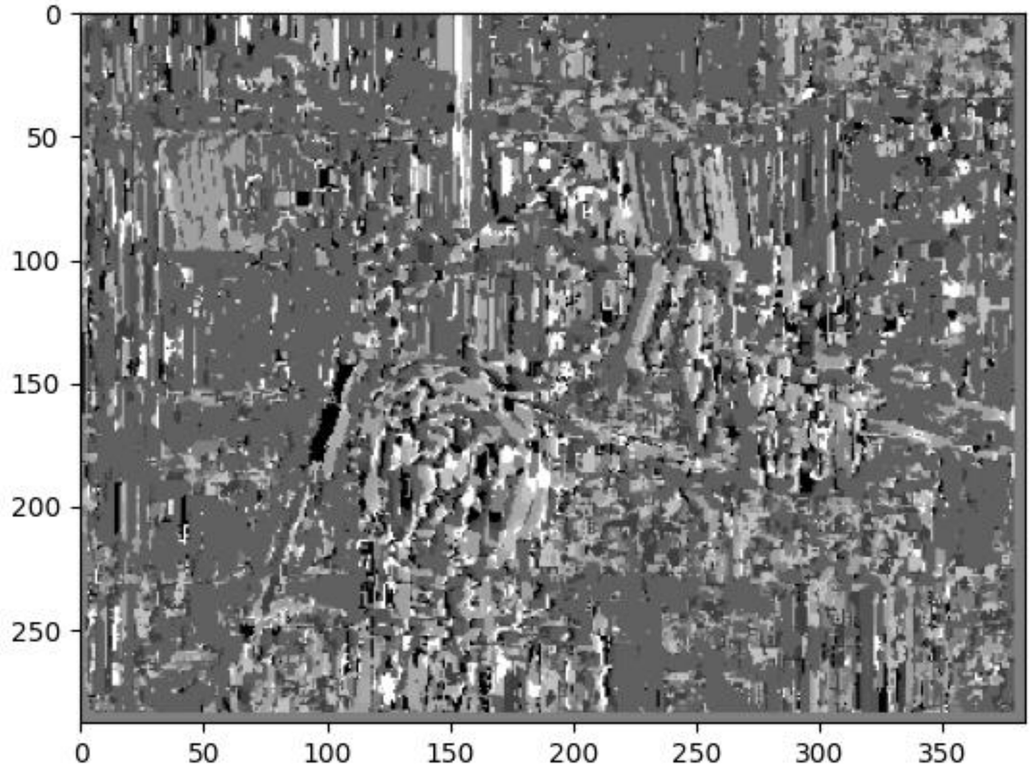
   I got the following result:

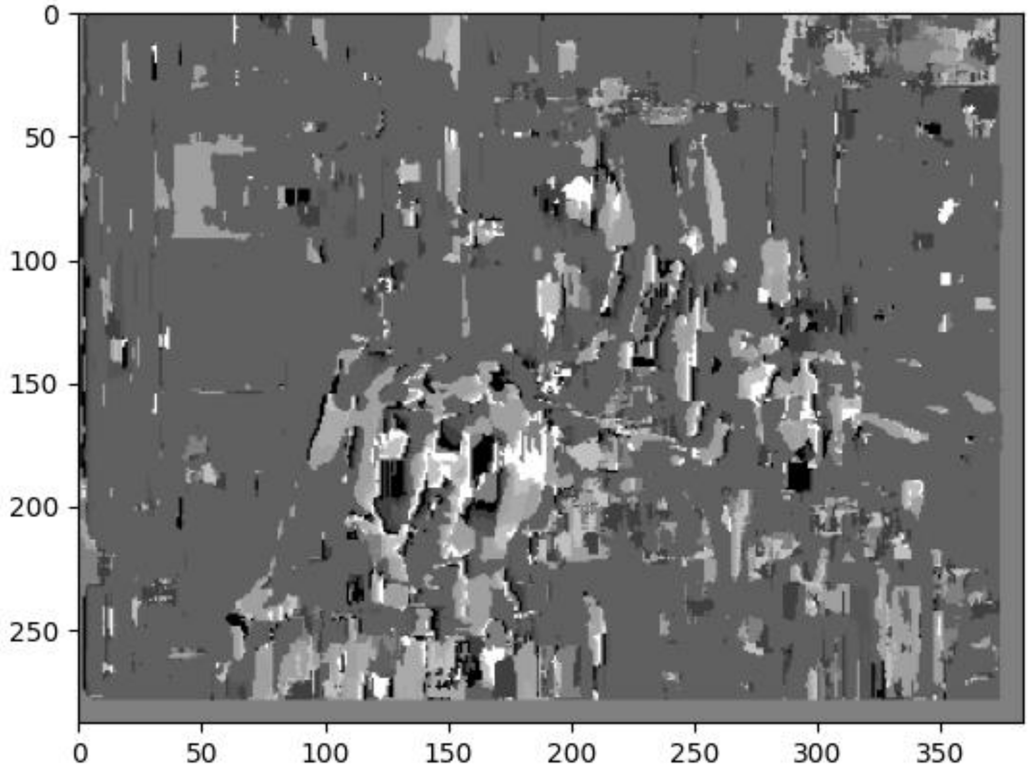


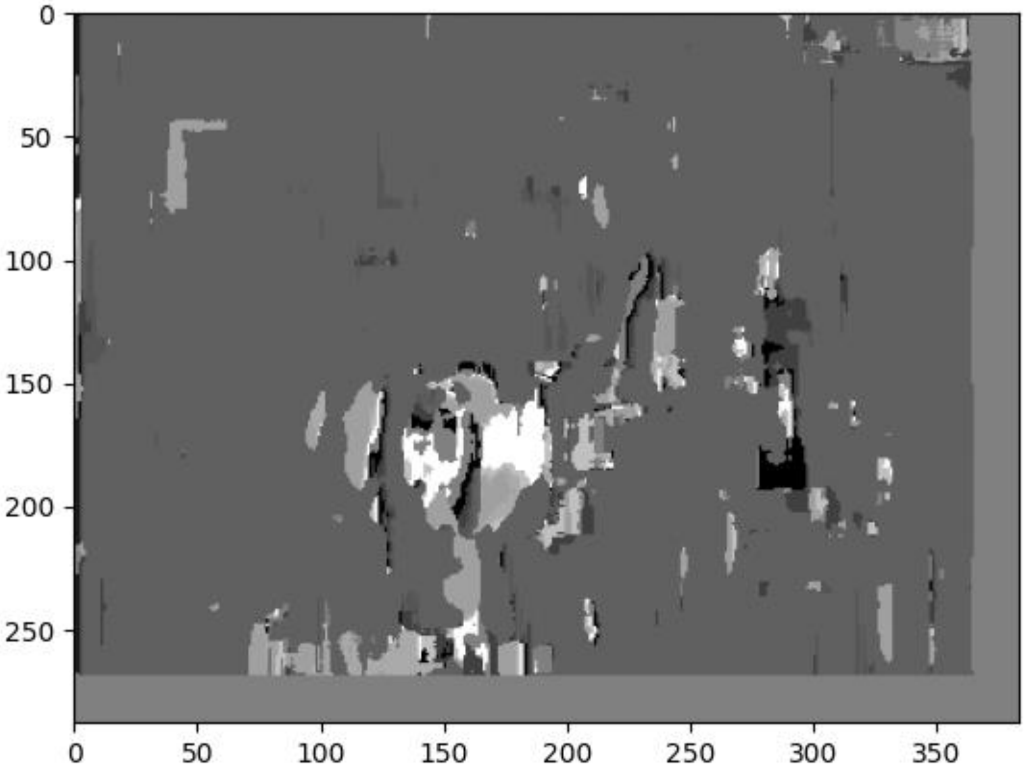
```
total time is 50.807626724243164 s
tsukuba_w: 5_d: 10
total time is 52.922001123428345 s
tsukuba_w: 10_d: 10
```

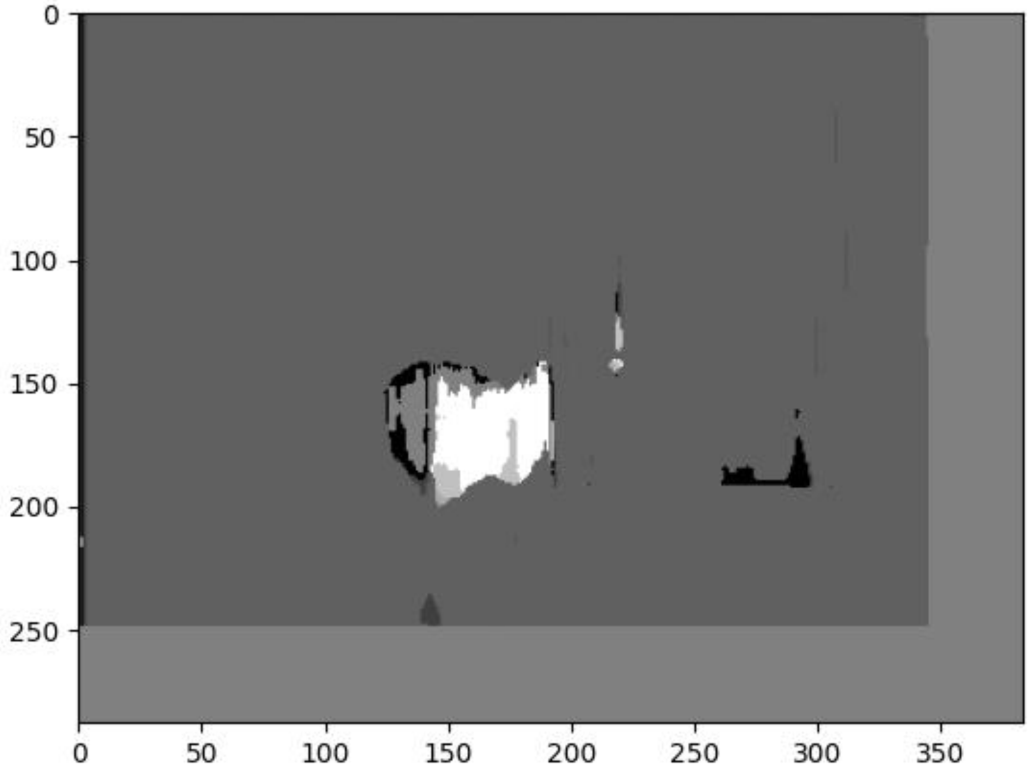


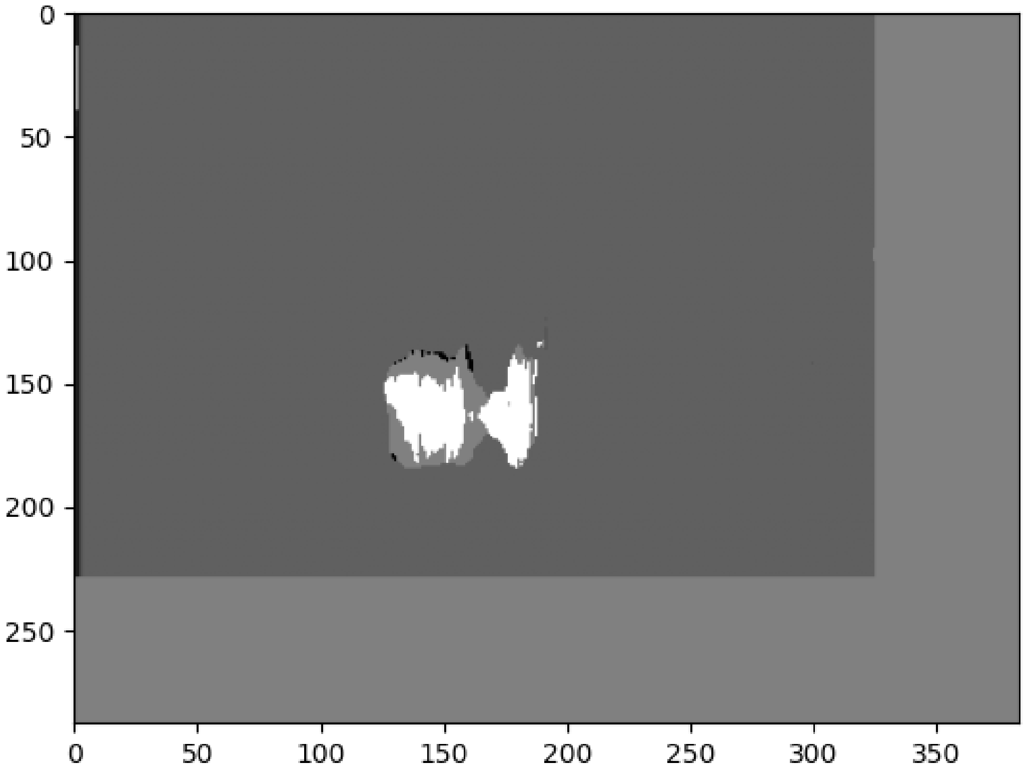
```
total time is 47.935020208358765 s
tsukuba_w: 20_d: 10
total time is 53.44917416572571 s
tsukuba_w: 40_d: 10
total time is 56.44277501106262 s
tsukuba_w: 60_d: 10
total time is 61.59318971633911 s
tsukuba_w: 80_d: 10
total time is 63.63213777542114 s
tsukuba_w: 100_d: 10
```

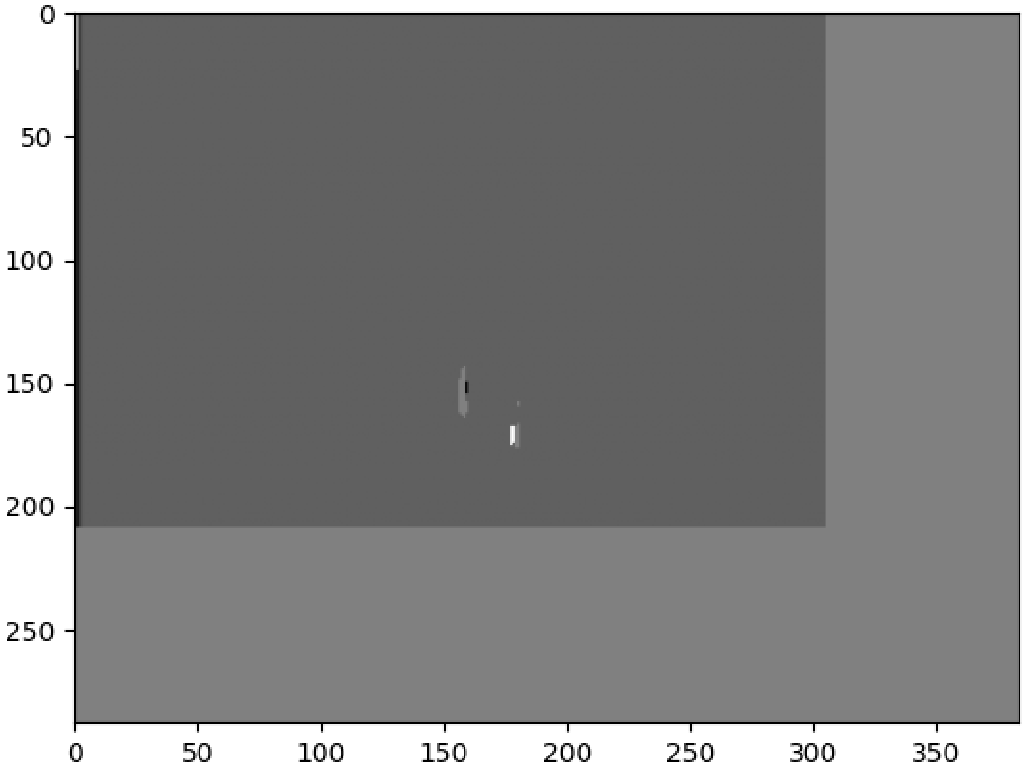

0
50
100
150
200
250
0
50
100
150
200
250
300
350

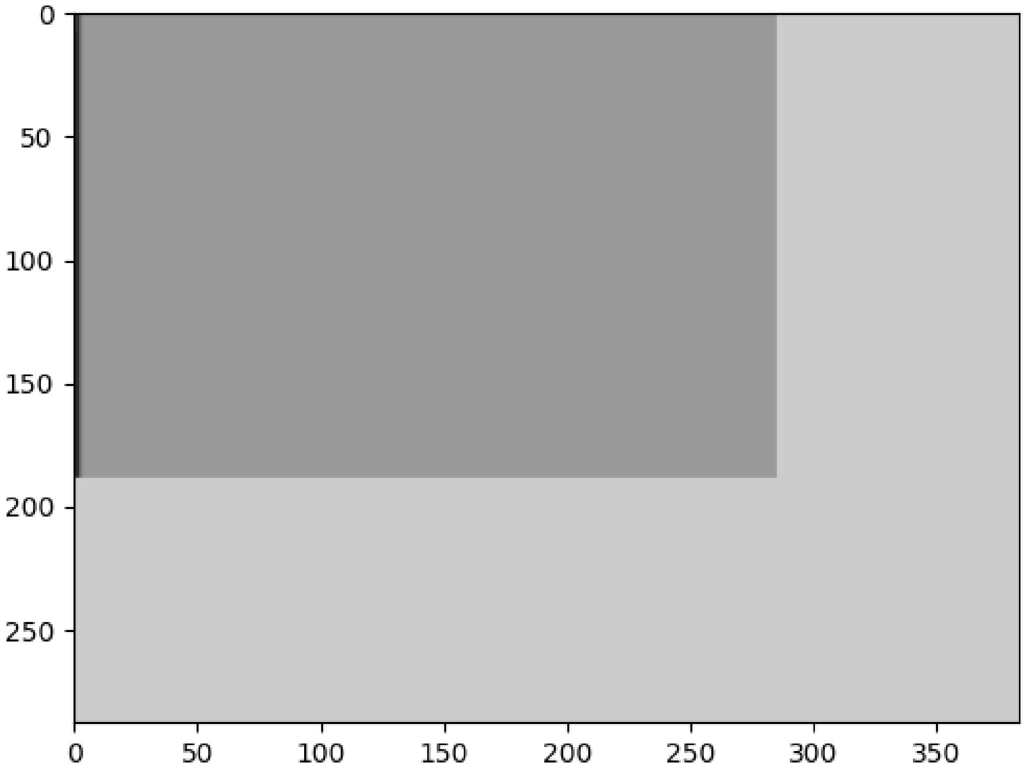

0
50
100
150
200
250
0
50
100
150
200
250
300
350


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
350


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

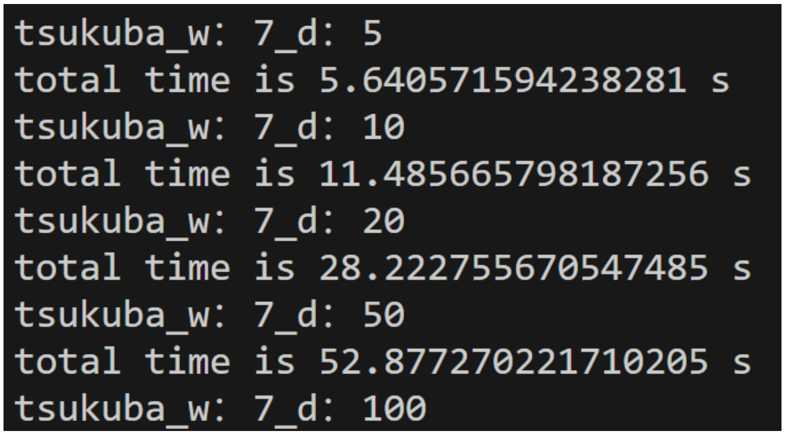

0
50
100
150
200
250
0
50
100
150
200
250
300
350



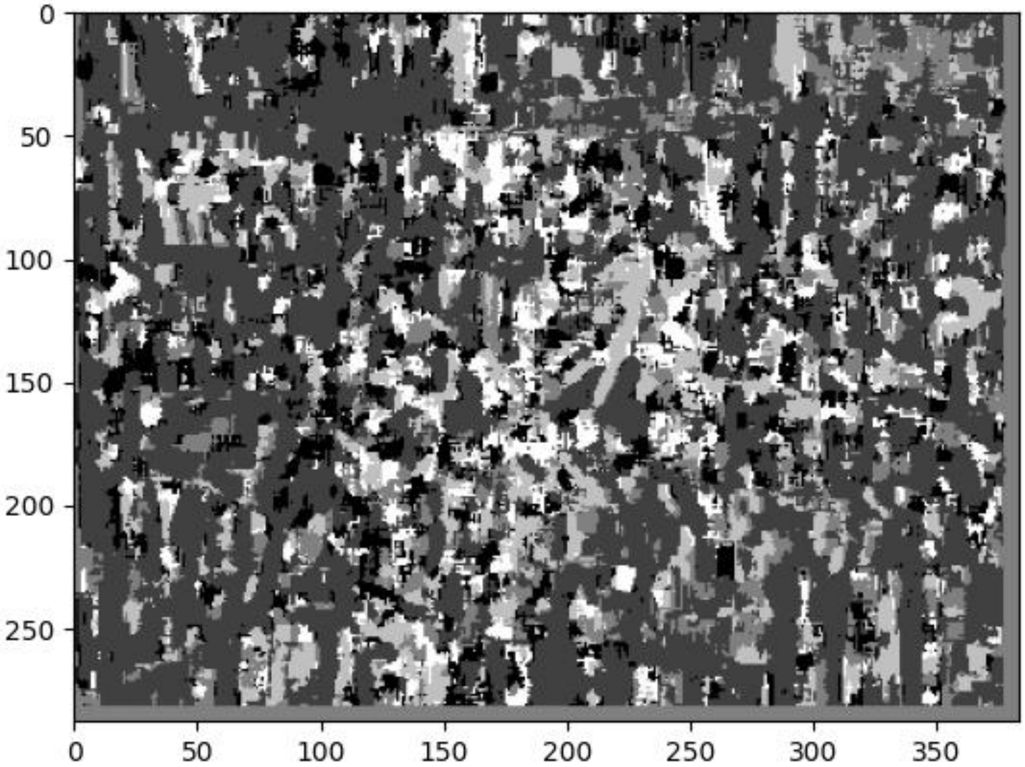
Comparing these results, window sizes of 5 or 10 is a good choice!

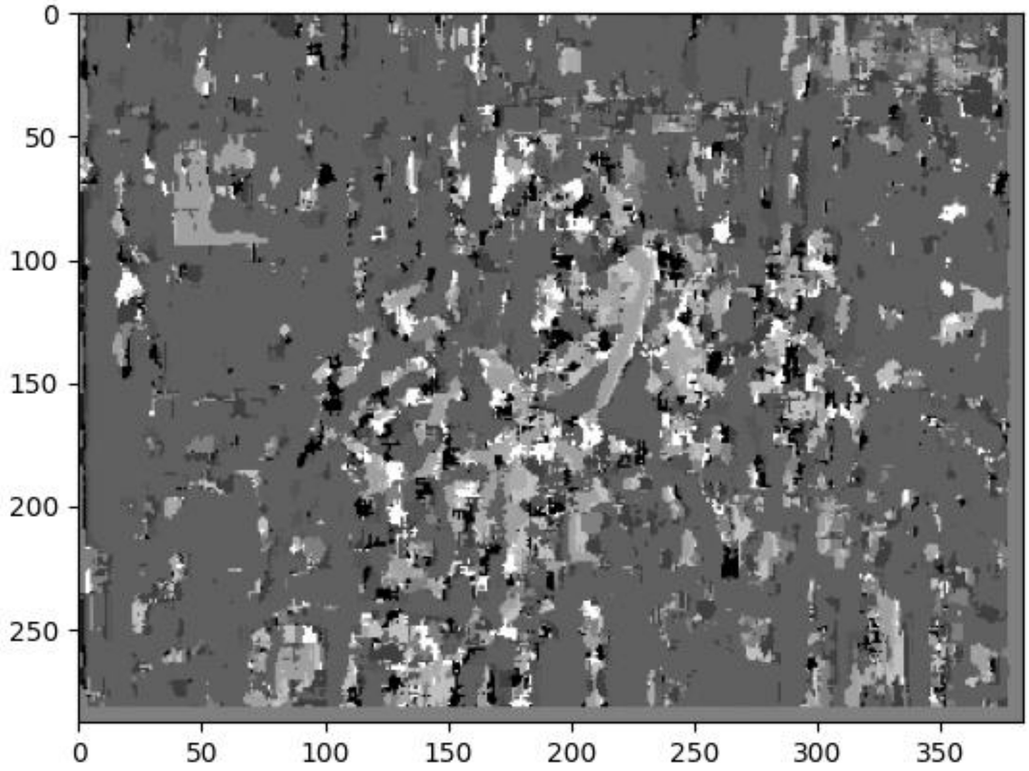
2. **Disparity range:** what is the range of the scanline in the second image that should be traversed in order to find a match for a given location in the first image? Examine the stereo pair to determine what is the maximum disparity value that makes sense, where to start the search on the scanline, and which direction to search in. Report which settings you ended up using.

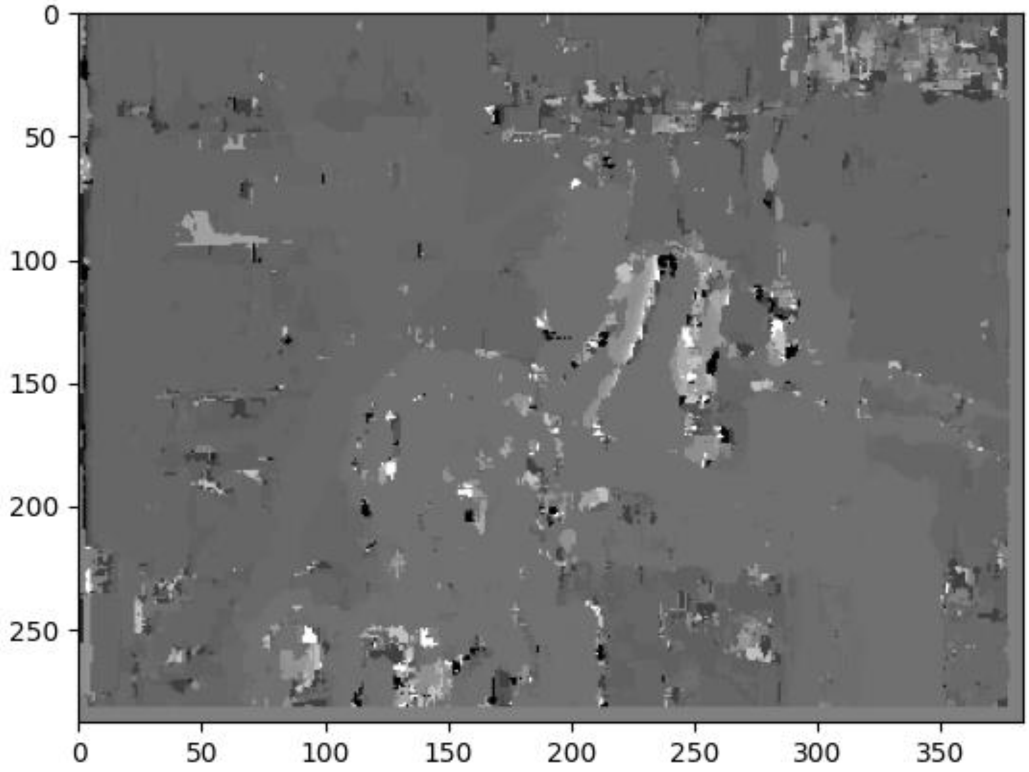
   Let's set window size to 10 and see what happens!
   As shown in the runtime console, disparity range influence time a lot,

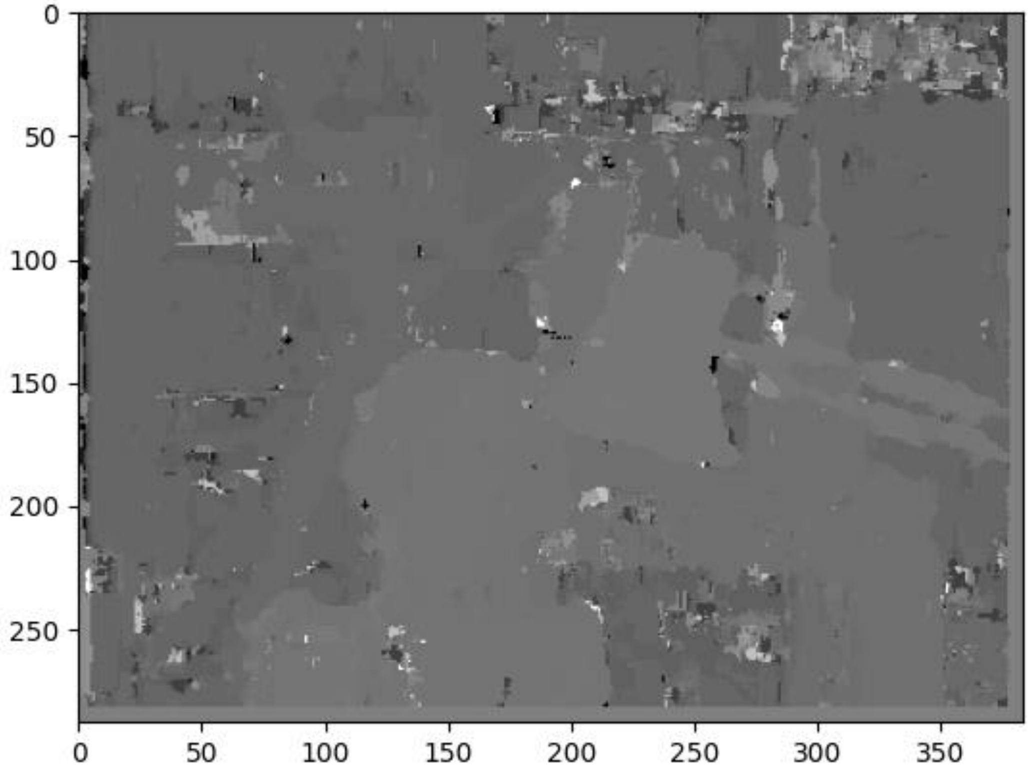


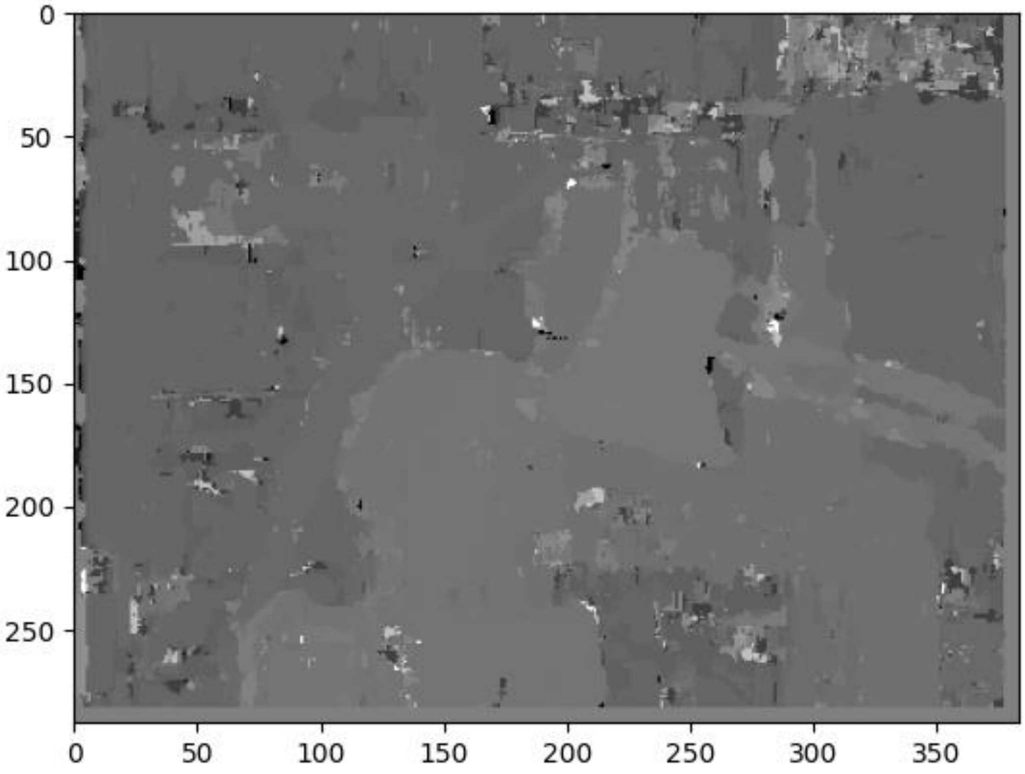
```
tsukuba_w: 7_d: 5
total time is 5.640571594238281 s
tsukuba_w: 7_d: 10
total time is 11.485665798187256 s
tsukuba_w: 7_d: 20
total time is 28.222755670547485 s
tsukuba_w: 7_d: 50
total time is 52.877270221710205 s
tsukuba_w: 7_d: 100
```




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
350

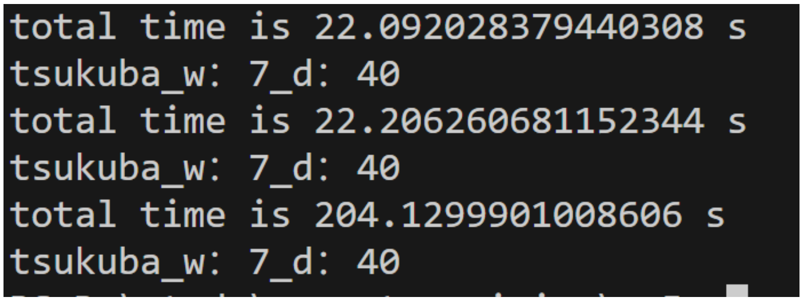

0
50
100
150
200
250
0
50
100
150
200
250
300
350



As we can see in the images, the bigger disparity range is, the better. However, considering time complexity, we should choose disparity=50 for this case.

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.

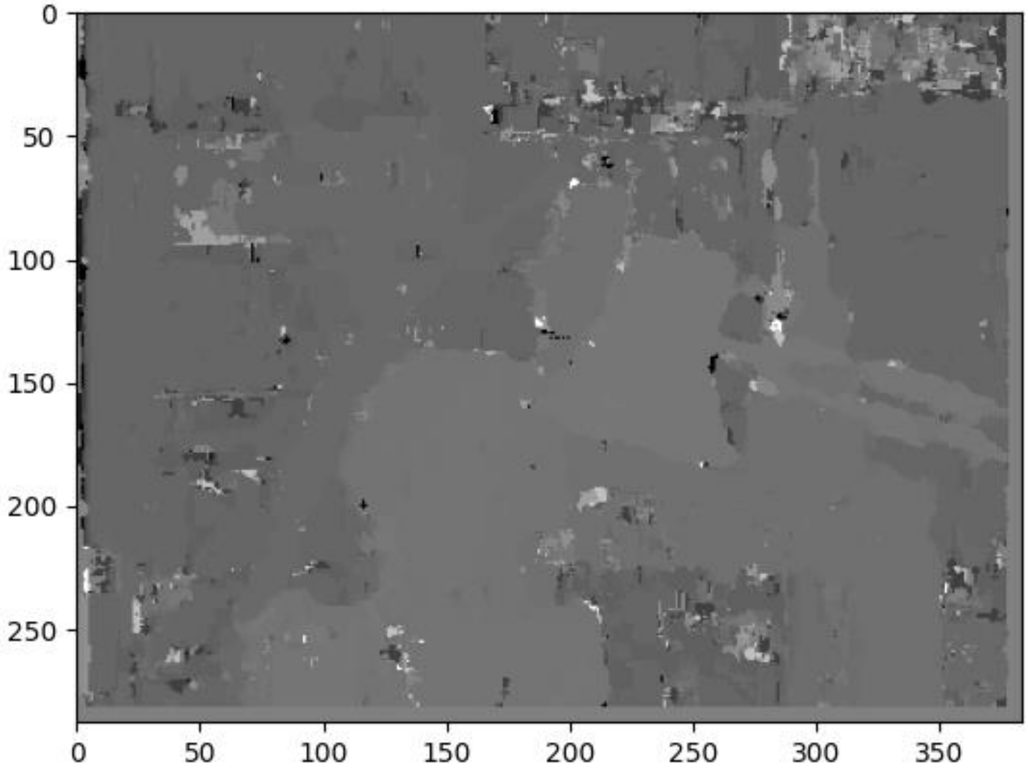
Choosing window size=7 and disparity range=40, I change the algorithm and compare.



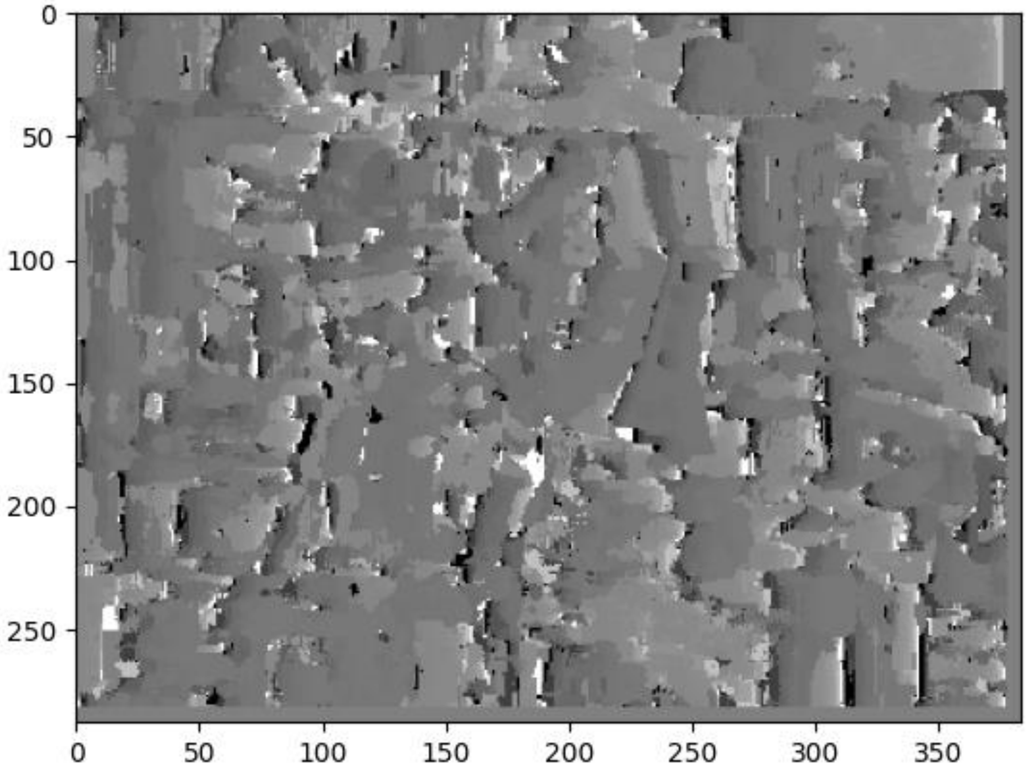
```
total time is 22.092028379440308 s
tsukuba_w: 7_d: 40
total time is 22.206260681152344 s
tsukuba_w: 7_d: 40
total time is 204.1299901008606 s
tsukuba_w: 7_d: 40
```

The normalized correlation algorithm is quite slow compared to others.

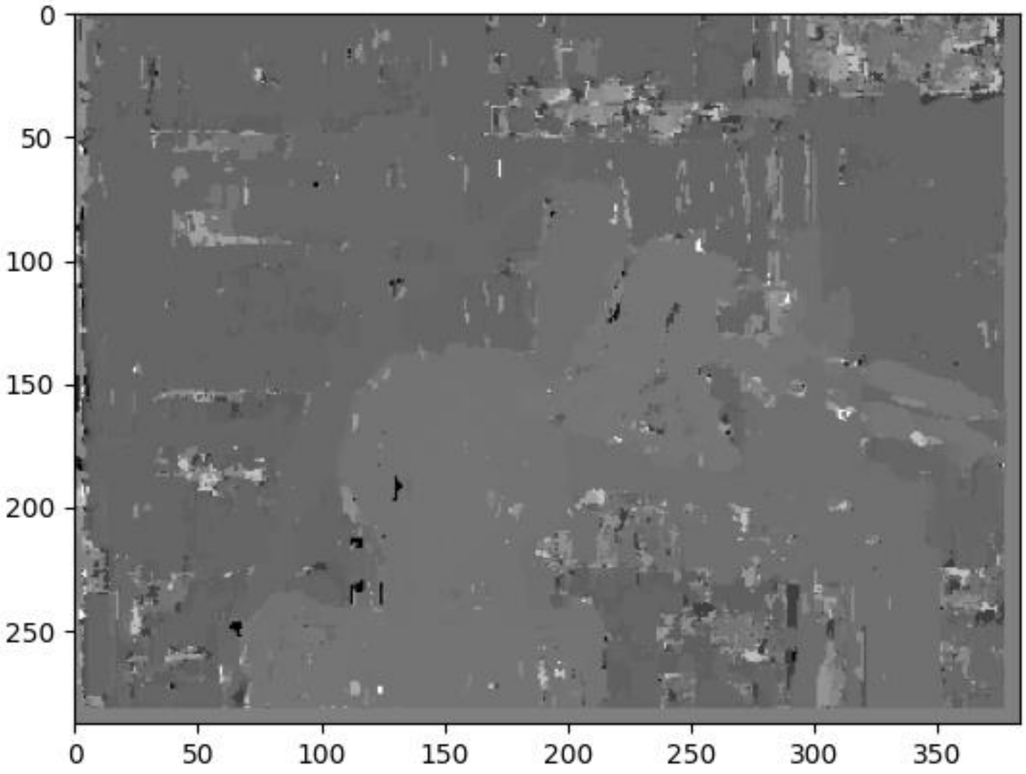
SSD:



SAD:



Normalized correlation:



The quality of NC algorithm is highest, but is most time consuming.

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.

one shortcoming of my implementation is the overhead of algorithm. However, if we try to use cuda for acceleration, it can be better.

Inside the output image, the center has better results than corners. A proper parameter can also make the result better.

For better running time performance, we should choose SSD and SAD instead of Normalized correlation.

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.