CV Assignment 3

WINDOW BASED STEREO MATCHING

JACOB KILVER

1. Question 1

A window size of 13 pixels was used for the left disparity.

Figure 1: ps3-1-a-1.png

A window size of 11 pixels was used for the right disparity.

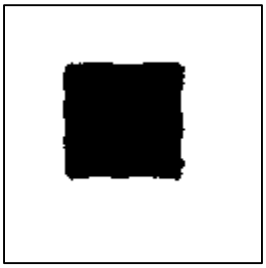

Figure 2: ps3-1-a-2.png

2. Question 2

a. Output images

Figure 3: ps3-2-a-1.png

Figure 4: ps3-2-a-2.png

b. Comparison to Ground Truth

First, I would like to point out that it seems that my left are right disparity outputs are opposite to those of the ground truth. For example, for the right disparity image (Figure 4) the subjects are more toward the left than in the ground truth image. However, in my output, the subjects are shifted left in the left disparity image. I did make sure to check my math, so this might have been a notation problem.

Regardless, this does not affect the actual results that were obtained. The ground truth image is definitely much cleaner and crisper than the result from SSD window based matching. While you can still identify many of the same features in the window-based matching output, it is much easier to do so in the ground truth image. Additionally, some of the finer features, such as the holes in the rings in the upper right, are lost because the window was too large. Finally, there are a number of details that should not have been found that were included in the window based matching results (see upper left corners).

3. Question 3

a. Results with noise

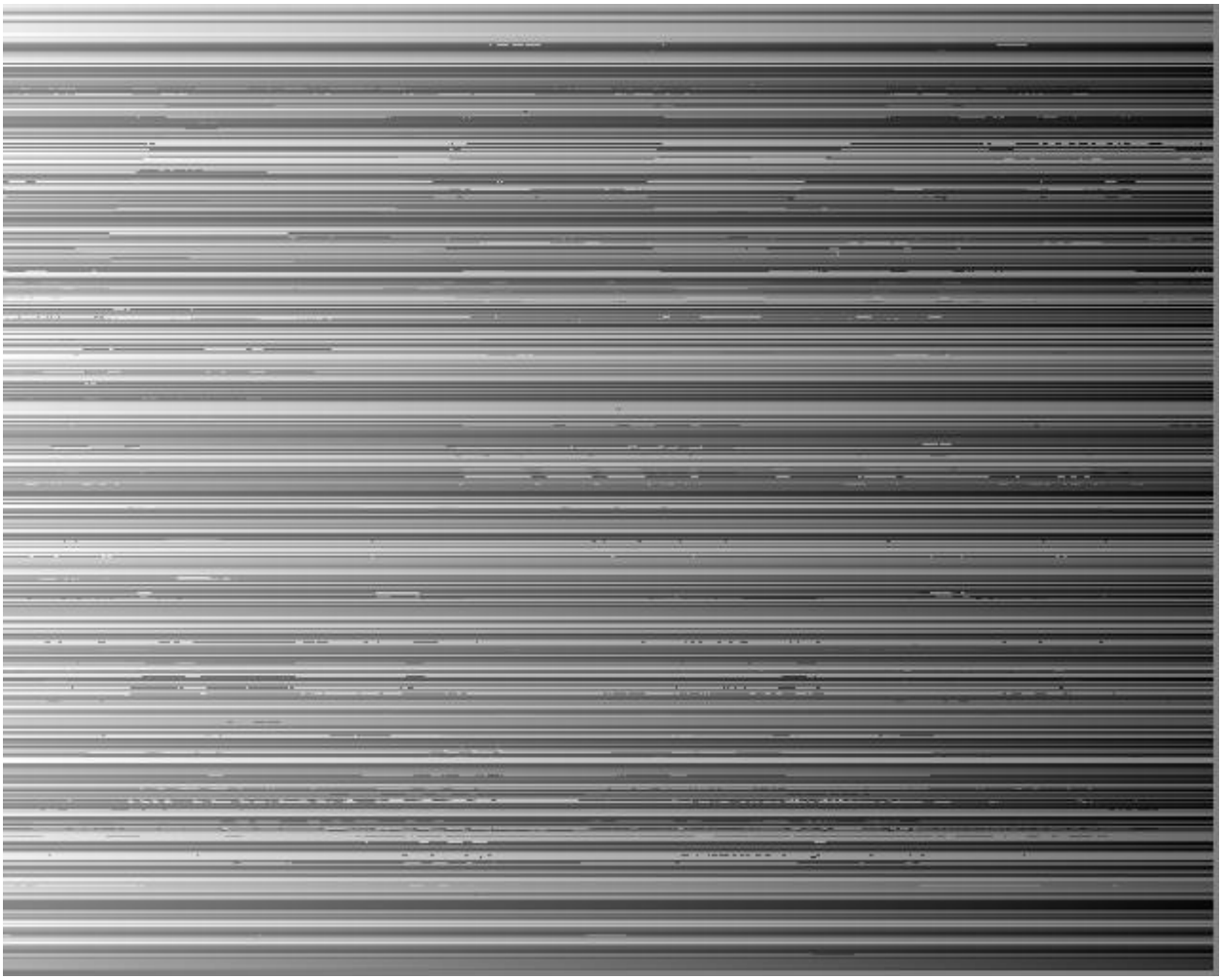

Figure 5: ps3-3-a-1.png

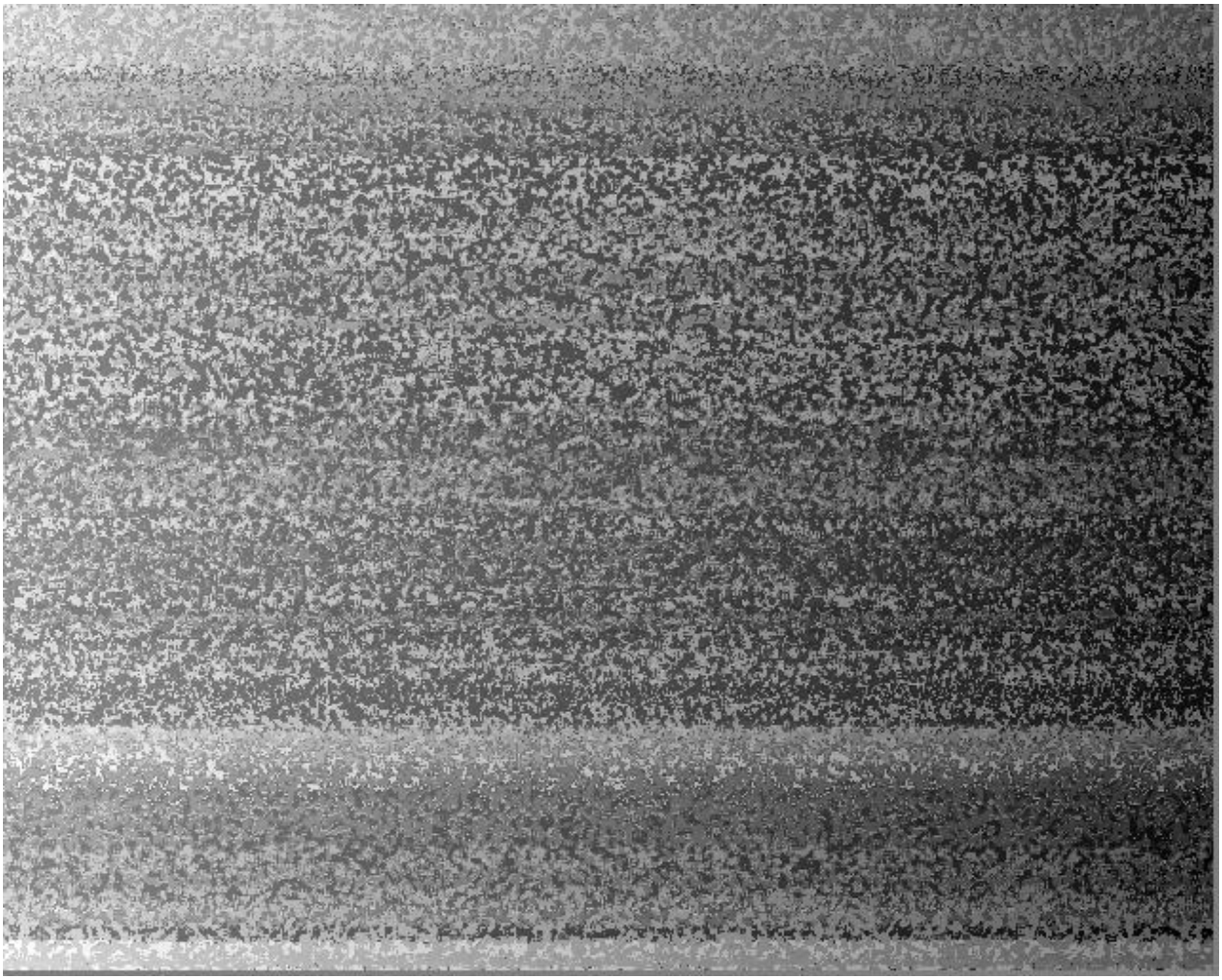

Figure 6: ps3-3-a-2.png

With the same window size (3 pixels) the output is pretty terrible. Basically all the information is lost from the original disparity images. Different window sizes could be tried, but my implementation of the window-based matching algorithm was very slow. Nevertheless, this example shows that even a little noise is enough to foil the window-based SSD stereo matching algorithm.

b. Results with heightened contrast

Figure 7: ps3-3-b-1.png

Figure 8: ps3-3-b-2.png

With heightened contrast, the window-based stereo matching output looks basically the same. There does seem to be a little more noise around the edges of the image, but at other places in the image, particularly the center, there is a heightened contrast. Again, many of the features seen in the ground truth images can still be pointed out, but not nearly as easily.

4. Question 4

a. Normalized correlation

Figure 9: ps3-4-a-1.png

Figure 10: ps3-4-a-2.png

The output using normalized correlation is very similar to that using SSD, but the output seems like it has “salt and pepper” noise added to it. Compared to the ground truth, many of the features are still identifiable, but the contrast between them is much less. Additionally, some of the smaller details were lost, even with a relatively small window size.

b. With noise and contrast

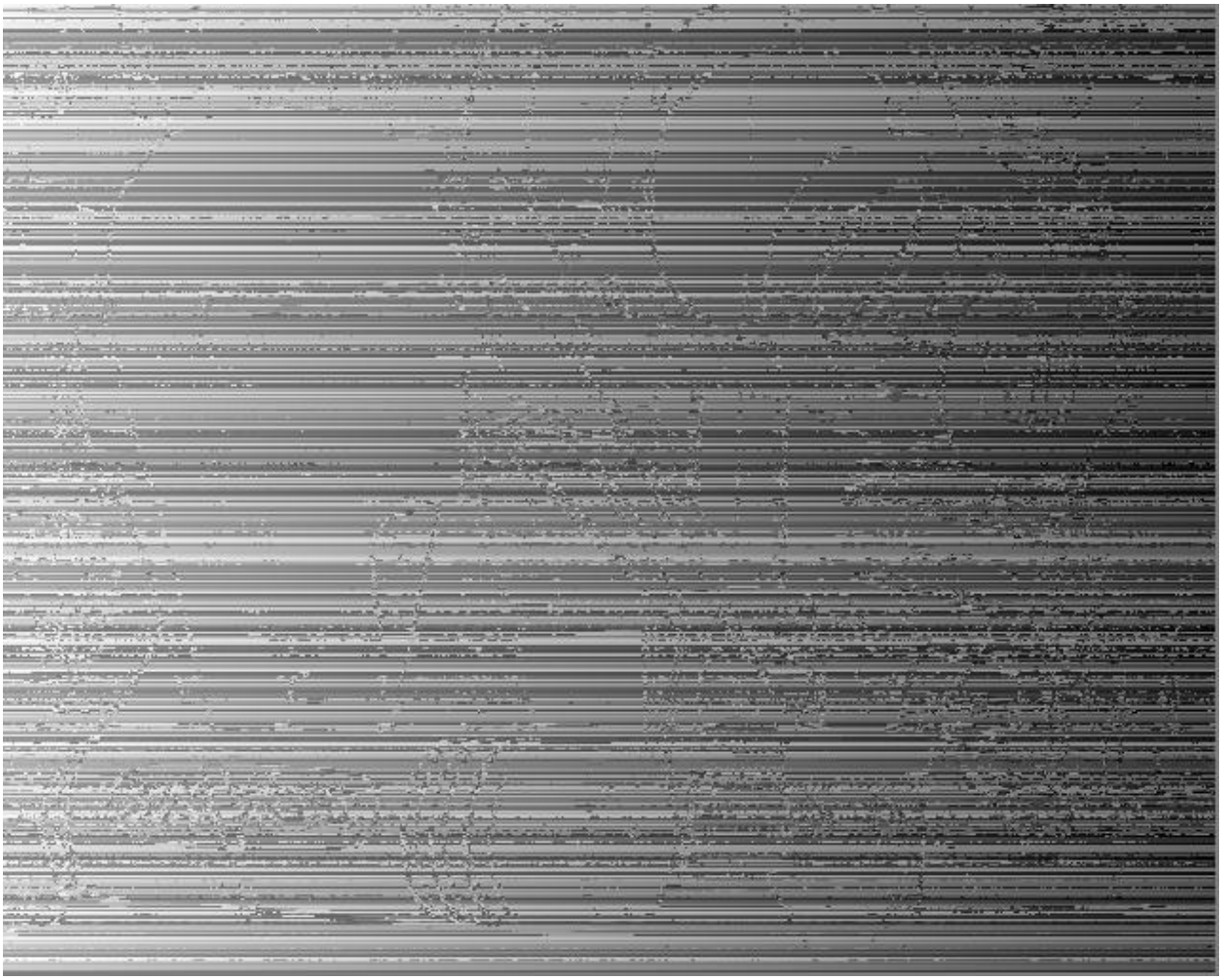

Figure 11: ps3-4-b-1.png

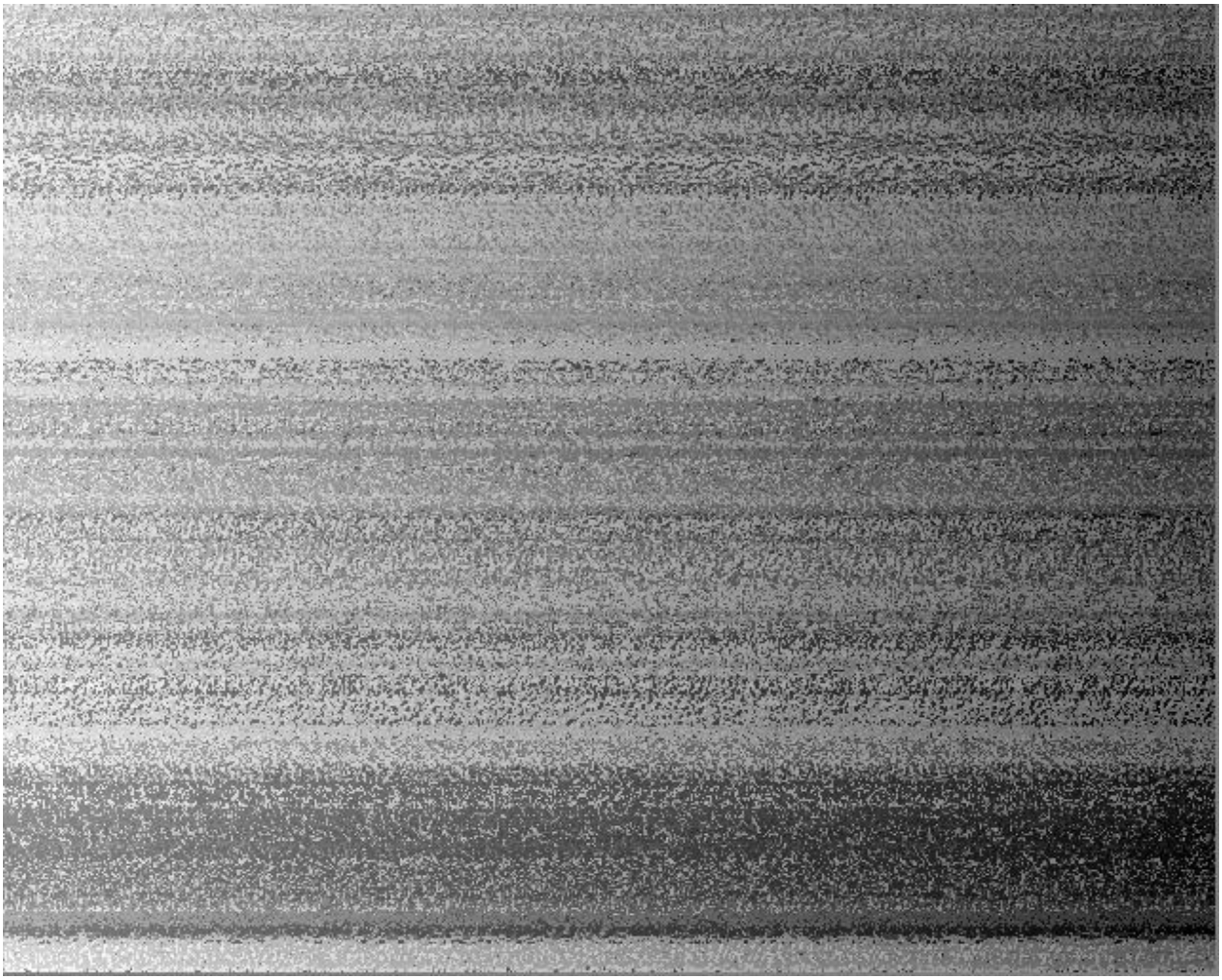

Figure 12: ps3-4-b-2.png

The results of using normalized correlation with noise were basically identical to those using SSD. The little bit of noise that was added was enough to completely confuse the window-based stereo matching algorithm. All the disparity information that was in the original output has been lost. Since this algorithm worked a little faster than the SSD implementation, various window sizes were used. None of them produced better results than the ones shown above with a window size of 3.

Figure 13: ps3-4-b-3.png

Figure 14: ps3-4-b-4.png

With heightened contrast, the results are very similar to the original normalized correlation output. Again, the subjects in the ground truth images are identifiable in the output images, just not as easily so. The difference now is that the image as a whole appears darker, especially on the right edge.

5. Question 5

Figure 15: ps3-5-a-1.png

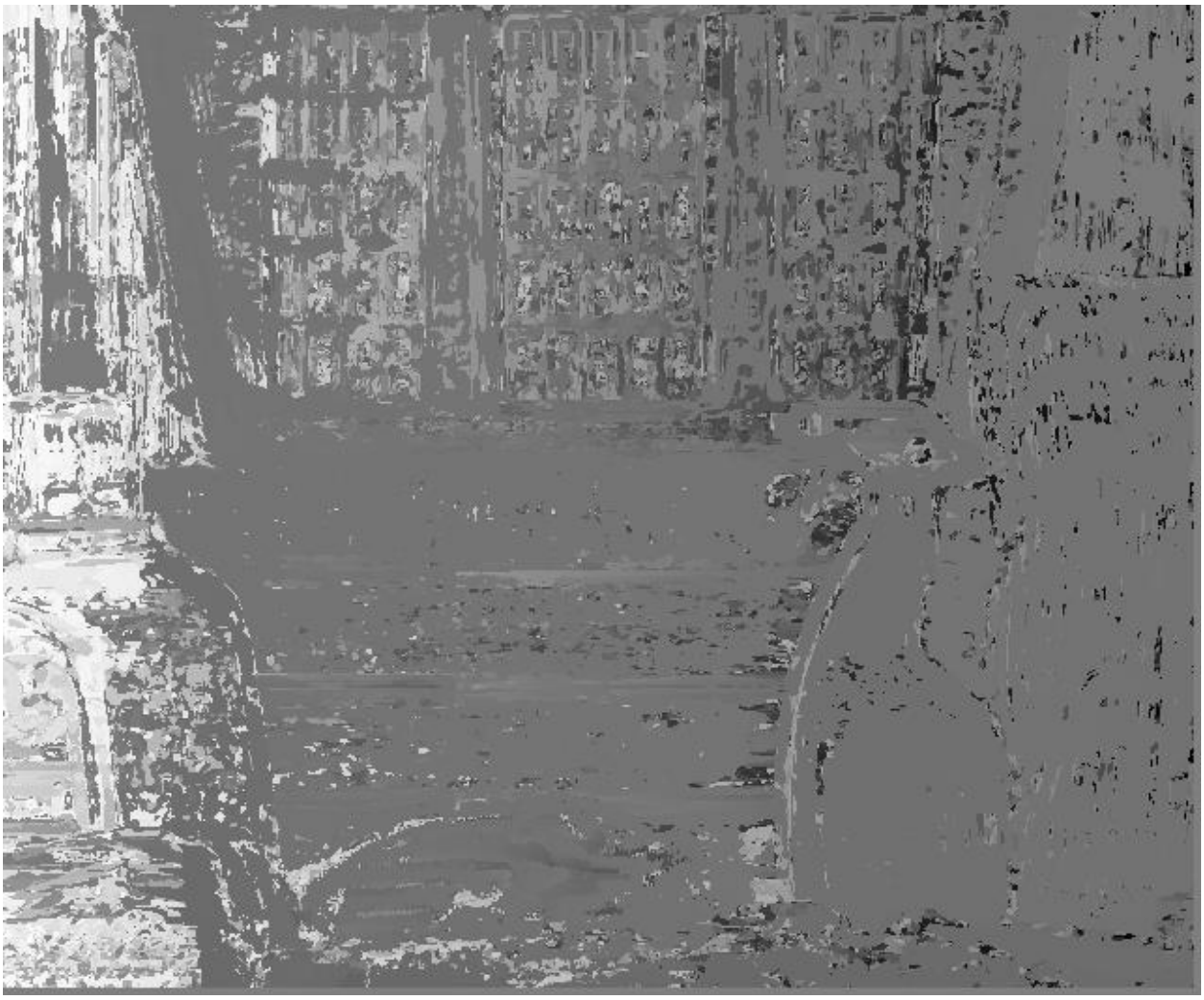

Figure 16: ps3-5-a-2.png

Once again I would like to acknowledge that I see that the left and right disparity images appear to be swapped according to the ground truth. This does not affect the actual results however.

Getting window-based stereo matching to work in a real world scenario is difficult. Since you are using windows, there is a correlation between the size of window and the amount of detail that can be detected. Larger windows lead to more accuracy but less detail, while smaller windows can detect more detail but are more susceptible to error. For the images above I first smoothed (blurred) the input images and then ran the normalized correlation window based matching algorithm with a window size of 5x5 pixels. The results were not very good as compared to the ground truth image, but they were the best I could obtain.

In conclusion, window based stereo matching will always be an approximation, so the output of these algorithms should be treated as such.