COMPSCI 527 Homework 6

Cody Lieu, Yixin Lin

December 7, 2015

1 Introduction

3D object reconstruction from a set of 2D images is fundamentally an ill-conditioned problem, due to its inverse nature. We investigate the effects of noise and other parameters on the success of reconstructing the original 3D object from stereo images. To this effect, we use the eight-point algorithm with the Hartley preprocessing[1] (by using the canonical homogeneous coordinates). We quantify the accuracy of the reconstructed 3D object by measuring three types of errors: motion, structure, and reprojection error.

Our goal in this paper is to investigate the following questions:

- Which types of error fail first when enough noise is added for the shape to be unrecognizable?
- What happens when we change the image resolution but keep sigma fixed?

2 Increasing levels of noise and the effect on error types

2.1 Introduction

The software provided creates a simulated world with a cubic object (with multiple colored squares on each side) and two cameras, adds Gaussian noise to the image coordinates, and runs the eight-point algorithm for 3-d reconstruction. Because reconstruction is a ill-conditioned/ill-posed problem (due to its inverse nature), small errors in the data will often result in large errors in the 3D reconstruction.

We investigate the effect of pseudo-random Gaussian noise on different types of errors, especially as the image become so distorted that the reconstructed shape becomes unrecognizable.

We have three broad categories of errors:

1. Motion error: the difference between true and computed motion, for both translation and rotation, measured in degrees
2. Structure error: the difference between true and computed structure, measured in RMS average percentage error of the overall size of the object
3. Reprojection error: the difference between the image of the original object and reprojected image of the reconstructed object, measured in RMS pixels per point

We pose the question: at the point where enough noise has been added so that the reconstructed object is unrecognizable, what types of errors fail first?

2.2 Methods

We ran the experiment according to the following procedure, in which a subroutine was applied on a range of sigmas from `0:.25:5`, with n = 150 for each sigma; we recorded the average errors for each, as shown in Figures 1-3.

Procedure

1. Create a 3D world, with object (Gaussian noise sigma = 0), two cameras, and the resulting images

2. Compute the true transformation between the camera reference frames

3. Compute the true structure

4. Run subroutine for each sigma in range `0:.25:4`

   - Add noise to the original image with sigma
   - Compute the resulting transformation figure
   - Measure and record motion, structure, and reprojection errors

After this, we also drew the reconstructed figures for choice sigma (Figures 4-9), to demonstrate the effects of noise on the resulting reconstruction and to find out when the reconstructed object became unrecognizable.

Experiment for random image equivalence

Additionally, we interpreted the second inflection points (when the error rates level off) as representing the sigma values that added a level of noise to the image such that it is no longer reconstructable. Essentially, this is the point where enough noise has been added so that the reconstructed image has no correlation with the original object. To confirm our interpretation, we devised the following experiment:

1. Take the `img` struct returned by `world` and setting all of the image points to random values using a (truncated) version of `randn`

2. Perform 3D reconstruction as usual on this randomly generated point cloud

3. Compute the rotation, translation, structure, and reprojection errors, as usual

4. Repeat 1,000 times and take the average errors

The results are summarized in Table 1.

2.3 Data

Figures 1.1, 1.2, 2, and 3. Motion (rotation), Motion (translation), Structure, and Reprojection Error as a function of Sigma Value

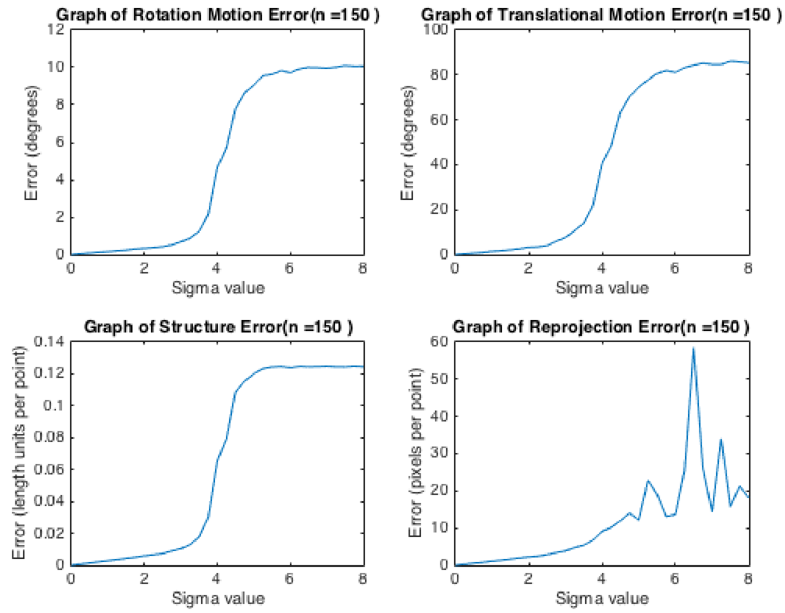


Figure 4. Reconstructed objects for `sigma`=0

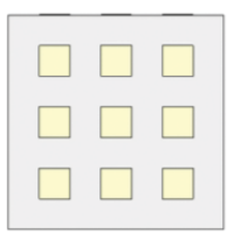


Figure 5. Reconstructed objects for `sigma`=1

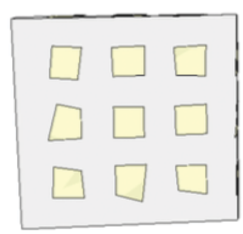


Figure 6. Reconstructed objects for `sigma`=2

Figure 7. Reconstructed objects for `sigma`=3

Figure 8. Reconstructed objects for `sigma`=4

Figure 9. Reconstructed objects for `sigma`=5

(We plotted rotation and translational separately despite them both measured in degrees, because rotational error had a low range and therefore the sigmoid characteristic was not easily observed.)

Figure 10, 11, 12. Reprojection error scatterplots (in pixels)

Error scatterplot for sigma = 0 Error scatterplot for sigma = 1 Error scatterplot for sigma = 6

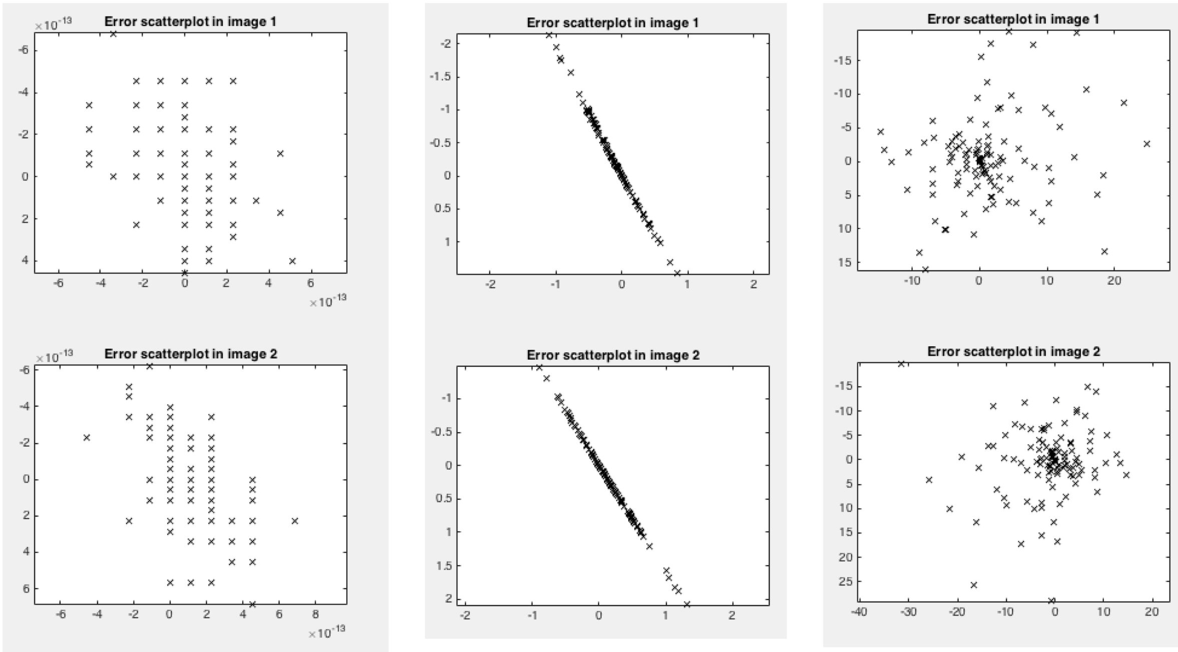


Table 1. Errors for Randomly Generated Image Points for n = 1000

Rotation Error	Translational Error	Structure Error	Reprojection Error
127.4189	64.2218	0.1249	4.2279e+04

2.4 Conclusions

Clearly, there is a severe drop in quality of the reconstructed object after sigma reaches around 3.25, and the reconstructed object for sigma=4 (Figure 8) is completely unrecognizable.

As we can see from the error plots, there is an inflection point around 3.25 where motion error (especially translation error) begins to climb drastically. This is followed immediately by structure error, which also has a similar inflection point. Finally, there is a slight inflection point in reprojection error, which is relatively delayed compared to the other errors (around sigma=4.25).

Why are the motion and structure graphs approximately linear before the inflection point? As we increase sigma, the Gaussian noise-generating function gets linearly scaled, making the slight errors increase linearly (on average).

Why is there an inflection point around sigma=3.25? That is the point where there is enough error for catastrophic failure in the reconstructed object such that it becomes unrecognizable. Again, this point exists because 3D reconstruction is an ill-posed problem, where minor errors in input result in massive failures in output. This inflection point is exactly that amount of errors. As obvious in figure 8 and 9, the object is completely unrecognizable after this inflection point.

After a certain point, however, the image has enough noise so that it is essentially a random image (one with random values for pixels), uncorrelated with the original object. This is essentially the maximum amount of average error attainable, and is the reason why the error rate tapers off (at around 10 degrees for rotation error, 80 degrees for translation error, and 0.12 unit lengths per point for structural error, according to figures 1-3). To confirm this, we devised the experiment in the section labeled "Errors for Randomly

Generated Image Points for n = 1000", where we took the average errors for *1000 randomly generated point clouds* after attempted 3D reconstruction. Evidently, the translational and structure error were clearly near-equal to the asymptotes of figure 2 and 3 respectively, which showed that at large enough sigma values (approx. 6 and above), the reconstructed figures were *no better than random noise* in terms of correlation with the original object.

Reprojection, however, does not follow this trend of smoothing out. In fact, this is the only error plot to not follow the sigmoid function; after a certain point, there are massive oscillations in the graph (especially see reprojection error in Figures 13, 14, 15). We hypothesize that this is caused by the reprojection down to 2D, which destroys information in an unbounded way (in fact, some points may not even be visible in the camera).

The reprojection error plots for different sigma are also insightful in showing the progression of reprojection error as the reconstructed image experiences failure. When sigma is 0 (Figure 10), the plot includes all points that are essentially 0– indicating an almost perfect reconstruction of the reprojected image. However, as sigma gets larger but below the inflection point talked about above (Figure 11), the errors all lie on a straight line– since the size of the displacements depend roughly *only on distance from the camera* (because of the nature of projecting a 3D object down onto 2D, where farther points seem smaller). Finally, when the sigma reaches the point where catastrophic failure occurs (Figure 12), the displacements become a random cluster around the origin (the reprojected image now no longer has much correlation to the original object).

3 Effects of modifying image resolution

3.1 Introduction

As we know, the Gaussian random noise generator uses sigma, which is defined in terms of pixels. We investigate the question of the effect of noise with varying image resolution but fixed sigma: because sigma is defined in pixels, modifying the number of pixels by resizing will have a noticeable effect on the effect of noise.

3.2 Methods

We modified world to accept optional arguments for resolution, and thereby resized the image. We tested both half the resolution and double the resolution of the image, running our subroutine with n=150 to gather information on the error plots for each resolution.

3.3 Data

Figure 13. Error Plots for Resolution of 960x540

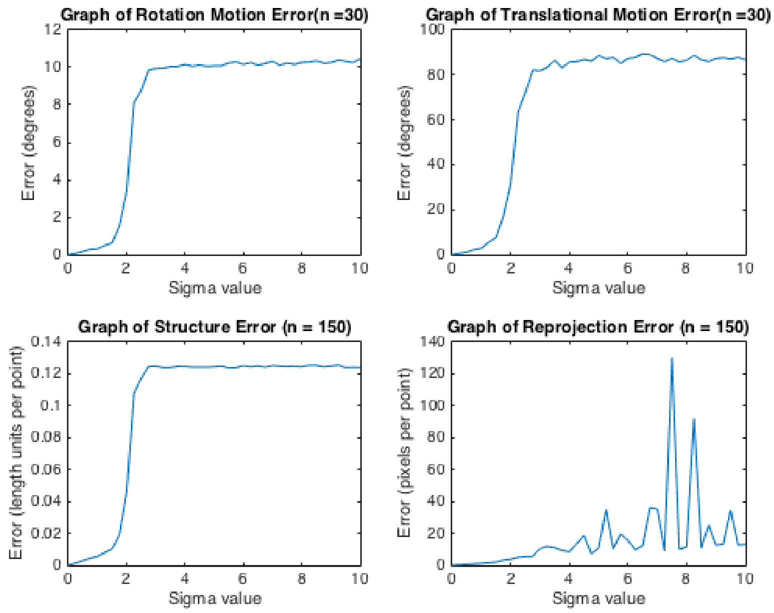


Figure 14. Error Plots for Resolution of 1920x1080

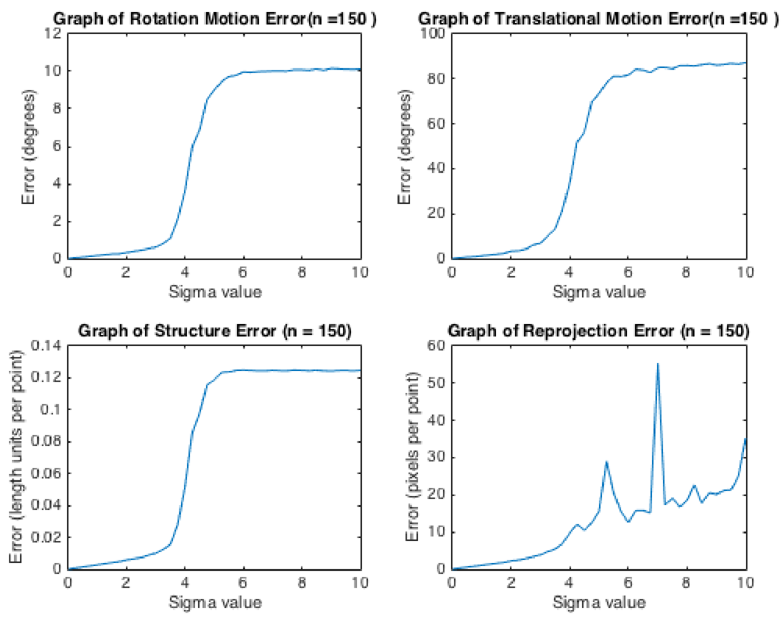
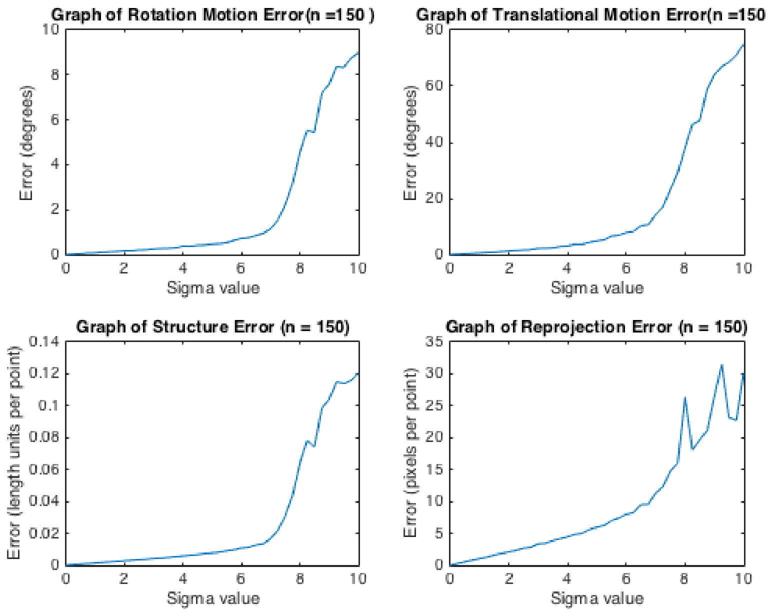


Figure 15. Error Plots for Resolution of 3840x2160



3.4 Conclusions

Clearly, a higher resolution leads to a lower error for fixed sigma. Since sigma is given in pixels, it's obvious that increasing the resolution of the image (i.e. scaling the number of pixels per point) will lead to a lessened effect from the noise-generating function.

It is interesting to note, however, that the relationship is linear. Observe that the error values for (for example) sigma=1 in a resolution of 960x540 are approximately equal to sigma=2 for resolution 1920x1080 and sigma=4 for resolution 3840x2160. In fact, each plot of a certain resolution is simply the horizontal scaling of the plot of the image half its resolution.

This is because the sigma value scales the randomly generated values in the same number of dimensions as the image (2), which means that an increase of sigma is offset by a proportional increase in both dimensions of resolution (i.e. in an image with double resolution in both dimensions, the equivalent sigma is also double).

4 Conclusion

In the first experiment, we learned that the rotation, translation, and structure error rates have a sigmoid relationship with respect to the amount of noise added to the image. In the second, we explored the relationship between resolution and different error rates, comparing values for fixed sigma.

References

[1] Hartley, Richard. "In defense of the eight-point algorithm." Pattern Analysis and Machine Intelligence, IEEE Transactions on 19, no. 6 (1997): 580-593.

[2] Chojnacki, Wojciech, and Michael J. Brooks. "On the consistency of the normalized eight-point algorithm." Journal of Mathematical Imaging and Vision 28, no. 1 (2007): 19-27.