Spring 2023

PROJECT 4

Perception for autonomous robots

March 16 2023

Student:
Nishant Awdeshkumar Pandey

Instructors:
Dr. Samer Charifa

Course code:
ENPM 673

Contents

1 Question

1.0.1 Question 1.1

In this section, calibration has to be done. To get the desired output following steps were executed:

- Read all the datasets and input all the given values.
- Find the Fundamental matrix by using 8 point algorithm: To execute 8 point algorithm Find the A matrix, and find its Singular value decomposition.

$$A = \begin{bmatrix} x_1'x_2 & x_1'y_2 & x_1' & y_1'x_2 & y_1'y_2 & y_1' & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n'x_{n+1} & x_n'y_{n+1} & x_n' & y_n'x_{n+1} & y_n'y_{n+1} & y_n' & x_{n+1} & y_{n+1} & 1 \end{bmatrix}$$

- Further RANSAC is applied to increase the accuracy.
- After Finding the F matrix essential matrix E is found.
- $$E = K^T F K$$
  ,
- where K is the camera matrix
- By decomposing E we get four different configurations of Rotation and translation.

  $C1 = U(:,3)R1 = UWV^T, C2 = U(:,3)R2 = UWV^T, C3 = U(:,3)R3 = UW^TV^T, C4 = U(:,3)R4 = UW^TV^T$
- Where $E = UDV^T$ and W= $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- It is important to note that If det(R)=1, the camera pose must be corrected i.e. C=C and R=R.
- Triangulation Check for Cheirality Condition needs to be applied to find the correct R and T.
- To check the cheirality condition, triangulate the 3D points (given two camera poses) using linear least squares to check the sign of the depth Z in the camera coordinate system w.r.t. camera centre. A 3D point X is in front of the camera if: $r3(XC) > 0$ where r3 is the third row of the rotation matrix (z-axis of the camera). Not all triangulated points satisfy this condition due to the presence of correspondence noise. The best camera configuration, (C, R, X) is the one that produces the maximum number of points satisfying the cheirality condition.
- Further the Homography matrices H1 and H@ can be obtained.

$$H1 = K_1 F K_2^{-1}, H2 = K_2 F K_1^{-1}$$

1.0.2 Results

```
VALUES FOR ART ROOM
Fundamental Matrix for art room
 [[ 7.90735252e-10 -1.30908997e-05  4.19782773e-03]
 [ 1.35665642e-05 -6.35193593e-07  2.74773078e-01]
 [-4.30565845e-03 -2.75919376e-01  1.00000000e+00]]
Essential Matrix for art room
 [[-1.37212615e-05 -7.89529269e-02 -1.02963795e-02]
 [ 8.21707653e-02 -1.57062796e-03  9.96565131e-01]
 [ 1.04095429e-02 -9.96821380e-01 -1.62342110e-03]]
First Rotation for art room
 [[ 0.08602776 -0.12654543  0.9882234 ]
 [-0.00161021  0.99188167  0.12715406]
 [-0.99629144 -0.01253003  0.08512559]]
Translation for art room
 [[ 0.99682517]
 [ 0.01017042]
 [-0.07896925]]

Left image Homogrphy matrix
 [[-2.82708156e-01 -3.70121525e-03  3.32431346e+01]
 [-4.19079969e-03 -2.76197983e-01  3.41359624e+00]
 [-1.32177784e-05  2.40703446e-07 -2.67588075e-01]]
Right image homogrpahy matrix
 [[ 1.02450109e+00  3.03759725e-02 -4.23915203e+01]
 [ 1.47018864e-02  1.00087535e+00 -8.77935913e+00]
 [ 4.61858224e-05  1.36938778e-06  9.73745044e-01]]
```

Figure 1: Sample result of part one for the first data set

1.1 Question 1.2

In this sub-question, Rectification needs to be done.

- To do so the H1 and H2 matrices are used to do perspective transforms.
- on these rectified images the matching points and parallel epilines are drawn.

1.1.1 Result

Figure 2: Sample result of part two for the second data set

Figure 3: Sample result of part two for the second data set

1.2 Question 1.3

In this part, the following steps were executed:

- Cross-correlation method is applied between the matching window in the first image and a patch of pixels in the second image.
- Formula is:
$$C(x,y) = \frac{\sum_i \sum_j (w(i,j) - \mu_w)(p(i+x, j+y) - \mu_p)}{\sigma_w \sigma_p}$$
- Where, n this formula, w(i,j) is the pixel value in the matching window centred at (x_0, y_0), p(i+x, j+y) is the pixel value in the patch centred at (x,y) in the second image,μ_w and σ_w

are the mean and standard deviation of the pixel values in the matching window, and μ_p and σ_p are the mean and standard deviation of the pixel values in the patch. The resulting value C(x,y) is a measure of the similarity between the matching window and the patch centred at (x,y) in the second image.

- $d(x, y) = \text{argmax}_{d_0} C(x, y, d_0)$
- In this formula, C(x,y,d_0) is the normalized cross-correlation between a matching window in the first image and a patch of pixels in the second image, displaced by a distance d_0 in the horizontal direction. The disparity d(x,y) is the horizontal offset that maximizes the cross-correlation, and it represents the difference in horizontal position between corresponding pixels in the two images. To compute the disparity for each pixel in the image, you would need to apply this formula at each pixel location (x,y) in the first image, searching for the value of d_0 that maximizes the cross-correlation for that location.
- save the disparity as a grayscale and colour image using heat map conversion.

1.2.1 Result

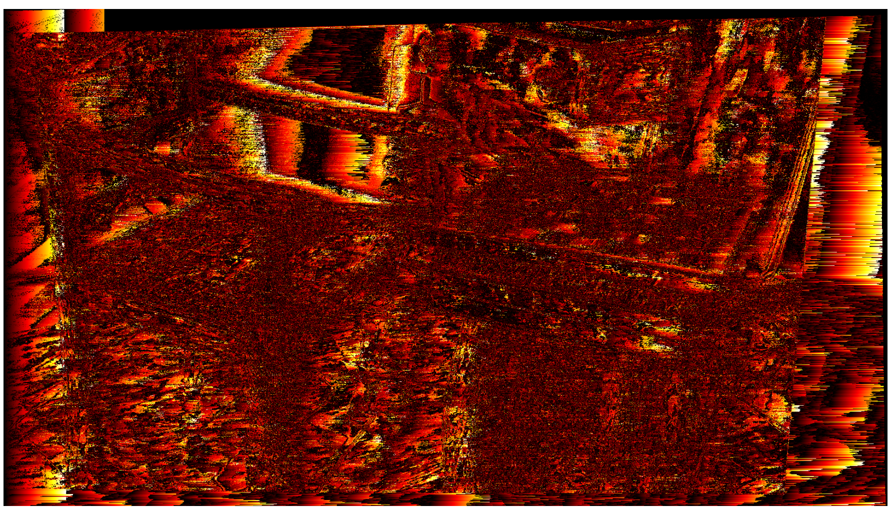

Figure 4: Sample result of part three for the second data set

Figure 5: Sample result of part three for the second data set

1.3 Question 1.4

I Compute the depth image in this part:

- To calculate depth the following formula is used

$$depth = (baseline * focallength)/disparity$$

- Focal length can be derived from the given camera matrix.
- Further the heat map and depth image is saved.

1.3.1 Result

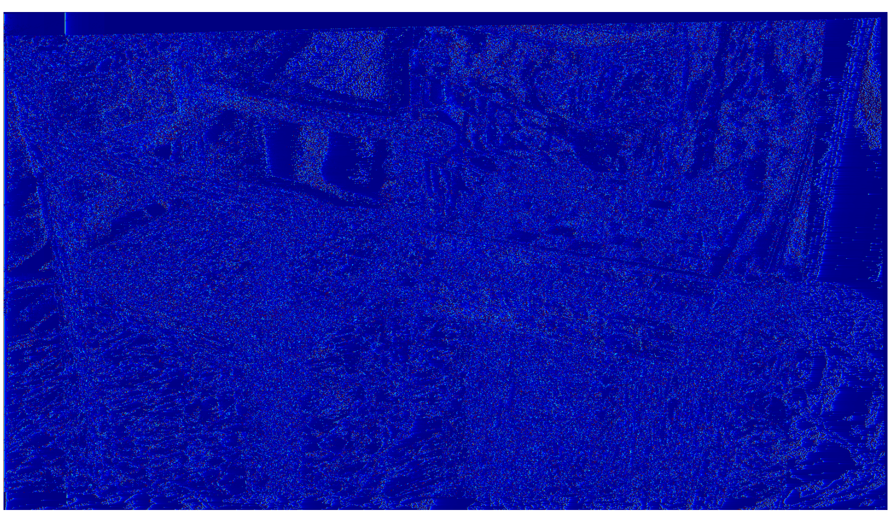

Figure 6: Sample result of part four for the second data set

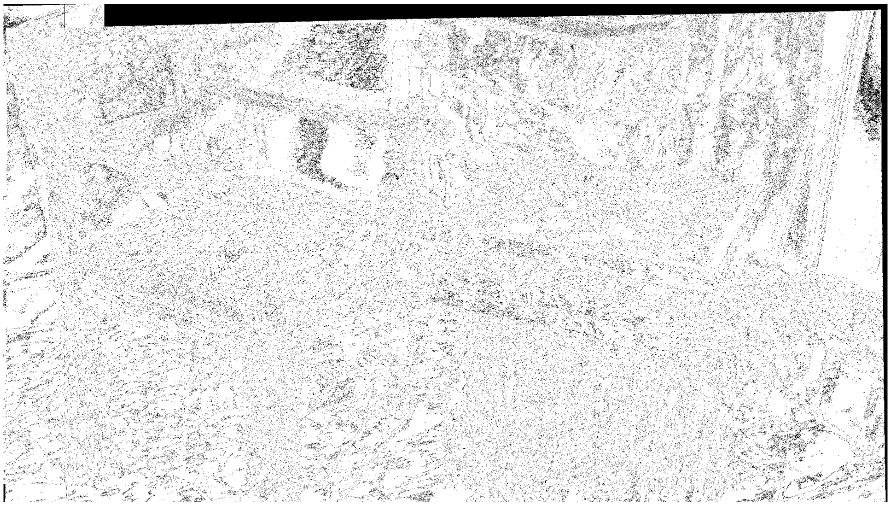

Figure 7: Sample result of part four for the second data set

2 Problems Encountered

- Finding RANSAC in the first part was very tricky, I was not getting the right inliers. I had to use various sources like GitHub to solve the problem.
- In the second part no issues were encountered.
- Coding the cross-correlation part was tricky.
- No issues were faced in the last part as well.

3 References

- https://people.scs.carleton.ca/~c_shu/Courses/comp4900d/notes/homography.pdf
- http://ai.stanford.edu/~birch/projective/node20.html#:~:text=Thus%20both%20the%20Essential%20and,latter%20deals%20with%20uncalibrated%20cameras.
- https://subscription.packtpub.com/book/data/9781789806311/8/ch08lvl1sec70/performing-mean-normal
- https://github.com/iamjadhav/stereo-vision-for-depthestimation/tree/07f596dc5a2a004f33e93a1419b8
- opencv documentation