673 Project 3

Perception for Autonomous Robots
Faculty: Samer Charifa

Prasanna Thirukudanthai Raghavan
UID: 118287546

Department of Robotics
University of Maryland
United States

Contents

1 Project Pipeline: 2
1.1 Calibration: . . . 2
1.2 Fundamental Matrix: . . . 2
1.3 RANSAC: . . . 2
1.4 Essential Matrix . . . 2
1.5 Camera Pose: . . . 3
1.6 Rectification: . . . 3
1.7 Correspondence: . . . 3
1.8 Dataset 1: . . . 4
1.9 Dataset 2: . . . 5
1.10 Dataset 3: . . . 6

Chapter 1

Project Pipeline:

1.1 Calibration:

Since we use SIFT which for the detection of the matching points we initially process the image before we feed into SIFT and then get the matching done by using the $cv2$.BFMatcher as that here can give us a better output than FLANN matcher even if it takes longer and the accuracy is maintained.

1.2 Fundamental Matrix:

From the camera intrinsic matrix the A matrix is obtained for this system with 9 unknowns and each returns one equation which is a scalar when a constraint for one of the equation is added and this can be obtained using the 8 point hartley algorithm and the resulting matrix will be:

$$Af = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \end{bmatrix} f = 0$$

Now we know that the x_1 and x'_1 are the features In which the $X_2^T.F.X_1 = 0$ is the constraint that allows us to solve using the SVD and this causes multiple errors in case the rank is higher than 2, which due to normalization needs to be redone using the epipolar constraint and best fit selected using the RANSAC.

1.3 RANSAC:

We use RANSAC to reject the outliers after normalization to get the best fit as it should not have a null value and to get the inliers of the epipolar lines and the threshold must be chosen in such a way that we are able to reduce all the erroneous values to decrease the processing requirement as well.

1.4 Essential Matrix

This is taken from the equation $E = K'^T.F.K$ from the intrinsic parameters that are known and then using which we obtain the pose information.

1.5 Camera Pose:

From the essential matrix we obtain the pose of the cameras with respect to each other. Where Rotation and translation are obtained as 4 possible mathematical solution after decomposing E using SVD as $E = UDV^T$. This is namely R_1 and R_2 which will be $R_1 = UWV^T$ and $R_2 = uWTV^T$ and as we know the matrix W is $\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ here if we take the third column of U to be t then the poses can be taken as (R_1,t),(R_1,-t),(R_2,t) and (R_2,-t) of first camera with respect to the second camera. In the chance that all the elements of t are zero then of the above possibilities only one is valid as the points need to be positive and the maximum positive value is considered.

1.6 Rectification:

The fundamental matrix we estimated earlier we an obtain the lines needed to make the images parallel. This process is using the epipolar constraint for the transformation. The H1 and the H2 homography matrices are calculated using the function and the rectification is obtained.

1.7 Correspondence:

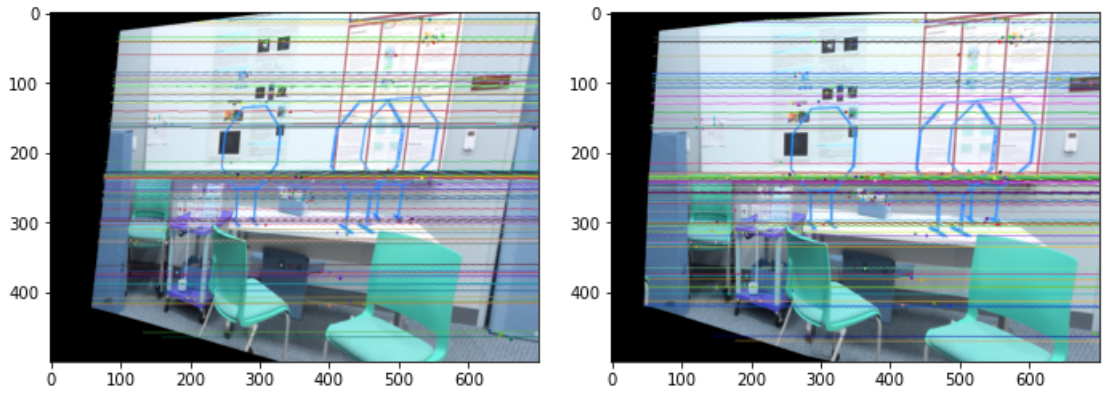
The left and the right images are taken and the points between them are matched to find the corresponding epipolar lines. A generalized block matching would yield acceptable results when implemented through Sum of Squared Differences and the disparity is calculated by using the image pixel difference with respect to the width of the image. Here the depth map is computed by using the formula of

$$z = \frac{baseline * f}{disparity}$$

Output

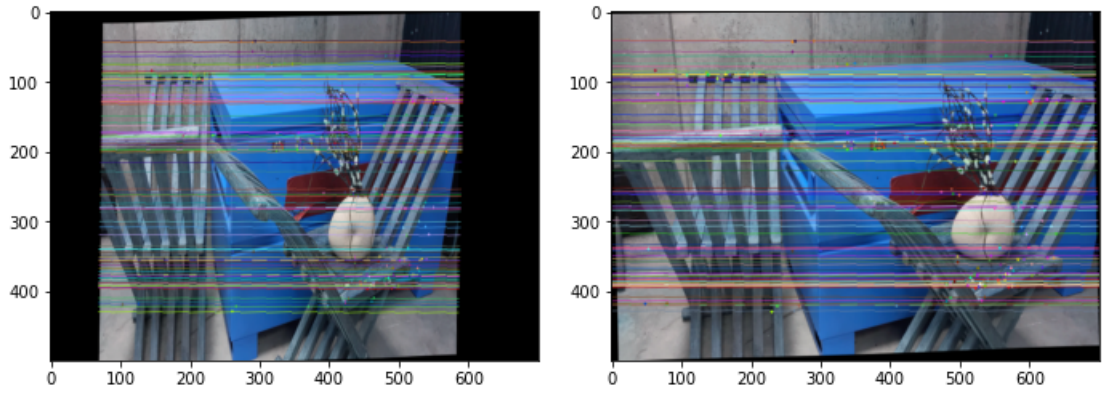
1.8 Dataset 1:

```
F matrix
 [[ 2.64146637e-07  1.13262877e-03 -1.94122561e-01]
 [-1.12851289e-03 -7.02216335e-07  2.16253909e+00]
 [ 1.91482943e-01 -2.15224092e+00  1.00000000e+00]]
E Matrix
 [[ 2.55280314e-04  8.33228372e-01  1.76935182e-01]
 [-8.29749262e-01 -5.39954429e-04  5.29170355e-01]
 [-1.76910703e-01 -5.24041424e-01  1.46672438e-03]]
Rotation
 [[-0.44536623 -0.18584889  0.87584765]
 [-0.18683176 -0.93739504 -0.29391229]
 [ 0.87563852 -0.29453477  0.38276161]]
Translation
 [[-0.52385528]
 [ 0.1774672 ]
 [-0.83311526]]
Homography Mat 1 :  [[ 1.49182419e+00 -2.02234799e-01  2.50772188e+02]
 [-2.31185899e-01  2.17274326e+00  7.06233748e+01]
 [-1.36496954e-03  5.43104758e-05  2.59538931e+00]]
Homography Mat 2 :  [[ 7.96928706e-01 -3.52382150e-02  7.34399857e+01]
 [-1.07396986e-01  1.00572595e+00  3.29928085e+01]
 [-6.31547283e-04  2.79254578e-05  1.19539302e+00]]
```



1.9 Dataset 2:

```
F matrix
 [[-2.15458850e-07  3.79266848e-05  2.86158725e-02]
 [-2.53031717e-05 -2.05699949e-05 -2.19548152e+00]
 [-3.19065136e-02  2.20444005e+00  1.00000000e+00]]
E Matrix
 [[-1.71160547e-04  2.96903424e-02  2.18481362e-02]
 [-2.01045933e-02 -7.59405649e-03 -9.99525307e-01]
 [-2.08263116e-02  9.99316484e-01 -7.82558379e-03]]
Rotation
 [[ 0.00946253 -0.70059632 -0.7134951 ]
 [-0.03656601  0.71280739 -0.70040599]
 [ 0.99928644  0.03271729 -0.01887305]]
Translation
 [[ 0.99932033]
 [ 0.0220748 ]
 [-0.0295227 ]]
Homography Mat 1 :  [[ 1.63181224e+00 -3.35902143e-02  1.64490507e+02]
 [-3.32765643e-02  2.20306859e+00  1.01400833e+01]
 [ 1.97643948e-05  1.57039340e-05  2.18357379e+00]]
Homography Mat 2 :  [[ 1.00529027e+00  1.71231515e-02 -5.80244362e+00]
 [-1.29540941e-02  9.99924404e-01  4.16345314e+00]
 [ 1.69853232e-05  2.89311723e-07  9.94495262e-01]]
```



1.10 Dataset 3:

```
F matrix
 [[ 4.67174786e-07 -3.72820997e-04  1.00167622e-01]
 [ 3.60147116e-04 -1.45849920e-05  1.05664671e+00]
 [-9.53223711e-02 -1.05522160e+00  1.00000000e+00]]
E Matrix
 [[-0.00319091 -0.56896263 -0.09923144]
 [ 0.55833497 -0.01263439  0.82422563]
 [ 0.08709148 -0.81703582 -0.00682217]]
Rotation
 [[ 0.99986907 -0.00805517  0.01403384]
 [ 0.0079157   0.99991901  0.00996542]
 [-0.01411298 -0.00985303  0.99985186]]
Translation
 [[ 0.81634825]
 [ 0.0935657 ]
 [-0.56993069]]
Homography Mat 1 :  [[-1.03310273e+00  8.42143410e-02  2.02927366e+00]
 [-8.94338143e-02 -1.06306411e+00  2.96491770e+01]
 [-3.33972784e-04  7.31537569e-06 -9.54638398e-01]]
Homography Mat 2 :  [[ 1.10118958e+00 -9.88056872e-03 -3.00093284e+01]
 [ 8.48946429e-02  9.99278525e-01 -2.69931317e+01]
 [ 3.16343217e-04 -2.83843124e-06  8.99451394e-01]]
```