Project 3: Stereo Vision System

ENPM 673 Spring 2021

Diane Ngo (UID 117554960)

Project Objective

This project implements the concept of Stereo Vision. Three datasets are given, with each two images at different camera angles. The project consists of roughly four steps: calibration, rectification, correspondence, and computation of depth image. The pipeline for this project is:

1. Compare 2 images in a dataset, use SIFT for feature matching
2. Estimate Fundamental Matrix through SVD, Ransac or Least Squares
3. Estimate Essential Matrix by multiplying the K intrinsic camera calibration, transpose of it, with the fundamental matrix
4. Decompose E to get the rotation and translation matrix (R and T)
5. Apply perspective transformation so that the epipolar lines are horizontal for each image
6. Print the homography matrices and plot the epipolar lines and feature points for both images
7. SSD/Cross correlation for each epipolar line
8. Calculate Disparity
9. Rescale disparity and save as gray scale and heat map conversion
10. Compute depth information for each pixel and save as color and heat map conversion

1. Calibration

The user is prompted to choose a dataset when the project is run. There are 3 datasets, labeled 1, 2, or 3. The images are then resized and a dictionary of all the data provided by the docx are included for each dataset. SIFT will be used to compute the feature matching. The images are converted to grayscale and the function orb.detectAndCompute(img, None) calculates the key points and descriptors for both images. Then, brute force match detection is used to find the matches between the two images. It is sorted and then the program draws the feature points and lines on the images. For this test, an array of 30 points of the key points are created, and then 8 points will be randomly selected. These key points are used to calculate the fundamental matrix.

To calculate the fundamental matrix, the formula below is used. It is given from the project description. In the program, k1 is x, y, and k2 is x', y'. The constraint for the fundamental matrix has to satisfy x'^T Fx = 0.

$$\begin{bmatrix} x_1x_1' & x_1y_1' & x_1 & y_1x_1' & y_1y_1' & y_1 & x_1' & y_1' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_mx_m' & x_my_m' & x_m & y_mx_m' & y_my_m' & y_m & x_m' & y_m' & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} = 0$$

The size of the key points passed to the fundamental matrix function are of size (8, 2). Here the fundamental matrix will be 8x9. Singular Value Decomposition is performed on the A matrix calculated from the above image. The last column of the V matrix is taken and reshaped to a 3x3 matrix as x because it corresponds to the least singular value. I used linear least squares to solve Ax=0 or seen as Af=0.

This project has an Output folder of all the images saved by datasets 1-3. Note that there is a color discrepancy between Matplotlib and OpenCV, so there may be a duplicate file with different colors.

The problem that I struggled with the most was the calculation of the fundamental matrices and having those inputted into the next different functions. For each dataset, contains a working feature matching, rectification, epipolar line calculation, and disparity for both grayscale and heatmap.

DATASET 1

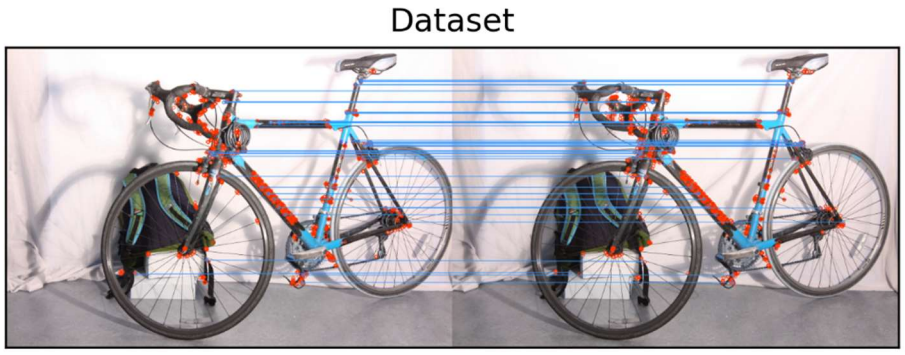

Figure 1. Feature Matching with SIFT

Figure 2. Rectification

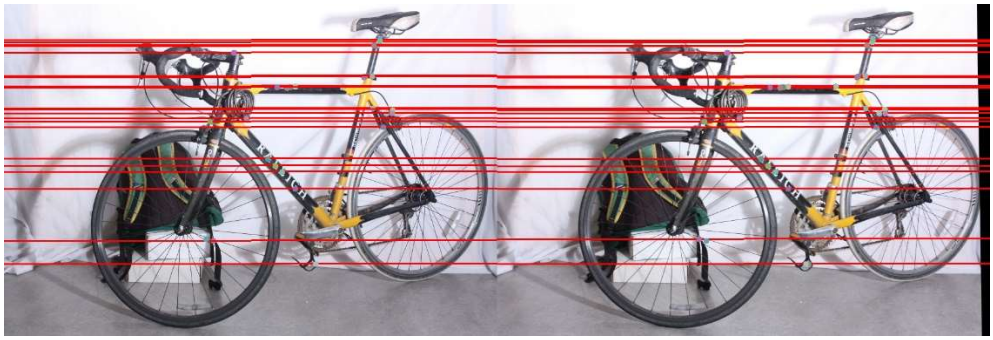

Figure 3. Epipolar Lines on Rectified Images

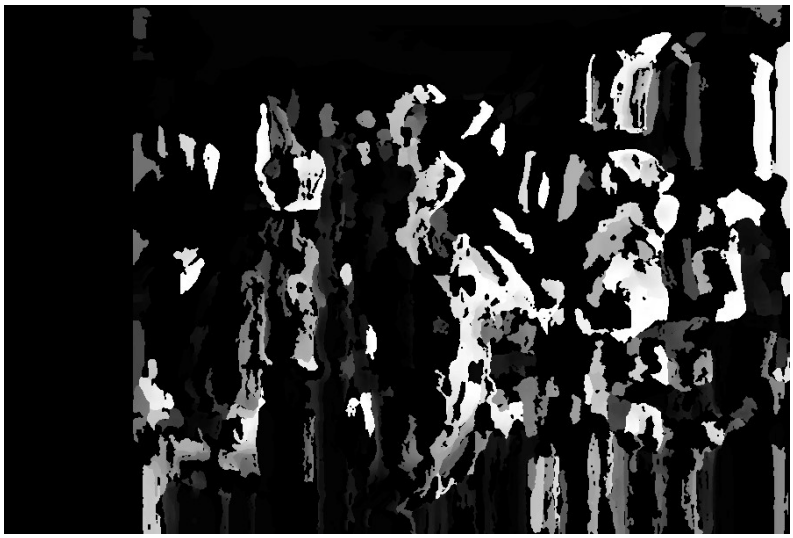

Figure 4. Grayscale of Disparity

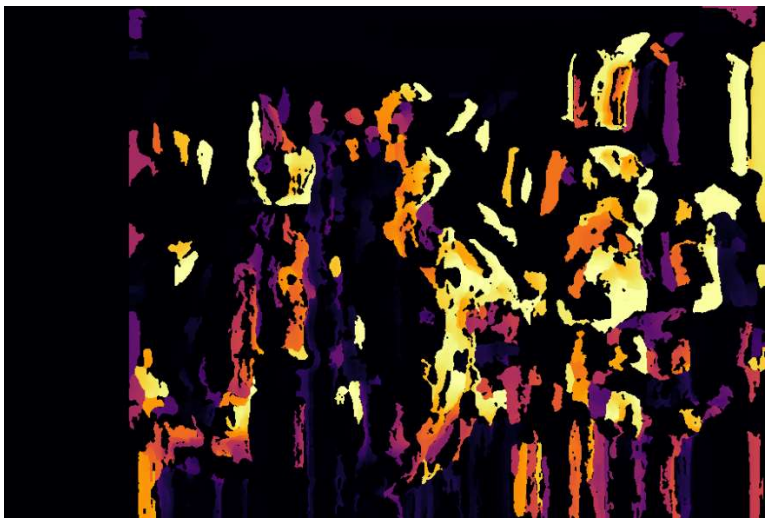

Figure 5. Heatmap

Fundamental Matrix

[[4.37622482e-09 -8.12227326e-05 1.45596322e-02]
 [8.50037908e-05 2.80078388e-06 -5.74323895e+11]
 [-1.53167070e-02 5.74323895e+11 1.00000000e+00]]

Essential matrix

[[1.27126831e-09 -1.36628793e+03 -1.82647572e+03]
 [-1.39094371e-09 -1.86446837e+03 1.49277245e+03]
 [-1.69637018e+03 1.82293283e+15 2.43719974e+15]]

Rotation

[[-9.99999999e-01 1.95663076e-05 2.61594941e-05]
 [3.26674077e-05 5.98955019e-01 8.00782670e-01]
 [-1.85082618e-13 8.00782670e-01 -5.98955019e-01]]

Translation

[9.94239760e+10 3.04352208e+15 -9.94239784e+10]

H1 Matrix

[[-5.93675085e+11 -1.33871552e+10 3.16954557e+13]
 [1.53167070e-02 -5.74323895e+11 3.40728205e+00]
 [8.50037908e-05 -6.10351562e-05 -5.74323895e+11]]

H2 Matrix

[[1.00000000e+00 -1.70760571e-14 5.11590770e-12]
 [1.70760571e-14 1.00000000e+00 -7.67386155e-12]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]

DATASET 2

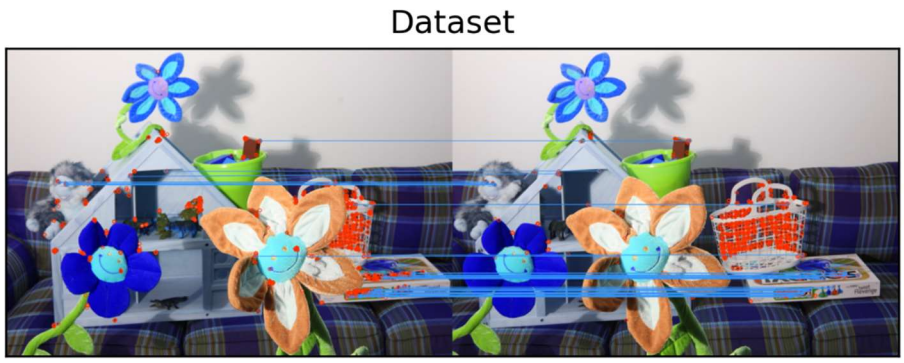

Figure 6. Feature Matching with SIFT

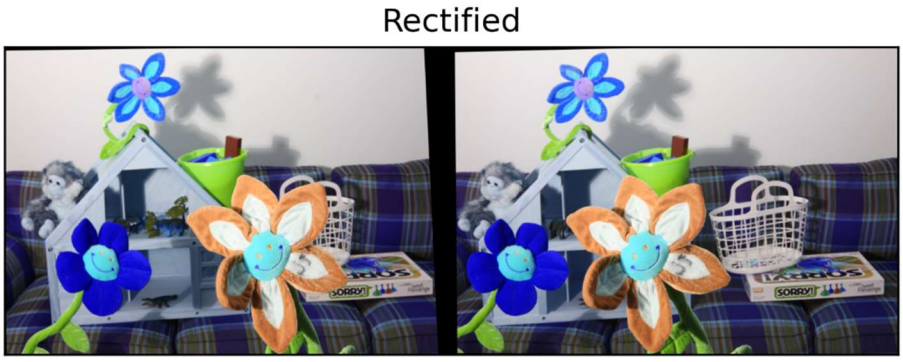

Figure 7. Rectification

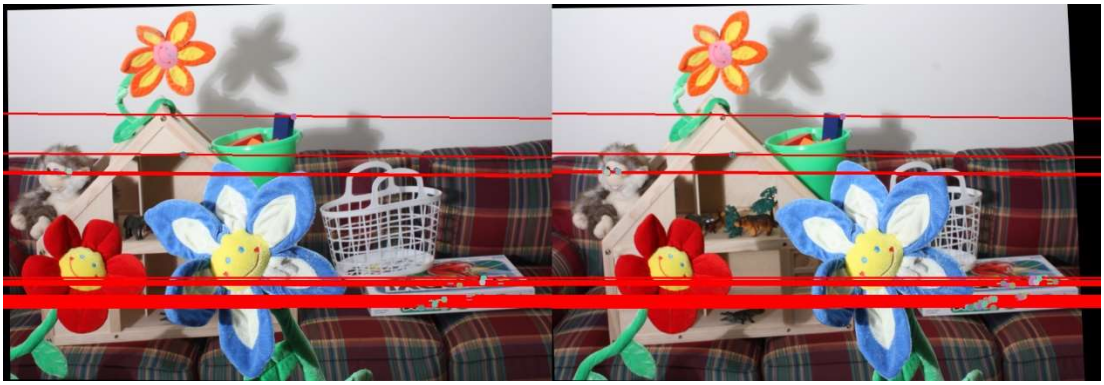

Figure 8. Epipolar

Figure 9. Grayscale Disparity

Figure 10. Heatmap

Fundamental Matrix

[[2.94510878e-09 1.77344569e-05 -8.25933189e-03]
 [-1.03983098e-05 -5.12519663e-06 3.60606760e-01]
 [4.78129780e-03 -3.60278262e-01 1.00000000e+00]]

Essential Matrix

[[32.92568891 -205.31093828 -272.59837205]
 [8.86499986 220.30467199 -40.19496035]
 [-141.94737711 917.97374247 1179.16813069]]

Rotation

[[-0.94859063 0.21546553 -0.23184136]
[-0.07338858 0.5628052 0.82332522]
[-0.30787973 -0.7980131 0.51805903]]

Translation

[-30.92541662 1549.8725754 -134.07414056]

H1 Matrix

[[3.43894839e-01 1.22723546e-02 -5.67978636e+00]
[-4.88676767e-03 3.60261799e-01 2.64128785e+00]
[-1.06247304e-05 -5.39031664e-06 3.68529499e-01]]

H2 Matrix

[[9.78384772e-01 8.28969612e-03 6.88402852e+00]
[-2.32583754e-02 9.99838829e-01 1.00953247e+01]
[-4.99521658e-05 -4.23236631e-07 1.02170461e+00]]

DATASET 3

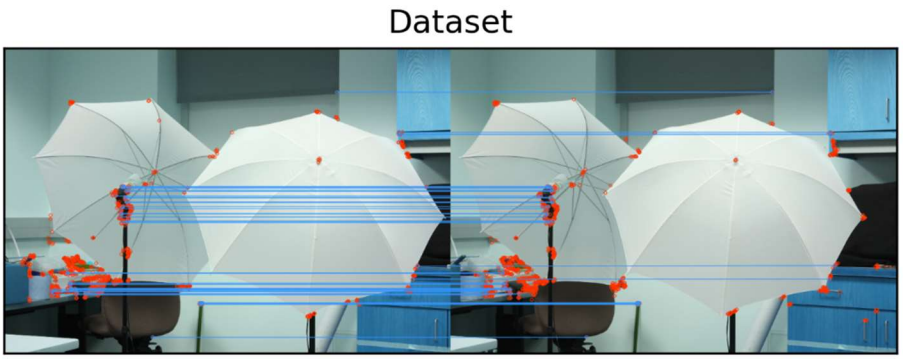

Figure 11. Feature Matching with SIFT

Figure 12. Rectification

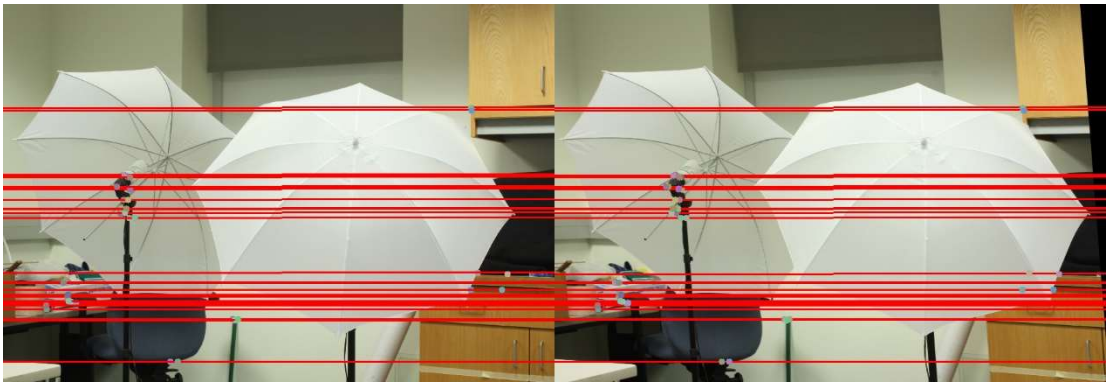

Figure 13. Epipolar

Figure 14. Grayscale Disparity

Figure 15. Heatmap

Fundamental Matrix

[[-3.20892207e-22 3.14731077e-18 -1.39236818e-15]
 [-3.29107048e-18 -3.20892207e-19 -4.32611806e-03]
 [1.46177977e-15 4.32611806e-03 -1.20792265e-13]]

Essential Matrix

[[-3.25939955e-22 8.31568994e-11 -6.59192779e-11]
 [-3.22105954e-22 -7.74005680e-11 -8.02421782e-11]
 [-7.71491829e-11 1.96867162e+01 -1.56025819e+01]]

Rotation

[[9.99999995e-01 -7.72987644e-05 6.12626355e-05]
 [-9.86316860e-05 -7.83711224e-01 6.21125355e-01]
 [-7.42659093e-12 -6.21125359e-01 -7.83711228e-01]]

Translation

[2.47761409e-03 2.51198595e+01 2.47761421e-03]

H1 Matrix

[[4.44394601e-03 3.38083118e-04 -2.85938119e-01]
 [2.21170960e-13 4.32611806e-03 -9.72555370e-14]
 [7.30804911e-16 7.28381202e-16 4.32611806e-03]]

H2 Matrix

[[1.00000000e+00 -1.02142457e-13 3.09228199e-11]
 [1.02142457e-13 1.00000000e+00 -4.53610483e-11]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]