Giorgio Mendoza
CS549-E24-E01
Dr. Alexandros Lioulemes

Week 7 Assignment

Objective: The objective of this project is to accurately reconstruct a 3D scene using stereo images captured from two different perspectives. The steps involved in achieving this goal include: Camera Calibration, Image Acquisition, Feature Extraction and Matching, Fundamental Matrix Calculation, Essential Matrix Computation, Recovery of Rotation and Translation, Projection Matrix Formation, 3D Point Triangulation and Re-projection Error Estimation.

Code Description: The program initiates by loading the following images that were previously undistorted using camera calibration.

Figure 1. Undistorted Images

The SIFT algorithm was used to detect keypoints and compute descriptors in both images, which represent distinctive features in the images. A brute-force matcher was also used to find matches between the descriptors of the two images. Then a ratio test was applied to retain only the best matches, ensuring reliable correspondences between the two images as shown below.

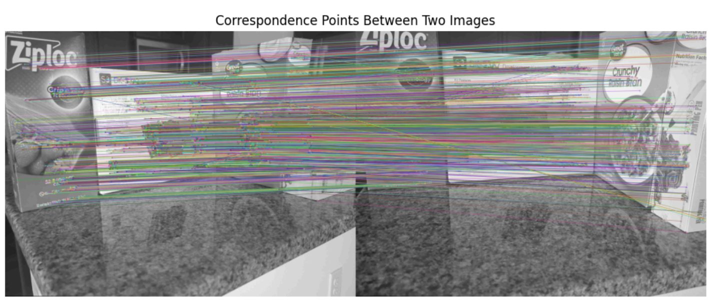

Figure 2. Correspondence Points

Using the matched keypoints, the code computes the Fundamental matrix with the RANSAC algorithm to model the relationship between the points in the two images. The F matrix was validated by filtering the matches that fit the epipolar constraint.

```
Fundamental Matrix F:
 [[-2.02134213e-08  8.20086139e-07 -3.89752132e-05]
 [-2.32957826e-06  4.08321880e-07  9.90021617e-03]
 [ 1.19923391e-04 -9.89220451e-03  1.00000000e+00]]
```

Figure 3. Fundamental Matrix F

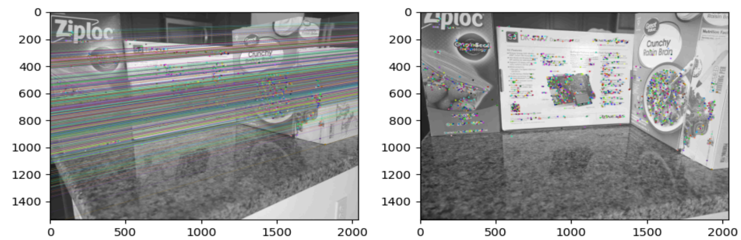

Figure 4. Epipolar Lines and Correspondence Points

The Essential matrix was derived from the Fundamental matrix using the camera's intrinsic parameters. Singular Value Decomposition was performed on the Essential matrix to recover the relative Rotation (R1, R2) and Translation between the two camera frames.

```
Essential Matrix E:
 [[ -0.90419433   6.03089435   1.51303541]
 [-24.2986711    4.12165038  30.19345352]
 [ -2.51227522 -34.0365086    0.83800499]]
Determinant of the Essential Matrix (should be close to 0): -3.7187881568153313e-13
Singular values of the Essential Matrix: [3.92116500e+01 3.44453304e+01 3.34737382e-15]
```

Figure 5. Essential Matrix E

Four possible projection matrices (P21, P22, P23, P24) for the second camera were generated using the possible combinations of Rotation and Translation obtained from the Essential matrix.

Then the program triangulated 3D points from the corresponding 2D points in both images using the LinearLSTriangulation function. This function applies the linear least squares method to compute the 3D coordinates of the points. The triangulation process is repeated for each of the four projection matrix options to identify the matrix that produces the most points in front of both cameras, ensuring a physically plausible 3D reconstruction.

```
Triangulated 3D Points:
 [[-0.47839151  0.18164472  1.95617687]
 [-0.47697741  0.20483133  1.95494037]
 [-0.47697741  0.20483133  1.95494037]
 [-0.4707525   0.19674074  1.95601648]
 ...
 [ 0.53428286  0.19491632  1.90308745]
 [ 0.53085272  0.2645603   1.89362574]
 [ 0.53518425  0.40432017  1.87320568]]
```

Figure 6. Triangulated 3D Points

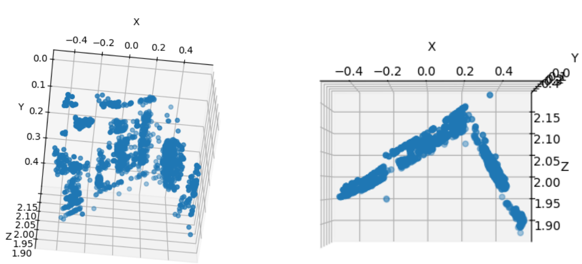

Figure 7. Triangulated 3D Points Representation

The triangulated 3D points were finally projected back onto the original image planes using the corresponding projection matrices.

The re-projection error, which measures how closely the re-projected points match the original 2D points, is calculated for both images to assess the accuracy of the 3D reconstruction.

```
Mean Re-projection Error in Image 1: 16.64055341868452
Mean Re-projection Error in Image 2: 20.74622949288274
```

Figure 8. Projection Error

Lastly, the program saved the 3D points and their corresponding colors into a PCD file using the SavePCDToFile function. Open3D was used to display the point cloud file from multiple views, offering a clear visualization of the 3D structure.

Conclusion: The 3D reconstruction process successfully generated a point cloud representation of the scene captured in the reference images. The projection matrices were computed based on the fundamental and essential matrices, leading to a relatively accurate reconstruction with mean re-projection errors of approximately 16.64 pixels for Image 1 and 20.75 pixels for Image 2.

The 3D point cloud visually corresponds well with the reference images, demonstrating the effectiveness of the camera calibration and feature matching processes. Despite some re-projection error present, the overall structure and spatial relationships of the objects in the scene are accurately captured in the reconstructed 3D model.

Figure 9. Reference Image and 3D Points Reconstruction