CSCI 5561 Homework 5: Stereo Reconstruction

Ashvin Pidaparti, 5421833

Input images are shown in figure 1. The key points of each image were detected and computed using SIFT. The K-Nearest Neighbors algorithm was used to find the corresponding points between the two input images and a bidirectional consistency check and ratio test were applied to determine distinctiveness and validity of each pair of matches. The final matches are visualized in figure 2. The fundamental matrix F was computed using the 8-point algorithm within RANSAC, with 5000 iterations and a RANSAC threshold of .1 compared against the result of v^tFu and vF^tu^t. The epipolar lines of the image pair were visualized in figure 3. It was found that the computation of F could be somewhat inconsistent, so the computation was run several times and the visualizations were evaluated until a relatively high quality fundamental matrix was computed, and the result was cached and reused for later processes. The cached file is attached in this submission. From the fundamental matrix and the provided intrinsic parameter, a provided function was used to calculate 4 possible rotation matrices and camera vectors corresponding to the positions of the two cameras (figure 4). One camera was defined to be the origin and the rotation matrix rotated the second camera relative to the first. For each of these rotation matrix-camera vector pairs, each key point was computed in 3 dimensional space. Point clouds and cameras are visualized in figure 5. The correct camera pose was disambiguated via cheirality and the correct camera poses are shown in figure 6. Homographies were computed to rectify the input images such that their epipolar lines were parallel to the x-axis. This is shown in figure 7. A disparity map was constructed from the warped images, where the disparity is the pixel distance between a pixel in one image and the corresponding pixel in the other image and the corresponding pixel minimizes feature distance between the 2 pixels. Rather than iterating along each row of each image and computing the disparity, K-Nearest Neighbors was used on a row by row basis. Intuitively, the nearest neighbor of a pixel in a row in the left image in the right image will be of minimum feature distance from the original pixel. As such, we can simply compute the pixel distance between a pair of nearest neighbors within a row. The visualization of this is shown in figure 8.

Figure 1. Input images

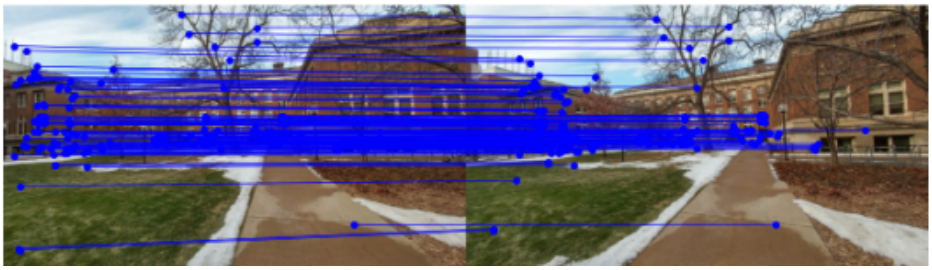


Figure 2. Correspondence of SIFT Key points after bidirectional consistency check and ratio test

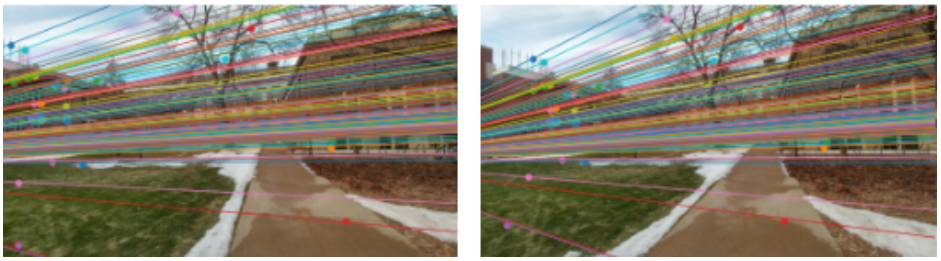
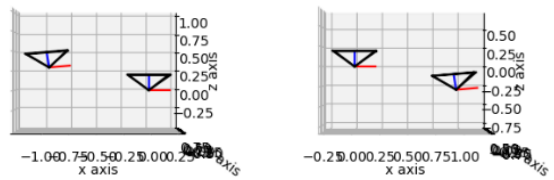


Figure 3. Epipolar lines based upon the fundamental matrix



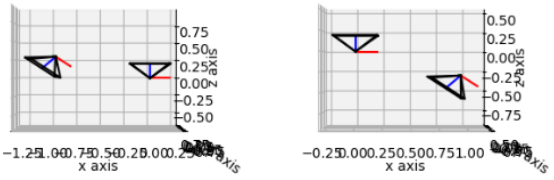
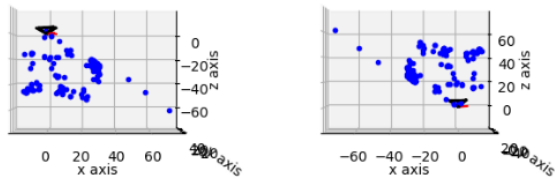


Figure 4. Valid configurations of camera poses



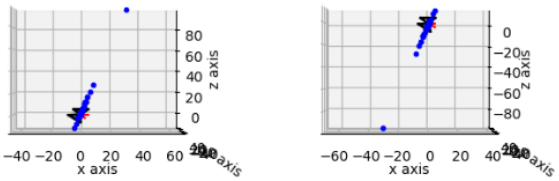


Figure 5. Valid configurations of camera poses with 3 dimensional point clouds

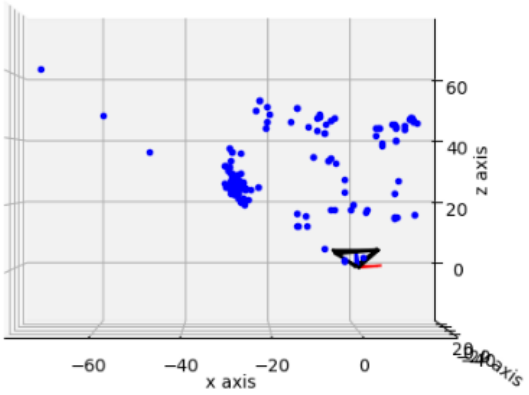


Figure 6. Disambiguated camera pose

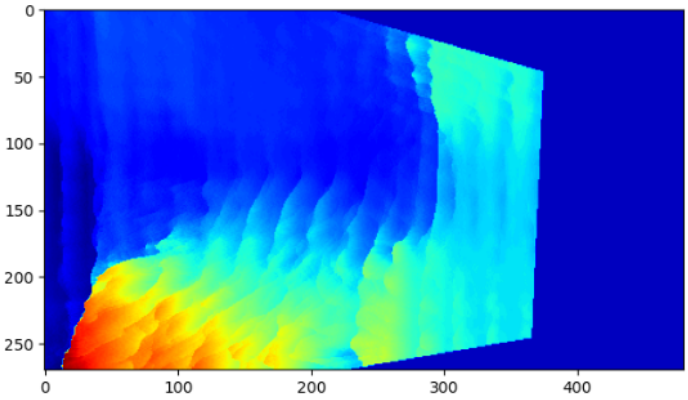


Figure 8. Disparity map

Figure 7. Rectified input images