jy3119 (01700345), sh1319 (01708685)

1. Data Collection

Data Selection Given the task of detecting interest points from an object, we immediately recognised the importance of selecting an appropriate object. The object chosen was a medicine box with multiple images, colors, and words, suggesting plentiful features for easy detection of keypoint using any algorithm. The medicine box is shown below in Figure 1.

Figure 1. Sample FD Image from Dataset

2. Keypoint Correspondences Between Images

Manual and automatic detection of correspondences between images contrast in both the quantity and quality of the keypoints. The advantages of manual detection include the accuracy and flexibility in choosing the keypoints, as they can be selected and adjusted manually. Manual detection is dependent on the human eye, which is much less likely to fail or cause any outliers, making this method more robust. On the contrary, manual detection of keypoint correspondences between images may lack in quantity. This is because this task has to be done by human one by one, hence is much more labour intensive and slower.

The strengths of automatic detection lie with the quantity of correspondences detected. The interest point detection algorithm used was detectSURFFeatures, with the parameter metric threshold set to 2000. Before the feature

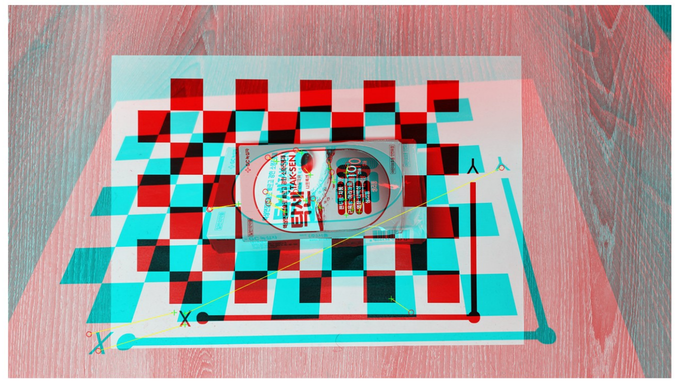

Figure 2. Automatic Correspondence with 21 out of 180 Keypoints Displayed

detection, the images were grayscaled. Then, we used the matchFeatures algorithm with the SAD Metric and a metric threshold of 5 to match the interest points of the 2 images. Using this algorithm, a total of 180 key points were detected from the algorithm, which definitely served as an advantage in terms of quantity. However, the quality of each keypoint may be of issue for the automatic detection algorithm. For instance, Figure 2 displays an outlier connecting the letter X to the letter Y, which is visibly an error. This problem can be mitigated by selection of the few best correspondences from the automatic detection algorithm.

In summary, both manual and automatic algorithms have their strengths and weaknesses, with manual detection being more ideal for images with fewer features, where automatic detection struggles. On the other hand, automatic matching is favorable in the case with many keypoints and features. Despite the potential of outliers for automatic detection, the sheer quantity of keypoints cover for this, making the trade-off worthy.

3. Camera Calibration

Method and Result 4 images taken at different positions out of 18 images taken were accepted by the Camera Calibrator application on Matlab. Extrinsic parameters can be extracted using the calibrator. For instance, the Rotation

matrix has been found: $\begin{bmatrix} 0.0229 & 0.9988 & -0.0423 \\ -0.9997 & 0.0227 & -0.0035 \\ -0.0025 & 0.0424 & 0.9991 \end{bmatrix}$
The intrinsic matrix has been calculated as well. Intrinsic matrix, K: $\begin{bmatrix} 971.17 & 0 & 0 \\ 0 & 970.85 & 0 \\ 651.90 & 368.31 & 1 \end{bmatrix}$

Focal lengths f_1 = 971.17, f_2 = 970.85 have been found. Focal length of the camera is in pixel dimensions, indicating how the axes are related to one another.
Parameters c_x = 651.90, c_y = 368.31 have been found as well, and they represent the coordinates of the principal point in pixel dimensions. For the actual images that were accepted, refer to Figure 9.

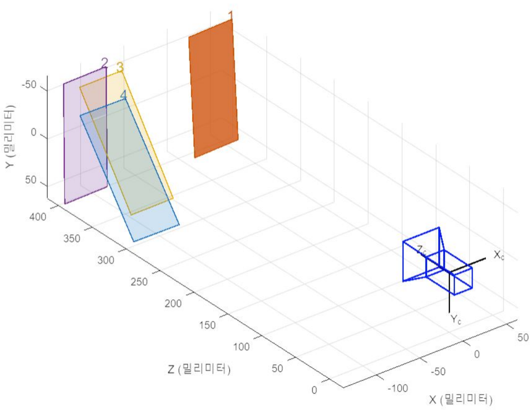

Figure 3. Extrinsic Parameters Visualization

Illustration of Distortion Distortion of images can be determined and measured with the MATLAB toolbox. Using the parameters obtained from the Camera Calibrator, the distortion vector can be found - for example, a FD image chosen has shown a distortion vector of [0.1386 -0.3827], indicating slight distortion. 'undistortImage' function from the toolbox has been used to fix the distortions in the image, as shown in Figure 4. The toolbox however, does not prove to be correcting the camera distortion effectively, with only slight distortions present in the image.

4. Transformation Estimation

HG Image The homography matrix retrieved from two HG images selected as shown in appendix Figure 10 was $\begin{bmatrix} 0.9949 & -0.0898 & 0.0000 \\ 0.1031 & 0.9987 & 0.0000 \\ -145.2231 & 202.3153 & 0.9999 \end{bmatrix}$. Using the homography matrix, we retrieved the image projecting the keypoints correspondences as shown below in Figure 5. Most keypoints shown on the figure is accurate; however, there are definitely some outliers, caused by the table background pattern outside of the object.

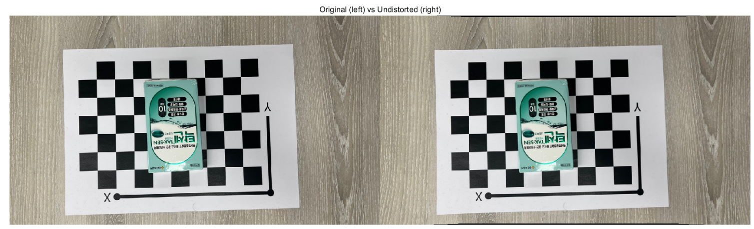

Figure 4. Using the MATLAB Toolbox to fix distortions

Figure 5. HG Image Keypoints and Correspondences

FD Image The fundamental matrix retrieved from two FD images selected as shown in appendix Figure 11 $\begin{bmatrix} 0.0000 & 0.0000 & -0.0028 \\ -0.0000 & -0.0000 & 0.0082 \\ -0.0000 & 0.0121 & 0.9999 \end{bmatrix}$. Here, the function estimateFundamentalMatrix from MATLAB was used, with the RANSAC method, 10000 number of trials, 0.1 distance threshold, and 99.99 confidence. The reason behind the choice of the RANSAC algorithm was because the least median square method does not remove outliers from the image while RANSAC does, hence suggesting RANSAC will produce a better matrix with less errors. Using the fundamental matrix from this function, we were able to retrieve the inliers and epipolar lines as illustrated in Figure 6.

Outlier Tolerance From the results of the FD image, we know that the original input had a total of 21 matches, while

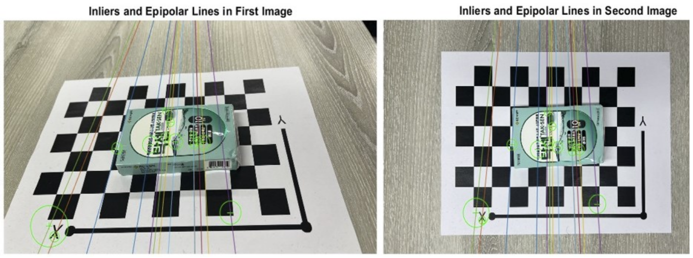

Figure 6. FD Image Inliers and Epipolar lines

the inliers that were kept were only limited to 13, suggesting that 8 of the 21 matches were considered outliers.

5. 3D Geometry

Stereo Rectification The FD images were stereo rectified using the estimateStereoRectification function and the rectifyStereoImages function. To begin with, the inliers point from the previous step was used to create transforms for both left and right images. Then, by applying the transforms to the each images, the images in Figure 7 were derived, along with the corresponding epipolar lines. Lastly, the images were illustrated in one image together, as shown in the appendix Figure 12.

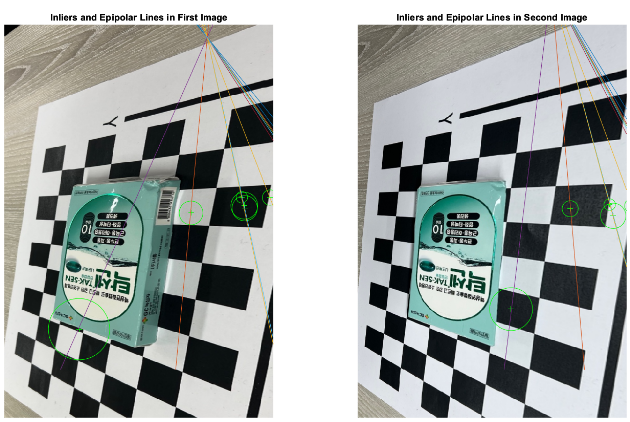

Figure 7. Stereo rectified pair of Images with Epipolar Lines

Depth map Another pair of images have been taken at different angles with a stereo camera setup. (Refer to Figure 13.) Stereo parameters have been computed with a stereo camera calibrator application on MATLAB afterwards. Refer to Figure 14 in Appendix for extrinsic parameter visualization of stereo camera setup. Stereo images have been rectified. Then, disparity map and depth map have been calculated on MATLAB based on the rectified images and stereo parameters, as shown in Figure 8, which illustrates disparity and depth map respectively.

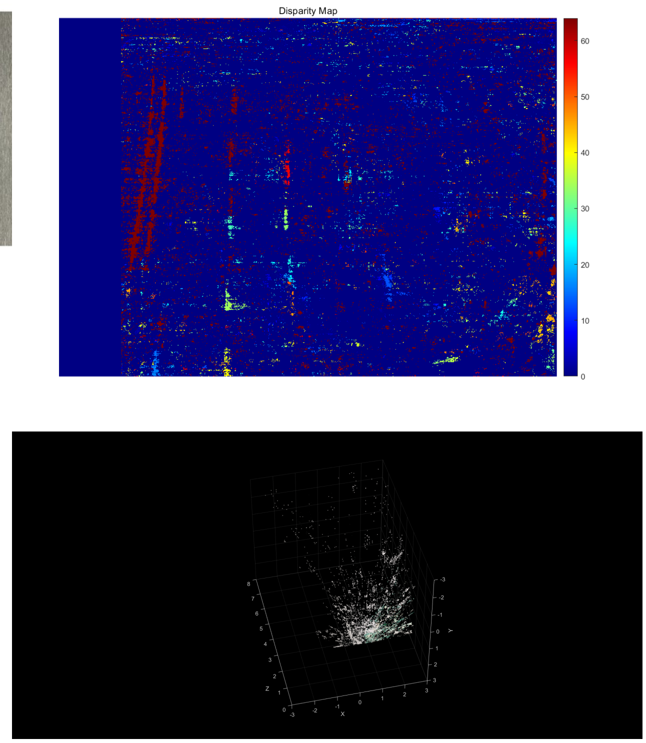

Figure 8. Disparity and depth map

6. Conclusions

In conclusion, manual and automatic keypoint correspondence detection have their own advantages and disadvantages. While automatic keypoint detection is useful and highly efficient in detecting many keypoints in a short period of time, manual keypoint detection is usually slower and more tedious. However, it can be more accurate, especially for objects that do not have sufficient or obvious keypoint features. Camera calibrator application on MATLAB allows one to compute intrinsic and extrinsic parameters. Distortion parameters can also be computed with the application. However, MATLAB is unable to fix distortions in a distorted picture effectively. Transformation estimation has also been computed by MATLAB using the RANSAC method. RANSAC method proves to be a better method to produce a fundamental matrix than using the least median square method, as the RANSAC method removes outliers and produces less error. Using stereo camera setup, images taken at different angles can be rectified and a depth map can be produced from the stereo camera parameters and rectified images from MATLAB. However the result shown was very faint, indicating that MATLAB may not be always efficient in rectifying images and producing an accurate depth map from it.

7. Appendix

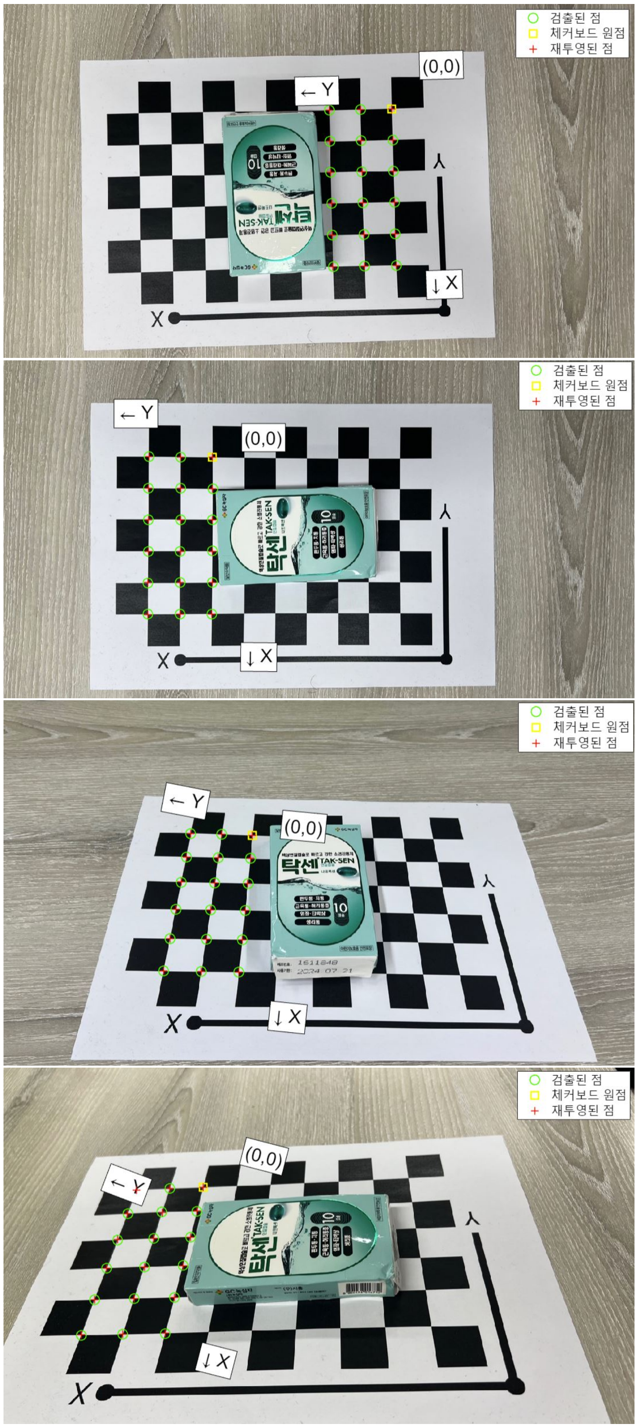

Figure 9. Illustration of Distortion of the camera

Figure 10. HG Image Pair Selected for Transformation Estimation

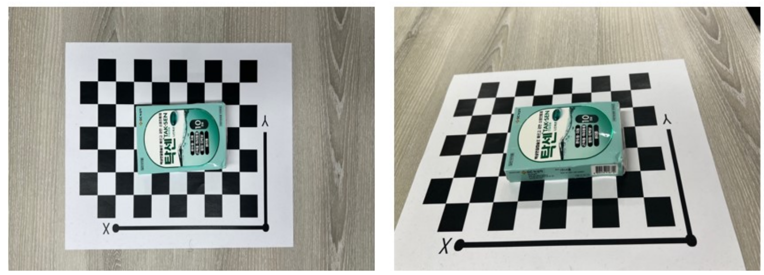

Figure 11. FD Image Pair Selected for Transformation Estimation

Figure 12. Stereo rectified pair of Images Combined

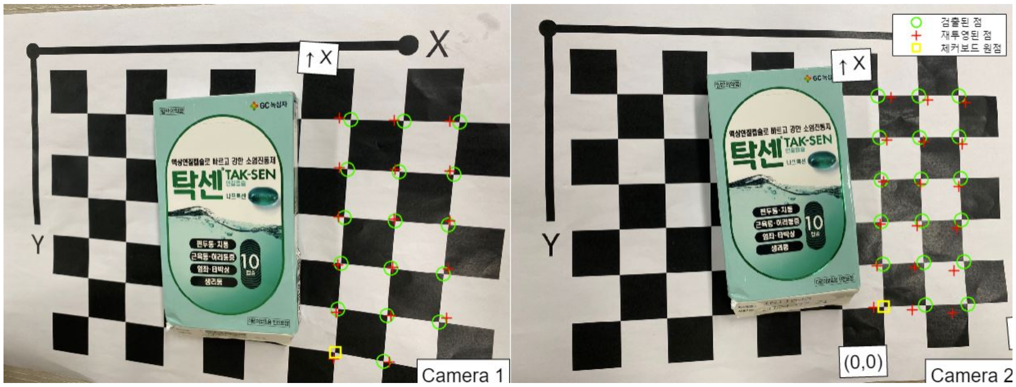

Figure 13. Stereo images used for rectification and building of depth map

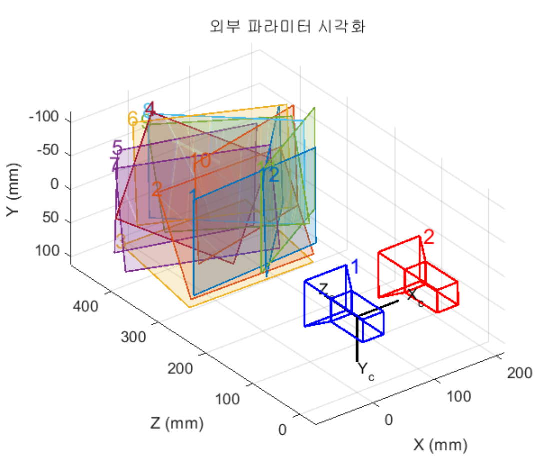

Figure 14. Extrinsic parameter visualization of stereo images