Computer Vision PA2 Report

20195125 이성민

1. SIFT Feature Extraction and Matching

Function ‘vl\_sift’ takes a gray color image as input and outputs frames and descriptors. Frames are consisted of 4XN matrix where N means the number of feature points. In each column, first and second elements mean the center of a feature point on the image plane and third is scale and fourth is orientation. Descriptors are consisted of 128XN matrix. It contains scale invariant informations of each feature point. So two descriptor matrices from two images are passed to ‘vl\_ubcmatch’ function to find corresponding feature points between two images. Outputs of ‘vl\_ubcmatch’ are matches and scores. Matches consists of 2XM matrix where M means the number of match points between two feature points.

1. Estimate Essential matrix Using RANSAC

To estimate essential matrix, randomly selected five matching points(x,y coordinates) are passed to ‘calibrated\_fivepoint’ function. Instead of passing the original matching points into the function, they should be normalized by calibration matrix to satisfy constraint essential matrix and corresponding points . The function outputs several essential matrices. For each essential matrix, we have to count the number of inliers of the essential matrix to find a best one. Fundamental matrix F is calculated by using a relation between essential matrix and fundamental matrix . And using this fundamental matrix F, we computed a ‘distance’ for all matching point and count the number of points that are smaller than a threshold. This method is originated from 8 point RANSAC algorithm.

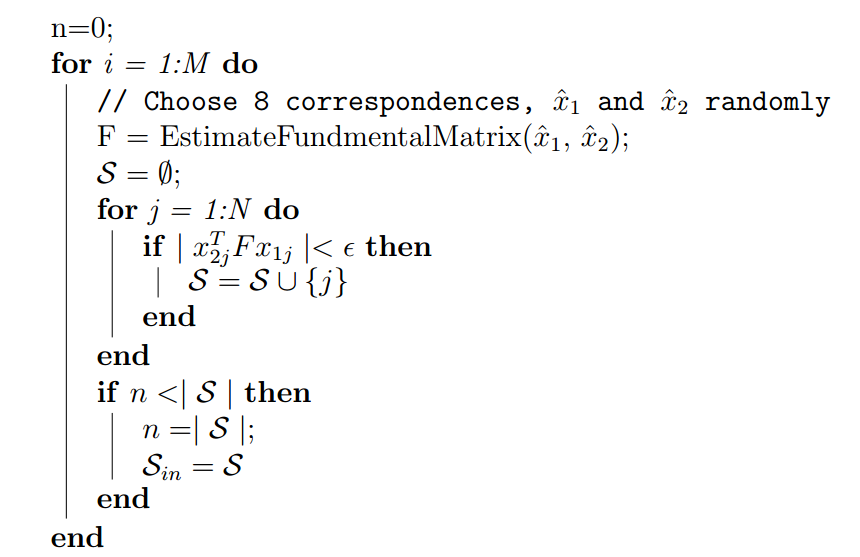


Figure . 8 point RANSAC with fundamental matrix

1. Essential matrix Decomposition

Using the best essential matrix E calculated by RANSAC, now we can estimate camera matrix R and t. By simply using singular value decomposition of E, we can find four possible camera matrices.

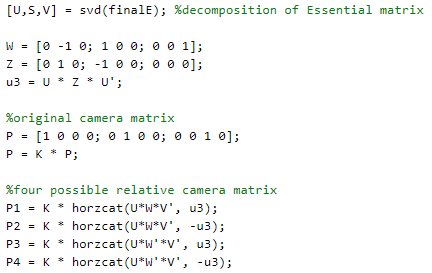


Figure . Relative camera pose matrices

1. Triangulation

According to the relation where x is a point on image plane and X is a world 3D point, we have to calculate 3D points X using known parameters and x. it can be easily solved by multiplying skew matrix of x which translates into where

This homogeneous linear system can be solved by SVD of A.

1. Results

Actually results generated by my own source code is depressing. Results generated by four possible camera pose can be found at ply files in the repository.

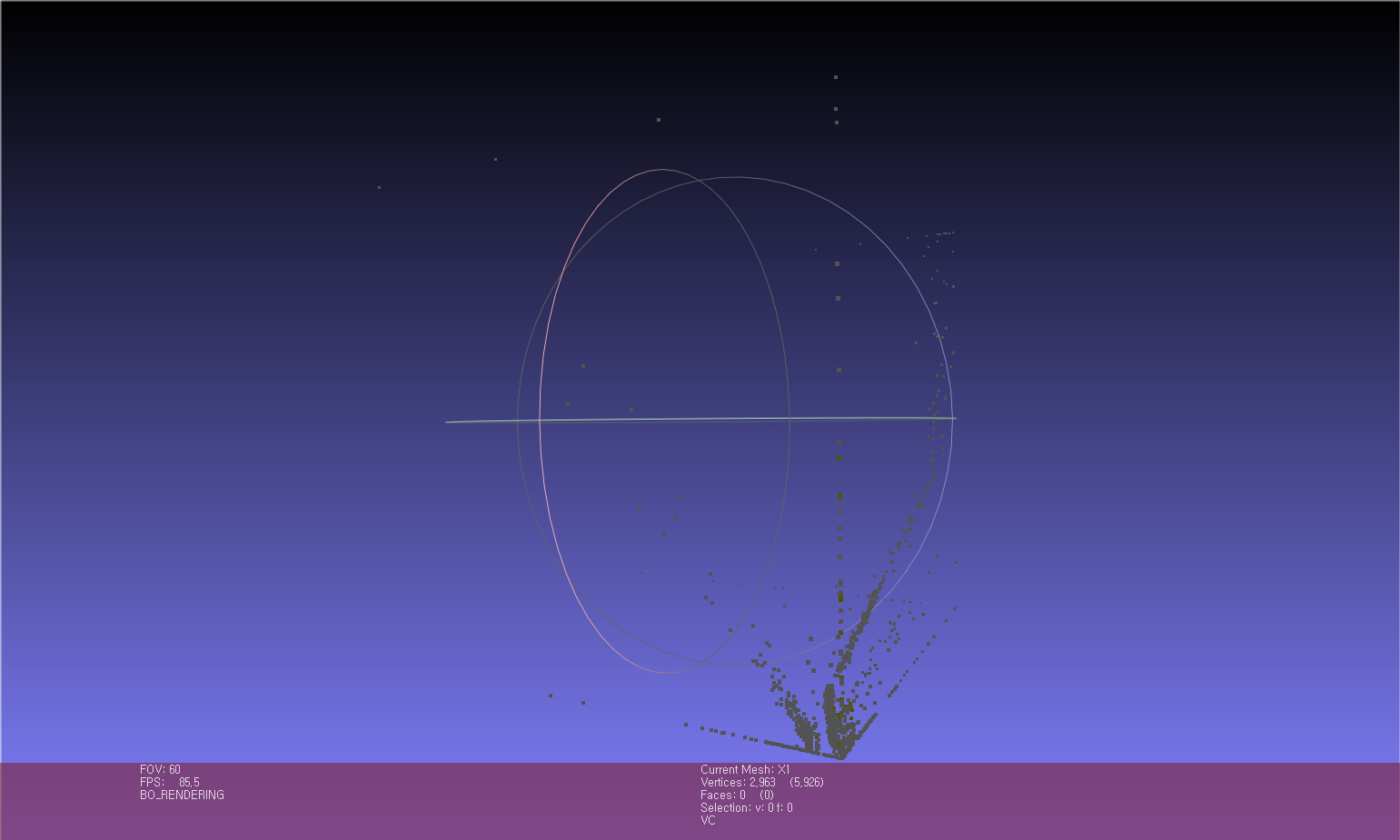


Figure . Example result of triangulation

Almost every points calculated by triangulation are located nearly at the original point. Structures in the given images cannot be found. In my opinion, there would be some mistakes when estimating essential matrix because other parts(feature extracting and matching, triangulation) are just using open functions.