

Report of Anton Maksimov (antonma, 16-952-137), Task 7 "Structure from motion"  
on ETHZ course "Computer Vision".

1. Using 8-point RANSAC we estimate fundamental matrix  $F$  of the first pair of images (the first and the last), matches are extracted with SIFT (RANSAC threshold  $10^{-4}$ ). After that we compute essential matrix as  $E = K^\top F K$  using provided camera matrix  $K$ . We obtain projection matrix using `decomposeE.m`. Then we triangulate corresponding inliers (with `linearTriangulation.m`), which were calibrated on cameras multiplying  $K^{-1}$ . and use as the identity projection matrix for the first image points and obtained one for the last image points.

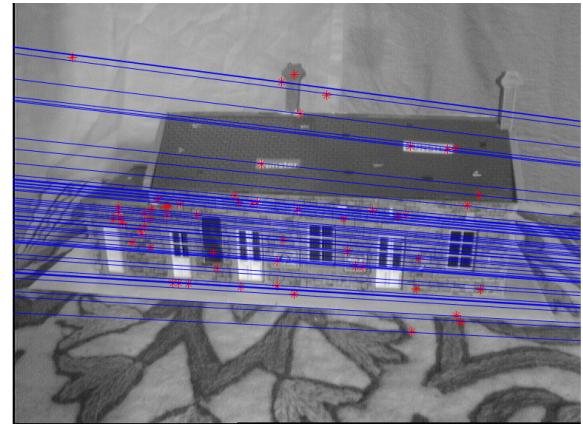
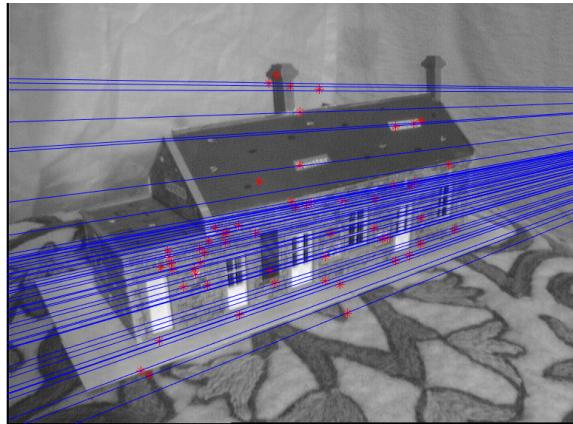


Figure 1: Epipolar lines corresponding to fundamental matrix calculated from the first and the last images. Relative position of the first camera is almost orthogonal to the view of the second, so the epipolar lines are almost parallel. Epipolar lines on the first image intersect at the adequate point — position of the second camera. To plot epipolar lines we used code from one of the previous tasks.

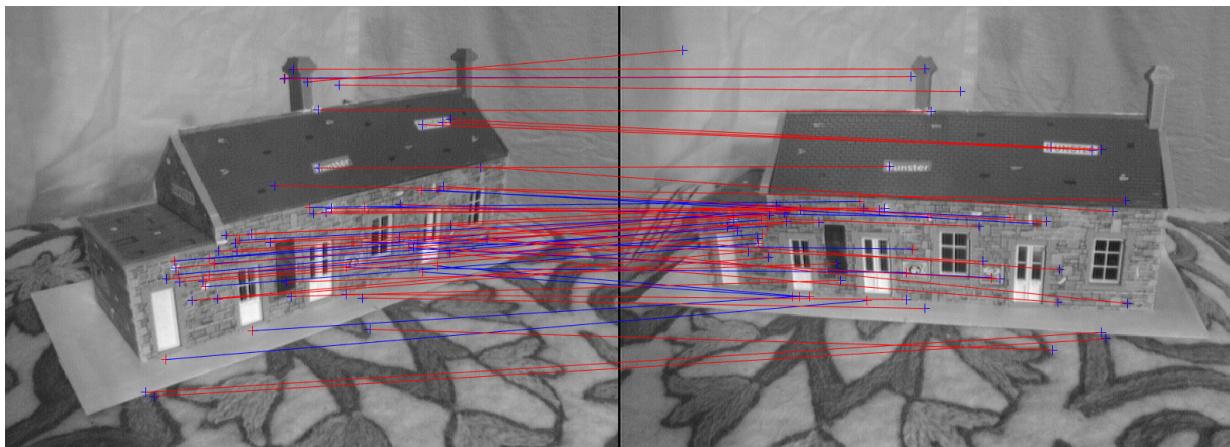


Figure 2: Initial search of matches between the first and the last (fifth) images using 8-point RANSAC, red lines show outliers, blue — inliers.

**2–3.** Using included function `ransacfitprojmatrix.m` (which is 6-point RANSAC with included 6-point DLT in order to have an error measure as reprojection error) we obtain projection matrices for other views relative to the first view. Similarly to previous we obtain triangulated 3D-point and plot them in the end. All this is implemented as a function `addView` which is called for every new view (fig. 3–6, resulting points and cameras at fig. 7).

**4.** In order to get dense reconstruction we use code from the previous task6 adapted for this task. We change almost only `exercise6.m` file in `dispairity` folder, which is called from `exercise8` in order to obtain disparity map and rectified images (using graph-cut as the best in quality method from two used in task6), which are then triangulated to produce 3D-image using as projection matrices ones obtained from the obligatory part of the task. (fig. 8)

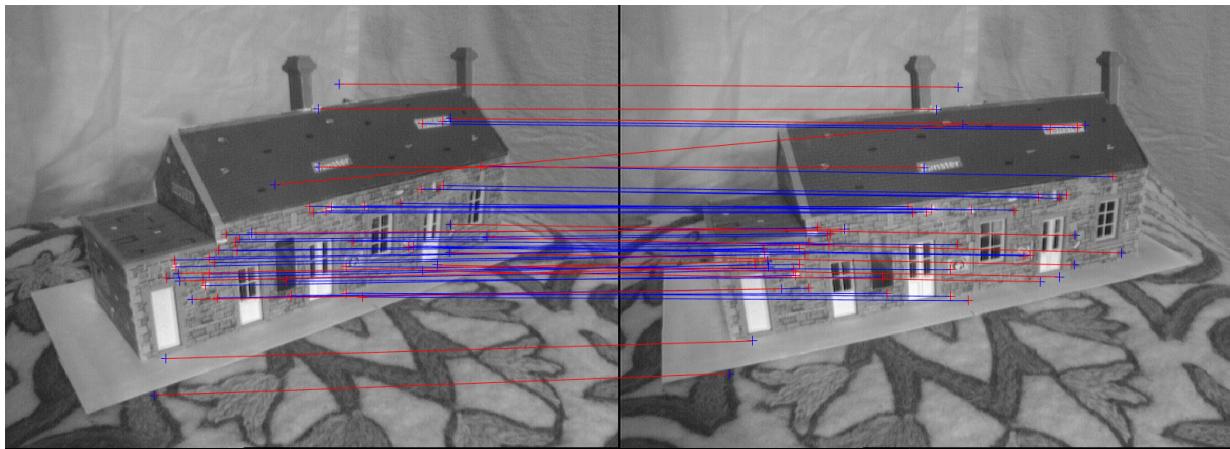


Figure 3: Between first and second images.

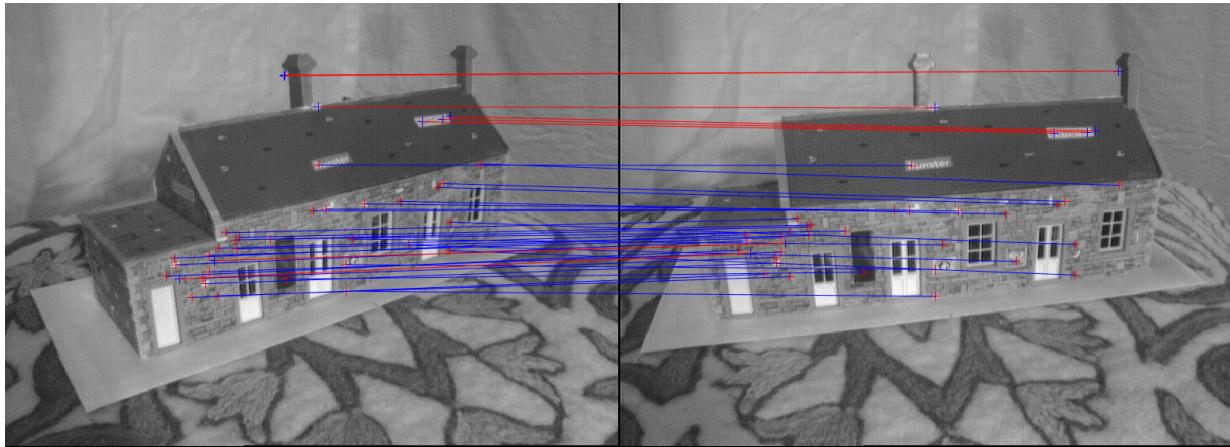


Figure 4: Between first and third images.

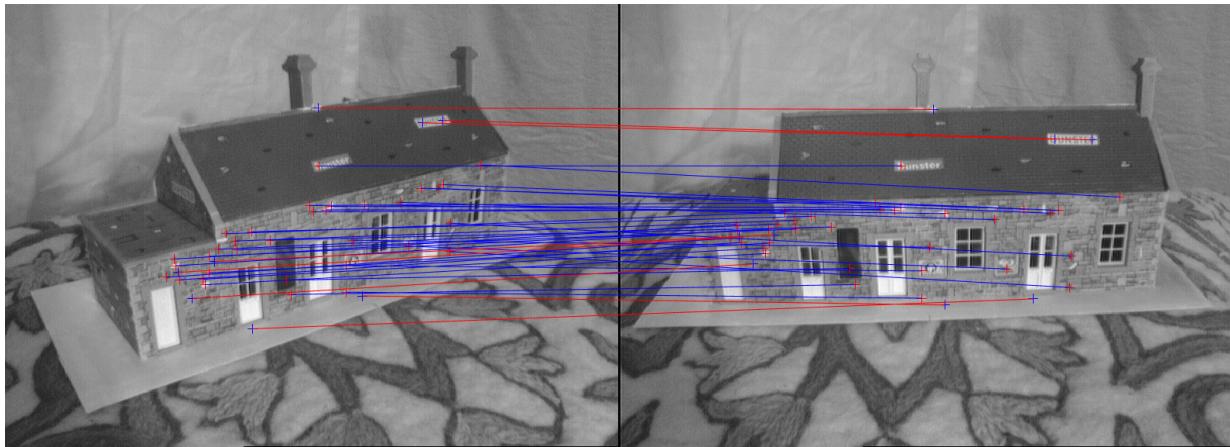


Figure 5: Between first and fourth images.

Figure 6: Matches between images obtained using using 6-point RANSAC and with selection of inliers using 6-point DLT (filtering using corresponding reprojection error), red lines show outliers, blue — inliers.

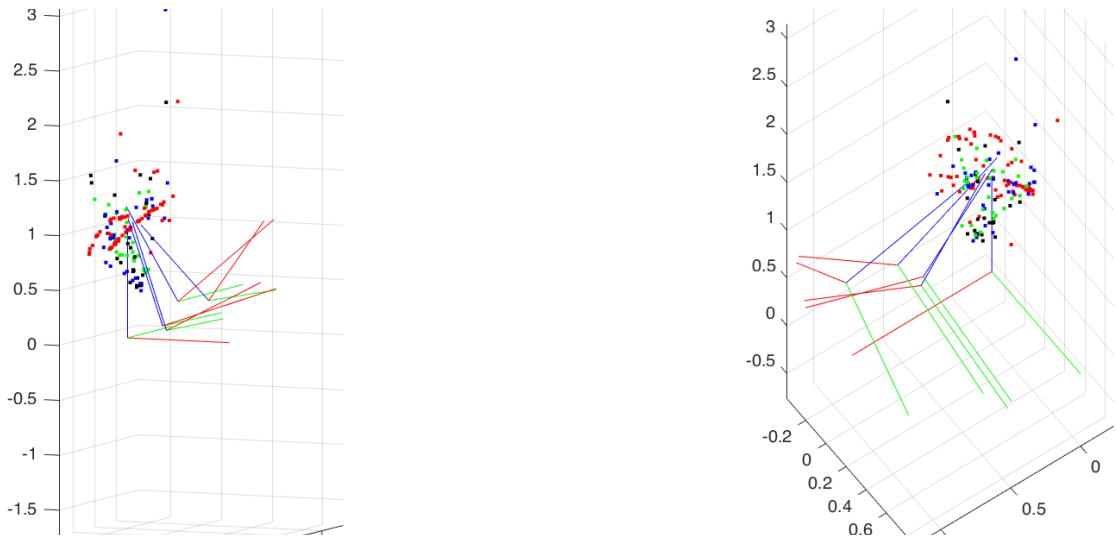


Figure 7: Triangulated 3D-points (sparse reconstruction) and camera poses extracted from images (red — from the first triangulation during initialization). Blue axes show depth of the camera (direction of view from the camera origin), and this corresponds to the true their arrangement (fan-like). There are much less points from other than the first views, and some are outliers, which could be caused by non-ideality of RANSAC + use of points which are at the background of house.

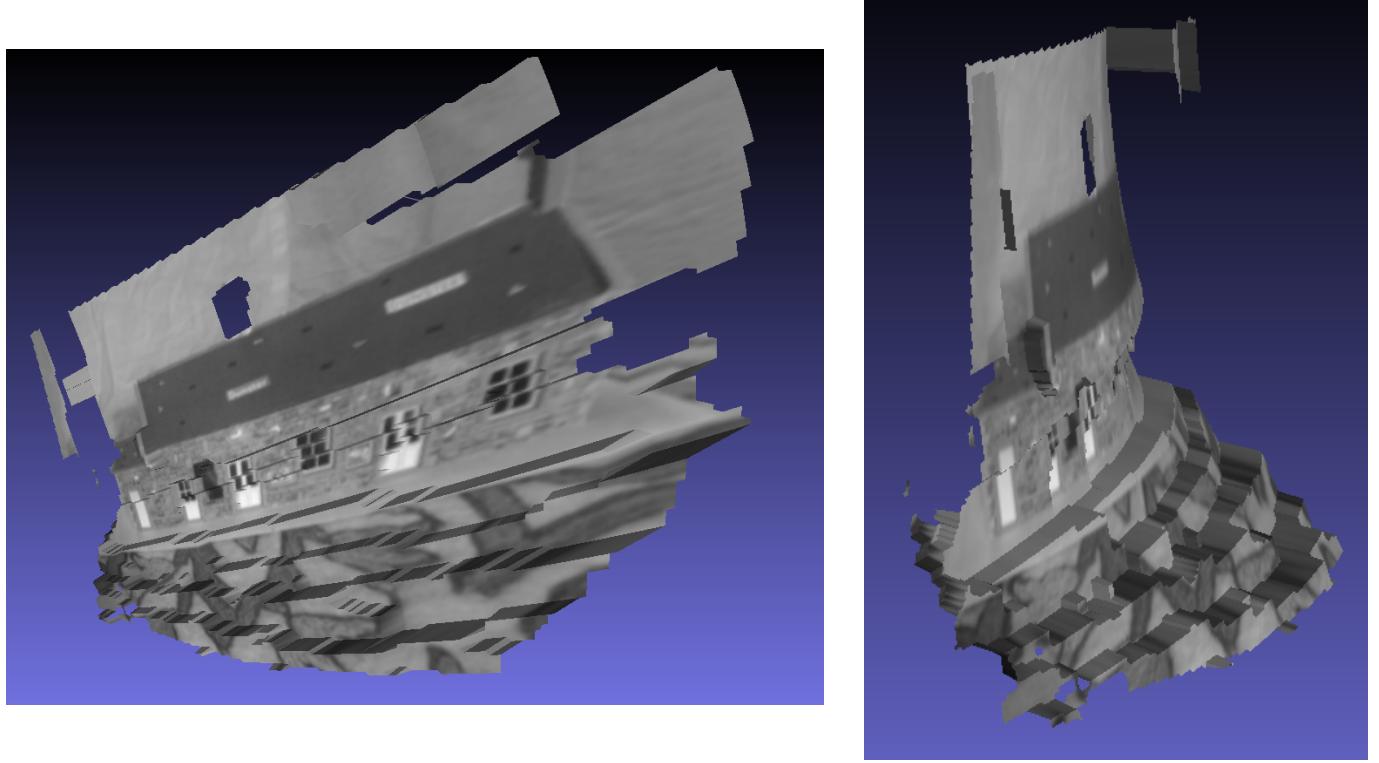


Figure 8: Dense 3D-reconstruction using the previous lab pipeline (rectification, graph-cut for disparity, triangulation using projection matrices, which were obtained using this lab results). A little «steppy», but shows right change in depth (floor is closer, house is in the middle and background is deeper). Steps could be better smoothed when changing parameters of graph-cut and using full-scale image (was used window size 20 on the 0.5 scale). As we have grayscale image, we assigned the same grayscale values for every RGB component.