Report: Structure from Motion

Rishabh Singh: 19-953-793

November 22, 2019

1 Feature extraction and initialization with epipolar geometry

Our objective is to develop a pipeline of structure from motion. We load the first and last images of the given object. This was done because these two images have large baseline as well as sufficiently many overlapping feature matches. SIFT features were extracted in both the images. We filter some features to avoid choosing from top of the image as the top portion consists of only background. The filtered features were matched as shown in Fig.1:

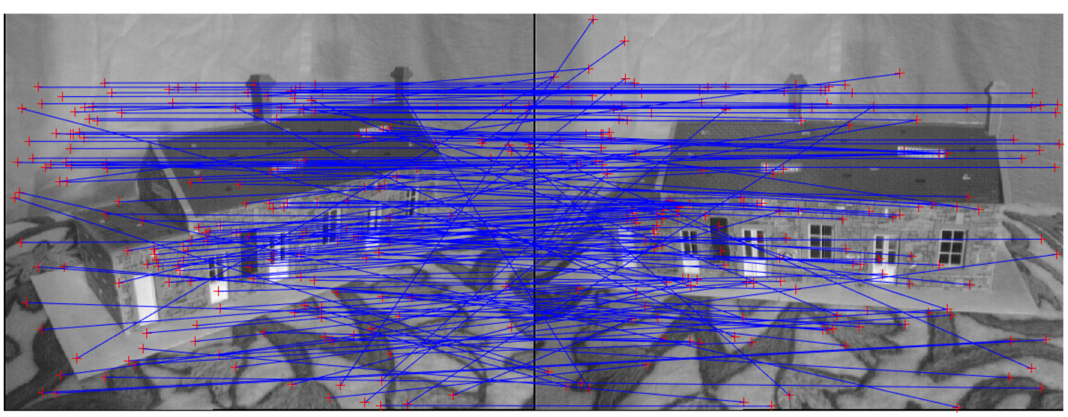


Figure 1: SIFT matches between first and last images

The matched points were obtained in both the images. These points were converted into homogeneous coordinates. These points were converted into calibrated points by taking into account the intrinsic matrix provided K for the images. We estimated the essential matrix for the image configuration using these calibrated matched points using 8-point RANSAC. The inliers and outliers obtained after this step are shown in Fig.2 and 3:

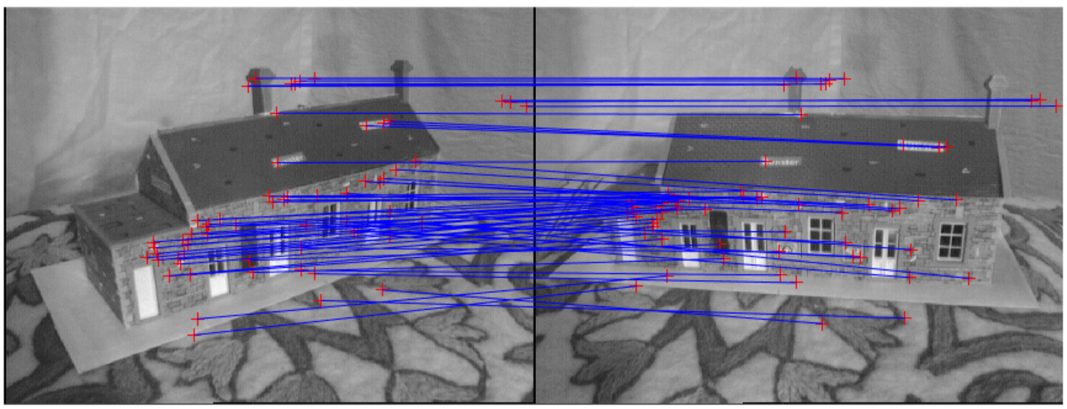


Figure 2: Task 1: Inlier matches after 8-point RANSAC (images 0 and 4).

The projection matrix for the first view is assumed to be $[I|0]$. The essential matrix obtained is decomposed into rotation and translation matrices to obtain the projection matrix of the second view. We fix the projection matrix if its determinant is negative. We triangulated the inliers in both the images using the projection matrices thus obtained to find the corresponding 3D points.

2 Triangulation and adding new views

The second image was added as an additional view. SIFT features were obtained for this image. These were matched against features from the first image that were triangulated. These matches were used to find the projection matrix for the added view from the image and 3D points using 6-point DLT algorithm. The wrong matches were filtered using the 6-point RANSAC algorithm. Similar to the above steps, new 3D points were obtained by triangulating the image points in the given pair of images. This whole process is repeated for

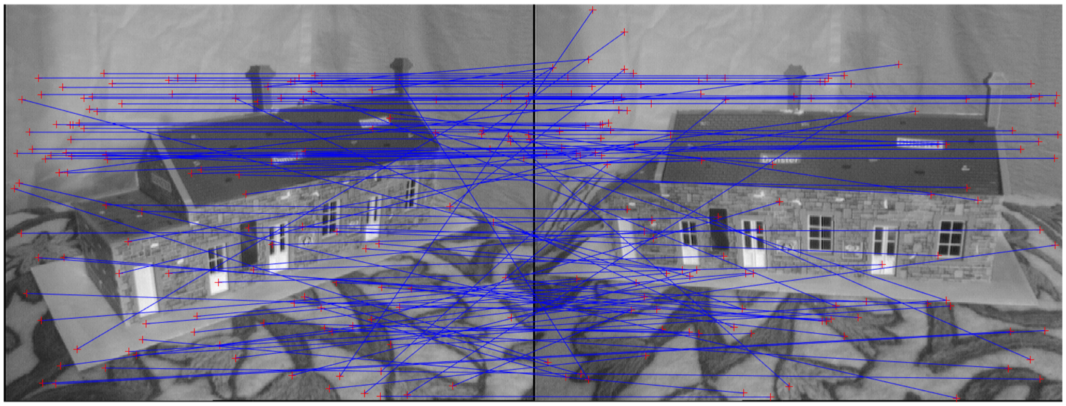
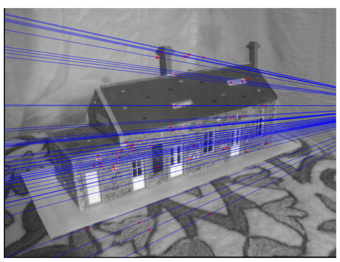
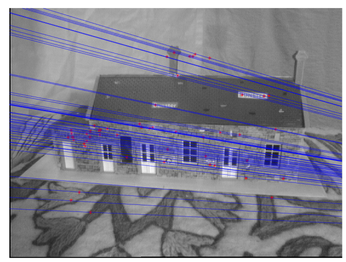


Figure 3: Task 1: Outlier matches after 8-point RANSAC (images 0 and 4).



(a) image 0



(b) image 4

Figure 4: Task 1: Epipolar geometry of the image pair used for initialization. The epipolar lines in each view are shown in blue.

the other two views added. The inlier and outlier matches after each added view is shown in Fig 5-10.

3 Plotting

The triangulated 3D points from all the inliers matches in each view are plotted in Fig.6 and 7. Camera poses for every view is also visualised in the same figure.

4 Dense Reconstruction

Images 0 and 1 were chosen for dense reconstruction since no outliers were found in their matching. SIFT matches were obtained in both images. These matches were used to estimate the fundamental matrix of the camera pair. 8-point RANSAC was used to eliminate the outlier matches. These images were then rectified based on the inlier points obtained

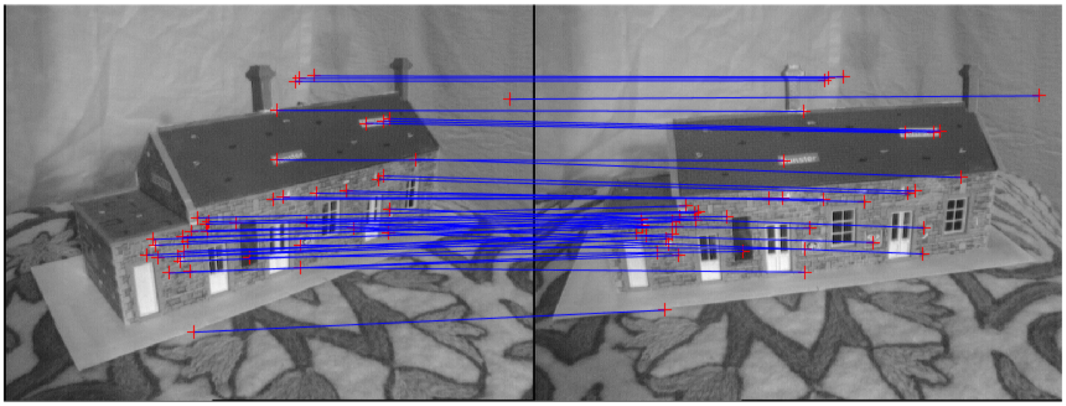


Figure 5: Task 2: Inlier matches after 6-point RANSAC (images 0 and 2)

in the previous step. The disparity map was obtained for each image using the rectified outputs. Winner takes all algorithm with a window size of 41 was used to compute these disparity maps. The obtained result is shown in Fig.13. Both the options of creating a dense reconstruction were utilised. The results obtained are shown in Fig 14-19.

5 Discussion

- No outliers were detected when images 0 and 1 were matched as these have very similar scenes.
- If the baseline is larger, the triangulation of the feature points is better.
- The result from plotting of 3D plotting is not as expected. Additional optimization might be required to obtain a better sparse reconstruction.
- Dense reconstruction obtained after creating a mesh of the scene is imperfect near depth discontinuities.
- Dense reconstruction obtained by grayscale coloring of the 3D points resembles the given image better.

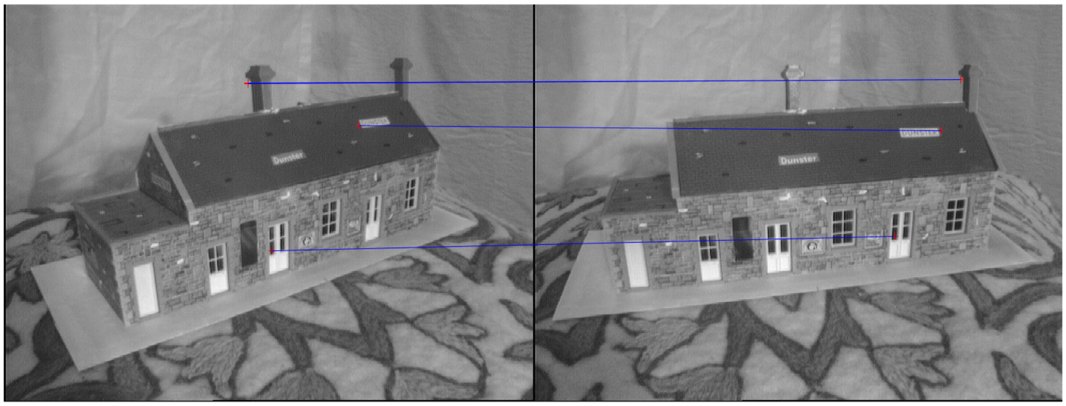


Figure 6: Task 2: Outlier matches after 6-point RANSAC (images 0 and 2)

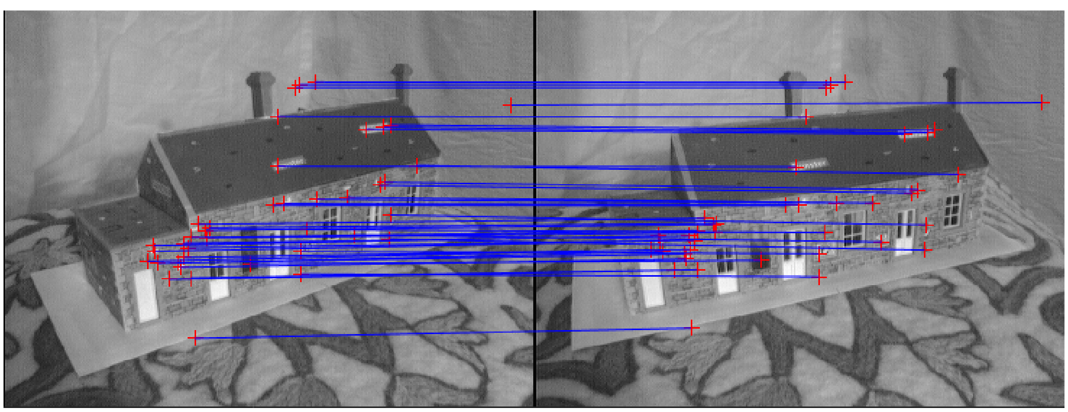


Figure 7: Task 2: Inlier matches after 6-point RANSAC (images 0 and 1)

Figure 8: Task 2: Outlier matches after 6-point RANSAC (images 0 and 1)

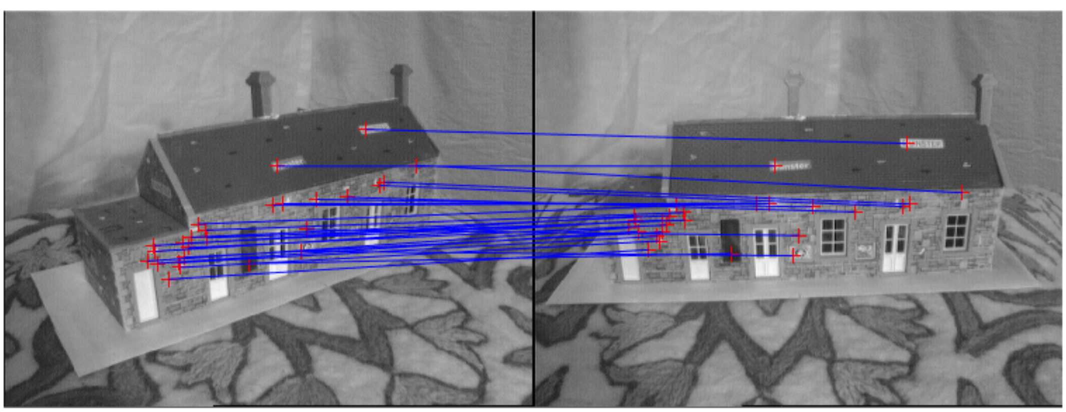


Figure 9: Task 2: Inlier matches after 6-point RANSAC (images 0 and 3)

Figure 10: Task 2: Outlier matches after 6-point RANSAC (images 0 and 3)

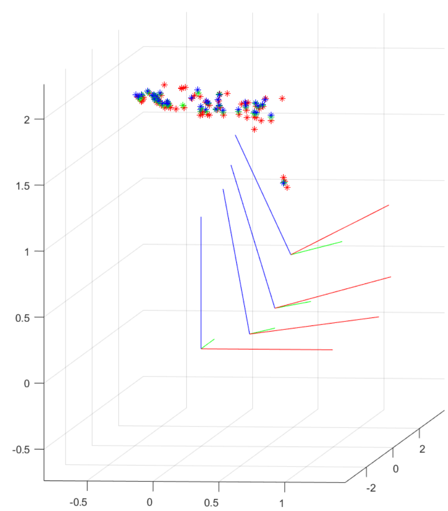


Figure 11: Task 3: Triangulated points and camera poses plotted in 3D. Red dots are the triangulated points from the initialization, green dots are the triangulated points from the first additional image and blue dots are from the second additional image.

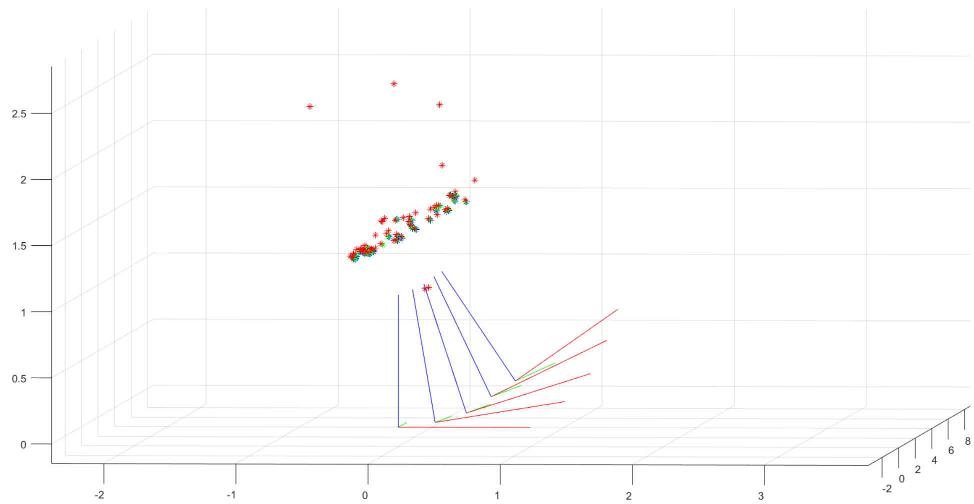


Figure 12: Task 3: Triangulated points and camera poses plotted in 3D. Red dots are the triangulated points from the initialization, green dots are the triangulated points from the first additional image, blue from the second and cyan from the third additional image.

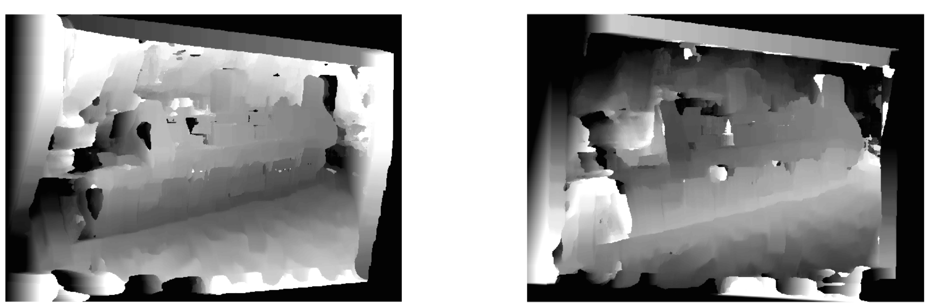


Figure 13: Task 4: Disparity map for images 0 and 1.

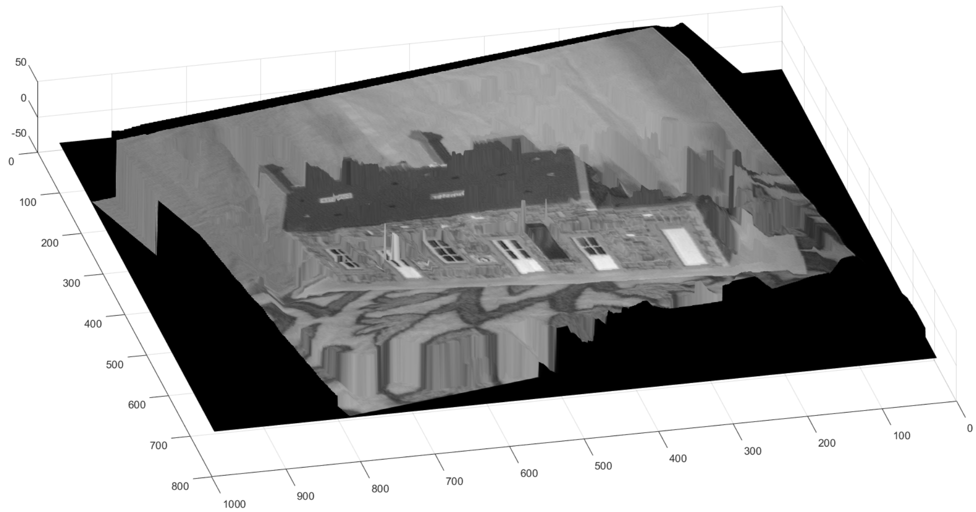


Figure 14: Task 4: Dense reconstruction of the scene using grayscale coloring of 3D points

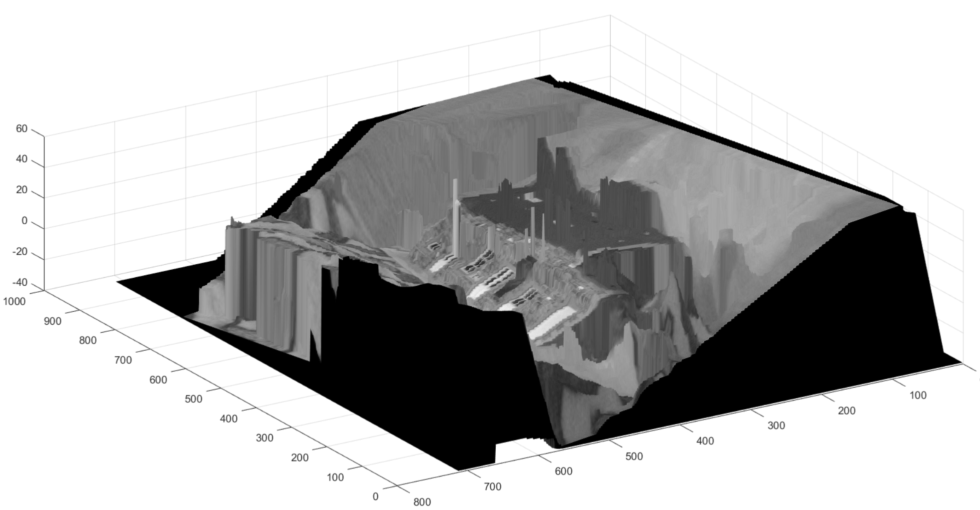


Figure 15: Task 4: Dense reconstruction of the scene using grayscale coloring of 3D points(same as above, from a different viewpoint)

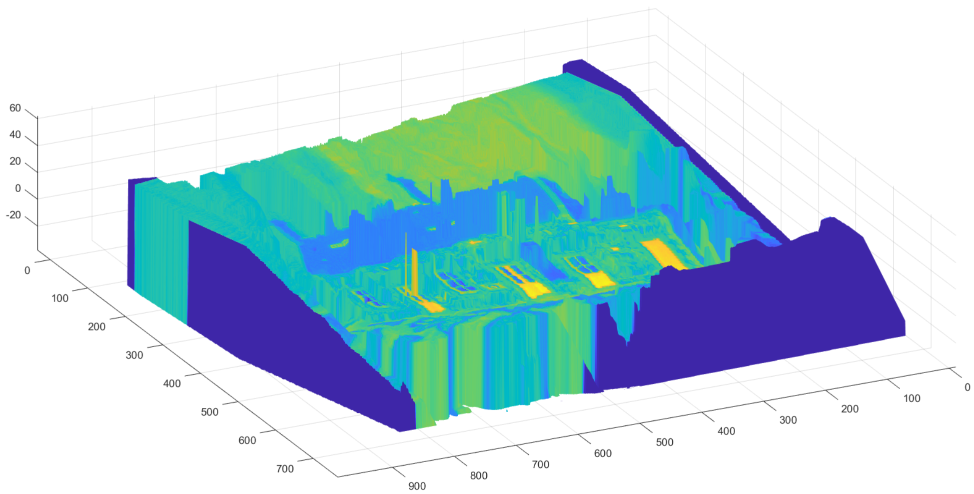


Figure 16: Task 4: Dense reconstruction of the scene using RGB coloring of 3D points

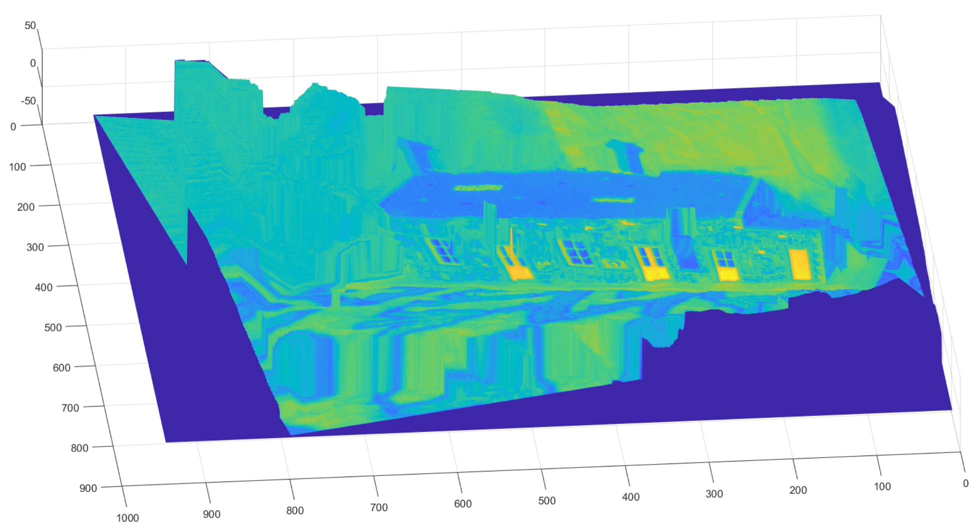


Figure 17: Task 4: Dense reconstruction of the scene using RGB coloring of 3D points (same as above, from a different viewpoint)

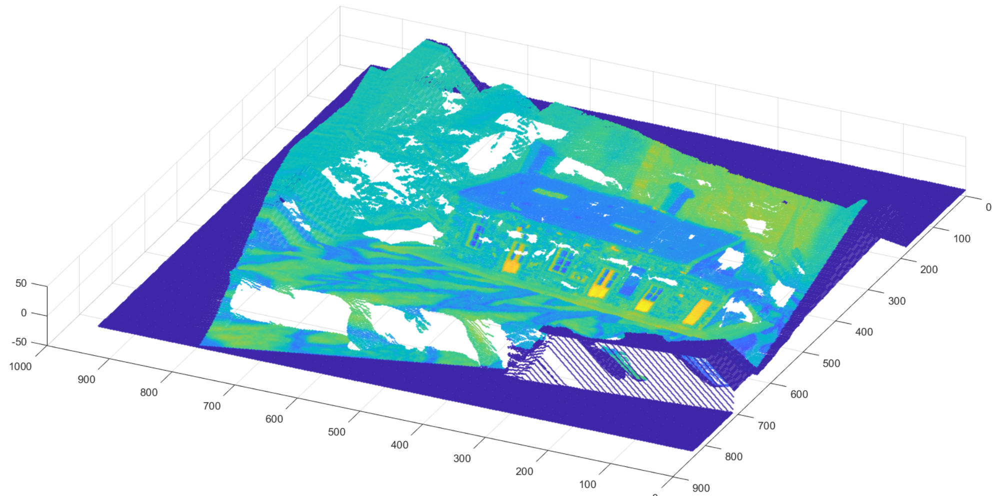


Figure 18: Task 4: Dense reconstruction of the scene using scatter plot mesh

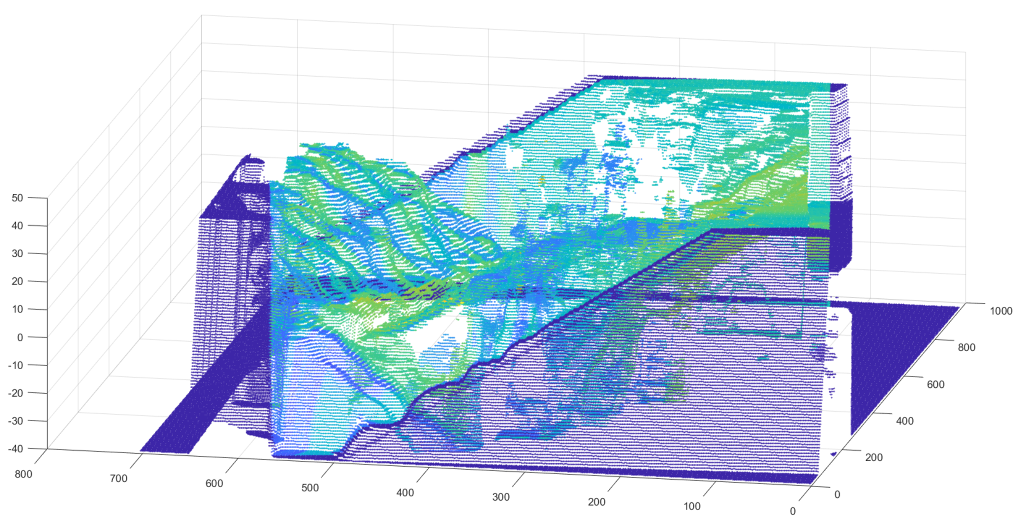


Figure 19: Task 4: Dense reconstruction of the scene using scatter plot mesh (same as above, from a different viewpoint)