CS364/AM792: Homework #4

Due on 7 May, 2020 at 17:00pm

Unathi K. Skosana

Problem 1

(a)

The removal of outliers was done with the threshold distance of 200 pixels. It is better here to use such a lenient threshold here, since the steps that follow involve overdetermined systems that approximate the various quantities of interest in the least-squares sense.

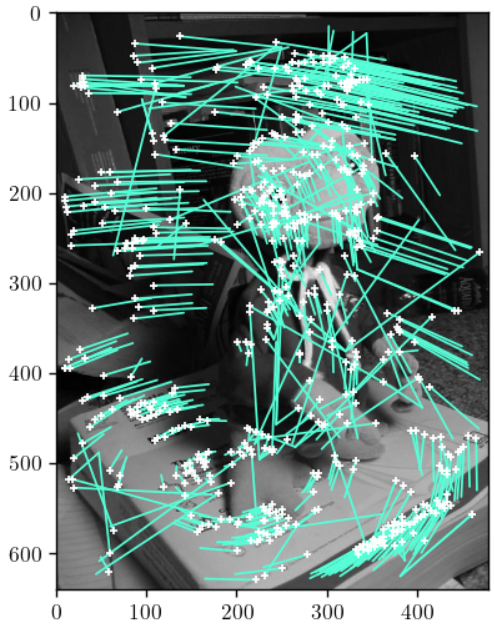


Figure 1: Filtered SIFT features

(b)

Following the RANSAC procedure to estimate F with a Simpson distance threshold of 1 pixel for $10N$ iterations where N is the number of matches. The size of the consensus set produced was 418 and the fundamental matrix F estimated from the consensus set is given below

$$F = \begin{bmatrix} -2.48079775\text{e}{-6} & -1.42955448\text{e}{-5} & 6.24619756\text{e}{-3} \\ 4.57085149\text{e}{-6} & -1.86170164\text{e}{-6} & 3.59836273\text{e}{-2} \\ -7.34248289\text{e}{-3} & -3.07400834\text{e}{-2} & -9.98832968\text{e}{-1} \end{bmatrix}$$

The image below shows the consensus set of matches

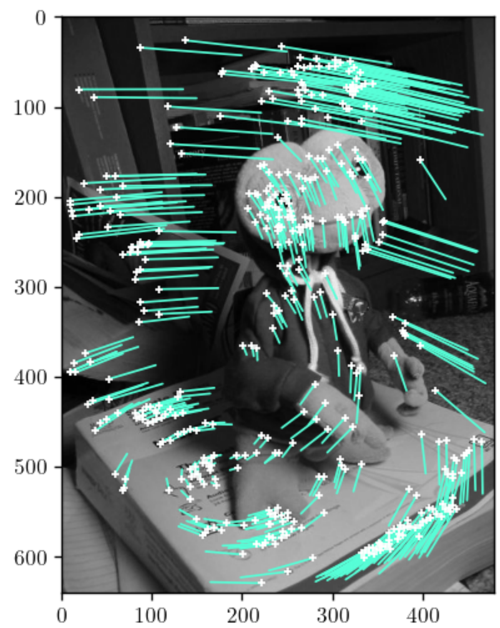


Figure 2: Consensus set SIFT matches

(c)

The essential matrix can be determined from the fundamental matrix and the calibration matrices of the cameras in a straight forward manner. Below are the singular values of the essential matrix in descending order.

$$\sigma_1 = 2.50637074\text{e}{+}1$$
$$\sigma_2 = 2.50226239\text{e}{+}1$$
$$\sigma_3 = 3.11422341\text{e}{-}15$$

and the recomputed essential matrix via the fix on handout is

$$E = \begin{bmatrix} 1.13914812 & 6.6127527 & -0.72360584 \\ -2.11435877 & 0.86752977 & -24.90670533 \\ 4.38710103 & 23.73844459 & 0.66624 \end{bmatrix}$$

Problem 1 continued on next page...

(d)

Using the calibration matrix provided for both cameras. The first camera matrix P_1 is easily determined as

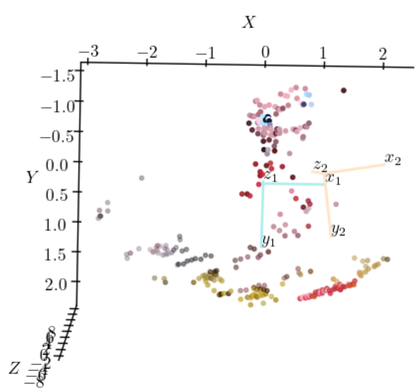
$$P_1 = \begin{bmatrix} 677.6328 & 0.0 & 240.5 & 0.0 \\ 0.0 & 682.6328 & 320.5 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \end{bmatrix}$$

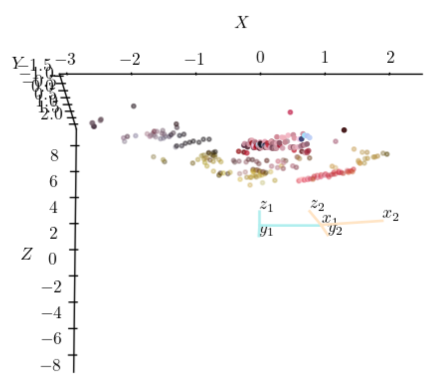
A match was randomly sampled from the consensus set to triangulate its 3-space correspondence, then this triangulated 3-space point was used to figure out which pair of camera matrices gives correct geometry. From this, P_2 was determined as

$$P_2 = \begin{bmatrix} 6.15208698\text{e}{+}2 & -9.88571941\text{e}{+}1 & 3.58848121\text{e}{+}2 & -5.88388075\text{e}{+}2 \\ 4.16584571\text{e}{+}1 & 6.76883118\text{e}{+}2 & 3.29850886\text{e}{+}2 & 1.09546197e + 2 \\ -1.81762243\text{e}{-}1 & 5.49667247\text{e}{-}3 & 9.83327145\text{e}{-}1 & 2.66992956\text{e}{-}1 \end{bmatrix}$$

(e)

Using the camera matrices determined above, we can triangulate all the matches in consensus set to produce a 3D reconstruction of the observed scene. Points with a z coordinate further than 9 were removed from the reconstruction. These points can be inaccurate due to a number of reasons i.e errors/noise match identification but the perhaps the most relevant here is the following: If a point in 3-space is very far away from the camera centres then the angle between the rays from the two camera centres pointing towards the point be quite narrow. In such a scenario the two rays will almost appear parallel (no intersection) and the SVD will have no exact solution for $A\mathbf{X} = \mathbf{0}$, thus produce an error-prone least-squares solution.





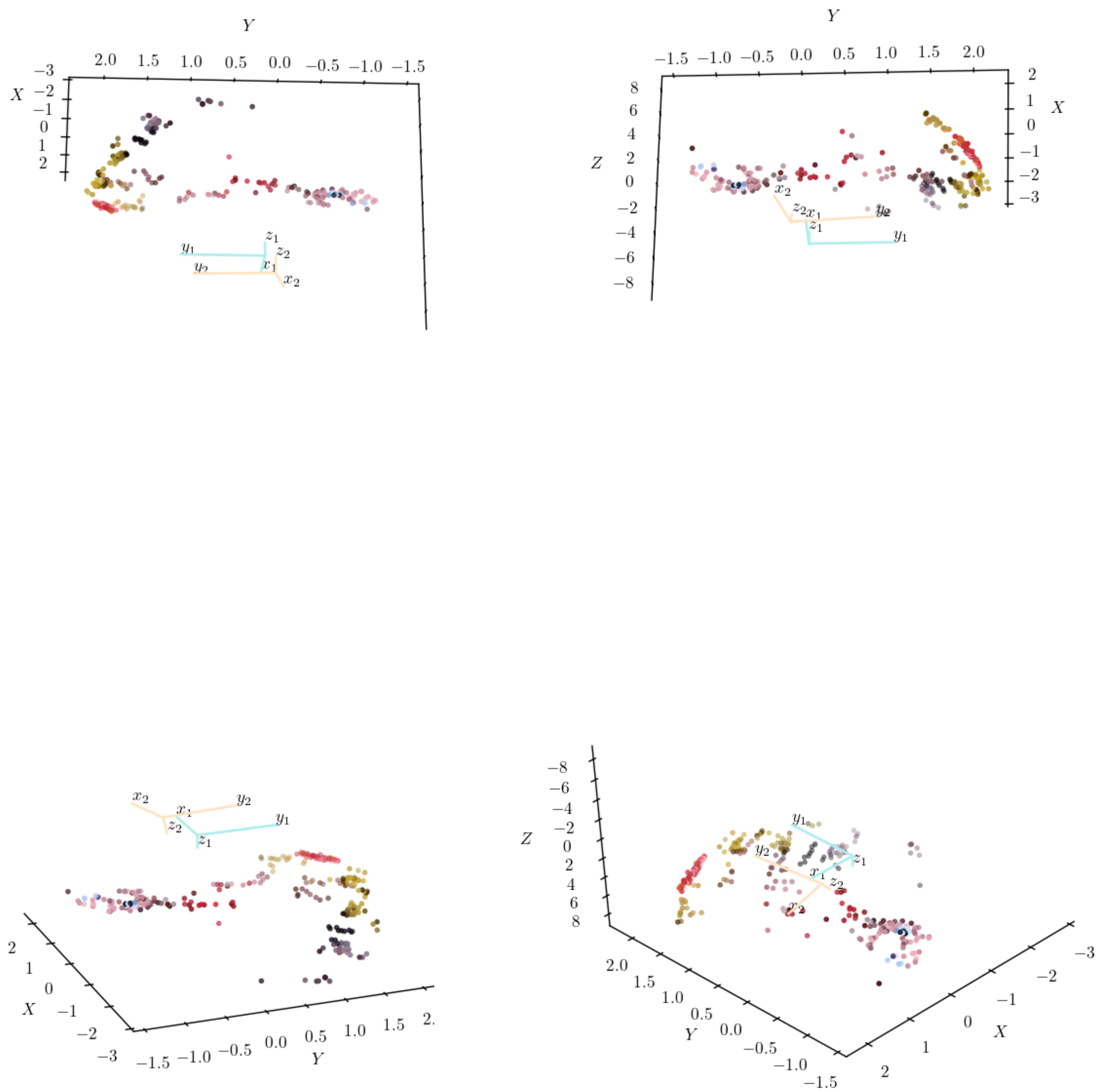


Figure 4: Different views of a figure that slightly resembles ET

Problem 2

(a)

Using the method already described to construct the transformation matrices T_1, T_2

$$T_1 = \begin{bmatrix} 1.01254555\text{e}{+}0 & -1.81277806\text{e}{-}1 & -9.53879428\text{e}{+}0 \\ 2.22458298\text{e}{-}1 & 9.76306317\text{e}{-}1 & -4.73204094\text{e}{+}1 \\ 1.36222826\text{e}{-}4 & -2.43882116\text{e}{-}5 & 9.70646051\text{e}{-}1 \end{bmatrix} \tag{1}$$

$$T_2 = \begin{bmatrix} 1.05843644\text{e}{+}0 & -2.66125883\text{e}{-}2 & -1.98319624\text{e}{+}2 \\ 1.55922518\text{e}{-}1 & 1.00836582\text{e}{+}0 & -7.13240170\text{e}{+}1 \\ 4.01122752\text{e}{-}4 & 2.84646210\text{e}{-}5 & 8.56560707\text{e}{-}1 \end{bmatrix} \tag{2}$$

we can warp the original images to produce

Figure 5: Warped ETs.

(b)

To construct epipolar lines we have to follow the same procedure as in assignment 3 with the newly constructed P_1', P_2'

$$P_1' = K_n R_n[\mathbb{I}| - \tilde{\mathbf{C}}_1] \tag{3}$$
$$P_2' = K_n R_n[\mathbb{I}| - \tilde{\mathbf{C}}_2] \tag{4}$$

The new fundamental matrix of the warped images is

$$F' = \begin{bmatrix} -2.46748640\text{e}{-6} & -1.43590530\text{e}{-5} & 6.19579186\text{e}{-3} \\ 4.73499398\text{e}{-6} & -1.96973184\text{e}{-6} & 3.55413979\text{e}{-2} \\ -7.25952489\text{e}{-3} & -3.02770122\text{e}{-2} & -9.98863866\text{e}{-1} \end{bmatrix} \tag{5}$$

From this we can thus produce the epipolar lines in the usual way with a minor modification of offsetting the calculated y of epipolar lines by miny used in from apply_homography. The reason why the top-left corner of a rectified image may no longer coincide with the origin of the image coordinate system is that the homographies obtained may translation components (first two entries of third column are non zero) while the cameras retain their orientations in the world coordinate system and positions.

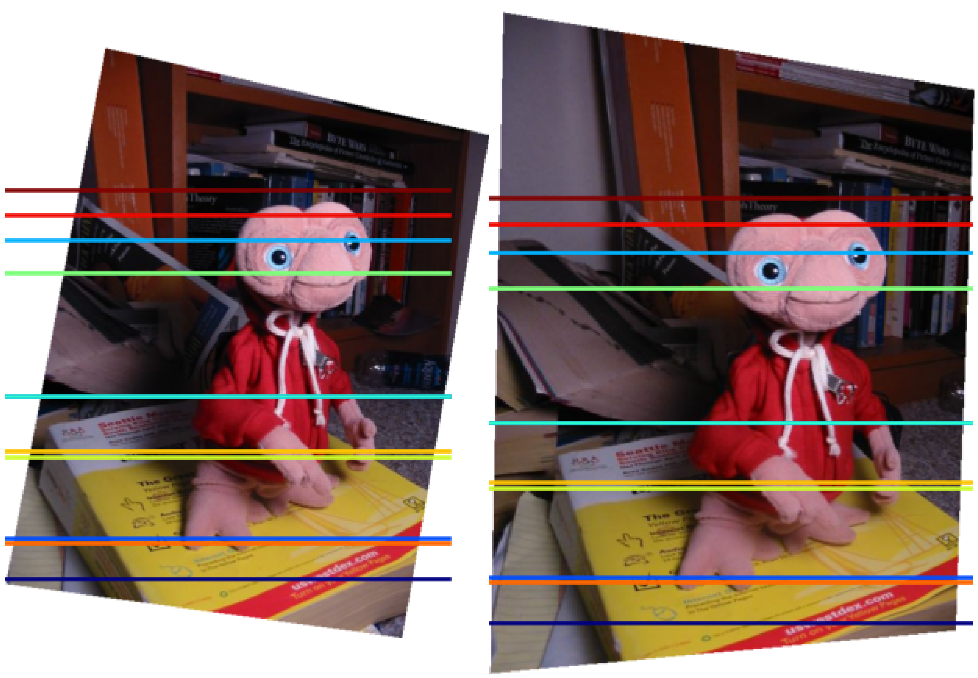


Figure 6: Epipolar lines