Lund University

Center for Mathematical Sciences

FMAN95

Computer Vision

Assignment 2

Author:
Anush Ghazayran

Instructor:
Viktor Larsson

VT24

Submitted: March 15, 2024

2 Calibrated vs. Uncalibrated Reconstruction

Exercise 1

In **Structure from Motion Problem**, given image projections x_{ij} (of scene point j in image i) we want to determine both $3D$ point coordinates X_j and camera matrices P_i such that

$$\lambda_{ij} x_{ij} = P_i X_j, \forall i, j.$$

If the calibration is unknown, that is P_i can be any non-zero 3x4 matrix then the solution to this problem is called a **projective reconstruction**. Such a solution can only be uniquely determined up to a projective transformation. To see this suppose that we have found cameras P_i and $3D$-points X_j such that

$$\lambda_{ij} x_{ij} = P_i X_j.$$

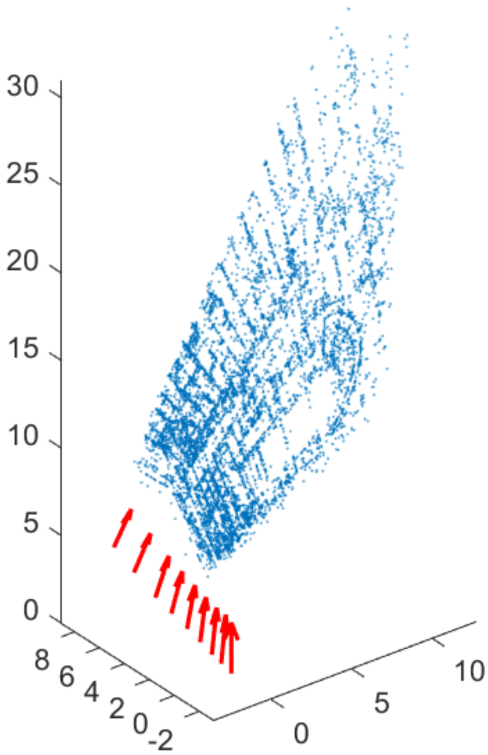
To construct a different solution we can take an unknown projective transformation T, and let $\widetilde{X}_j = TX_j$ and $\widetilde{P}_i = P_i T^{-1}$. The new cameras and scene points also solve the problem since

$$\lambda_{ij} x_{ij} = P_i X_j = P_i T^{-1} T X_j = \widetilde{P}_i \widetilde{X}_j$$

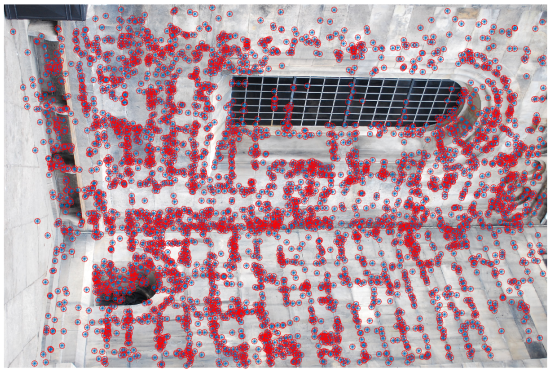
This means that given a solution we can apply any projective transformation to the $3D$ points and obtain a new solution.

Computer Exercise 1

The $3D$ reconstruction and the nine cameras in Figure 1: (a) does not look reasonable as it does not preserve angles and distances of the real figure.
For projecting $3D$ points I chose the third camera. From the second plot in Figure 1: (b), we can see that the projections appear to be close to the corresponding image points (as the blue stars are close to red circles).
Figure 2: shows the $3D$ reconstructions using T_1 and T_2 projective transformations. As we can see they are different than the one in Figure 1: (a), however, the cameras are still pointing to the $3D$ model, but, the angles are different now. Fugure 2: (b) shows a better reconstruction of the 3D figure as the angles look more accurate.
Lastly, from Figure 3: we can see that the projections of new $3D$ points are the same as before, since the camera matrices have also changed and the overall product is the same as before, as we saw in Exercise 1.

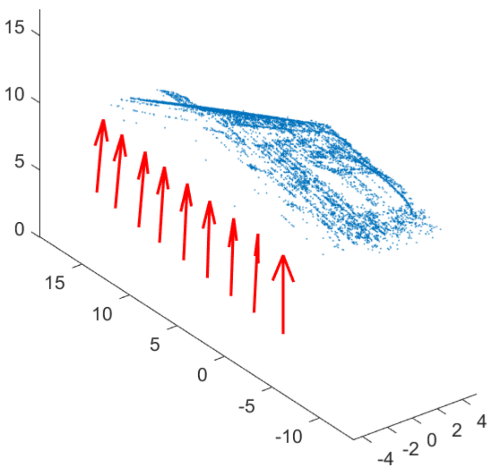


(a) 3D reconstructions.

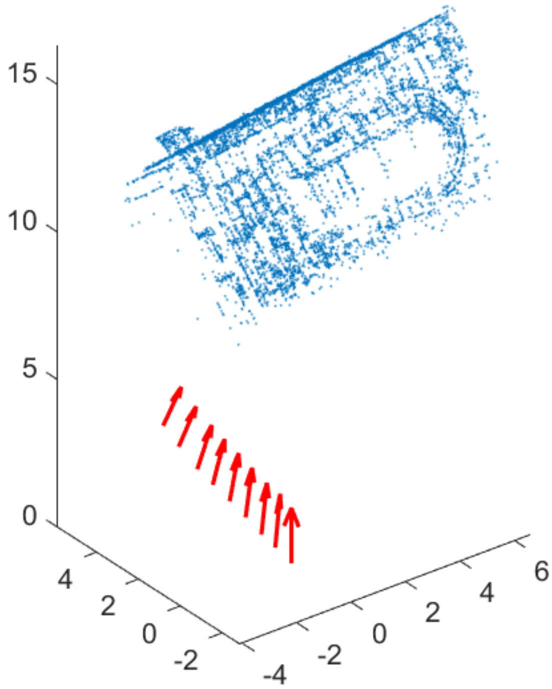


(b) Projection of 3D points.

Figure 1: 3D points of the reconstruction.

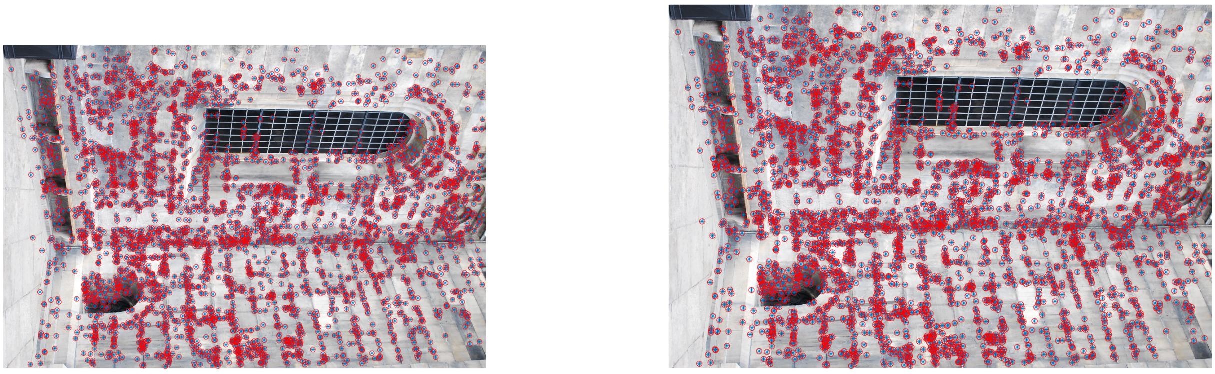


(a) Using T_1 transformation.



(b) Using T_2 transformation.

Figure 2: New 3D points and cameras.



(a) Using T_1 transformation.

(b) Using T_2 transformation.

Figure 3: Projection of the new 3D points.

Exercise 2

The solution of this problem is called a **Euclidean Reconstruction**. Given a solution $[R_i\ t_i]$ and X_j we can try to do the same trick as in the projective case. However when multiplying $[R_i\ t_i]$ with H the result does not necessarily have a rotation matrix in the first 3×3 block. To achieve a valid solution we need H to be a similarity transformation,

$$H = \begin{bmatrix} sQ & v \\ 0 & 1 \end{bmatrix},$$

where Q is a rotation matrix.

3 Camera Calibration

Exercise 3

First, let's verify the inverse of K,

$$KK^{-1} = \begin{pmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/f & 0 & -x_0/f \\ 0 & 1/f & -y_0/f \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

The matrix can be further factorized into

$$K^{-1} = \begin{pmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{pmatrix} = AB$$

Here the transformation A scales the coordinates by $1/f$, and B translates the coordinates

by $\left(-x_0, -y_0\right)$.
To normalize the image points we multiply both sides of the camera equation by K^{-1}. If we let $\tilde{x} = K^{-1}x$, we get

$$\lambda\tilde{x} = KK^{-1}[R \quad t]X = [R \quad t]X$$

The interpretation of this operation is to move from image space to camera space.
By this, we get a new camera matrix $[R \quad t]$ which is **normalized** (calibrated) **camera**. After normalization, the principal point $\left(x_0, y_0\right)$ ends up at the origin and the points with distance f get a new distance of 1 in the normalized image.
To normalize the points we need to multiply by $K^{-1} = \begin{pmatrix} 1/320 & 0 & -1 \\ 0 & 1/320 & -3/4 \\ 0 & 0 & 1 \end{pmatrix}$.
We get

$$\begin{pmatrix} 1/320 & 0 & -1 \\ 0 & 1/320 & -3/4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 240 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1/320 & 0 & -1 \\ 0 & 1/320 & -3/4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 640 \\ 240 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

For getting the view rays of the points we need to subtract the camera center from the points. Since the center of the image is at the origin, after normalization we will get $a = \left(-1, 0, 1\right)$ and $b = \left(1, 0, 1\right)$ viewing rays projecting the points. The angle between the viewing rays projecting to these points is $\theta = \cos^{-1}\left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|}\right) = 90°$.
Let's show that both $K[R\ t]$ and $[R\ t]$ have the same camera centers and principal axis. The camera center is the origin of the camera coordinate system. If C is a vector containing the coordinates of the camera center in the world coordinate system we will have

$$0 = K(RC + t) \Leftrightarrow 0 = K^{-1}K(RC + t) \Leftrightarrow C = -R^{-1}t = -R^T t \quad \text{(where } K \text{ is invertible).}$$

Which is the same as in the normalized version.
The viewing direction can be represented by the vector that goes from the camera center to the point $\left(0, 0, 1\right)$ in the camera coordinate system. In normalized camera, if X contains

the world coordinates of this point we have

$$\begin{pmatrix}0\\0\\1\end{pmatrix} = RX + t \Leftrightarrow X = R^T \begin{pmatrix}0\\0\\1\end{pmatrix} - R^T t = R_3^T + C$$

where R_3 is the third row of R. The viewing direction in the world coordinate system is therefore $R_3^T + C - C = R_3^T$. Since the K matrix doesn't change the last row of the R, it would be the same result for $K[Rt]$ camera.

Exercise 4

$$K = \begin{pmatrix}1000 & 0 & 500\\0 & 1000 & 500\\0 & 0 & 1\end{pmatrix}, \quad K^{-1} = \begin{pmatrix}1/1000 & 0 & -1/2\\0 & 1/1000 & -1/2\\0 & 0 & 1\end{pmatrix}.$$

Therefore,

$$\begin{aligned}K^{-1}P &= \begin{pmatrix}1/1000 & 0 & -1/2\\0 & 1/1000 & -1/2\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}1000 & -250 & 250\sqrt{3} & 500\\0 & 500(\sqrt{3}-\frac{1}{2}) & 500(1+\frac{\sqrt{3}}{2}) & 500\\0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1\end{pmatrix}\\&= \begin{pmatrix}1 & 0 & 0 & 0\\0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0\\0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1\end{pmatrix}.\end{aligned}$$

To normalize the points we get

$$\begin{pmatrix}1/1000 & 0 & -1/2\\0 & 1/1000 & -1/2\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}-1/2\\-1/2\\1\end{pmatrix},$$

$$\begin{pmatrix}1/1000 & 0 & -1/2\\0 & 1/1000 & -1/2\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\1000\\1\end{pmatrix} = \begin{pmatrix}-1/2\\1/2\\1\end{pmatrix},$$

$$\begin{pmatrix}1/1000 & 0 & -1/2\\0 & 1/1000 & -1/2\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}1000\\0\\1\end{pmatrix} = \begin{pmatrix}1/2\\-1/2\\1\end{pmatrix},$$

$$\begin{pmatrix} 1/1000 & 0 & -1/2 \\ 0 & 1/1000 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1000 \\ 1000 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} 1/1000 & 0 & -1/2 \\ 0 & 1/1000 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 500 \\ 500 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

4 RQ Factorization and Computation of K

Exercise 5

$$KR = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix} \begin{pmatrix} R_1^\top \\ R_2^\top \\ R_3^\top \end{pmatrix} = \begin{pmatrix} aR_1^\top + bR_2^\top + cR_3^\top \\ dR_2^\top + eR_3^\top \\ fR_3^\top \end{pmatrix}.$$

Since the matrix R is orthogonal R_3 has to have a length 1. Therefore, $f = \|A_3\|$ and $R_3 = \frac{1}{\|A_3\|} A_3$

In this case $f = \| \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \| = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1$ and $R_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

$A_2 = dR_2 + eR_3$ tells us that A_2 is a linear combination of two orthogonal vectors (both of length one). Hence, the coefficient e can be computed from the scalar product

$$e = A_2^\top R_3 = \begin{pmatrix} -\frac{700}{\sqrt{2}} & 1400 & \frac{700}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 700.$$

When e is known we can compute R_2 and d from

$$dR_2 = A_2 - eR_3 = \begin{pmatrix} -\frac{700}{\sqrt{2}} \\ 1400 \\ \frac{700}{\sqrt{2}} \end{pmatrix} - 700 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 1400 \\ 0 \end{pmatrix}.$$

$$d = 1400, R_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Similarly, $A_1 = aR_1 + bR_2 + cR_3$.

$$b = R_2^\top A_1 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{800}{\sqrt{2}} \\ 0 \\ \frac{2400}{\sqrt{2}} \end{pmatrix} = 0.$$

$$c = R_3^\top A_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{800}{\sqrt{2}} \\ 0 \\ \frac{2400}{\sqrt{2}} \end{pmatrix} = 800.$$

$$aR_1 = A_1 - bR_2 - cR_3 = \begin{pmatrix} \frac{800}{\sqrt{2}} \\ 0 \\ \frac{2400}{\sqrt{2}} \end{pmatrix} - 800 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1600}{\sqrt{2}} \\ 0 \\ \frac{1600}{\sqrt{2}} \end{pmatrix}.$$

$$a = 1600, R_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}.$$

We got $K = \begin{pmatrix} 1600 & 0 & 800 \\ 0 & 1400 & 700 \\ 0 & 0 & 1 \end{pmatrix}$, hence the focal length is 1400, skew is 0, aspect ratio is 8/7, and finally, the principal point is $\begin{pmatrix} 800, 700 \end{pmatrix}$.

Computer Exercise 2

$K1 = \begin{pmatrix} 2391.48 & 225.66 & 942.71 \\ 0 & 623.87 & 813.60 \\ 0 & 0 & 1 \end{pmatrix}$,

$K2 = \begin{pmatrix} 2393.95 & 0 & 932.38 \\ 0 & 2398.12 & 628.26 \\ 0 & 0 & 1 \end{pmatrix}$

They do not represent the same transformation as the matrices are not equal.

5 Direct Linear Transformation DLT

Exercise 7

If we multiply both sides of the last equation (17) with N we will get

$$Nx \sim NPX,$$

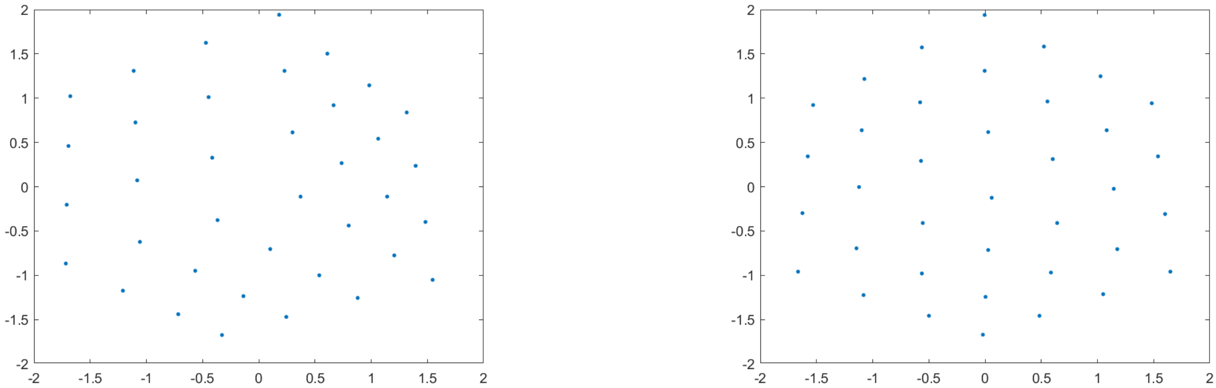
where $Nx \sim \tilde{x}$ (15), therefore from equation (16)

$$\tilde{x} \sim \tilde{P}X \sim NPX.$$

We get that $\tilde{P} \sim NP$, and after multiplying by the inverse matrix we get $P \sim N^{-1}\tilde{P}$.

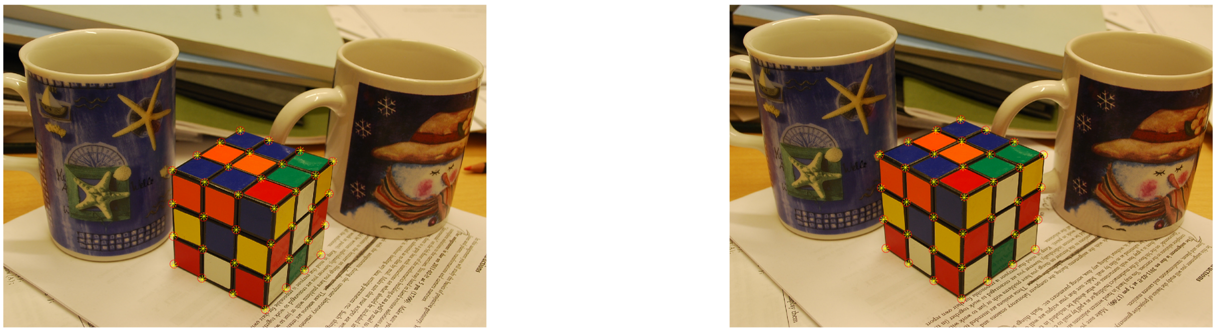
Computer Exercise 3

As you can see from Figure 4: the points are centered around $\big(0,0\big)$ with mean distance 1 to $\big(0,0\big)$. From Figure 5: you can see the camera center and the projected points from two solutions we got.



(a) For image 1. (b) For image 2.

Figure 4: Normalized measured points.



(a) For image 1. (b) For image 2.

Figure 5: Model and measured points.

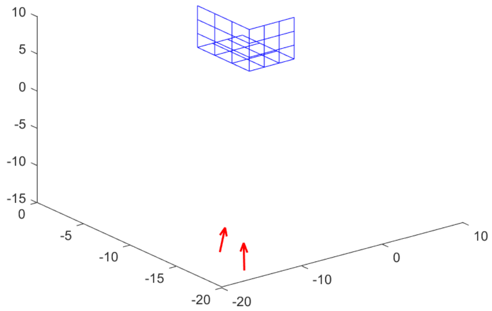


Figure 6: Model points with camera center

For the inner parameters for the first camera I got

$$K1 = \begin{pmatrix} 2448.60 & -18.09 & 959.80 \\ 0 & 2446.85 & 675.89 \\ 0 & 0 & 1 \end{pmatrix}.$$

From the matrix we can see that we get the "true" parameters, as it has the form $K = \begin{pmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix}$, where f is the focal length and x_0 and y_0 are the center of the image.
The skew is close to zero as $s = \frac{-18.09}{2446.85} \approx 0$.
There is no ambiguity as we estimate the projection using given $3D$ points and $2D$ projections. For Exercise 1, we use projection equation to estimate the projection matrix and $3D$ points at the same time and there is no definite solution, since we can apply any projective transformation to the $3D$ points and obtain a new solution.

6 Feature Extraction and Matching using SIFT

Computer Exercise 4

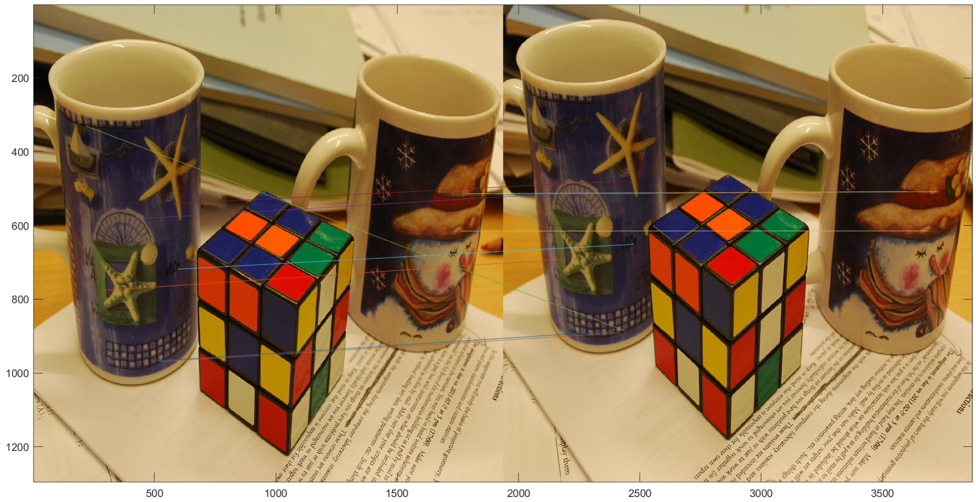


Figure 7: Matching points between images

7 Triangulation using DLT

Computer Exercise 5

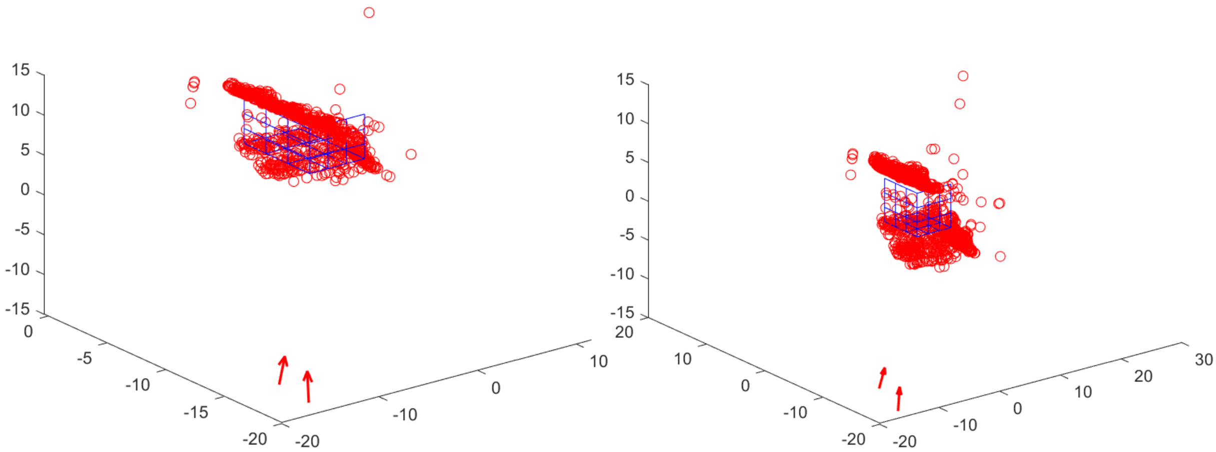
(a) For image 1.

(b) For image 2.

Figure 8: Comparing projected points to SIFT points Un-normalized case.

(a) For image 1. (b) For image 2.

Figure 9: Comparing projected points to SIFT points Normalized case.



(a) Un-normalized case. (b) Normalized case.

Figure 10: 3D plots for Computer Exercise 5