Assignment 2

Computer Vision
Calibration and DLT

Devosmita Chatterjee
Name of the collaborator: Ara Jafarzadeh

2 Calibrated vs. Uncalibrated Reconstruction.

Exercise 1.

The camera equations are given by

$$\lambda x = PX$$

where $x = [x_1, x_2, 1]^T$, $X = [X_1, X_2, X_3, 1]^T$ and P is a 3×4 matrix.
Then for any projective transformation $\tilde{X} = TX$, the above equations can be written as

$$\lambda x = PT^{-1}TX = \tilde{P}\tilde{X}$$

where $\tilde{P} = PT^{-1}$ is a valid camera.
Therefore, if X is the estimated 3D-points, then a new solution can always be obtained from TX for any projective transformation T of 3D space.

Computer Exercise 1.

See m-file-compEx1.m
Yes, figure. 1 looks like a reasonable reconstruction
Yes, the projections appear to be close to the corresponding image points in figure. 2.
First projection appear reasonable shown in figure. 3.
No, the projections are not changed in figure. 4.

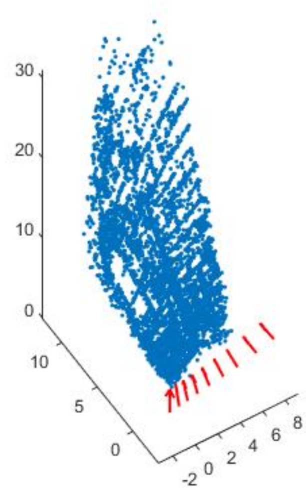

Figure. 1

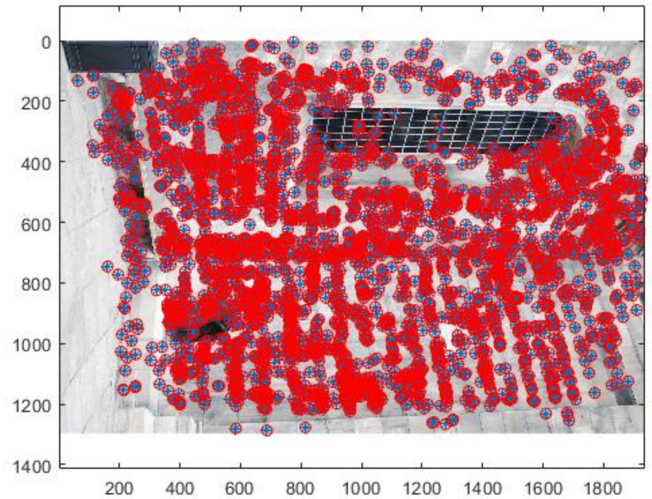

Figure. 2

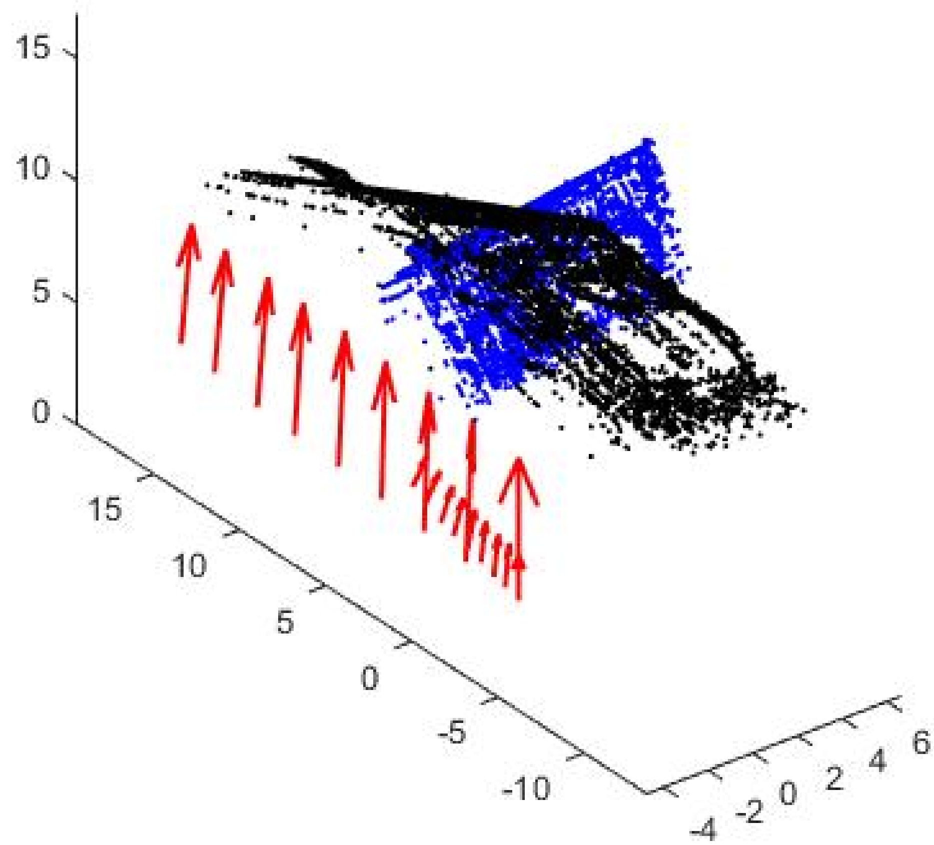

Figure. 3

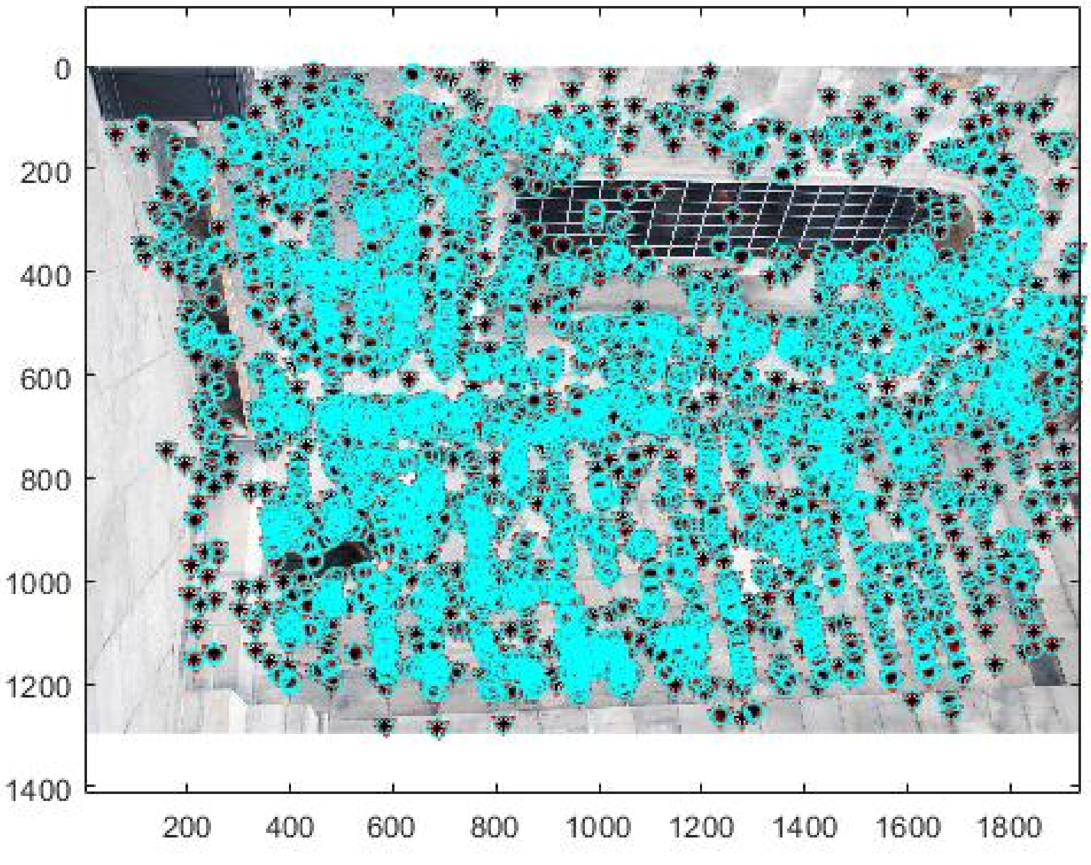

Figure. 4

Exercise 2.

We can not get the same projective distortions as in Computer Exercise 1 when we use calibrated cameras because by multiplying the camera matrix $K[R|t]$ by K^{-1}, we get the camera matrix $[R|t]$ where R is a similarity matrix.
Under the assumption of calibrated cameras, there is always similarity transformation.
For calibrated cameras, the similarity ambiguity is the only ambiguity.

3 Camera Calibration

Exercise 3.

To find the inverse of $K = \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}$

Now, $\begin{matrix} k_{11} = f & k_{12} = 0 & k_{13} = 0 \\ k_{21} = 0 & k_{22} = f & k_{23} = 0 \\ k_{31} = -fx_0 & k_{32} = -fy_0 & k_{33} = f^2 \end{matrix}$

Also, $\left|K\right| = f^2$

Then, $K^{-1} = \frac{1}{|K|}\begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix}^T = \frac{1}{f^2}\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ -fx_0 & -fy_0 & f^2 \end{bmatrix}^T = \frac{1}{f^2}\begin{bmatrix} f & 0 & -fx_0 \\ 0 & f & -fy_0 \\ 0 & 0 & f^2 \end{bmatrix} =$
$\begin{bmatrix} \frac{1}{f} & 0 & -\frac{x_0}{f} \\ 0 & \frac{1}{f} & -\frac{y_0}{f} \\ 0 & 0 & 1 \end{bmatrix}$

Now, $AB = \begin{bmatrix} \frac{1}{f} & 0 & 0 \\ 0 & \frac{1}{f} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix} = K^{-1}$

Thus, the matrix K^{-1} can be factorized into AB
$K^{-1} = AB$.
The geometric interpretation of the transformation A is that it scales by a factor $\frac{1}{f}$.
The geometric interpretation of the transformation B is that it shifts to the principal point (x_0, y_0).
When normalizing the image points of a camera with known inner parameters we apply the transformation K^{-1}. The interpretation of this operation is to transform the image coordinates with respect to the camera matrix.
The principal point (x_0, y_0) end up $(0, 0)$.
A point with distance f to the principal point end up to unit distance from $(0, 0)$.

$K = \begin{bmatrix} 320 & 0 & 320 \\ 0 & 320 & 240 \\ 0 & 0 & 1 \end{bmatrix}$

$f = 320, x_0 = 320, y_0 = 240$

$K^{-1} = \begin{bmatrix} \frac{1}{f} & 0 & -\frac{x_0}{f} \\ 0 & \frac{1}{f} & -\frac{y_0}{f} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{320} & 0 & -1 \\ 0 & \frac{1}{320} & -\frac{3}{4} \\ 0 & 0 & 1 \end{bmatrix}$

$x1_{normalized} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$x2_{normalized} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

The angle between the viewing rays $a = (-1, 0, 1)$ and $b = (1, 0, 1)$ projecting to these points is given by $a.b = ||a||||b||cos(\theta) \Rightarrow cos(\theta) = \frac{a.b}{||a||||b||} = \frac{(-1,0,1).(1,0,1)}{\sqrt{2}\sqrt{2}} = 0 \Rightarrow \theta = 90^0$.
To show that the camera $P_1 = K[R|t]$ and the corresponding normalized version $P_2 = [R|t]$ have the same camera center and principal axis.

$\bar{0} = P_1\begin{bmatrix} C_1 \\ 1 \end{bmatrix} \Rightarrow \bar{0} = K\begin{bmatrix} R & t \end{bmatrix}\begin{bmatrix} C_1 \\ 1 \end{bmatrix} \Rightarrow K(RC_1 + t) = \bar{0} \Rightarrow RC_1 + t = K^{-1}\bar{0} = \bar{0}$
$\Rightarrow C_1 = -R^{-1}t$

$V_1 = \begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \end{bmatrix}$

$\bar{0} = P_2\begin{bmatrix} C_2 \\ 1 \end{bmatrix} \Rightarrow \bar{0} = \begin{bmatrix} R & t \end{bmatrix}\begin{bmatrix} C_2 \\ 1 \end{bmatrix} \Rightarrow RC_2 + t = \bar{0} \Rightarrow C_2 = -R^{-1}t$

$V_2 = \begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \end{bmatrix}$

4 RQ Factorization and Computation of K

Exercise 4.

$$P = [A, a] = \begin{bmatrix} \frac{800}{\sqrt{2}} & 0 & \frac{2400}{\sqrt{2}} & 4000 \\ -\frac{700}{\sqrt{2}} & 1400 & \frac{700}{\sqrt{2}} & 4900 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 3 \end{bmatrix}$$

$$A = \begin{bmatrix} A_1^T \\ A_2^T \\ A_3^T \end{bmatrix} = \begin{bmatrix} \frac{800}{\sqrt{2}} & 0 & \frac{2400}{\sqrt{2}} \\ -\frac{700}{\sqrt{2}} & 1400 & \frac{700}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = KR = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix} = \begin{bmatrix} aR_1^T + bR_2^T + cR_3^T \\ dR_2^T + eR_3^T \\ fR_3^T \end{bmatrix}$$

Using $A_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, $||R_3|| = 1$ and the third row of the camera matrix $A_3^T = fR_3^T$,

$f = ||fR_3^T|| = ||A_3^T|| = \sqrt{(-\frac{1}{\sqrt{2}})^2 + (0)^2 + (\frac{1}{\sqrt{2}})^2} = 1$,

$\therefore R_3 = \frac{1}{f} A_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.

Using $A_2 = \begin{bmatrix} -\frac{700}{\sqrt{2}} \\ 1400 \\ \frac{700}{\sqrt{2}} \end{bmatrix}$, $R_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $e = A_2^T R_3$,

e = 700.

Using the second row of the camera matrix $A_2 = dR_2 + eR_3$,

$dR_2 = A_2 - eR_3 = \begin{bmatrix} 0 \\ 1400 \\ 0 \end{bmatrix}$.

Using $||R_2|| = 1$, $d = ||dR_2^T|| = 1400$.

$\therefore R_2 = \frac{1}{d} \begin{bmatrix} 0 \\ 1400 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Using $b = A_1^T R_2$,

$b = 0$.

Using $c = A_1^T R_3$,

$c = 800$.

Using the first row of the camera matrix $A_1^T = aR_1^T + bR_2^T + cR_3^T$, $aR_1^T = A_1^T - bR_2^T - cR_3^T = [\frac{1000}{\sqrt{2}}, 0, \frac{1000}{\sqrt{2}}]$

Using $||R_1|| = 1$, $a = ||aR_1^T|| = 1600$.

$\therefore R_1 = \frac{1}{a} \begin{bmatrix} \frac{1600}{\sqrt{2}} \\ 0 \\ \frac{1600}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

$$R = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$K = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} 1600 & 0 & 800 \\ 0 & 1400 & 700 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \gamma f & sf & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Focal length $f = 1400$

Skew $s = 0$
Aspect ratio $\gamma = 1.14$
Principal point of the camera $(x_0, y_0) = (800, 700)$

Computer Exercise 2.

See m-file-compEx1.m
$K = \begin{bmatrix} 2394 & 0 & 932.4 \\ 0 & 2398.1 & 628.3 \\ 0 & 0 & 1 \end{bmatrix}$
$K1 = \begin{bmatrix} 2394 & 0 & 932.4 \\ 0 & 599.5 & 628.3 \\ 0 & 0 & 1 \end{bmatrix}$
$K2 = \begin{bmatrix} 2394 & 0 & 932.4 \\ 0 & 2398.1 & 628.3 \\ 0 & 0 & 1 \end{bmatrix}$
They represent the same transformations with different aspect ratios.

5 Direct Linear Transformation DLT

Exercise 5. (OPTIONAL.)

Norm of a vector can never be negative. It is always positive or zero. Thus the minimum value of $||Mv||^2$ is always zero. Therefore, the linear least squares system $min_v ||Mv||^2$ always has the minimum value 0.
If M has the singular value decomposition $M = U \sum V^T$, then U & V are orthogonal matrices $(UU^T = U^T U = I)$ & $(VV^T = V^T V = I)$ and
$||Mv||^2 = (Mv)^T (Mv) = v^T M^T M v = v^T (U \sum V^T)^T (U \sum V^T) v = v^T (V \sum^T U^T)(U \sum V^T) v$
$= v^T V \sum^T \sum V^T v$= $(\sum V^T v)^T (\sum V^T v)$ = $|| \sum V^T v||^2$.
Again if $||v||^2 = 1$, $||V^T v||^2 = ||V^T||^2 ||v||^2 = 1.1 = 1$.
The new problem gives the same minimal value as the previous problem because we choose the smallest eigenvector for both which is the same.
The second solution is $\tilde{v} = V^T v$. Since V is orthogonal, $v = V\tilde{v}$.
There is always at least two solutions to this problem since both v & -v satisfies the problem.
If rank(M) $<$ n, that is, the last singular value σ_n is zero which means the eigenvalue selected will be zero, then there is an exact nonzero solution to $Mv = 0$.

Exercise 6.

$\tilde{x} \sim Nx$ and $\tilde{x} \sim \tilde{P}X$ gives $Nx \sim \tilde{P}X \Rightarrow x \sim N^{-1}\tilde{P}X \Rightarrow P = N^{-1}\tilde{P}$

Computer Exercise 3.

See m-file-compEx3.m
Yes, it looks like the points are centered around (0, 0) with mean distance 1 to (0, 0) in figure. 5.

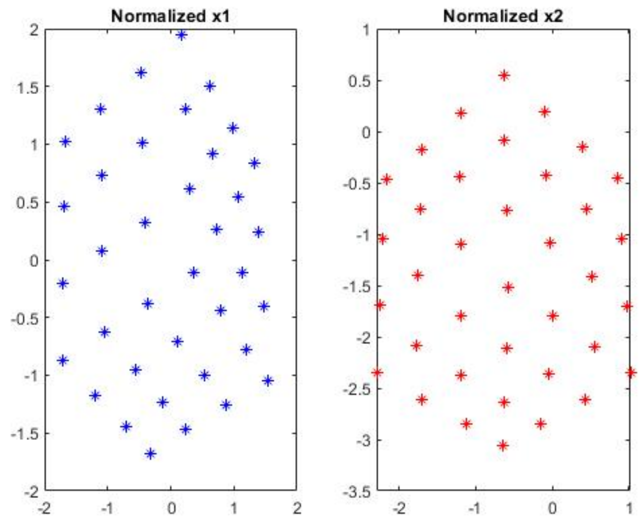

Figure. 5

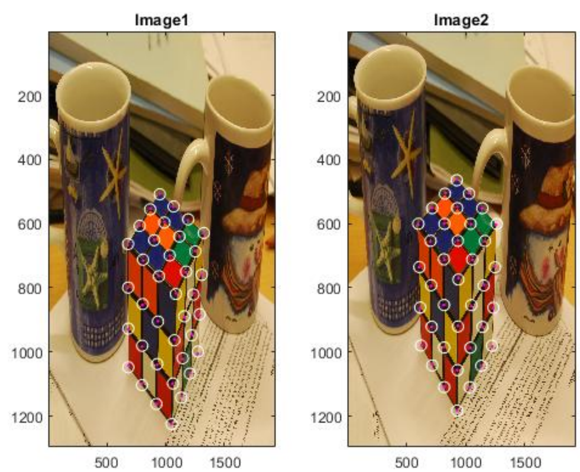

Figure. 6

Smallest singular value is close to zero. Same for $||Mv||$.
Yes, the result seen in figure. 6 seems reasonable since the projected points and image points are close to each other.
The inner parameters of the cameras are

$$K_1 = \begin{bmatrix} 2448.1 & -18 & 960.1 \\ 0 & 2446.4 & 675.7 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_2 = \begin{bmatrix} 2448.1 & -18.0 & 960.1 \\ 0 & 2446.4 & 675.7 \\ 0 & 0 & 1 \end{bmatrix}$$

There is no ambiguity because we do not know the transformations.

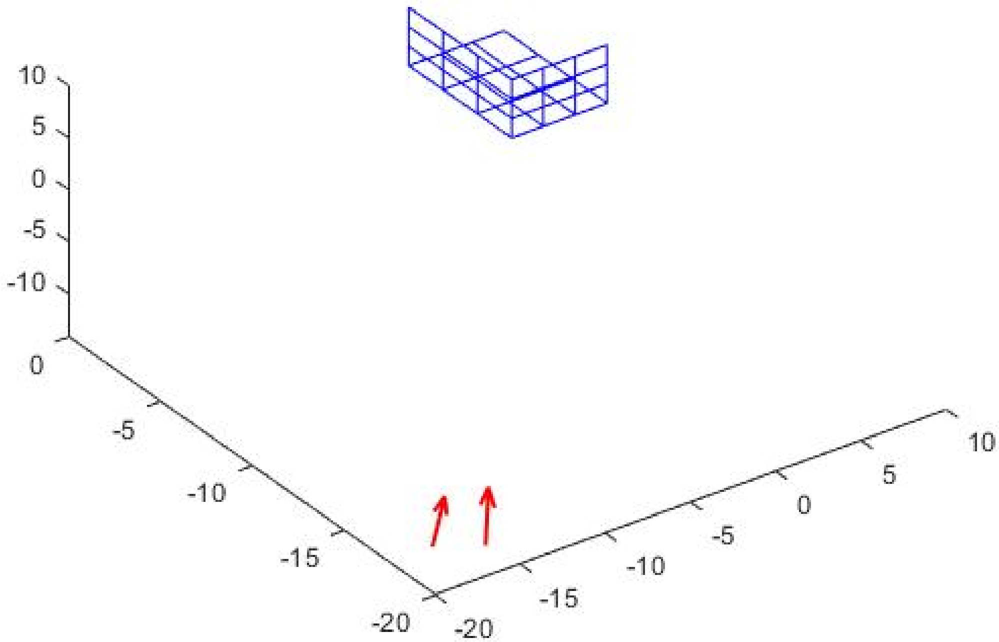

Figure. 7

6 Feature Extraction and Matching using SIFT

Computer Exercise 4.

See m-file-compEx5.m

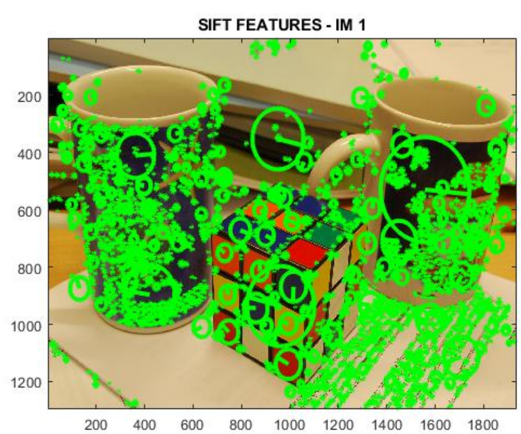

Figure. 8

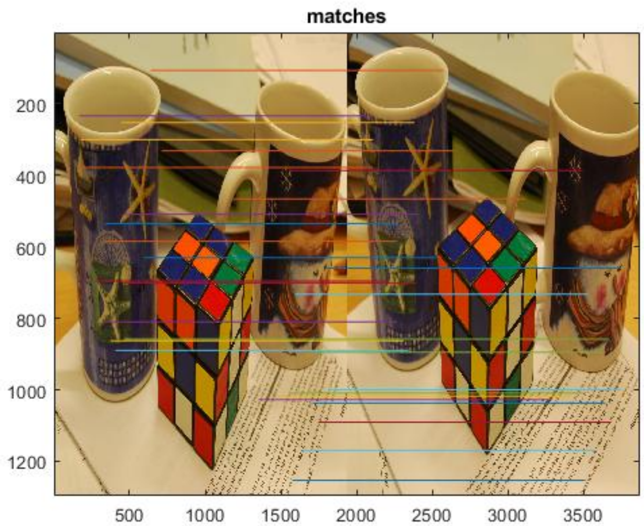

Figure. 9

7 Triangulation using DLT

Computer Exercise 5.

See m-file-compEx5.m
The SIFT projections is shown in figure. 10.
The filtered model is shown in figure. 11.

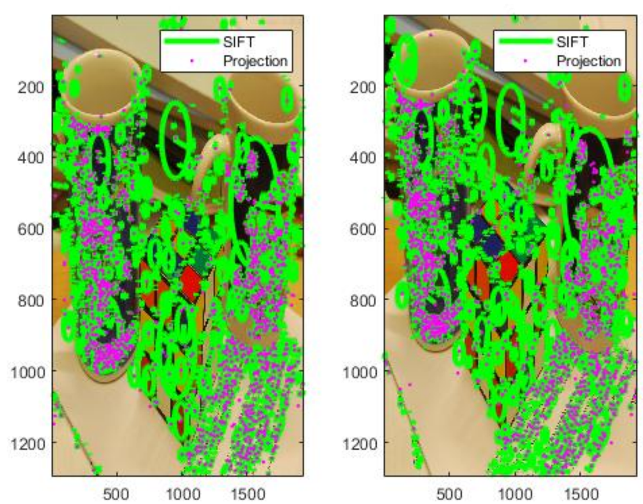

Figure. 10

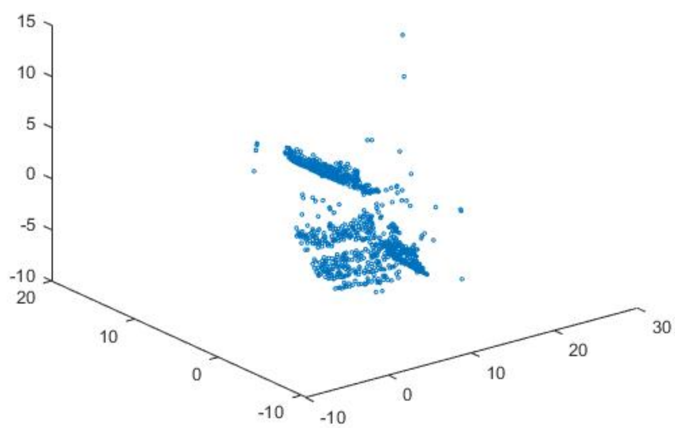

Figure. 11