Datorseende Assignment 2

er8251da-s

January 2024

1 Introduction

Solution to the exercises in assignment two of spring term course in Computer Vision 2024.

2 Calibrated vs. Un-calibrated Reconstruction

2.1 Exercise 1

With projective reconstruction, equal accuracy is achieved regardless of any transformations applied to the camera equation.

$$\hat{X} = H^{-1}X \rightarrow \lambda\hat{X} = PHH^{-1}X = \hat{P}\hat{X}$$

Because $\hat{P} = PH$ is also a valid camera, transformations do not affect the accuracy of our projections in un-calibrated camera projections.

2.2 Computer Exercise 1

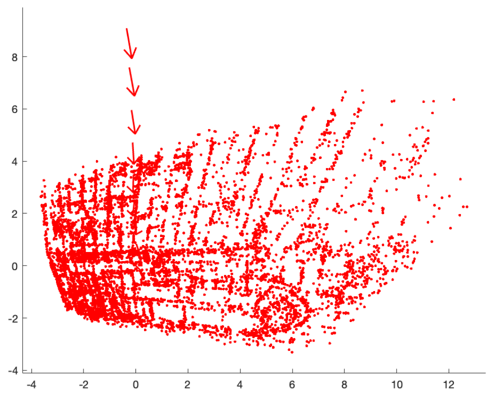

Figure 1: Projection from Camera 1

The construction looks reasonable.

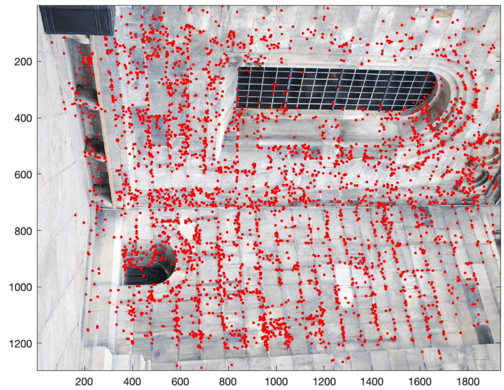

Figure 2: Projection from Camera 1 with Image

Yes, the projection points seem to match corresponding points in the image.

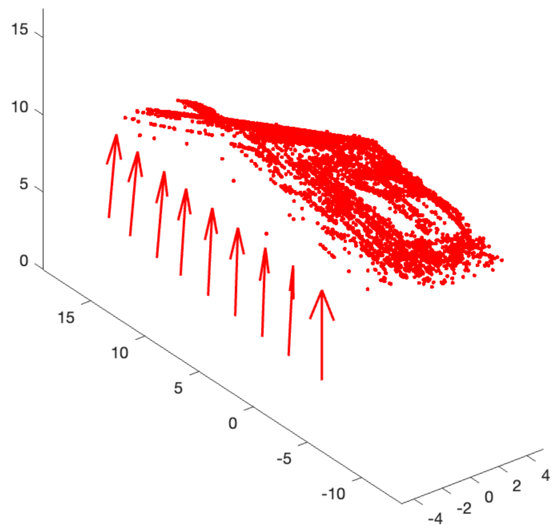

Figure 3: Projection with Transformation 1

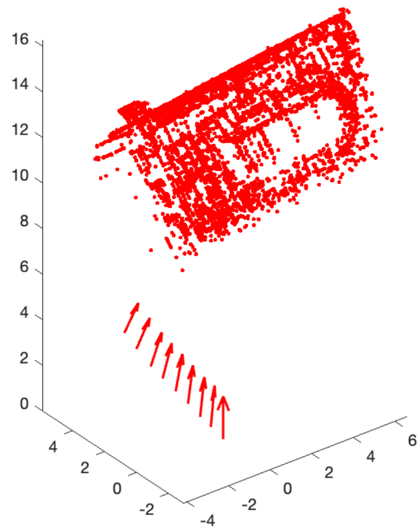

Figure 4: Projection with Transformation 2

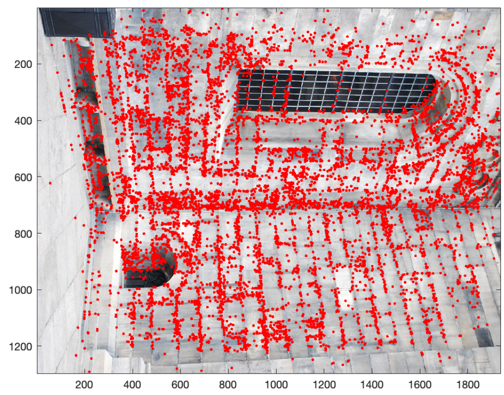

Figure 5: Projection with Transformation 1 with image

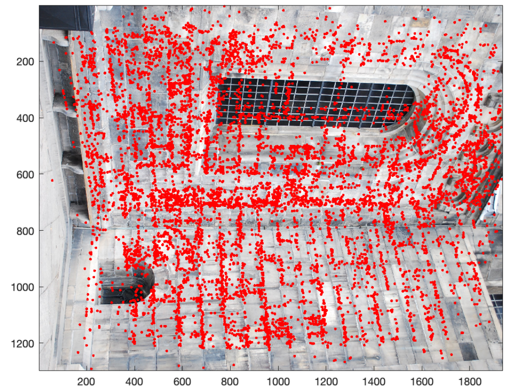

Figure 6: Projection with Transformation 2with image

As we can tell from figures 3 and 4, the projections are initially distorted but that does not affect the reconstruction in figure 5 and 6.

2.3 Exercise 2

In the case of euclidean projections, T needs to be a similarity matrix to achieve a valid solution. In euclidean projections we try to solve:

$$\lambda_{ij} x'_{ij} = [R_i \quad t_i] X_j$$

When applying a transformation to the rotational part of [Ri ti], a valid solution is only preserved if T is a similarity matrix as:

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

where Q is a rotation. Multiplication with T yields:

$$T[R_i \quad t_i] X_j = s(R_i Q + t_i X_j)$$

The solution is valid only since $R_i Q$ is a rotational matrix.

In essence, since euclidean projections are euclidean transformations from 3D to 2D points, a transformation T can only have valid solutions as long as T itself is a euclidean transformation. Euclidean transformations are scaling, rotation, or mirroring which none of distort projections. These euclidean transformations preserve angles and lines.

3 Camera Calibration

3.1 Exercise 3

For an upper triangular 3x3 matrix, the inverse has:

1. Diagonal elements as the reciprocals

2. Other elements such that $AA^{-1} = I$

Which yields:

$$K^{-1} * K = \begin{bmatrix} 1/f * f & 0 & -x_0/f * f + x_0 \\ 0 & 1/f * f & -y_0/f * f + y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

We verify A and B by multiplication:

$$AB = \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} = K$$

A re-scales by a factor $1/f$ and B shifts the principal point by (x_0, y_0)

When normalising with K^{-1} the effects are:

- Principal point is shifted to (0,0).
- Image is re-scaled by $\frac{1}{f}$, meaning that a point f away from the principal point will be 1 unit distance away from the principal point.

The given points will be scaled to fit the normalised image. In K the principal point is (320, 240), and image size of (640,480) and corners in (0,0), (640,0), (640,480), and (0,480). Since the principal point is centred at (0,0) and the range of coordinates are [-1,1], the new points will be:

- $(0, 240) \rightarrow (-1, 0)$
- $(640, 240) \rightarrow (1, 0)$

We can calculate the angle of viewing rays between the two points (in homogeneous coordinates) by examining the dot product:

$$(-1, 0, 1) \cdot (1, 0, 1) = 0$$

and we get $a \cdot b = abcos(\theta) \rightarrow cos(\theta) = 0 \rightarrow \theta = \pm 90°$

The angle between the two viewing rays are 90°.

Lastly the camera $K[R\ t]$ and the normalised $[R\ t]$ have the same camera centre and principal axis since K^{-1} cancels out the K from $K[R\ t]$. Left is [R t] which will have the same nullspace as $K[Rt]$.

Intuitively, we can say that in the formulation $K[R\ t]$, K holds the inner camera parameters while R and t contain its spatial information. While the inner parameters of K affect its view of the world, it should not affect the nullspace (being an effect of the camera's location).

Because the last row of the K-matrix is also the principal axis of the camera the last row will be preserved.

3.2 Exercise 4

If

$$N = \begin{bmatrix} s & 0 & -sx_i \\ 0 & s & -sy_i \\ 0 & 0 & 1 \end{bmatrix}$$

where s is a scaling factor, and x_i and y_i represent shifts of the focal point in P_2, to normalise, s is set to the inverse of the focal point (i.e 1/1000). With an s of 1/1000 the coordinates will be shifted by $-\frac{x_i}{1000}$ and $-\frac{y_i}{1000}$. Since the focal point of a normalised camera should be (0,0), x_i and y_i will both be 500.

The normalised camera is calculated by:

$$\hat{P} = NP = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1000 & -250 & 250\sqrt{3} & 500 \\ 0 & 500(\sqrt{3}-\frac{1}{2}) & 500(1+\frac{\sqrt{3}}{2}) & 500 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix}$$

Similarly to the translation of points in Exercise 3 the new points will be:

- $(0,0) \to (-0.5,-0.5)$
- $(0,1000) \to (-0.5,0.5)$
- $(1000,0) \to (0.5,-0.5)$
- $(1000,1000) \to (0.5,0.5)$
- $(500,500) \to (0,0)$

4 RQ-Factorization and Computation of K

4.1 Exercise 5

With

$$K = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}, \quad R = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix}$$

simple matrix multiplication yields:

$$KR = \begin{bmatrix} aR_1^T + bR_2^T + cR_3^T \\ dR_2^T + eR_3^T \\ fR_3^T \end{bmatrix}$$

To find R and f, we extract A from the matrix $P = [A \quad a]$ and establish the matrix equation:

$$A = KR \to \begin{bmatrix} A_1^T \\ A_2^T \\ A_3^T \end{bmatrix} = \begin{bmatrix} aR_1^T + bR_2^T + cR_3^T \\ dR_2^T + eR_3^T \\ fR_3^T \end{bmatrix}$$

4.1.1 R_3 and f

Since R_3 is orthogonal and thus has length 1, R_3 and f can be calculated with the third row of A:

$$f = \|A_3\| \rightarrow R_3 = \frac{1}{\|A_3\|} A_3$$

$$f = 1 \quad \text{and} \quad R_3 = A_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

4.1.2 R_2 and e

Since A_2 is a linear combination of the two vectors $d * R_2$ and $e * R_3$, e can be computed as $e = A_2^T R_3$ which yields:

$$e = A_2^T R_3 = \begin{bmatrix} -\frac{700}{\sqrt{2}} & 1400 & \frac{700}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 350 + 0 + 350 = 700$$

With e, d is found by solving for:

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

And because R_2 is orthogonal with length 1:

$$d = 1400 \rightarrow R_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

4.1.3 R_1, a, b, and c

To finally calculate R_1, a, b, and c we solve the equation:

$$A_1 - bR_2 - cR_3 = aR_1$$

From which we get the equations:

$$c = A_1^T R_3 = \begin{bmatrix} \frac{800}{\sqrt{2}} & 0 & \frac{2400}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = -400 + 0 + 1200 = 800$$

$$b = A_1^T R_2 = \begin{bmatrix} \frac{800}{\sqrt{2}} & 0 & \frac{2400}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0 + 0 + 0 = 0$$

And we can finally solve for a and R_1:

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

With this:

$$a = 1600, \quad b = 0, \quad c = 800, \quad \text{and} \quad R_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

4.1.4 The K Matrix

With a-f known, we construct the K matrix:

$$K = \begin{bmatrix} \lambda f & sf & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1600 & 0 & 800 \\ 0 & 1400 & 700 \\ 0 & 0 & 1 \end{bmatrix}$$

1. Focal length = 1400
2. Skew = 0
3. Aspect ratio = $\frac{1600}{1400} = \frac{8}{7}$
4. Principal point = (800, 700)

4.2 Computer Exercise 2

$$K_{T_1} = \begin{bmatrix} 2391.5 & 225.7 & 942.7 \\ 0 & 623.9 & 813.6 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_{T_2} = \begin{bmatrix} 2394.0 & 0.0 & 932.4 \\ 0 & 2398.1 & 628.3 \\ 0 & 0 & 1 \end{bmatrix}$$

As we can tell, even taking scaling in mind, the transformations are not the same.

5 Direct Linear Transformation DLT

5.1 Exercise 7

$$P = N^{-1}\hat{P}$$

5.2 Computer Exercise 3

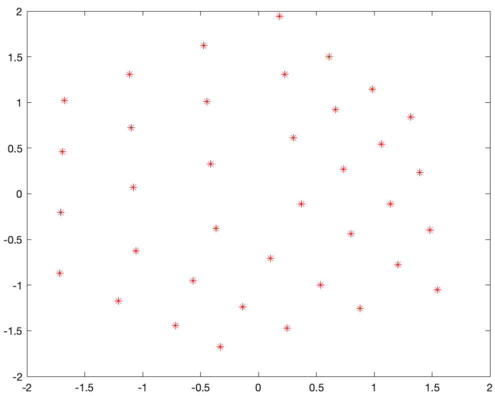

Figure 7: Normalized Points from Camera 1

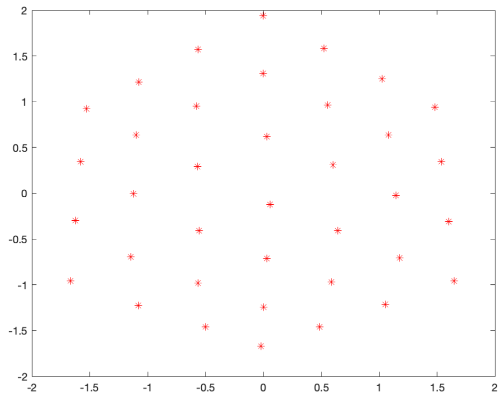

Figure 8: Normalized Points from Camera 2

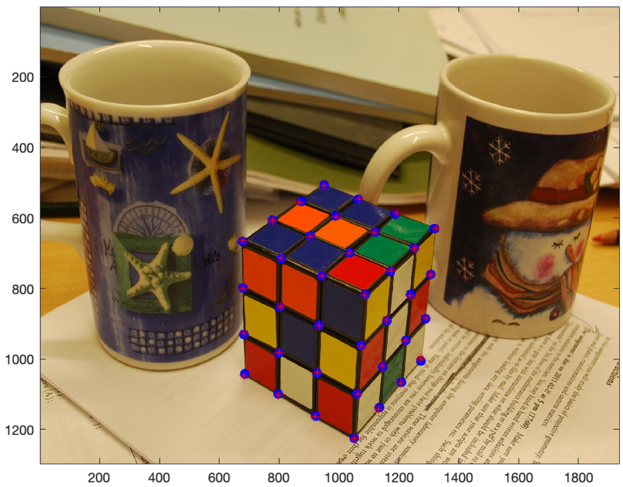

Figure 9: Points and projection on Image - Camera 1

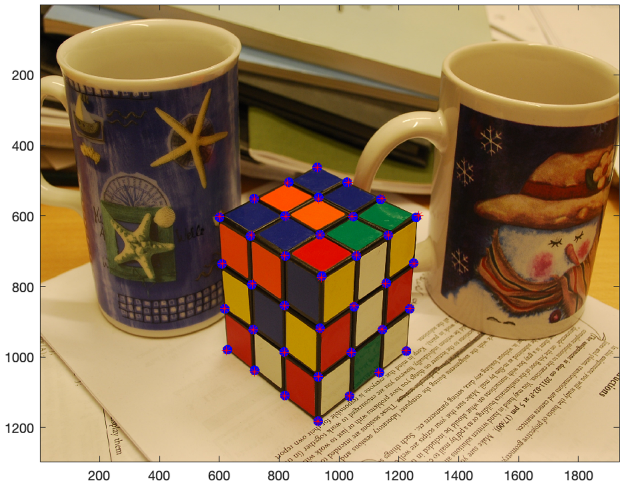

Figure 10: Points and projection on Image - Camera 2

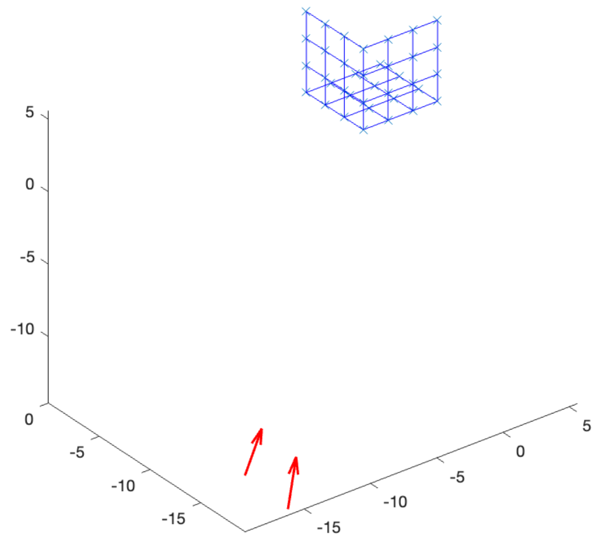

Figure 11: Reconstruction of Cube

The two camera matrices are:

$$K_1 = \begin{bmatrix} 2448.6 & -18.1 & 959.8 \\ 0 & 2446.8 & 675.9 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_2 = \begin{bmatrix} 2389 & -24.5 & 814.2 \\ 0 & 2401.0 & 790.5 \\ 0 & 0 & 1 \end{bmatrix}$$

In CE3, we have retrieve the true parameters because we have a known 3D model to project points onto provided 2D images. This difference ensures that the parameters we calculate from the projections are true, as they are based on a 3D structure. There's no ambiguity here because the process is direct. We use a definitive 3D model and 2d projections to relate the camera to two known truths.

In E1, we're trying to infer a 3D projection from two 2D images without a definitive model. This situation becomes ambiguous because multiple 3D projections could result in the same 2D projections. Without a ground truth, distinguishing between these possibilities a ground truth can not be determined.

In essence, the difference lies in creating a 3D-model based on only 2D-image measurements, or translating a ground truth 3D-model onto 2D projections of that same model. The camera matrices retrieved from the model to image projections must hence accurately translate from a ground truth in both images and in the model.

6 Triangulation using DLT

6.1 Computer Exercise 4-5

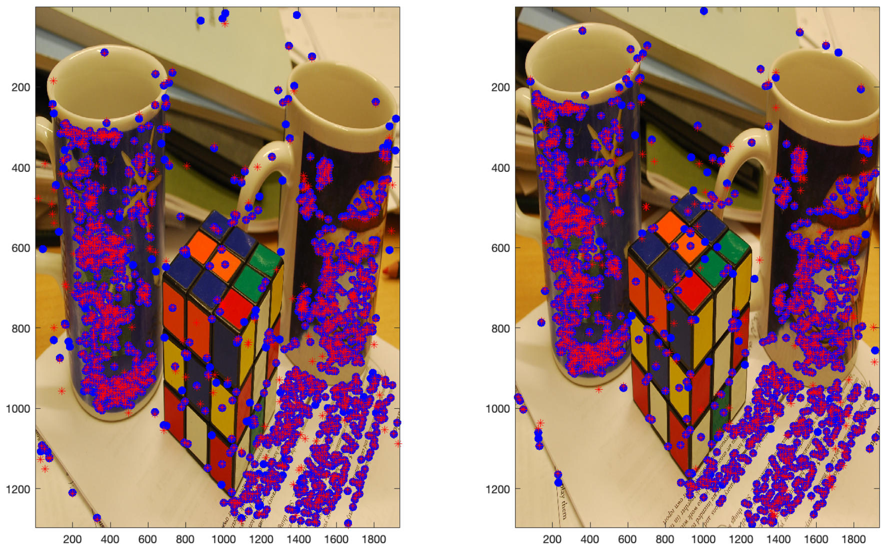

Figure 12: Measured and Calculated Matched Points by DLT with Normalisation in Image 1

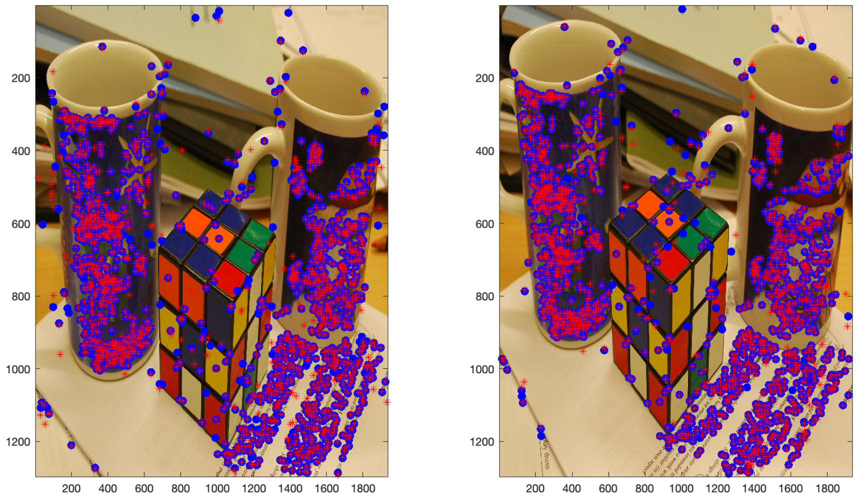

Figure 13: Measured and Calculated Matched Points by DLT with Normalisation in Image 2

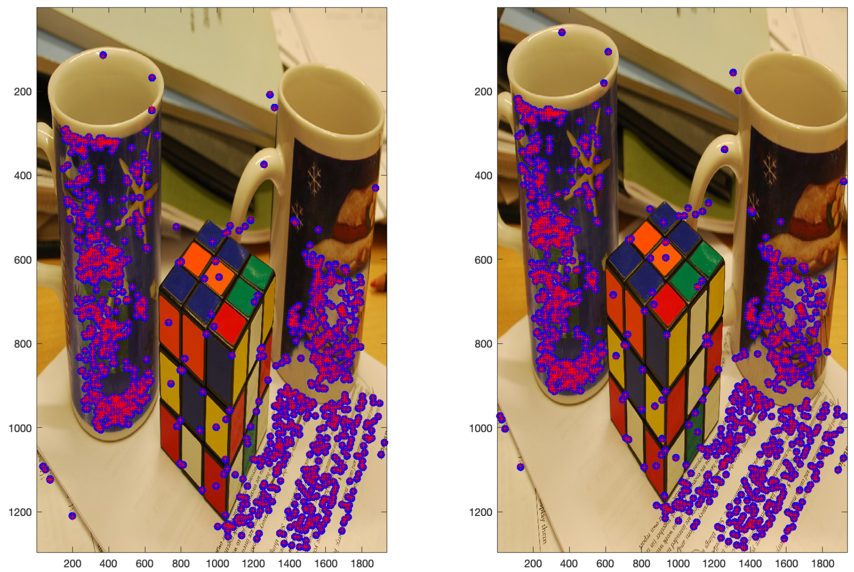

Figure 14: Measured and Calculated Matched Points by DLT without Normalization Filtered by Good Points in Image 1

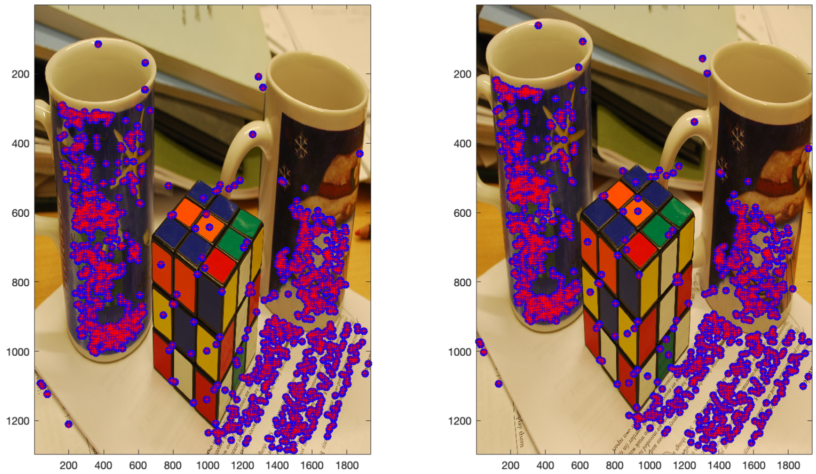

Figure 15: Measured and Calculated Matched Points by DLT with Normalisation Filtered by Good Points in Image 2

In the above figure we can see an improvement when using normalized DLT calculation. The error between measured and calculated points have decreased and as a result more points are present after filtering by pixel distance.

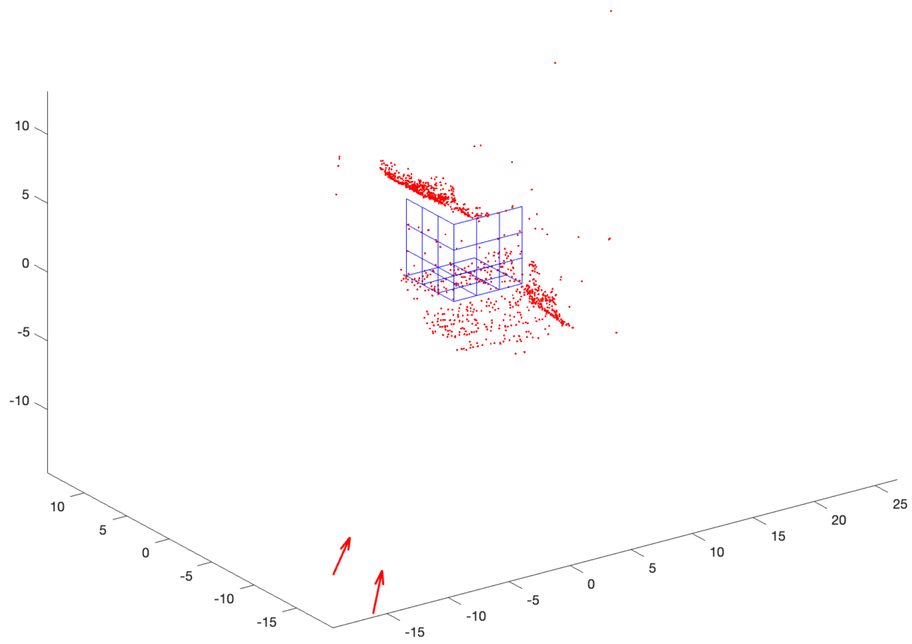

Figure 16: 3D model of cube and matched good points