Assignment 3 EEN020 Computer vision

4 December 2023

Erik Norlin

The Fundamental Matrix

Theoretical Exercise 1

We have the camera pair

$$P_1 = [I \mid 0], \text{ and } P_2 = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

For these camera matrices the fundamental matrix is calculated as

$$F = [t]_\times M$$

where $M = P_{2_{1:3,1:3}}$ and $t = P_{2_{1:3,4}}$

$$F = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 2 & 0 & 0 \end{pmatrix}$$

Given a projected point $x_1 = (0{,}2)$ in P_1 from the 3D points X, the epipolar line in the second image generated from x_1 is calculated as

$$l_2 = Fx_1 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$$

To find which of the points (1,0), (3,2) and (1,1) correspond to x_1 in P_2 we compute $x_2^T F x_1 = x_2^T l_2$. The corresponding point will satisfy $x_2^T F x_1 = x_2^T l_2 = 0$.

$$x_2^T F x_1 = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} = 2 \neq 0$$

$$x_2^T F x_1 = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} = 2 \neq 0$$

$$x_2^T F x_1 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} = 0$$

Thus, $x_2 = (1{,}1)$ could be a projection of the same point X in P_2.

Theoretical Exercise 2

We have the camera pair

$$P_1 = [I \mid 0], \text{ and } P_2 = [M \mid t] = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 3 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \end{pmatrix}$$

and we want to compute the epipoles by projecting the camera centers.

$$e_1 = P_1 \begin{pmatrix} -M^{-1}t \\ 1 \end{pmatrix} = [I \mid 0] \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$$

$$e_2 = P_2 \begin{pmatrix} -I^{-1} \cdot 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 3 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

We compute the fundamental matrix as

$$F = [t]_\times M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 3 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & -6 \\ 6 & 3 & -1 \end{pmatrix}$$

We check that $e_2^T F = 0$ and $F e_1 = 0$

$$e_2^T F = \begin{pmatrix} 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & -6 \\ 6 & 3 & -1 \end{pmatrix} = \mathbf{0}$$

$$F e_1 = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & -6 \\ 6 & 3 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \mathbf{0}$$

We can observe that F does not have full rank because the first two rows of it are linearly dependent. This implies that the determinant of F is zero.

Theoretical Exercise 3

We have the camera pair

$$P_1 = [I \mid 0], \text{ and } P_2 = [A \mid t]$$

and want to compute the epipoles by projecting the camera centers

$$e_1 = P_1 C_2 = \begin{bmatrix} I \mid 0 \end{bmatrix} \begin{bmatrix} -A^{-1}t \\ 1 \end{bmatrix} = -A^{-1}t$$

$$e_2 = P_2 C_1 = \begin{bmatrix} A \mid t \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = t$$

The fundamental matrix is $F = [t]_\times A$ and we verify that the epipoles always fulfill $e_2^T F = 0$ and $Fe_1 = 0$.

$$e_2^T F = t^T [t]_\times A = \begin{pmatrix} t_1 & t_2 & t_3 \end{pmatrix} \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} A = \begin{pmatrix} t_2 t_3 - t_3 t_2 \\ -t_1 t_3 + t_3 t_1 \\ t_1 t_2 - t_2 t_1 \end{pmatrix}^T A = \mathbf{0}$$

$$Fe_1 = -[t]_\times A A^{-1} t = -[t]_\times t = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} -t_3 t_2 + t_2 t_3 \\ t_3 t_1 - t_1 t_3 \\ -t_2 t_1 + t_1 t_2 \end{pmatrix} = \mathbf{0}$$

The determinant of the fundamental matrix is zero because it is not full rank. $[t]_\times$ is of rank 2 (not full rank) and A is of rank 3 (full rank and hence invertable). Multiplying $[t]_\times$ with A gives F with rank 2 (not full rank). Hence, the determinant of F must be zero.

Theoretical Exercise 4

We have that

$$\tilde{x}_1 \sim N_1 x_1, \; \tilde{x}_2 \sim N_2 x_2, \text{ and } \tilde{x}_2^T \tilde{F} \tilde{x}_1 = 0$$

Expanding $\tilde{x}_2^T \tilde{F} \tilde{x}_1$ we obtain

$$\tilde{x}_2^T \tilde{F} \tilde{x}_1 = (N_2 x_2)^T \tilde{F} N_1 x_1 = x_2^T N_2^T \tilde{F} N_1 x_1 = 0$$

where $F = N_2^T \tilde{F} N_1$ is the fundamental matrix for the original (un-normalised) points, and thus $x_2^T F x_1 = 0$.

Computer Exercise 1

Normalised case

The minimum singular value is equal to $||Mv||$ and is approximately 0.0497, the determinant of F is equal to zero, and the epipolar constraints $\tilde{x}_2^T \tilde{F} \tilde{x}_1$ are roughly zero for all points. The computed fundamental matrix for this case is

$$F \approx \begin{pmatrix} -3.39e-08 & -3.72e-06 & 5.77e-03 \\ 4.67e-06 & 2.89e-07 & -2.67e-02 \\ -7.19e-03 & 2.63e-02 & 1.00 \end{pmatrix}$$

Figure 1 and 2 show the two images, and their corresponding epipolar lines. We can see from the images that the point-to-line distance in pixels is small, in fact the mean error is 0.35 pixels in total for both images.

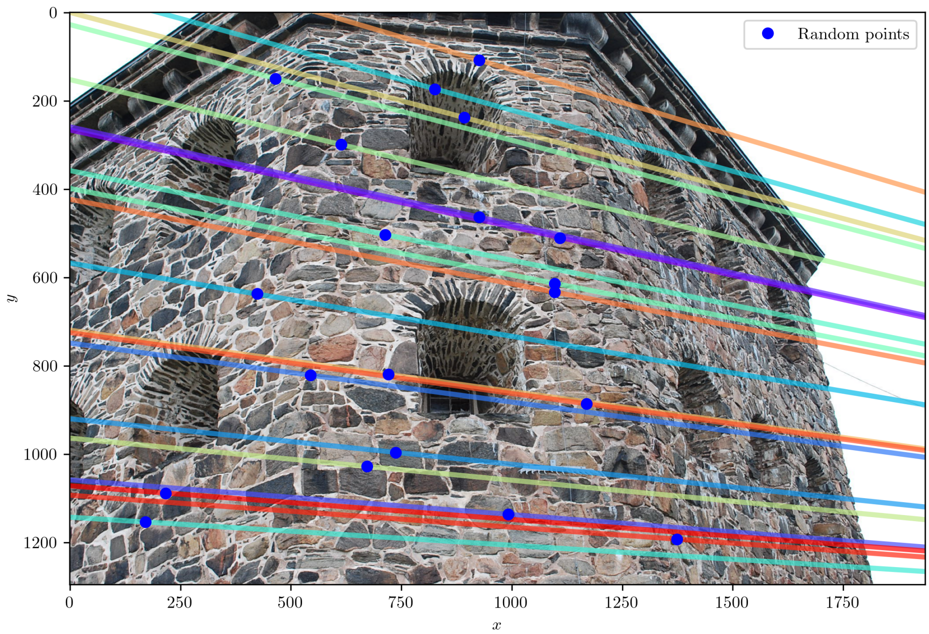

Figure 1: View from the first camera, 20 random points and their corresponding epipolar lines.

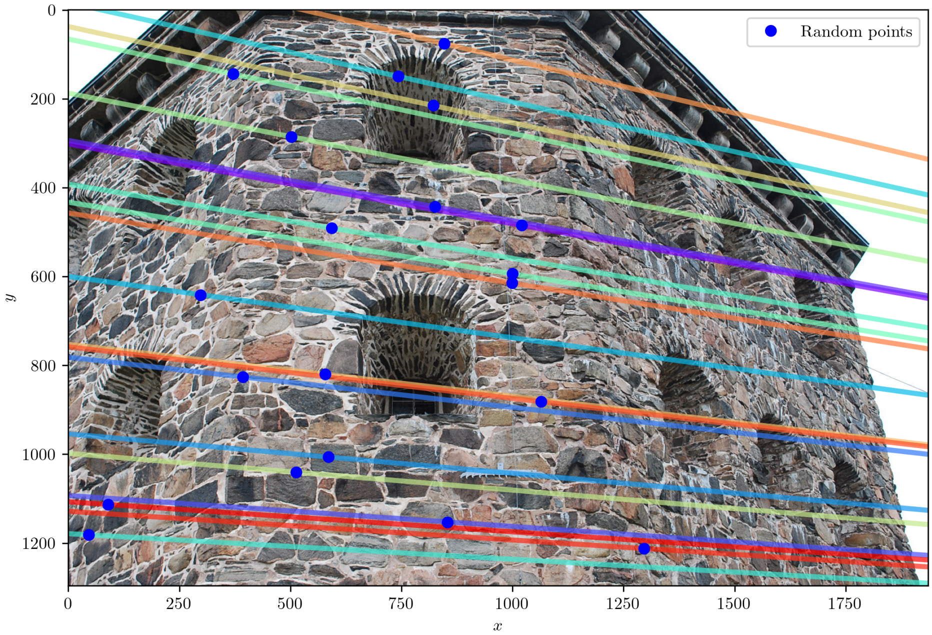

Figure 2: View from the second camera, 20 random points and their corresponding epipolar lines.

Figure 3, 4 and 5 show histograms of the distance (error) between the points and their corresponding epipolar lines.

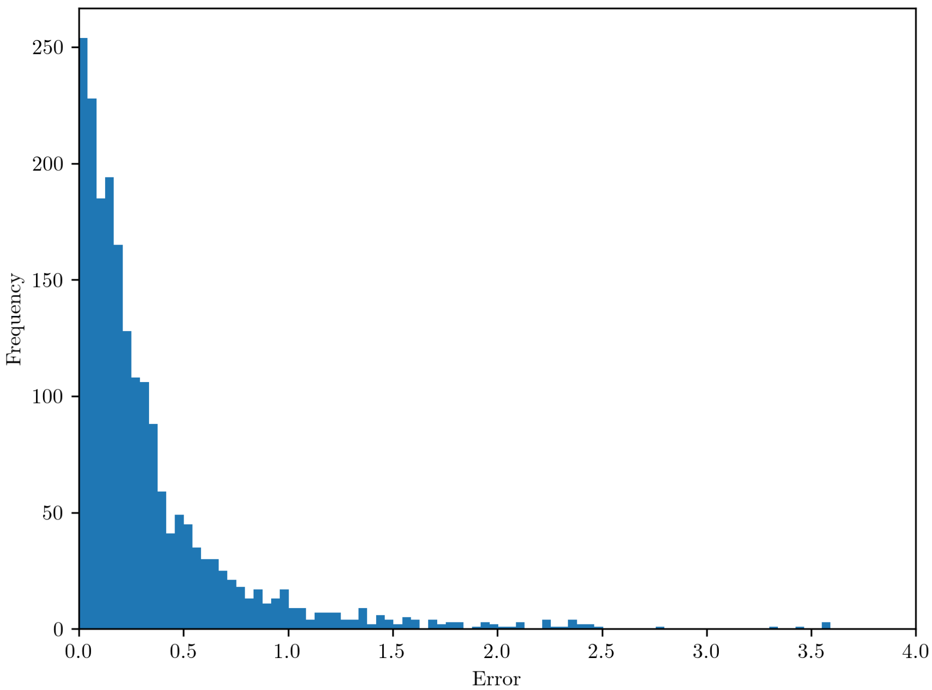

Figure 3: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the first image. The mean point-to-line distance is ≈ 0.34.

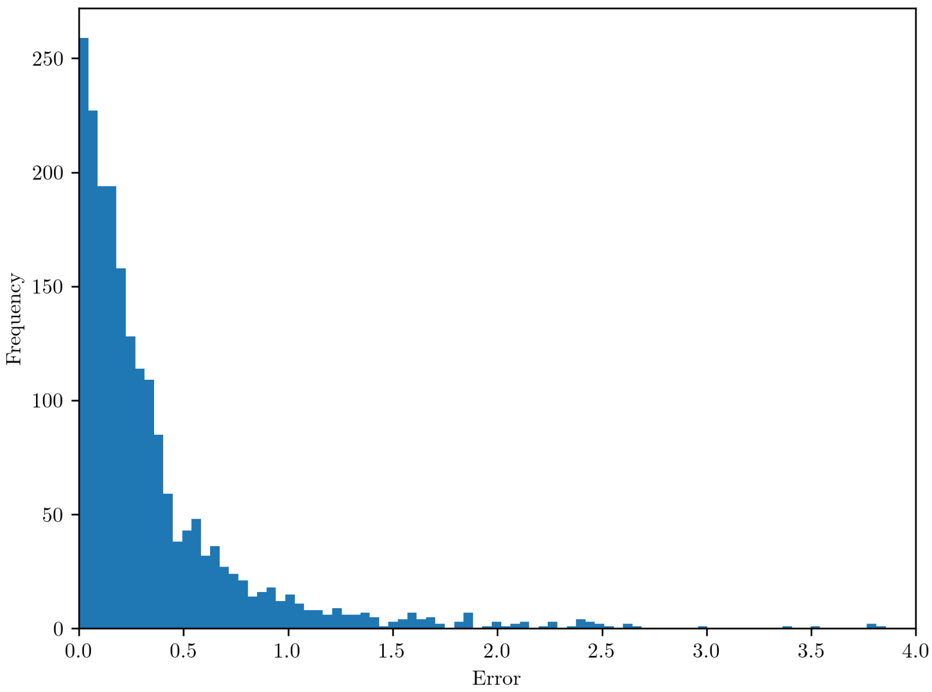

Figure 4: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the second image. The mean point-to-line distance is ≈ 0.36.

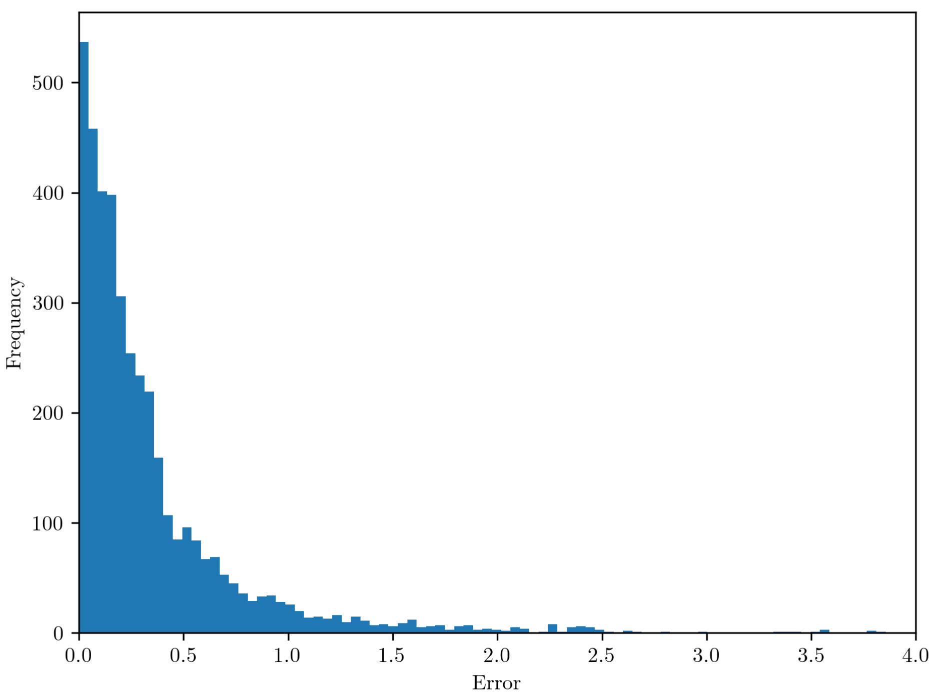

Figure 5: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in both images. The mean point-to-line distance is ≈ 0.35.

Unnormalised case

The minimum singular value is equal to $||Mv||$ and is approximately 0.5629, the determinant of F is equal to zero, and the epipolar constraints $x_2^T F x_1$ are roughly zero for all points. The computed fundamental matrix for this case is

$$F \approx \begin{pmatrix} -3.16e-08 & -3.85e-06 & 5.83e-03 \\ 4.79e-06 & 2.82e-07 & -2.66e-02 \\ -7.24e-03 & 2.62e-02 & 1.00 \end{pmatrix}$$

Figure 6 and 7 show the two images, the same 20 random points as before and their corresponding epipolar lines. We can see that the point-to-line distance is small in the images, in fact the mean error is 0.47 pixels in total for both images, about 35% larger than the total mean error for the normalised case. This shows that normalising the image points before estimating the fundamental matrix gives more accurate results in this case.

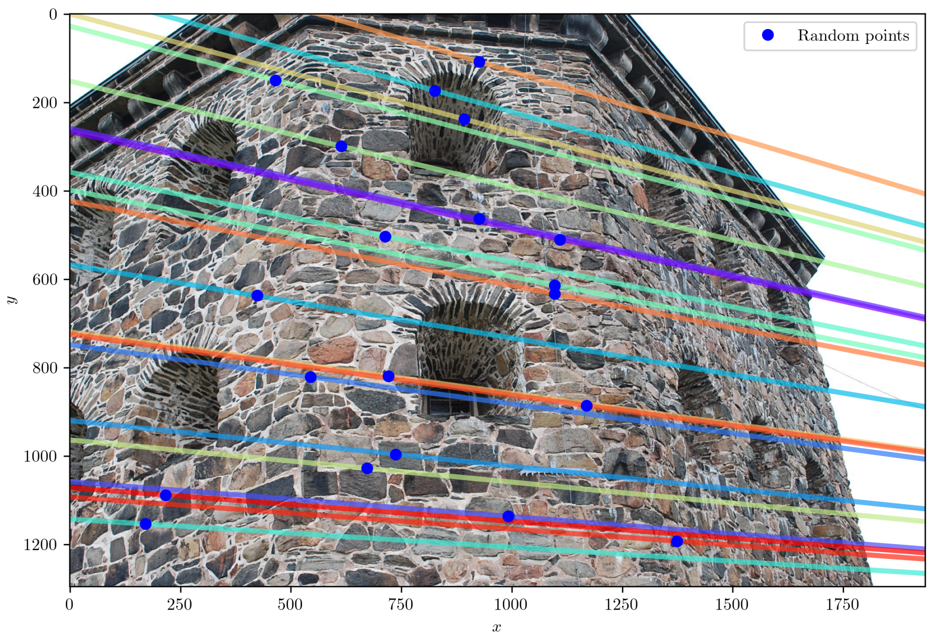

Figure 6: View from the first camera, 20 random points and their corresponding epipolar lines.

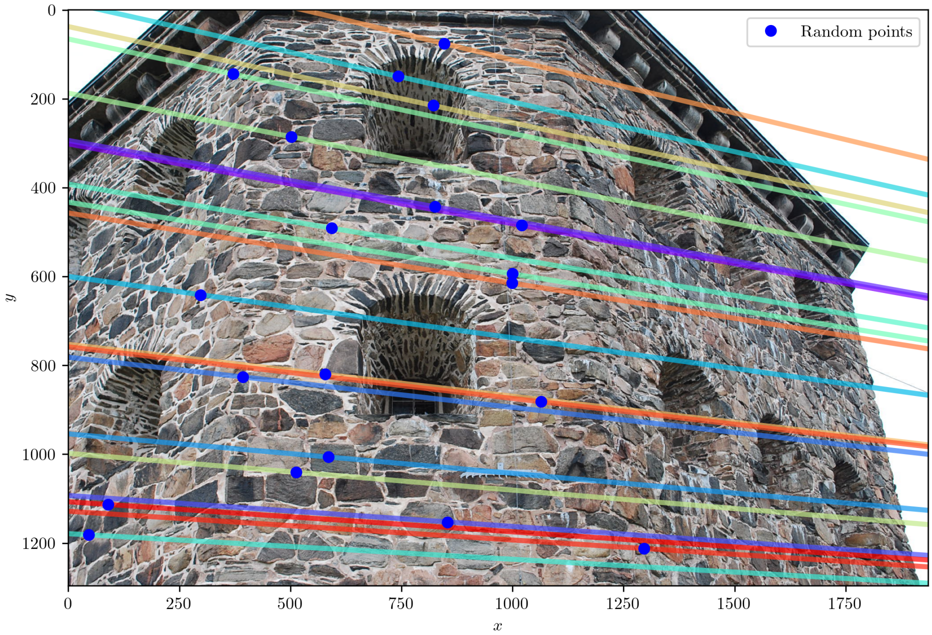

Figure 7: View from the second camera, 20 random points and their corresponding epipolar lines.

Figure 8, 9 and 10 show histograms of the distance (error) between the points and their corresponding epipolar lines. The errors are larger for the unnormalised case for both images.

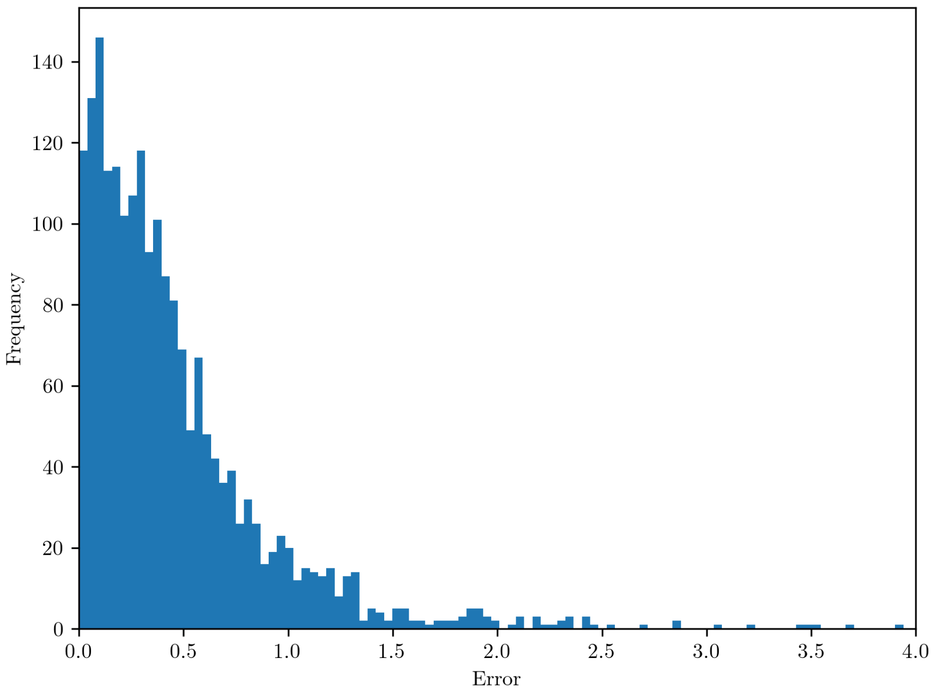

Figure 8: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the first image. The mean point-to-line distance is ≈ 0.46.

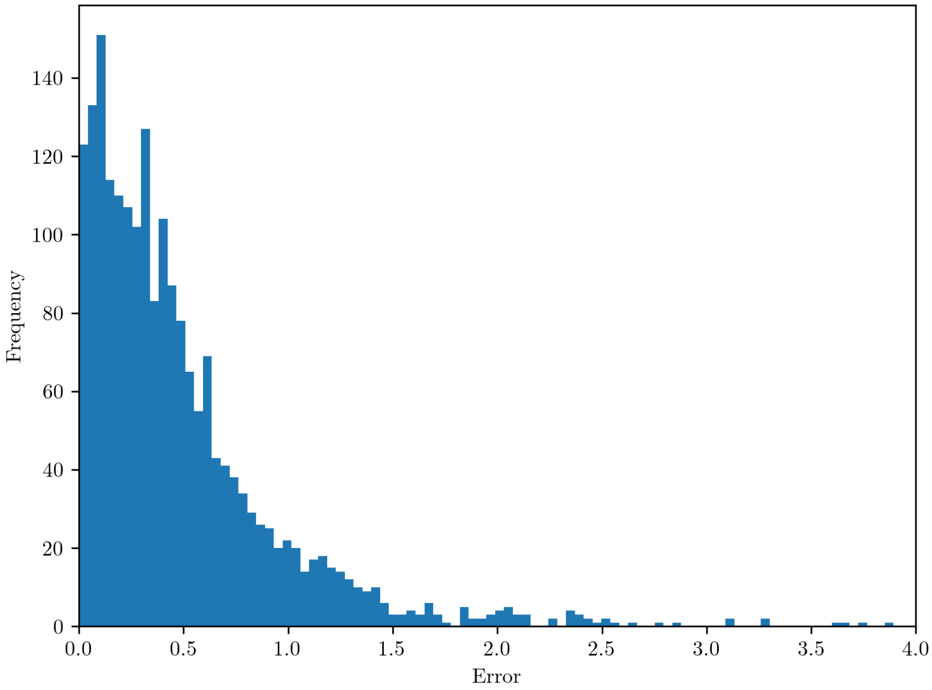

Figure 9: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the second image. The mean point-to-line distance is ≈ 0.49.

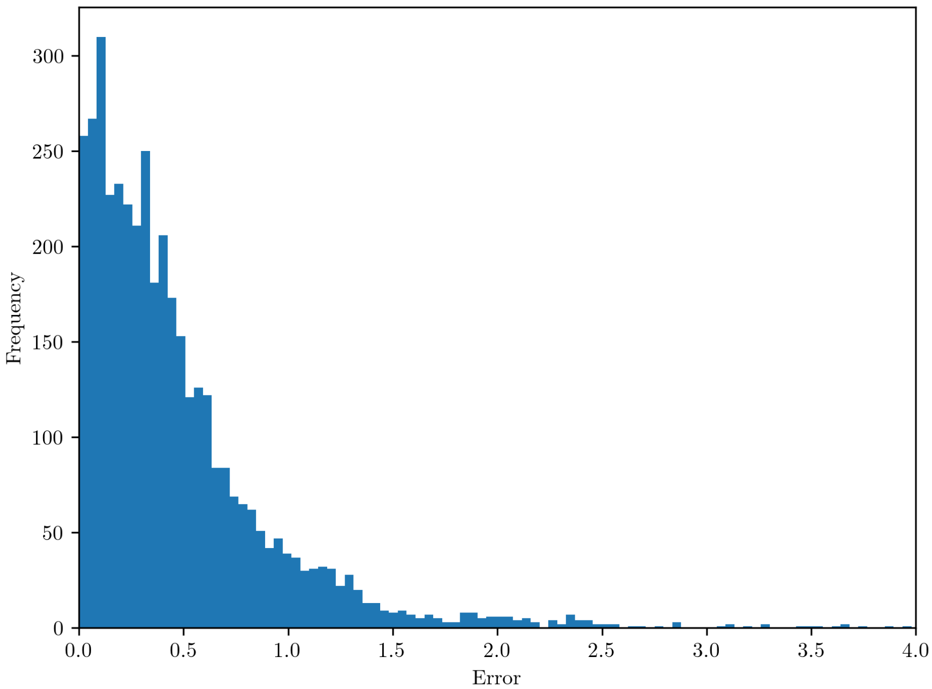

Figure 10: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in both images. The mean point-to-line distance is ≈ 0.47.

Theoretical Exercise 5

Considering the fundamental matrix

$$F = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix}$$

and the camera pair

$$P_1 = [I \mid 0], \text{ and } P_2 = [[e_2]_\times F \mid e_2]$$

where $e_2 \sim \text{null}(F^T) = (-1{,}0{,}1)^T$. Hence,

$$P_2 = [[e_2]_\times F \mid e_2] = \left[\begin{pmatrix} 0 & -e_{2_3} & e_{2_2} \\ e_{2_3} & 0 & -e_{2_1} \\ -e_{2_2} & e_{2_1} & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} e_{2_1} \\ e_{2_2} \\ e_{2_3} \end{pmatrix} \right] = \begin{pmatrix} -2 & 0 & -4 & -1 \\ 0 & 2 & 2 & 0 \\ -2 & 0 & -4 & 1 \end{pmatrix}$$

We verify that the projection of the scene points $X_1 = (0{,}3{,}1)^T$ and $X_2 = (-1{,}2{,}0)^T$ in both cameras fulfill the epipolar constraint $x_2^T F x_1 = 0$. For X_1 we have

$$x_1 \sim P_1 X_1 = [I \mid 0] \begin{pmatrix} 0 \\ 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \Rightarrow x_1 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$$

$$x_2 \sim P_2 X_1 = \begin{pmatrix} -2 & 0 & -4 & -1 \\ 0 & 2 & 2 & 0 \\ -2 & 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ 8 \\ -3 \end{pmatrix} \Rightarrow x_2 = \begin{pmatrix} 5/3 \\ -8/3 \\ 1 \end{pmatrix}$$

$$x_2^T F x_1 = \begin{pmatrix} 5/3 & -8/3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = 0$$

For X_2 we have

$$x_1 \sim P_1 X_2 = [I \mid 0] \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \Rightarrow x_1 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

$$x_2 \sim P_2 X_1 = \begin{pmatrix} -2 & 0 & -4 & -1 \\ 0 & 2 & 2 & 0 \\ -2 & 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix} \Rightarrow x_2 = \begin{pmatrix} 1/3 \\ 4/3 \\ 1 \end{pmatrix}$$

$$x_2^T F x_1 = \begin{pmatrix} 1/3 & 4/3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = 0$$

To compute the camera center of P_2 we have that

$$P_2 \begin{pmatrix} C_2 \\ \mu \end{pmatrix} = 0 \Leftrightarrow [e_2]_\times F C_2 + \mu e_2 = 0$$

The nullspace of F and hence $[e_2]_\times F$ is $C_2 = e_1$, and this leaves us with $\mu = 0$. Thus,

$$C_2 = \begin{pmatrix} -2 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

which is a point at infinity.

The Essential Matrix

Theoretical Exercise 6

We show that the eigenvalues of $[t]_\times^T [t]_\times$ are the squared singular values of the skew symmetric matrix $[t]_\times$ with the SVD USV^T as following

$$[t]_\times^T [t]_\times = (USV^T) USV^T = VS^T U^T USV^T = VS^T SV^T$$

Since S is a diagonal matrix we can write

$$[t]_\times^T [t]_\times = VS^2 V^T$$

This shows that S^2 diagonalizes $[t]_\times^T [t]_\times$ and contains therefore the eigenvalues of $[t]_\times^T [t]_\times$, which are the squared singular values of $[t]_\times$.

If w satisfies $-t \times (t \times w) = \lambda w$ for some λ we have that

$$-t \times (t \times w) = -[t]_\times [t]_\times w = [t]_\times^T [t]_\times w = \lambda w$$

since $[t]_\times + [t]_\times^T = 0$. From this we have that w is an eigenvector of $[t]_\times^T [t]_\times$ with eigenvalue λ.

From theory we find the formula

$$u \times (v \times w) = (u \cdot w)v - (u \cdot v)w$$

Then in this case we get

$$-t \times (t \times w) = (-t \cdot w)t - (-t \cdot t)w$$

If $w = t$ we get that

$$[t]_\times^T [t]_\times t = -t \times (t \times t) = (-t \cdot t)t - (-t \cdot t)t = 0 \cdot t$$

In other words, $w = t$ is an eigenvector of eigenvalue zero. If w is perpendicular to t, i.e., $t \cdot w = 0$ we get

$$[t]_\times^T [t]_\times w = -t \times (t \times w) = (-0)t - (-t \cdot t)w = ||t||^2 w$$

Thus, any vector w perpendicular to t is an eigenvector with eigenvalue $||t||^2$ to $[t]_\times^T [t]_\times$.

Since $S^2 \in \mathbb{R}^{3\times 3}$ it has three elements along the diagonal. Hence, $[t]_\times^T [t]_\times$ has three eigenvalues, and these are $||t||^2$ with $\pm t$ and 0. Furthermore, since S contains the singular values of $[t]_\times$ and S^2 contains the eigenvalues of $[t]_\times^T [t]_\times$ then the singular values of $[t]_\times$ are the square root of S^2, which are $||t||$ with $\pm t$ and 0.

We find an SVD of $E = [t]_\times R$ with $[t]_\times$ having the SVD as demonstrated above as following

$$E = [t]_\times R = US(V^T R)$$

Both V^T and R are orthonormal and the product of two orthonormal matrices is also orthonormal. $US(V^T R)$ is therefore a valid SVD of E, and its singular values are $S = \mathrm{diag}(||t||, ||t||, 0)$.

Computer Exercise 2

The minimum singular value is equal to $||Mv||$ and is approximately 0.0066, the determinant of F is equal to zero, and the epipolar constraints $\tilde{x}_2^T E \tilde{x}_1$ are roughly zero for all points. The computed essential matrix (normalised with the last element) is

$$E \approx \begin{pmatrix} -8.89 & -1005.81 & 377.08 \\ 1252.52 & 78.37 & -2448.17 \\ -472.79 & 2550.19 & 1.00 \end{pmatrix}$$

and scaled such that its singular values are 1, 1 and 0.

Figure 11 and 12 show the two images, the same 20 random points as in Computer Exercise 1 and their corresponding epipolar lines. We can see that the point-to-line distance in pixels is small here as well. However, the total mean point-to-line distance for both images this time is 2.08 pixels, which is about a 592% larger error than for the normalised case in Computer Exercise 1. The reason for this increase in error is because we are fundamentally changing the estimated E when we set its singular values to 1, 1 and 0. Since E is defined up to scale, the scaling in of itself is not a problem. When estimating E two singular values are approximately the same but not precisely and the last singular value is approximately zero, so we will never get a perfect solution for E. Then when we enforce the singular values to be "perfect" we are destroying the solution. However, performing this enforcement is important for camera reconstruction later, and since the error is still relatively small this is overlooked.

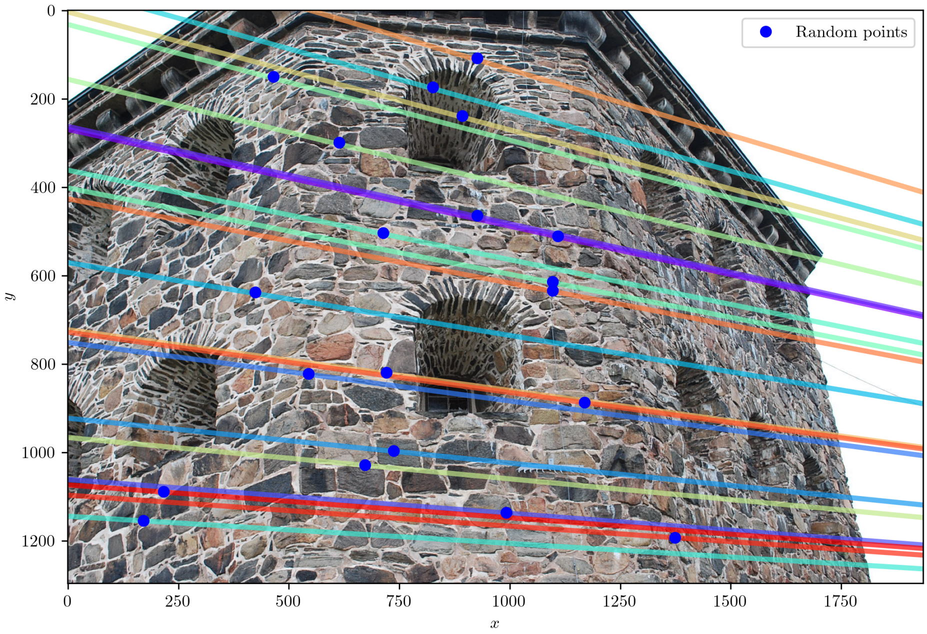

Figure 11: View from the first camera, 20 random points and their corresponding epipolar lines.

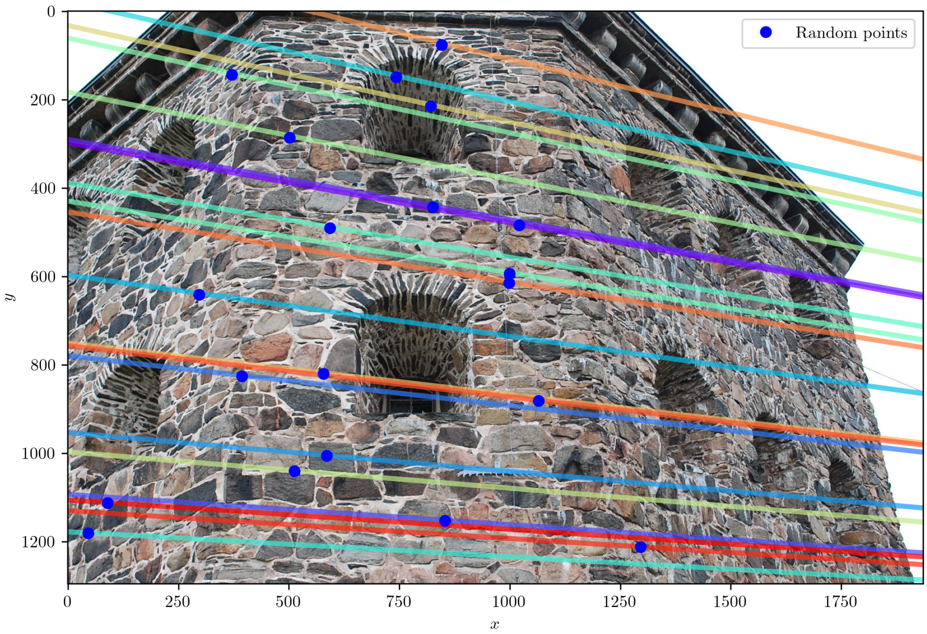

Figure 12: View from the second camera, 20 random points and their corresponding epipolar lines.

Figure 13, 14 and 15 show histograms of the distance (error) between the points and their corresponding epipolar lines. The errors are larger than for both cases in Computer Exercise 1.

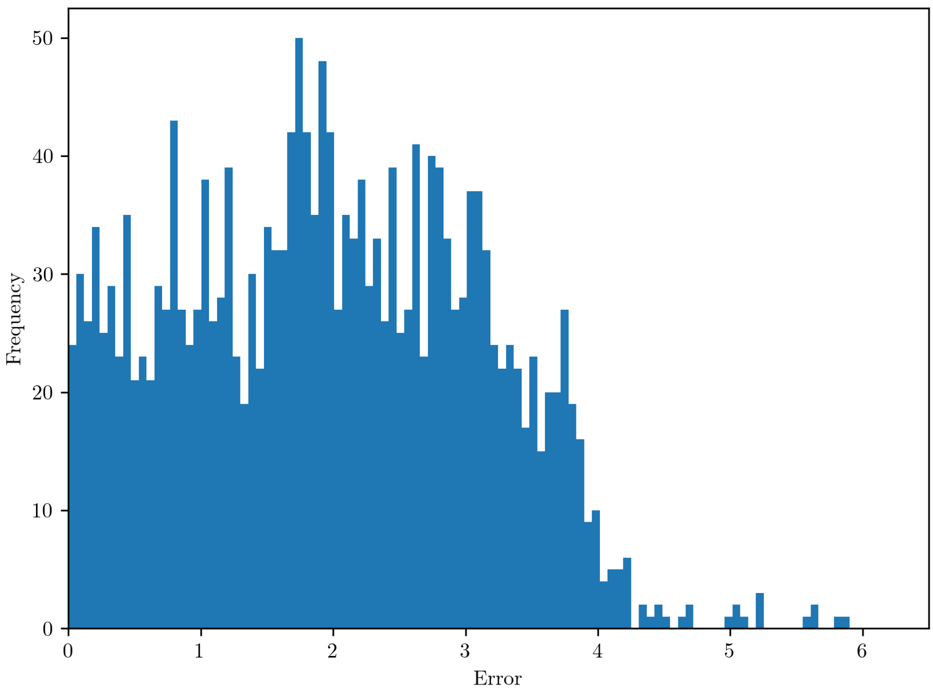

Figure 13: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the first image. The mean point-to-line distance is ≈ 1.98, about 582% larger than the normalised case in Computer Exercise 1.

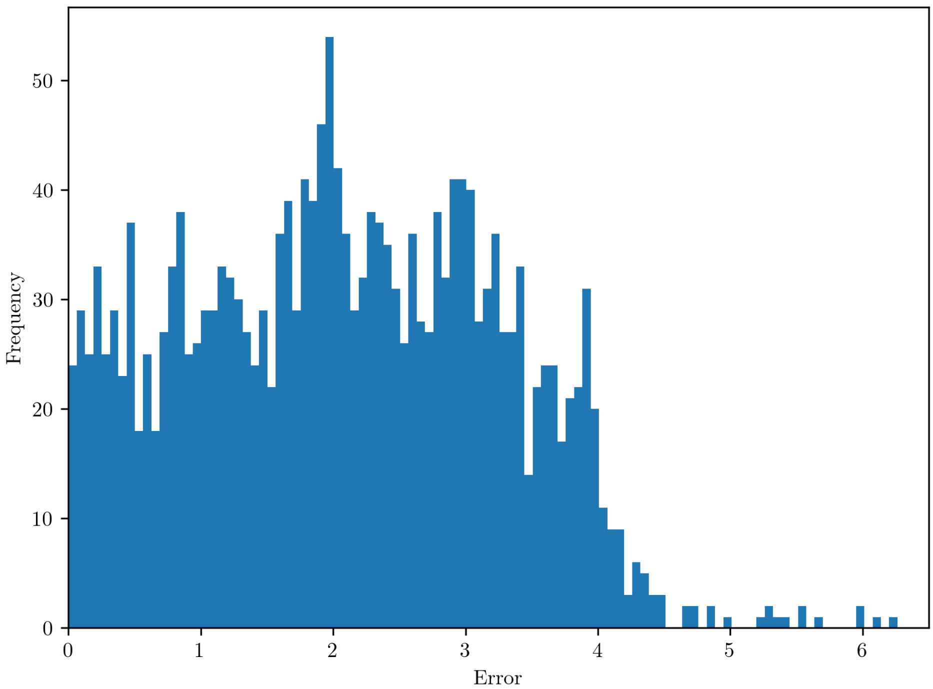

Figure 14: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in the second image. The mean point-to-line distance is ≈ 2.08, about 578% larger than the normalised case in Computer Exercise 1.

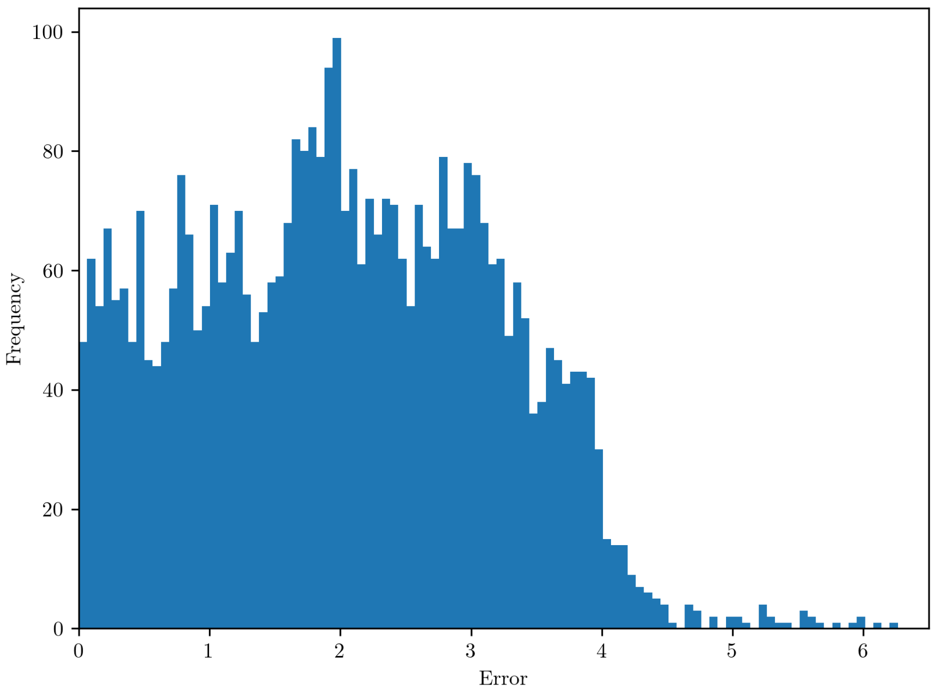

Figure 15: Histogram with 100 bins of the distance (error) from the random points and their corresponding epipolar lines in both images. The mean point-to-line distance is ≈ 2.03, about 580% larger than the normalised case in Computer Exercise 1.

Theoretical Exercise 7

Considering the essential matrix

$$E = U\text{diag}([1,1,0])V^T$$

where

$$U = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and } V = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

We verify that $\det(UV^T)$ is 1 by utilizing the formula

$$det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei - afh - bdi + bgf + cdh - ceg$$

$$det(UV^T) = det \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} = 0 + 1/2 + 0 + 0 + 1/2 + 0 = 1$$

$$E = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}^T = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Verifying that $x_1 = (2{,}0)^T$ and $x_2 = (-1{,}3)^T$ is a plausible correspondence we compute

$$x_2^T E x_1 = \begin{pmatrix} -1 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = 0$$

which shows the case.

If $x_1 = (2{,}0)^T$ is the projection of X in $P_1 = [I \mid 0]$ we want to show that X must be be one of the points

$$X(s) = \begin{pmatrix} 2 \\ 0 \\ 1 \\ s \end{pmatrix}$$

$$x_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \sim P_1 X = [I \mid 0] \begin{pmatrix} X_1/s \\ X_2/s \\ X_3/s \\ 1 \end{pmatrix} = \begin{pmatrix} X_1/s \\ X_2/s \\ X_3/s \end{pmatrix} = \frac{1}{s} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

which shows that X that projects to x_1 must be one of the points $X(s)$.

For each of the solutions

$P_2 = [UWV^T \mid u_3]$ or $P_2 = [UWV^T \mid -u_3]$ or $P_2 = [UW^TV^T \mid u_3]$ or $P_2 = [UW^TV^T \mid -u_3]$

where

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and u_3 is the third column of U, we want to compute s such that $X(s)$ projects to $x_2 = (1, -3)^T$. For $P_2 = [UWV^T \mid u_3]$ we have that

$$x_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix} \sim P_2 X = [UWV^T \mid u_3] \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} 1/(\sqrt{2}s) \\ -3/(\sqrt{2}s) \\ 1 \end{pmatrix}$$

$$\Rightarrow s = 1/\sqrt{2}$$

For $P_2 = [UWV^T \mid -u_3]$ we have that

$$x_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix} \sim P_2 X = [UWV^T \mid -u_3] \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} 1/(\sqrt{2}s) \\ -3/(\sqrt{2}s) \\ -1 \end{pmatrix}$$

$$\Rightarrow s = -1/\sqrt{2}$$

For $P_2 = [UW^TV^T \mid u_3]$ we have that

$$x_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix} \sim P_2X = [UW^TV^T \mid u_3] \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2}s) \\ 3/(\sqrt{2}s) \\ 1 \end{pmatrix}$$

$$\Rightarrow s = -1/\sqrt{2}$$

Finally, for $P_2 = [UW^TV^T \mid -u_3]$ we have that

$$x_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix} \sim P_2X = [UW^TV^T \mid -u_3] \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2/s \\ 0 \\ 1/s \\ 1 \end{pmatrix} = \begin{pmatrix} -1/(\sqrt{2}s) \\ 3/(\sqrt{2}s) \\ -1 \end{pmatrix}$$

$$\Rightarrow s = 1/\sqrt{2}$$

For $X(s)$ to be in front of a pair of calibrated cameras, such that each camera $P = [R \mid t]$ has det(R)=1, the depth of the scaled projected point in both cameras must be positive. P_1 and all solutions for P_2 are calibrated cameras with det(R)=1. The solutions for P_2 that gives positive z-coordinates after projection in the second camera are $P_2 = [UWV^T \mid u_3]$ and $P_2 = [UW^TV^T \mid u_3]$. However, only $s = 1/\sqrt{2}$ yields a positive z-coordinate in P_1. Thus, the only choice of P_2 that yields $X(s)$ in front of both P_1 and P_2 is $P_2 = [UWV^T \mid u_3]$.

Computer Exercise 3

Figure 16 and 17 show the two images, the image points and the projected 3D points from triangulation using the valid extracted camera from the estimated E. We can observe that the projected points fall close to the image points. Furthermore, the average pixel error in the first image is 0.99, 1.04 in the second image and 1.02 for both images.

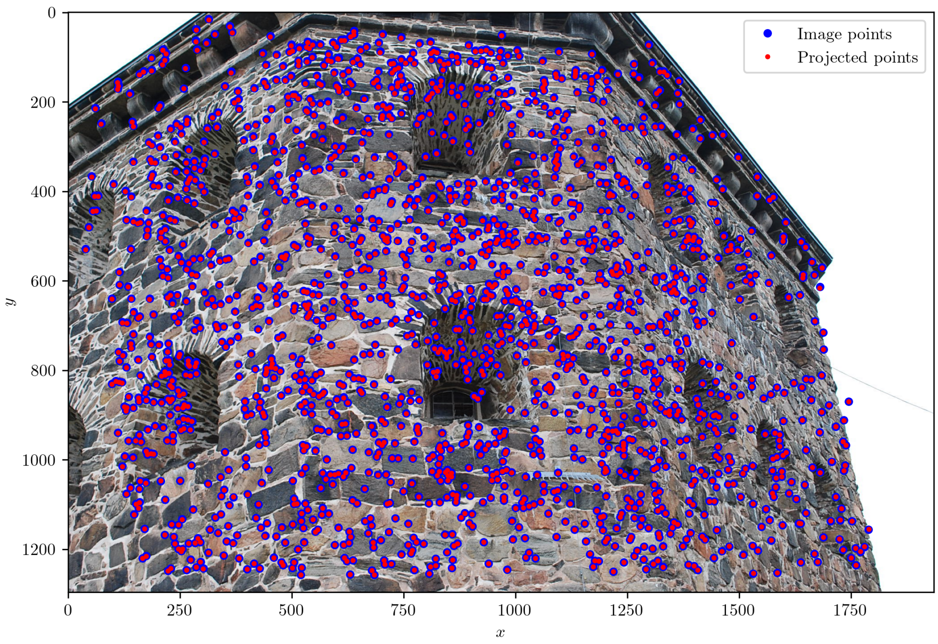

Figure 16: View from the first camera, image points and projected image points. The average pixel error in the first image is 0.99.

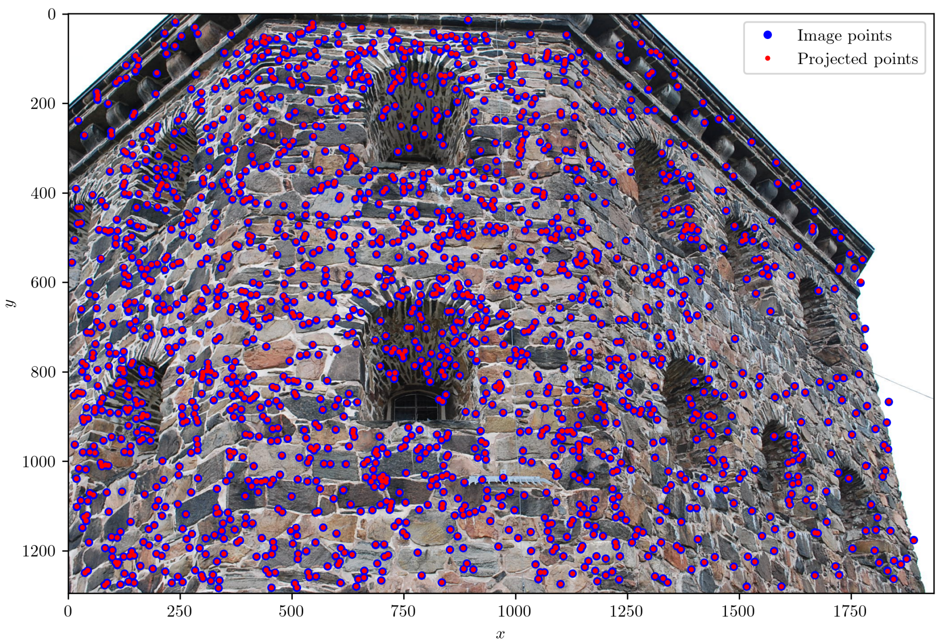

Figure 17: View from the second camera, image points and projected image points. The average pixel error in the first image is 1.04.

Figure 18, 19 and 20 show a 3D reconstruction from triangulation using the only valid extracted camera from the estimated essential matrix. The 3D reconstruction looks reasonable and similar to the object in the images. Camera 1 faces the object a little bit from the left and below, and camera 2 faces the object from the front and below as expected from the images.

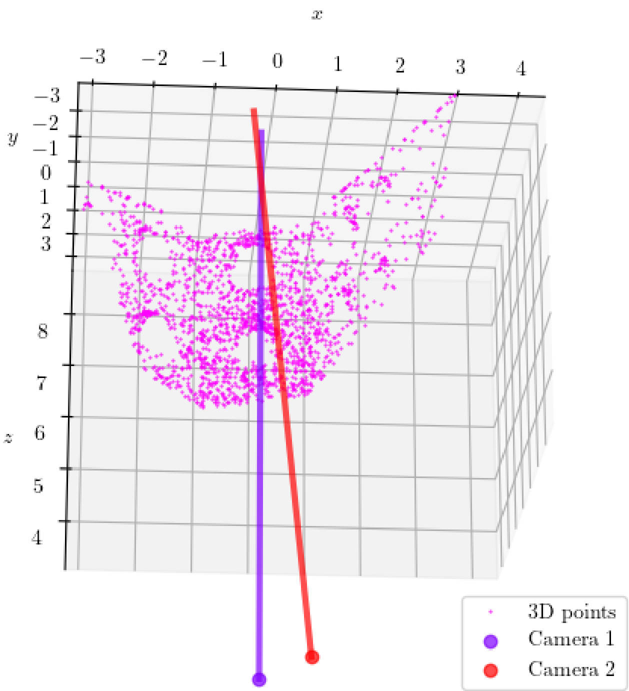

Figure 18: 3D reconstruction of the points from triangulation using the extracted camera from the estimated E. View from in front of the object.

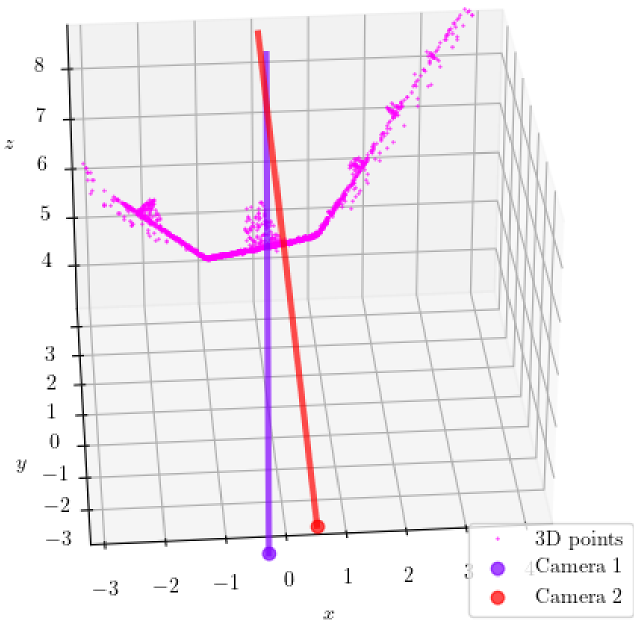

Figure 19: 3D reconstruction of the points from triangulation using the extracted camera from the estimated E. View from above the object.

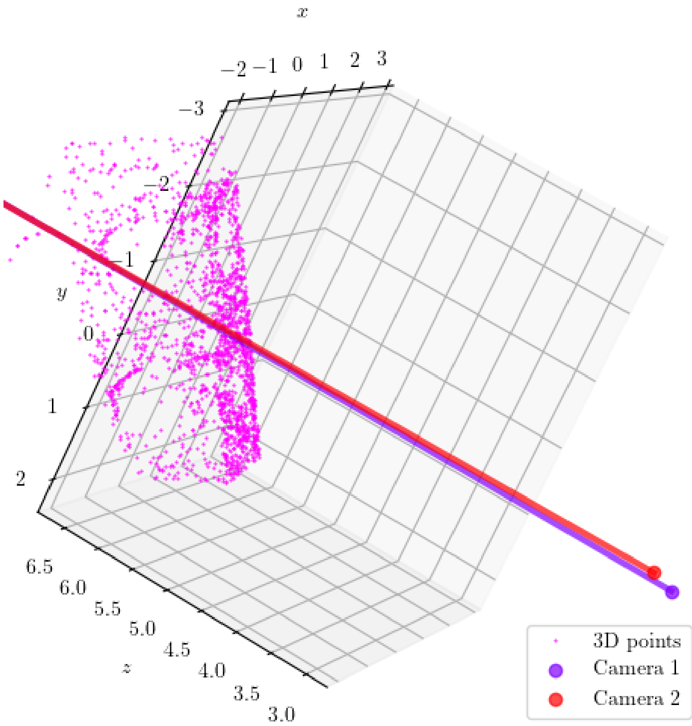

Figure 20: 3D reconstruction of the points from triangulation using the extracted camera from the estimated E. View from the left of the object.