Multiple View Geometry - Assignment 2

Itai Antebi
204817498

May 4, 2021

Calibrated vs. Uncalibrated Reconstruction

Exercise 1

When estimating structure and motion, we are given m points on n images:
$\{x_{ij}\}$ where $1 \leq i \leq n$, $1 \leq j \leq m$ and x_{ij} is the j-th point on the i-th image plane.
Let us assume that we have some solution:
$(\{P_i\}, \{X_j\})$ where $1 \leq i \leq n$ and P_i is the i-th camera matrix, and where $1 \leq j \leq m$ and X_j is the j-th 3D coordinate.
This would be a solution iff $\forall i, j\ x_{ij} \propto P_i X_j$.

However, notice that for any projective transformation T (an invertible 4×4 matrix) the following is also a solution:
$\left(\left\{P_i T^{-1}\right\}, \{TX_j\}\right)$ where $1 \leq i \leq n$ and $P_i T^{-1}$ is the i-th camera matrix, and where $1 \leq j \leq m$ and TX_j is the j-th 3D coordinate.
This is because $x_{ij} \propto P_i X_j \rightarrow x_{ij} \propto \left(P_i T^{-1}\right)(TX_j)$.
We shall note that this is under the assumption of uncalibrated cameras, and therefore there are not constraints on the camera matrix and any 3×4 matrix is a valid camera matrix.
Thus, when estimating structure and motion simultaneously, under the assumption of uncalibrated cameras, there is always an unknown projective transformation of 3D space that cannot be determined using only image projections.

Computer Exercise 1

a. The original 3D points DOES NOT look like a reasonable reconstruction. For example, the walls do not meet at a normal 90-degrees angle.
b. The projections DO appear to be close to the corresponding image points.
c. The first solution (T1) also DOES NOT look like a reasonable reconstruction. However, the second solution (T2) DOES look like a reasonable reconstruction.
d. After applying the first projection (T1), even though it doesn't look like a correct solution (see c.), the projections DO appear to be close to the corresponding image points and DO NOT appear to have changed. This shows us

that even though the solution is wrong (both the cameras and the 3D reconstruction) - the projection can still be aligned. This is due to the fact that we can only obtain a correct solution up to projective transformation (See exercise 1).

All plots can be viewed in the attached matlab_results.pdf output of the Matlab code.

Exercise 2

Since the cameras are now calibrated cameras, we know the calibration matrices K_i.
Thus, we can normalize the points $x'_{ij} = K_i^{-1} x_{ij}$ and look at the camera matrices $P_i = [R_i \ t_i]$.
A camera matrix $P \in R^{4\times3}$ is valid iff $P = [R \ t]$ in which $R \in R^{3\times3}$ is a valid rotation matrix ($RR^T = I$ and $det\left(R\right) = 1$).
Let us assume that we have some solution of a calibrated matrix: $\left(\left\{P_i\right\}, \left\{X_j\right\}\right)$ - $\forall i, j \ x'_{ij} \propto P_i X_j$, $P_i = [R_i \ t_i]$.
Given a projective transformation T (an invertible 4×4 matrix) that has an inverse of the form: $T^{-1} = \begin{bmatrix} sQ & v \\ 0 & 1 \end{bmatrix}$ where Q is a rotation matrix ($Q^T Q = I$, $det\left(Q\right) = 1$), v is a vector and $s > 0$ is a scalar, then the following is also a solution: $\left(\left\{P_i T^{-1}\right\}, \left\{T X_j\right\}\right)$.
Obviously, same as in exercise 1, it holds that $x'_{ij} \propto P_i X_j \rightarrow x'_{ij} \propto \left(P_i T^{-1}\right)\left(T X_j\right) \rightarrow x'_{ij} \propto \left(\frac{1}{s} P_i T^{-1}\right)\left(s T X_j\right)$.
But now we can also show that the new camera matrix $\frac{1}{s} P_i T^{-1}$ is a valid camera matrix: $\frac{1}{s} P_i T^{-1} = \frac{1}{s}\left[R_i \ t_i\right] \begin{bmatrix} sQ & v \\ 0 & 1 \end{bmatrix} = \frac{1}{s}\left[sR_i Q \ R_i v + t\right] = \left[R_i Q \ \frac{1}{s}\left(R_i v + t\right)\right]$
where $R_i Q \left(R_i Q\right)^T = R_i Q Q^T R_i^T = R_i R_i^T = I$ and $det\left(R_i Q \left(R_i Q\right)^T\right) = det\left(R_i\right) det\left(Q\right) det\left(Q^T\right) det\left(R_i^T\right) =$ 1.

To sum up, the corresponding statement for calibrated cameras is:
"When estimating structure and motion simultaneously, under the assumption of calibrated cameras, there is always an unknown transformation of 3D space T that cannot be determined using only image projections. The constrains on T are that it has an inverse is a similarity transformation".

Camera Calibration

Exercise 3

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

A has a geometric interpretation of scaling down the x, y coordinates by the focal length.
B has a geometric interpretation of shifting in order to align the center of the image.
When normalizing the image points of a camera with known inner parameters we apply the transformation K^{-1}.
The interpretation of this is to shift and scale each point so that it lends in the location where it should have been on the image - if the camera was aligned.
The principal point (x_0, y_0) end up at: $K^{-1}\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{f} & 0 & -\frac{x_0}{f} \\ 0 & \frac{1}{f} & -\frac{y_0}{f} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ - the origin, and a point with distance f to the principal point end up at a distance of 1 from the origin.

$K = \begin{bmatrix} 320 & 0 & 320 \\ 0 & 320 & 240 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow K^{-1} = \begin{bmatrix} \frac{1}{320} & 0 & -1 \\ 0 & \frac{1}{320} & -\frac{3}{4} \\ 0 & 0 & 1 \end{bmatrix}$
Normalizing $\begin{bmatrix} 0 \\ 240 \\ 1 \end{bmatrix}$ results in $K^{-1}\begin{bmatrix} 0 \\ 240 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{320} & 0 & -1 \\ 0 & \frac{1}{320} & -\frac{3}{4} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 240 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
Normalizing $\begin{bmatrix} 640 \\ 240 \\ 1 \end{bmatrix}$ results in $K^{-1}\begin{bmatrix} 640 \\ 240 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{320} & 0 & -1 \\ 0 & \frac{1}{320} & -\frac{3}{4} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 640 \\ 240 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.
The camera center for the camera $K[R\ t] = [KR\ Kt]$ is $-(KR)^{-1}Kt = -R^{-1}K^{-1}Kt = -R^{-1}t$ - the same as the camera center for the camera $[R\ t]$.
The principal axis for the camera $K[R\ t] = [KR\ Kt]$ is $det(KR)(KR)^3$.
Since the third row of K is $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$, then the third row of KR is the third row of R - $(KR)^3 = R^3$.
We also know that $det(KR) = det(K)\,det(R) = \gamma f^2 det(R)$.
So, the principal axis for the camera $K[R\ t]$ is $det(KR)(KR)^3 = \gamma f^2 det(R)\,R^3$ which is the same as the principal axis of the camera $[R\ t]$ $(det(R)\,R^3)$ after scaling (by γf^2).

Exercise 4

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

By inserting the calibration parameters and inverting the matrix we get:

$$K = \begin{bmatrix} 1000 & 0 & 500 \\ 0 & 1000 & 500 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow K^{-1} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

We can normalize the camera P by multiplying it to the left by K^{-1}:

$$K^{-1}P = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1000 & -250 & 250\sqrt{3} & 500 \\ 0 & 500\left(\sqrt{3}-\frac{1}{2}\right) & 500\left(1+\frac{\sqrt{3}}{2}\right) & 500 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix}$$

We can also normalize the corners by multiplying it to the left by K^{-1}:

$$(0,0) \Rightarrow K^{-1}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \Rightarrow \left(-\frac{1}{2}, -\frac{1}{2}\right)$$

$$(0,1000) \Rightarrow K^{-1}\begin{bmatrix} 0 \\ 1000 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1000 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} \Rightarrow \left(-\frac{1}{2}, \frac{1}{2}\right)$$

$$(1000,0) \Rightarrow K^{-1}\begin{bmatrix} 1000 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1000 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \Rightarrow \left(\frac{1}{2}, -\frac{1}{2}\right)$$

$$(1000,1000) \Rightarrow K^{-1}\begin{bmatrix} 1000 \\ 1000 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1000 \\ 1000 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} \Rightarrow \left(\frac{1}{2}, \frac{1}{2}\right)$$

and the center:

$$(500,500) \Rightarrow K^{-1}\begin{bmatrix} 500 \\ 500 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1000} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1000} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 500 \\ 500 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow (0,0)$$

This shows us that the normalization brought the center of the image to $(0,0)$ and the entire image is now of size 1×1.

RQ Factorization and Computation of K

Exercise 5

a.

$$K = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix},\ R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix} \Rightarrow R_j^T =$$

$$\begin{aligned} KR &= \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \\ &= \begin{bmatrix} R_{11}a + R_{21}b + R_{31}c & R_{12}a + R_{22}b + R_{32}c & R_{13}a + R_{23}b + R_{33}c \\ R_{21}d + R_{31}e & R_{22}d + R_{32}e & R_{23}d + R_{33}e \\ R_{31}f & R_{32}f & R_{33}f \end{bmatrix} \\ &= \begin{bmatrix} aR_1^T + bR_2^T + cR_3^T \\ dR_2^T + eR_3^T \\ fR_3^T \end{bmatrix} \end{aligned}$$

b.

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

We know form (a) that $fR_3^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$.
Since R is a rotation matrix, we know that:

$$\begin{aligned} R_3^T R_3 = 1 &\Rightarrow \left(fR_3^T\right)(fR_3) = f^2 \\ &\Rightarrow \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{2} + \frac{1}{2} = 1 = f^2 \\ &\Rightarrow f = 1 \\ &\Rightarrow R_3^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

Computer Exercise 2

Are K1 of the first camera and K1 of the camera with the projective transformation T1 the same?
false

Are K1 of the first camera and K1 of the camera with the projective transformation T2 the same?

true

This means that apllying the projective transformation $T2$ didn't change the calibraction parameters of the camera, but $T1$ did.

Direct Linear Transformation (DLT)

Exercise 6

.

a.

$$M = U\Sigma V^T$$

$$\begin{aligned}
\left|V^T v\right|^2 &= \left(V^T v\right)^T V^T v \\
&= v^T \left(V^T\right)^T V^T v \\
&= v^T V V^T v \\
&= v^T v = |v|^2
\end{aligned}$$

The second-to-last equation is correct because V is unitary (since it came from as SVD decomposition).
Thus, in particular, $|v|^2 = 1 \iff \left|V^T v\right|^2 = 1$.

b.
Let $\widetilde{v} = V^T v$.
We already know that $|\widetilde{v}|^2 = \left|V^T v\right|^2 = |v|^2$, and thus $|\widetilde{v}|^2 = 1 \iff |v|^2 = 1$.
We also know that $|Mv|^2 = \left|U\Sigma V^T v\right|^2 = \left\langle U\Sigma V^T v, U\Sigma V^T v\right\rangle = \left\langle \Sigma V^T v, U^T U\Sigma V^T v\right\rangle = \left\langle \Sigma V^T v, \Sigma V^T v\right\rangle = \left|\Sigma V^T v\right|^2$.
The second-to-last equation is correct because U is unitary (since it came from as SVD decomposition).
This leads us to infer the following:
$\min_{|v|^2=1} |Mv|^2 = \min_{|v|^2=1} \left|\Sigma V^T v\right|^2$ - because we proved that the functions $|Mv|^2$ and $\left|\Sigma V^T v\right|^2$ are the same (as a function of v).
$\min_{|v|^2=1} \left|\Sigma V^T v\right|^2 = \min_{|\widetilde{v}|^2=1} \left|\Sigma V^T v\right|^2$- because we proved that the domains of $|\widetilde{v}|^2 = 1$ and $|v|^2 = 1$ are equivalent.
$\min_{|\widetilde{v}|^2=1} \left|\Sigma V^T v\right|^2 = \min_{|\widetilde{v}|^2=1} |\Sigma\widetilde{v}|^2$ - because we simply denoted $\widetilde{v} = V^T v$.
Thus, we obtain $\min_{|v|^2=1} |Mv|^2 = \min_{|\widetilde{v}|^2=1} |\Sigma\widetilde{v}|^2$, as required.

Let us assume that we have a solution $\widetilde{v}$ that minimizes $|\Sigma\widetilde{v}|^2$ (and $|\widetilde{v}|^2 = 1$). Notice that $\widetilde{v} = V^T v \Rightarrow V\widetilde{v} = VV^T v = v$. The last equation is correct because V is unitary.
Thus, to find v we simply compute $v = V\widetilde{v}$ and we already know that $|v|^2 = 1$ and that it also minimizes$|Mv|^2$.

So, to sum up, if $\widetilde{v}$ is a solution to the second problem then $v = V\widetilde{v}$ is a solution to the first problem.

Let us assume that v is a solution to the first problem and let us observe $-v$.
We know that $-v$ is a different vector than v since$|v|^2 = 1 \Rightarrow v \neq 0 \Rightarrow -v \neq v$.
It holds that $|-v|^2 = 1$ since $|-v|^2 = |v|^2 = 1$, and, in addition, $|M(-v)|^2 = |-(Mv)|^2 = |Mv|^2$.
Thus, we have a different vector of norm 1 which produces the same (minimal) value for $|Mv|^2$.
So we have proven that there are always at least two solutions to this problem.

Assuming $rank(M) < n$:

$$\Sigma\widetilde{v} = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n \end{bmatrix} \begin{bmatrix} \widetilde{v}_1 \\ \vdots \\ \widetilde{v}_n \end{bmatrix} = \begin{bmatrix} s_1\widetilde{v}_1 \\ \vdots \\ s_n\widetilde{v}_n \end{bmatrix}$$

$$|\Sigma\widetilde{v}|^2 = \sum_{i=1}^{n} s_i^2 \widetilde{v}_i^2$$

Choosing $\widetilde{v_0} = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ will result in $|\Sigma\widetilde{v_0}|^2 = \sum_{i=1}^{n} s_i^2 \widetilde{v_0}_i^2 = 0$. This is because $\forall i \neq n$ $\widetilde{v_0}_i = 0$ and for $i = n$ $s_i = 0$ from out assumption.
Obviously this is the minimum possible value of $|\Sigma\widetilde{v}|^2$ and obviously $|\widetilde{v_0}|^2 = 1$.
Thus, $\widetilde{v_0}$ is a valid solution to the second problem.
As we have previously shown, to achieve a solution to the first problem we only have to compute $v = V\widetilde{v_0} = \begin{bmatrix} V_0 & \ldots & V_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = V_n$.
Thus, the solution to the original problem will be given by $v = V_n$ - the last column of V.

Exercise 7

$$\widetilde{x} \sim Nx \Rightarrow \lambda_1\widetilde{x} = Nx$$

$$\widetilde{x} \sim \widetilde{P}X \Rightarrow \lambda_2\widetilde{x} = \widetilde{P}X$$

$$\frac{\lambda_2}{\lambda_1}Nx = \frac{\lambda_2}{\lambda_1}\lambda_1\widetilde{x} = \lambda_2\widetilde{x} = \widetilde{P}X$$

$$\frac{\lambda_2}{\lambda_1}Nx = \widetilde{P}X$$

$$\frac{\lambda_2}{\lambda_1}x = N^{-1}\widetilde{P}X$$

$$x \sim \left(N^{-1}\widetilde{P}\right)X$$

Thus, setting $P = N^{-1}\widetilde{P}$ will result in $x \sim PX$.

Computer Exercise 3

It DOES looks like the points are centered around (0, 0) with mean distance 1 to (0, 0)
a.
The smallest singular value of M1 is: 0.015065
The smallest singular value of M2 is: 0.012149
Norm of M1*v1 is: 0.015065
Norm of M2*v2 is: 0.012149
The smallest singular values in both cases is at 10^-2 scale - close to 0.

b.
The projected points (using the calculated camera matrix) ARE close to the measured projections.
The camera centers and viewing directions DO look reasonable.
Computing the inner parameters of the first camers - we can know that these are the "true" parameters since we are given both the precise point model and the measured projections x. The ambiguity in Exercise 1 came from the fact that we didn't know the point model (3D-points), we only knew the measured projections. Thus, we could alter the 3D-points and the cameras and reach a different, yet still consistent, result.

c.
The RMSE for camera 1 WITH normalization is: 3.6176
The RMSE for camera 1 WITHOUT normalization is: 4.9026
The RMSE for camera 2 WITH normalization is: 3.2872
The RMSE for camera 2 WITHOUT normalization is: 3.8968
It is clear that normalization lowerd the RMSE and impored our results.

All plots can be viewed in the attached matlab_results.pdf output of the Matlab code.

Triangulation using DLT

Computer Exercise 4

a.
There is a small improvement when using the normalized points and matrices. For example, it is noticeable that there are less "fleeing" points. I should note that the improvement was not drastic.

b.
YES, I can distinguish the cup, the paper, and even the bottom of the second cup.

All plots can be viewed in the attached matlab_results.pdf output of the Matlab code.