

Camera calibration

Morten R. Hannemose, mohan@dtu.dk

February 25, 2022

02504 Computer vision course lectures,
DTU Compute, Kgs. Lyngby 2800, Denmark



Learning objectives

After this lecture you should be able to:

- implement and use the direct linear transformation to calibrate a pinhole camera
- implement and use Zhang's method (2000) to calibrate a pinhole camera using checkerboards

Presentation topics

Direct linear transformations

Zhang's method (2000)

Non-linear calibration

Practical remarks

Culmination of previous weeks

- Pinhole camera model
- Homogeneous coordinates
- Homographies
- Linear algorithms
- Calibration

Direct linear transformations

Direct linear transformation

Start from the projection equation

$$\mathbf{q}_i = \mathcal{P} \mathbf{Q}_i$$
$$\begin{bmatrix} sx_i \\ sy_i \\ s \end{bmatrix} = \mathcal{P} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}$$

then rearrange into the form $\mathbf{B}^{(i)} \text{flatten}(\mathcal{P}^\top) = \mathbf{0}$.

I use the form explained in LN: 2.7 Camera Resection.

Direct linear transformation i

$$\begin{aligned} \mathbf{q}_i &= \mathcal{P} \mathbf{Q}_i, \\ 0 &= \mathbf{q}_i \times \mathcal{P} \mathbf{Q}_i, \\ &= [\mathbf{q}_i]_{\times} \mathcal{P} \mathbf{Q}_i, \\ &= \mathbf{B}^{(i)} \text{flatten}(\mathcal{P}^T), \end{aligned}$$

where (continue to next slide) ...

Direct linear transformation ii

$$0 = \mathbf{B}^{(i)} \text{flatten}(\mathcal{P}^\top),$$

$$\mathbf{B}^{(i)} = \begin{bmatrix} 0 & -X_i & X_i y_i & 0 & -Y_i & Y_i y_i & 0 & -Z_i & Z_i y_i & 0 & -1 & y_i \\ X_i & 0 & -X_i x_i & Y_i & 0 & -Y_i x_i & Z_i & 0 & -Z_i x_i & 1 & 0 & -x_i \\ -X_i y_i & X_i x_i & 0 & -Y_i y_i & Y_i x_i & 0 & -Z_i y_i & Z_i x_i & 0 & -y_i & x_i & 0 \end{bmatrix},$$

$$= \mathbf{Q}_i \otimes [\mathbf{q}_i]_\times$$

$$\text{flatten}(\mathcal{P}^\top) = [\mathcal{P}_{11} \ \mathcal{P}_{21} \ \mathcal{P}_{31} \ \mathcal{P}_{12} \ \mathcal{P}_{22} \ \mathcal{P}_{32} \ \mathcal{P}_{13} \ \mathcal{P}_{23} \ \mathcal{P}_{33} \ \mathcal{P}_{14} \ \mathcal{P}_{24} \ \mathcal{P}_{34}]^\top$$

Direct linear transformation

Now let

$$\mathbf{0} = \mathbf{B} \text{ flatten}(\mathcal{P}^T),$$

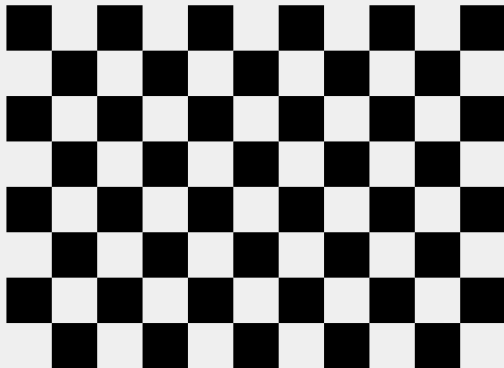
where

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(1)} \\ \mathbf{B}^{(2)} \\ \vdots \end{bmatrix},$$

and solve using SVD on \mathbf{B} .

Zhang's method (2000)

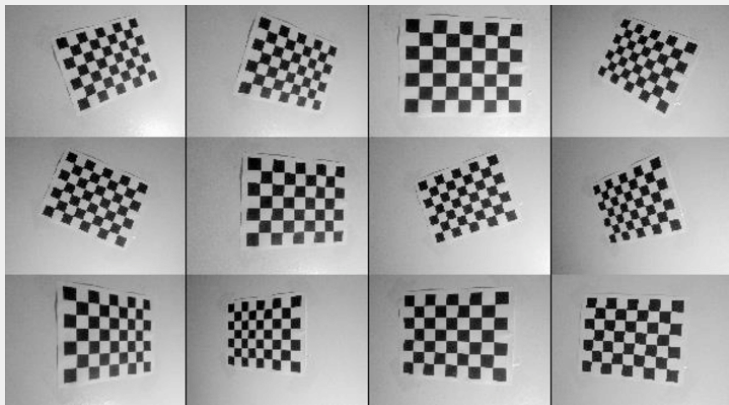
Using a checkerboard



www.calib.io | 8x11 | Checker Size: 15 mm.

Using a checkerboard

View the checkerboard in different poses:



Using a checkerboard

Problem:

Each view i has a different rotation \mathbf{R}_i and translation \mathbf{t}_i

How do we find all \mathcal{P}_i ?

Zhang's method

First, assume all checkerboard corners are in the $Z = 0$ plane:

$$\mathbf{Q}_j = \begin{bmatrix} X_j \\ Y_j \\ 0 \end{bmatrix}$$

Simplifying the projection equation

Let $\mathbf{r}_i^{(c)}$ is the c^{th} column of \mathbf{R}_i . Now projection is

$$\mathbf{q}_{ij} = \mathcal{P}_i \mathbf{Q}_j = \mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{r}_i^{(3)} & \mathbf{t}_i \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{bmatrix},$$

Simplifying the projection equation

Let $\mathbf{r}_i^{(c)}$ is the c^{th} column of \mathbf{R}_i . Now projection is

$$\begin{aligned} \mathbf{q}_{ij} = \mathcal{P}_i \mathbf{Q}_j &= \mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{r}_i^{(3)} & \mathbf{t}_i \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{bmatrix}, \\ &= \mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{t}_i \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}. \end{aligned}$$

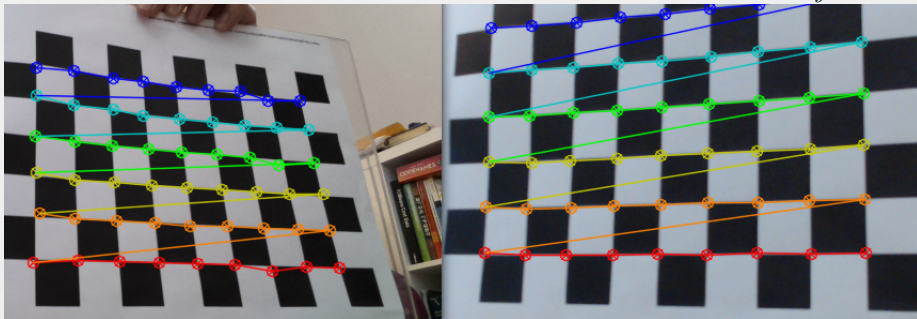
From projections to homographies

$$\begin{aligned} \mathbf{q}_{ij} &= \underbrace{\mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{t}_i \end{bmatrix}}_{\mathbf{H}_i} \underbrace{\begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}}_{\tilde{\mathbf{Q}}_j} \\ &= \mathbf{H}_i \tilde{\mathbf{Q}}_j \end{aligned}$$

The homographies \mathbf{H}_i can be determined from the plane-plane correspondence $\tilde{\mathbf{Q}}_j$ to \mathbf{q}_{ij} (week 2).

Corner correspondences

Need to find the **unique** positions of corners q_{ij} .



**Find all H_i from corners \tilde{Q}_j and
projections q_{ij} with $q_{ij} = H_i \tilde{Q}_j$
For example, using SVD.**

From homographies to the camera matrix

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_i^{(1)} & \mathbf{h}_i^{(2)} & \mathbf{h}_i^{(3)} \end{bmatrix} = \lambda \mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{t}_i \end{bmatrix}.$$

$\mathbf{r}_i^{(1)}$ and $\mathbf{r}_i^{(2)}$ are orthonormal, i.e.

$$\begin{aligned} \mathbf{r}_i^{(1)\top} \mathbf{r}_i^{(1)} &= \mathbf{r}_i^{(2)\top} \mathbf{r}_i^{(2)} = 1, \\ \mathbf{r}_i^{(1)\top} \mathbf{r}_i^{(2)} &= \mathbf{r}_i^{(2)\top} \mathbf{r}_i^{(1)} = 0. \end{aligned}$$

From homographies to the camera matrix

Express $\mathbf{r}_i^{(\alpha)}$ using $\mathbf{h}_i^{(\alpha)}$:

$$\begin{aligned}\mathbf{h}_i^{(\alpha)} &= \lambda \mathbf{K} \mathbf{r}_i^{(\alpha)}, \Leftrightarrow \\ \mathbf{K}^{-1} \mathbf{h}_i^{(\alpha)} &= \lambda \mathbf{r}_i^{(\alpha)}.\end{aligned}$$

From homographies to the camera matrix

Express $\mathbf{r}_i^{(\alpha)}$ using $\mathbf{h}_i^{(\alpha)}$:

$$\begin{aligned}\mathbf{h}_i^{(\alpha)} &= \lambda \mathbf{K} \mathbf{r}_i^{(\alpha)}, \Leftrightarrow \\ \mathbf{K}^{-1} \mathbf{h}_i^{(\alpha)} &= \lambda \mathbf{r}_i^{(\alpha)}.\end{aligned}$$

Now the constraints from the previous slide are:

$$\begin{aligned}\mathbf{h}_i^{(1)\top} \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_i^{(2)} &= 0, \\ \mathbf{h}_i^{(1)\top} \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_i^{(1)} &= \mathbf{h}_i^{(2)\top} \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_i^{(2)} = \lambda^2.\end{aligned}$$

We have found constraints on the camera matrix! 🥳

Number of constraints

- Two constraints doesn't seem that impressive?

Number of constraints

- Two constraints doesn't seem that impressive?
- Homography has eight degrees of freedom
- Pose of checkerboard has six (3 rotation, 3 translation)
- One homography can only fix two degrees of freedom of camera matrix.

Define some new variables i

How to put into practice?

Define the matrix:

$$\mathbf{B} = \mathbf{K}^{-\top} \mathbf{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix},$$

$$\mathbf{b} = [B_{11} \quad B_{12} \quad B_{22} \quad B_{13} \quad B_{23} \quad B_{33}]^{\top},$$

Define some new variables ii

Now define $\mathbf{v}_i^{(\alpha\beta)}$ such that

$$\mathbf{v}_i^{(\alpha\beta)} \mathbf{b} = \mathbf{h}_i^{(\alpha)T} \mathbf{B} \mathbf{h}_i^{(\beta)}.$$

Then $\mathbf{v}_i^{(\alpha\beta)}$ must be a 1×6 vector given by

$$\mathbf{v}_i^{(\alpha\beta)} = [H_i^{(1\alpha)} H_i^{(1\beta)}, \quad H_i^{(1\alpha)} H_i^{(2\beta)} + H_i^{(2\alpha)} H_i^{(1\beta)}, \quad H_i^{(2\alpha)} H_i^{(2\beta)}, \quad \dots \\ H_i^{(3\alpha)} H_i^{(1\beta)} + H_i^{(1\alpha)} H_i^{(3\beta)}, \quad H_i^{(3\alpha)} H_i^{(2\beta)} + H_i^{(2\alpha)} H_i^{(3\beta)}, \quad H_i^{(3\alpha)} H_i^{(3\beta)}],$$

where $H_i^{(rc)}$ is the element in row r and column c of \mathbf{H}_i .

From homographies to the camera matrix

Recall our constraints:

$$\begin{aligned} \mathbf{h}_i^{(1)\top} \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_i^{(1)} &= \mathbf{h}_i^{(2)\top} \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_i^{(2)} \\ \mathbf{h}_i^{(1)\top} \mathbf{B} \mathbf{h}_i^{(1)} - \mathbf{h}_i^{(2)\top} \mathbf{B} \mathbf{h}_i^{(2)} &= 0 = \\ &\left(\mathbf{v}_i^{(11)} - \mathbf{v}_i^{(22)} \right) \mathbf{b} = 0. \end{aligned}$$

Now we can express our constraints in matrix-form

$$\begin{bmatrix} \mathbf{v}_i^{(12)} \\ \mathbf{v}_i^{(11)} - \mathbf{v}_i^{(22)} \end{bmatrix} \mathbf{b} = \mathbf{0}$$

From homographies to the camera matrix

For all checkerboard poses

$$V\mathbf{b} = \begin{bmatrix} \mathbf{v}_1^{(12)} \\ \mathbf{v}_1^{(11)} - \mathbf{v}_1^{(22)} \\ \mathbf{v}_2^{(12)} \\ \mathbf{v}_2^{(11)} - \mathbf{v}_2^{(22)} \\ \vdots \end{bmatrix} \mathbf{b} = \mathbf{0}$$

When \mathbf{b} is found then \mathbf{K} can be determined using the formulas in Zhang's paper.

Find the camera matrix K through b using $Vb = 0$ and SVD, where V is built from the homographies H_i .

From homographies to poses

Recall: $\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_i^{(1)} & \mathbf{h}_i^{(2)} & \mathbf{h}_i^{(3)} \end{bmatrix} = \lambda \mathbf{K} \begin{bmatrix} \mathbf{r}_i^{(1)} & \mathbf{r}_i^{(2)} & \mathbf{t}_i \end{bmatrix}$

Now we can recover \mathbf{R}_i and \mathbf{t}_i :

$$\mathbf{r}_i^{(1)} = \frac{1}{\lambda} \mathbf{K}^{-1} \mathbf{h}_i^{(1)},$$

$$\mathbf{r}_i^{(2)} = \frac{1}{\lambda} \mathbf{K}^{-1} \mathbf{h}_i^{(2)},$$

$$\mathbf{t}_i = \frac{1}{\lambda} \mathbf{K}^{-1} \mathbf{h}_i^{(3)},$$

$$\mathbf{r}_i^{(3)} = \mathbf{r}_i^{(1)} \times \mathbf{r}_i^{(2)},$$

where $\lambda = \|\mathbf{K}^{-1} \mathbf{h}_i^{(1)}\|_2 = \|\mathbf{K}^{-1} \mathbf{h}_i^{(2)}\|_2$.

Last words on the poses

$$\mathbf{t}_i = \frac{1}{\lambda} \mathbf{K}^{-1} \mathbf{h}_i^{(3)}$$

What happens if $t_{iz} < 0$? And what does that mean?

Can we ensure correctness?

Last words on the poses

$$\mathbf{t}_i = \frac{1}{\lambda} \mathbf{K}^{-1} \mathbf{h}_i^{(3)}$$

What happens if $t_{iz} < 0$? And what does that mean?

Can we ensure correctness?

Run the code again with $-\mathbf{H}_i$ (flipped sign).

Non-linear calibration

Least-squares method

With projection equation

$$\mathbf{q}_{ij} = \mathbf{K} \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \end{bmatrix} \mathbf{Q}_j$$

then

$$E = \sum_{i,j} \left\| \text{dist} \left(\pi \left(\mathbf{K} \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \end{bmatrix} \mathbf{Q}_j \right) \right) - \pi(\mathbf{q}_{ij}) \right\|^2$$

Solution: minimize E w.r.t. \mathbf{K} , \mathbf{R}_i , \mathbf{t}_i and lens distortion.

Practical remarks

What images should we take?

- Our two constraints per image are based on $\mathbf{r}_i^{(1)}$ and $\mathbf{r}_i^{(2)}$.
- Two parallel checkerboards express the same constraints.

What images should we take?

- Our two constraints per image are based on $\mathbf{r}_i^{(1)}$ and $\mathbf{r}_i^{(2)}$.
- Two parallel checkerboards express the same constraints.
 - at least when we don't consider lens distortion
- Make sure to rotate the checkerboards so they are not parallel.
- Make the checkerboards take up as as much of the frame as possible!

How many images?

- Without lens distortion: In theory at least three.
- In practice: It depends...
 - Some people use 2000+ images for a single calibration¹
 - Look at the reprojection error

¹Schops, Thomas, et al. "Why having 10,000 parameters in your camera model is better than twelve." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

Learning objectives

After this lecture you should be able to:

- implement and use the direct linear transformation to calibrate a pinhole camera
- implement and use Zhang's method (2000) to calibrate a pinhole camera using checkerboards

Exercise time!