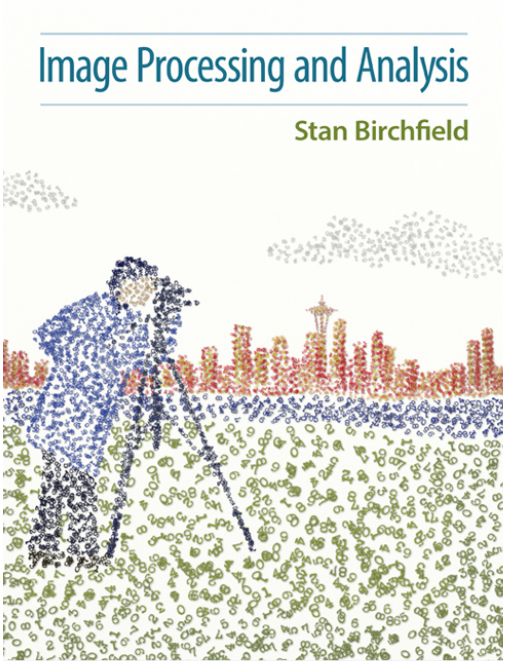
University
of Stavanger

Prof. Kjersti Engan

ELE510 Image processing and computer vision

Camera and calibration, (chap 13.4, 13.5 Birchfield) 2020



Parts in course presentations is material from Cengage learning. It can be used for teaching, and it can be share it with students on access controlled web-sites (Canvas) for use in THIS course. **But it should not be copied or distributed further in any way** (by you or anybody).

(13.5) Camera calibration

- Goal of calibration is to estimate intrinsic and extrinsic parameters from one or several images.
- Calibrating a pinhole camera involves estimating 11 parameters because the matrix $\mathbf{P_M}$ contains 12 elements, but it is unique only up to an unknown nonzero scaling factor. 3D-> 2D means we loose depth, therefore one free parameter.
- It is common to set Pm34 = 1 as a normalization of $\mathbf{P_M}$

2D image [pixels]

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \propto \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\mathbf{P}_{\{3\times 4\}}} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

3D world [m]

11 param.

$p_{34} = 1$

Camera calibration

$$\mathbf{P}_{\{3\times 4\}} = \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}_{\{3\times 3\}}} [\mathbf{R}_{\{3\times 3\}} \ \mathbf{t}_{\{3\times 1\}}]$$

- 6 values for the Euclidean rotation and translation (3 values each), and 5 values for the internal calibration matrix **K.**

- These are what we called the **extrinsic** and **intrinsic parameters**, respectively.

Camera calibration

- Camera calibration is the operation to *measure* and *compute* an estimate of the

1. **Intrinsic camera parameters**, and
2. **Extrinsic camera parameters**

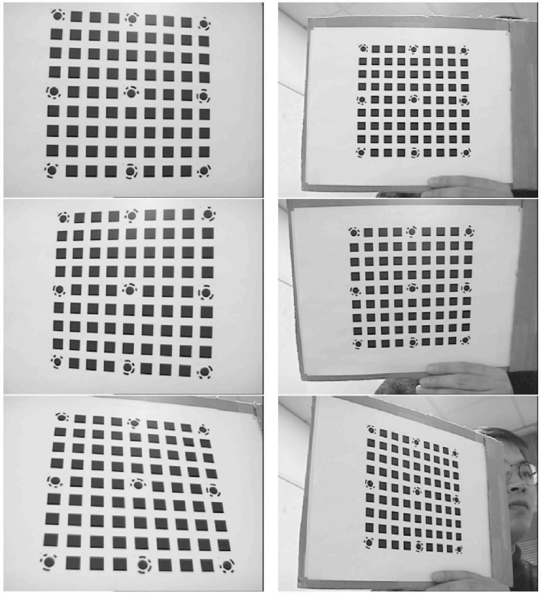
- Camera calibration is *necessary in machine-* and *robot vision* because we want to
- **Reconstruct** a world model of the scene or objects in the scene, and or
- **Measure** physical distances in the scene, and or
- **Interact** with objects in the scene (robot, hand-eye coordination).

- The most popular technique for calibrating the intrinsic camera parameters is **Zhang's algorithm**.

Zhangs algorithm - overview

- Require a planar calibration target with known coordinates /dimensions in meter. Typically used checkerboard
- Require 6-10 (more than 3) different images captured of the target at different positions.
- Find feature points (corners in the chekerboard, use Harris/SIFT) in the images (x,y)coordinates in the image, along with knowing the physical distance of the corner points in meters.
- From this we can estimate the intrinsic parameters K. (will look at some more details)
- Can also find extrinsic parameters. This is not always intersting

Checkerboard calibration pattern



(a) (b)

“Checkerboard” pattern suggested by Zhengyou Zhang (Microsoft Research) 1999

A plane viewed from different unknown orientations.

Zhangs Calibration Algorithm

- The projection of a target point onto the image plane is given by:

planar object

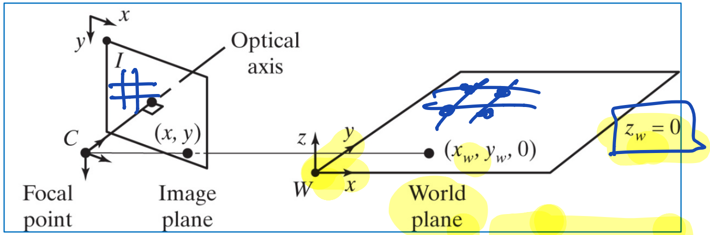
$$\begin{bmatrix}x\\y\\1\end{bmatrix} \propto \underbrace{\begin{bmatrix}\alpha&\gamma&u_0\\0&\beta&v_0\\0&0&1\end{bmatrix}}_{\mathbf{K}_{\{3\times3\}}} \underbrace{[\mathbf{r}_1\quad \mathbf{r}_2\quad \mathbf{r}_3\quad \mathbf{t}]}_{[\mathbf{R}_{\{3\times3\}}\quad \mathbf{t}_{\{3\times1\}}]} \begin{bmatrix}x_w\\y_w\\0\\1\end{bmatrix} = \underbrace{\mathbf{K}\,[\mathbf{r}_1\quad \mathbf{r}_2\quad \mathbf{t}]}_{\mathbf{H}_{\{3\times3\}}} \begin{bmatrix}x_w\\y_w\\1\end{bmatrix}$$

$[h_1\quad h_2\quad h_3]$

$$\boxed{\mathbf{K}\mathbf{r}_i \propto \mathbf{h}_i \implies \mathbf{r}_i \propto \mathbf{K}^{-1}\mathbf{h}_i}$$

$$\Rightarrow \boxed{\mathbf{r}_1^T\mathbf{r}_2 = 0}$$

$$\Rightarrow \boxed{\mathbf{r}_1^T\mathbf{r}_1 = \mathbf{r}_2^T\mathbf{r}_2} = 1$$



Define image of the absolute conic IAC: $\mathbf{w}_\infty = \mathbf{K}^{-T}\mathbf{K}^{-1} = (\mathbf{K}\mathbf{K}^T)^{-1}$

Combine with equations in red frame from prev. slide:

$r_1^T \cdot r_2 = 0$

$$\mathbf{h}_1^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_2 = 0 \Rightarrow \boxed{\Longrightarrow \mathbf{h}_1^T\mathbf{w}_\infty\mathbf{h}_2 = 0}.$$

$$\mathbf{h}_1^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_2 \quad \boxed{\Longrightarrow \mathbf{h}_1^T\mathbf{w}_\infty\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{w}_\infty\mathbf{h}_2}$$

$r_1^T \cdot r_1 = r_2^T \cdot r_2$

$r_i = K^{-1} h_i$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \propto \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}_{\{3\times3\}}} \underbrace{[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}]}_{[\mathbf{R}_{\{3\times3\}} \quad \mathbf{t}_{\{3\times1\}}]} \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} = \underbrace{\mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}]}_{\mathbf{H}_{\{3\times3\}}} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}$$

$$[h_1 \quad h_2 \quad h_3]$$

Define image of the absolute conic IAC: $\mathbf{w}_\infty = \mathbf{K}^{-T}\mathbf{K}^{-1} = (\mathbf{K}\mathbf{K}^T)^{-1}$

Combine with equations in red frame from prev. slide:

1x3 3x3 3x1

$$\mathbf{h}_1^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_2 = 0 \qquad \boxed{\Longrightarrow \ \mathbf{h}_1^T\mathbf{w}_\infty\mathbf{h}_2 = 0} \quad (1)$$

$$\mathbf{h}_1^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{h}_2 \quad \boxed{\Longrightarrow \ \mathbf{h}_1^T\mathbf{w}_\infty\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{w}_\infty\mathbf{h}_2} \quad (2)$$

$a_{11}^T \cdot w$ $a_{22}^T \cdot w$

Rewrite! **aij** are known from **H** matrix, **wi** contains unknown from $\mathbf{w}_\infty$

$$\mathbf{a}_{12}^T\mathbf{w} = 0 \quad (1)$$

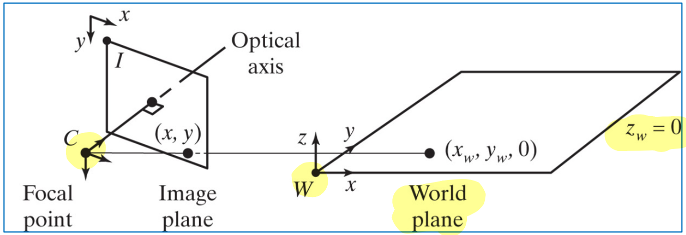
$$(\mathbf{a}_{11} - \mathbf{a}_{22})^T\mathbf{w} = 0 \quad (2)$$

From these we can estimate $\mathbf{w}_\infty$ when we have a number of different H´s. We have as many H´s as we have images.

From $\mathbf{w}_\infty$ we can find **K**

How to estimate the homography matrix H?

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \propto \mathbf{H} \begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix} = \boxed{[\mathbf{h}_1 \mathbf{h}_2 \mathbf{h}_3]} \begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix}$$



Using SIFT / Harris on calibration images we find corresponding pairs of points:

$$\boxed{(x_i\,, y_i) \leftrightarrow (x_{w_i}\,, y_{w_i})}$$

From these we can estimate **H**.
Normalized Direct Linear Troansfomration (DLT) algorithm can be used.
We will not look at those details, but with a number of points this can be done.

Have as many **H** estimate as we have images.

When we have K and Hj, we can find Rj and tj for each image j

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \propto \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}_{\{3\times 3\}}} \underbrace{[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}]}_{[\mathbf{R}_{\{3\times 3\}} \quad \mathbf{t}_{\{3\times 1\}}]} \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} = \underbrace{\mathbf{K}\,[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}]}_{\mathbf{H}_{\{3\times 3\}}} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}$$

$$\underbrace{K \cdot [r_1 \quad r_2 \quad t] \cdot \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}}$$

$\Downarrow$

[

Thereafter we can do a reprojection , i.e. With these parameters where would the physical point be on the image plane? Is that far from the true image point? -> Reprojection error:

reprojection

$$RepErr = \sum_i \sum_j ||\mathbf{x}_{ij} - g(\mathbf{K}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{x}_w^j, \phi)||^2$$

-> lens distortion

⇒ Can update iteratively the intrinsic and extrinsic parameters minimizing the reprojection error -> bundle adjustment

ALGORITHM 13.12 Zhang's camera calibration routine

CALIBRATECAMERAZHANG

Input: $n \geq 3$ images of a known planar target at different orientations and directions
Output: intrinsic parameters $\mathbf{K}$, lens distortion parameters ϕ, and extrinsic parameters

for each image **do**
    Detect image features corresponding to known points on target.
    Compute the homography using the normalized DLT algorithm.

$\underset{\sim}{H_j}$ imag j

Stack the entries from the homographies into matrix $\mathbf{A}$.
Solve $\mathbf{A}\mathbf{w} = \mathbf{0}$ for $\mathbf{w}$, then reshape into ω_∞.
Compute the five intrinsic parameters of $\mathbf{K}$ from $\boxed{\omega_\infty}$ using either
    Cholesky decomposition, or
    Equation (13.143) for v_0, then Equation (13.142) for λ, then Equations (13.144)–(13.147).
Compute extrinsic parameters $\mathbf{R}_i$ and $\mathbf{t}_i$ for each image $i = 1, \ldots, n$ using Equations (13.148)–(13.151).
Using these results as a starting point, minimize Equation (13.152) to perform bundle adjustment.

reprojection error