Camera calibration - 3D calibration rig

consider simpler case first (no distortion)

$$\underline{\mathbf{x}} = \underline{\mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}}\,\underline{\mathbf{X}}$$

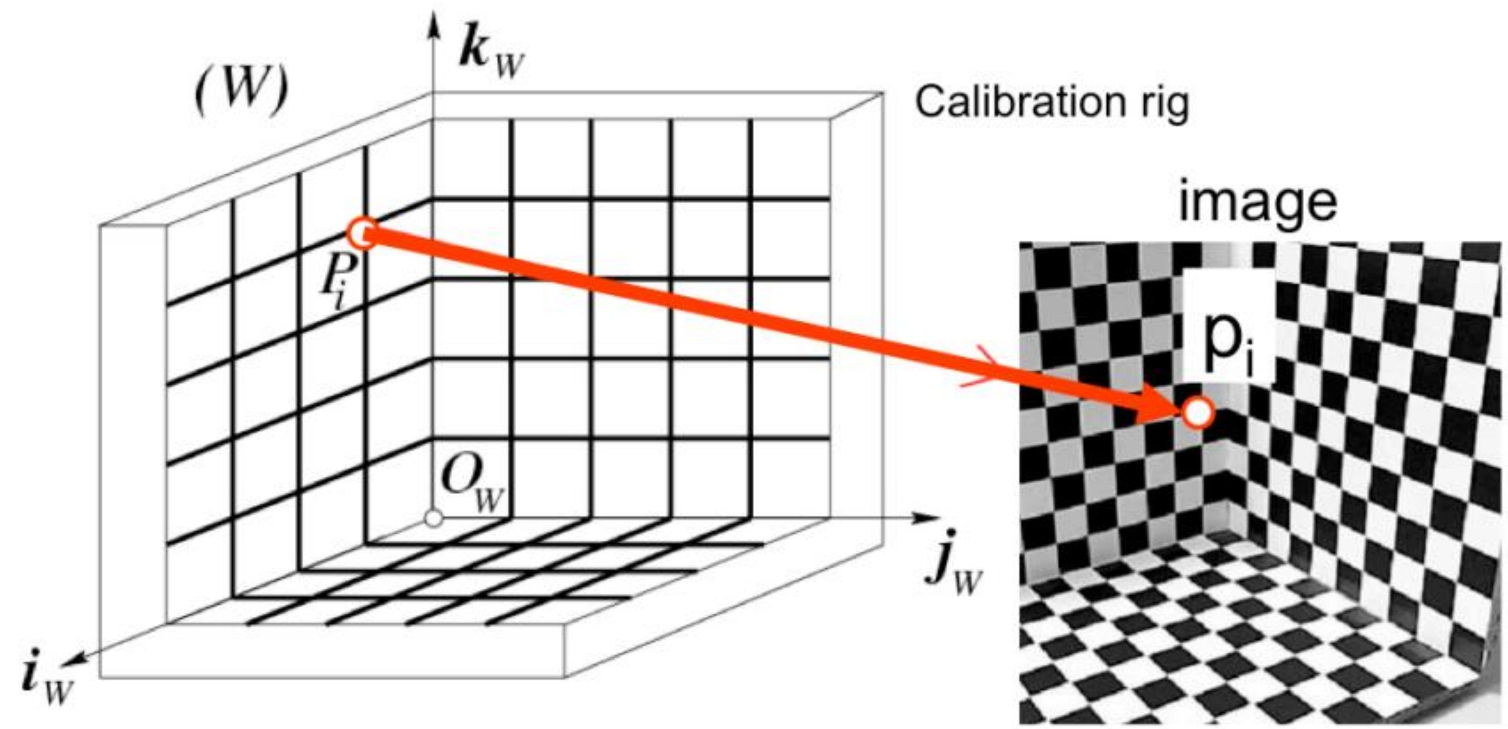
$$\underline{p_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix}} = \underline{M}\,\underline{P_i} = \begin{bmatrix} \frac{m_1 P_i}{m_3 P_i} \\ \frac{m_2 P_i}{m_3 P_i} \end{bmatrix}$$

known from calibration rig

(m_1, m_2, m_3 are rows of M)

$$u_i(m_3 P_i) - m_1 P_i = 0$$
$$v_i(m_3 P_i) - m_2 P_i = 0$$

each correspondence gives 2 equations, while M has 11 DoF



$$u_1(m_3 P_1) - m_1 P_1 = 0$$
$$v_1(m_3 P_1) - m_2 P_1 = 0$$
$$\vdots$$
$$u_n(m_3 P_n) - m_1 P_n = 0$$
$$v_n(m_3 P_n) - m_2 P_n = 0$$

we need at least 6 correspondences, but often we use more as measurement are often noisy

Camera calibration - 3D calibration rig

$$\begin{bmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ & \vdots & \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{bmatrix} \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix} = \mathbf{P}m = 0$$

$$\underset{m}{\text{minimize}} \quad \|\mathbf{P}m\|^2$$

$$\text{subject to} \quad \|m\|^2 = 1$$

to avoid trivial solution m=0

$\forall k \in \mathbb{R},\ km$ is also a solution

can be solved by singular value decomposition (SVD)

$$P = UDV^T$$

set m equal to the last column of V

$$M = K\,[R \mid t] = [KR \mid KRT] = [\underset{3\text{x}3}{C} \mid \underset{3\text{x}1}{CT}]$$

reformatting m into matrix M

Camera calibration - 3D calibration rig

$$\begin{bmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ & \vdots & \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{bmatrix} \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix} = \mathbf{P}m = 0$$

$$\underset{m}{\text{minimize}} \quad \|\mathbf{P}m\|^2$$

$$\text{subject to} \quad \|m\|^2 = 1$$

can be solved by singular value decomposition (SVD)

$$P = UDV^T$$

set m equal to the last column of V

$$\begin{array}{c} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{3x3} \quad \mathbf{3x1} \\ M = K\,[R \mid t] = [KR \mid KRT] = [C \mid CT] \end{array}$$

reformatting m into matrix M

$$C = KR$$

where K is upper triangular matrix

R is orthogonal matrix ($R^T R = 1$)

we can use QR decomposition to get K (intrinsic) and R (rotation)

to get T (translation), simply multiply C^{-1} to CT

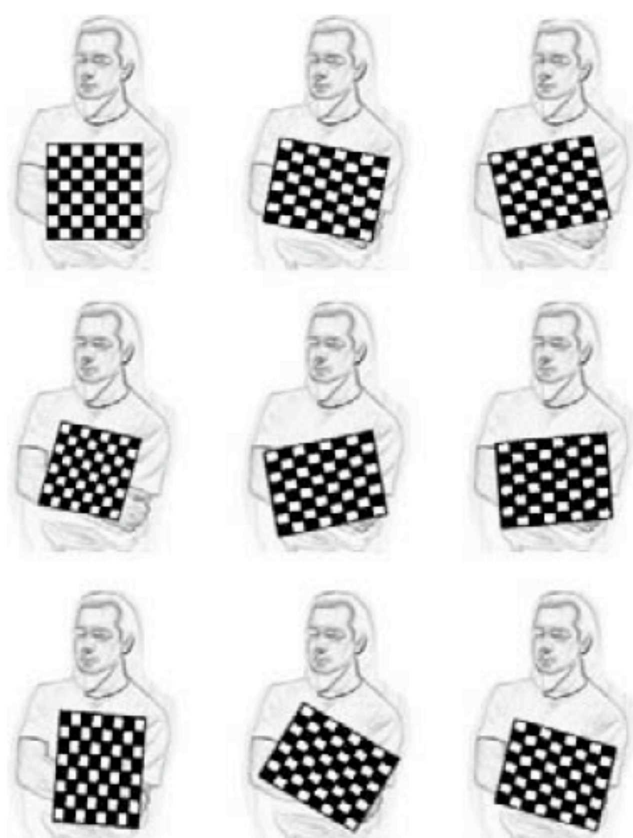
many other calibration methods and also we need to consider distortion!

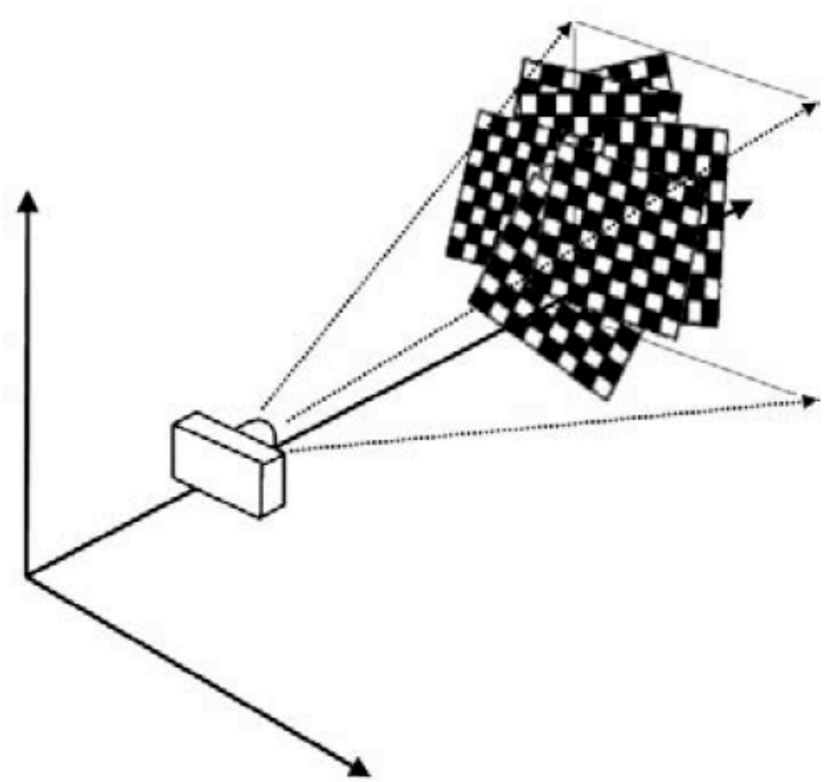
Camera calibration - 2D calibration chessboard

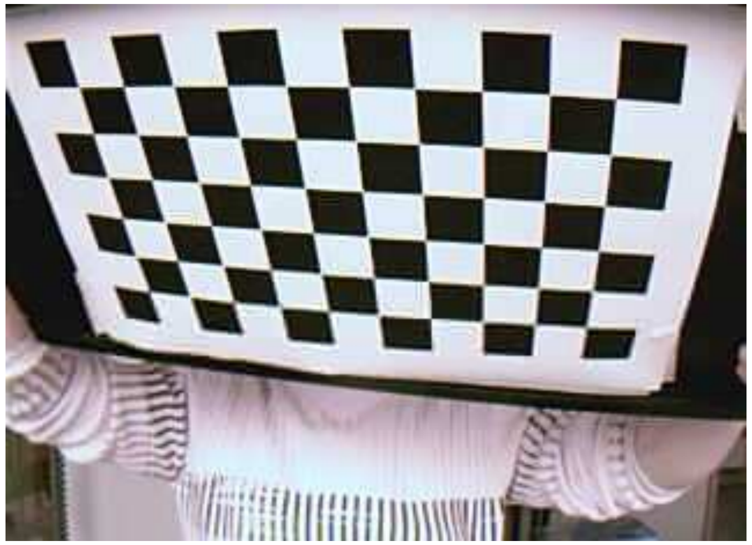
- 3D calibration rig hard to make (perpendicularity across 3 planes)

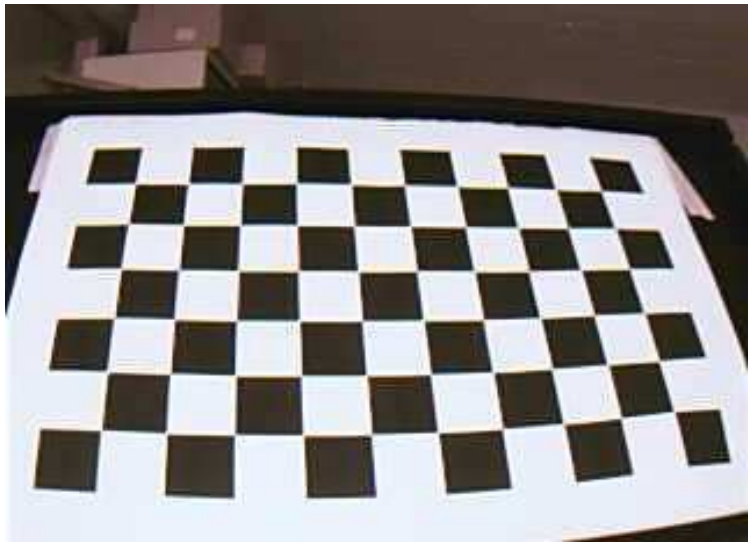
- Use a 2D pattern (e.g. a chessboard)

- Trick: set the world coordinate system to the corner of chessboard









Camera calibration - 2D calibration chessboard

- Trick: set the world coordinate system to the corner of chessboard

- Now: all points on the chessboard line in on plane!

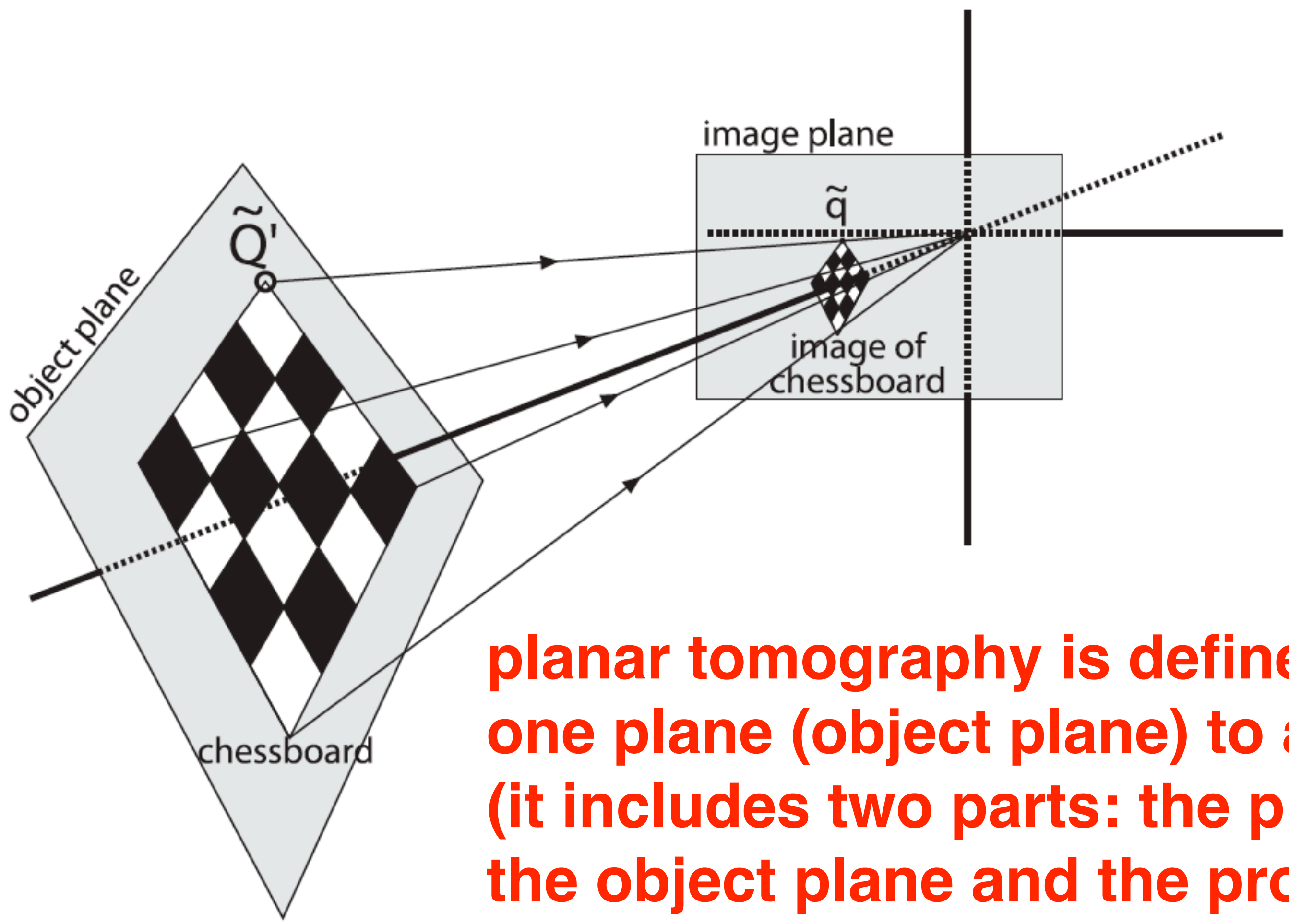
$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} f/s_x & 0 & o_x \\ 0 & f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cancel{r_{13}} & t_1 \\ r_{21} & r_{22} & \cancel{r_{23}} & t_2 \\ r_{31} & r_{32} & \cancel{r_{33}} & t_3 \end{bmatrix} \begin{bmatrix} U \\ V \\ \boxed{W} \\ 1 \end{bmatrix}$$

since all points lie in a plane, their W component is 0 in world coordinates

we can thus delete 3rd column of extrinsic matrix

Camera calibration - 2D calibration chessboard

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \underbrace{\begin{bmatrix} f/s_x & 0 & o_x \\ 0 & f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix}}_{\textbf{Homography } H} \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$$



$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\rho}_{\text{up to a scale}} H \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$$

planar tomography is defined as a perspective mapping from one plane (object plane) to another (image plane)
(it includes two parts: the physical transformation that locates the object plane and the projection that has camera intrinsic matrix)

Camera calibration - 2D calibration chessboard

$$\boldsymbol{H} = (\mathbf{h}_1,\ \mathbf{h}_2,\ \mathbf{h}_3) = \underbrace{\begin{bmatrix} f/s_x & 0 & o_x \\ 0 & f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix}}_{(\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{t})}$$

$$\boldsymbol{H} = (\mathbf{h}_1,\ \mathbf{h}_2,\ \mathbf{h}_3) = \mathbf{K}(\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{t})$$

$\mathbf{r}_1 = \mathbf{K}^{-1}\mathbf{h}_1$ **and** $\mathbf{r}_2 = \mathbf{K}^{-1}\mathbf{h}_2$

➔ **note that** $(\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{r}_3)$ **form an orthonormal basis**

thus $\mathbf{r}_1^\top \mathbf{r}_2 = \mathbf{0}$, $||\mathbf{r}_1|| = ||\mathbf{r}_2|| = 1$

➔ $\mathbf{h}_1^\top \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_2 = 0$

$\mathbf{h}_1^\top \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_1 = \mathbf{h}_2^\top \mathbf{K}^{-\top} \mathbf{K}^{-1} \mathbf{h}_2$

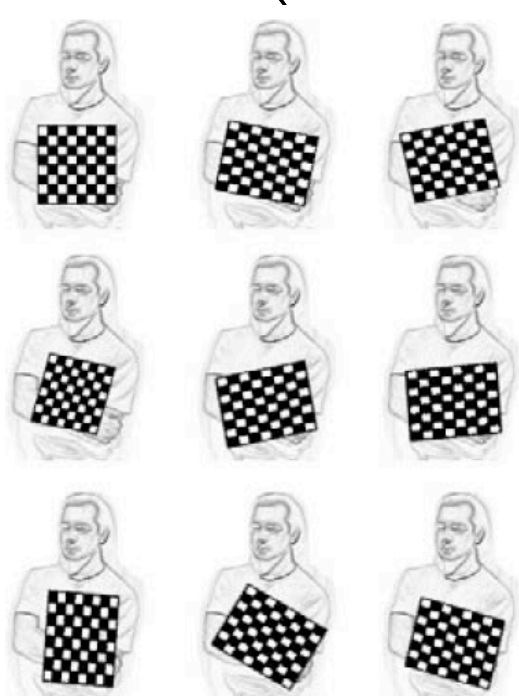
Camera calibration - 2D calibration chessboard

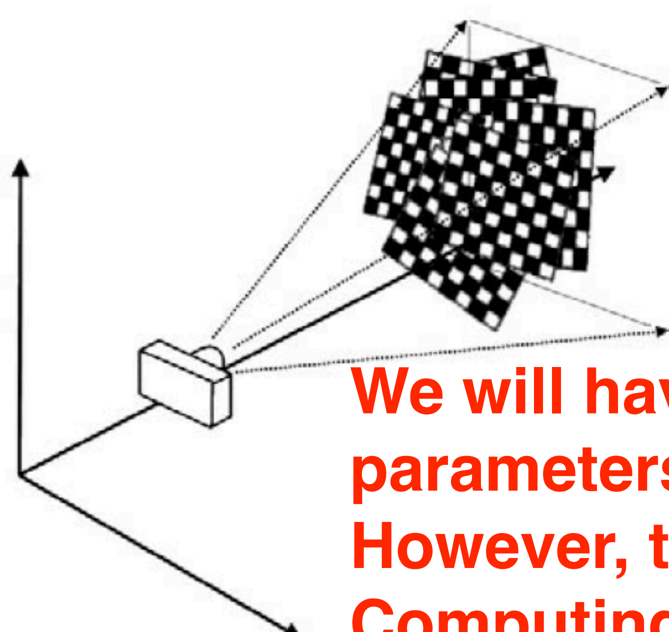
$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2 = 0$$
$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2$$

$\mathrm{B} := \mathrm{K}^{-\top}\mathrm{K}^{-1}$ is symmetric and positive definite

$$\mathrm{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

Note: K can be calculated from B using Cholesky factorization





We will have one set of new extrinsic parameters (define homography) for each view. However, the intrinsic parameters stay the same. Computing multiple H's from multiple views can lead us to a unique set of intrinsic parameters.

Camera calibration - 2D calibration chessboard

$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2 = 0$$

$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2$$

$$\mathrm{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

➔ **define** $\mathrm{b}=(b_{11},\ b_{12},\ b_{13},\ b_{22},\ b_{23},\ b_{33})$

leads to the system of $\mathrm{Vb} = 0$

- **Each plane gives 2 equations, B has 6 DoF, we need at least 3 planes from different views**

➔ **real measurements might be corrupted with noise, find solution that minimize least-squares error**

$$\mathbf{b} = \underset{\mathbf{b}}{argmin}\mathbf{Vb}$$

$$\mathbf{r}_1 = \lambda\mathbf{K}^{-1}\mathbf{h}_1$$

$$\mathbf{r}_2 = \lambda\mathbf{K}^{-1}\mathbf{h}_2$$

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

$$\mathbf{t} = \lambda\mathbf{K}^{-1}\mathbf{h}_3$$

$$H = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3\end{bmatrix} = \rho\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}\end{bmatrix} \Rightarrow \lambda = 1/||\mathbf{K}^{-1}\mathbf{h}_1||$$

Camera calibration - 2D calibration chessboard

$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2 = 0$$
$$\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_1 = \mathbf{h}_2^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2$$

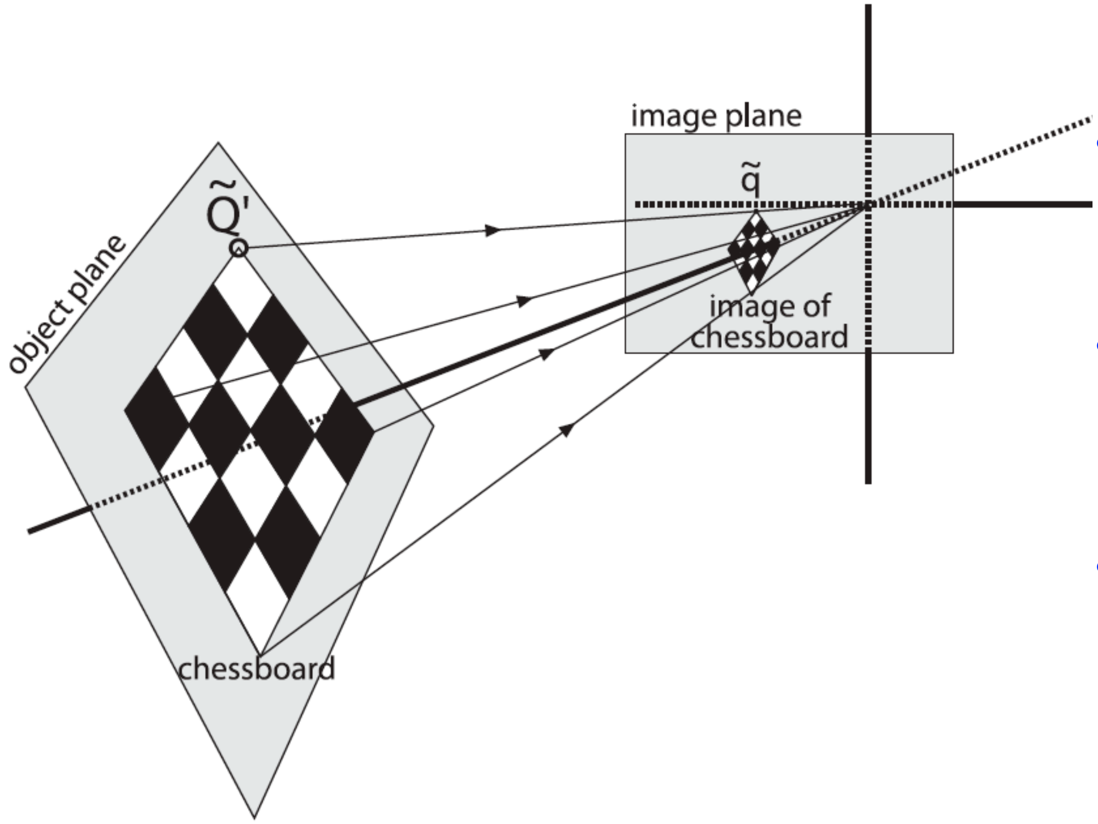
$$\mathrm{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

→ **define** $\mathrm{b}=(b_{11},\ b_{12},\ b_{13},\ b_{22},\ b_{23},\ b_{33})$

leads to the system of $\mathrm{Vb} = \mathbf{0}$

- **Each plane gives 2 equations, B has 6 DoF, we need at least 3 planes from different views**
- N **corners and** M **images of the chessboard (in different positions)**
- M **images of the chessboard provide** $2NM$ **constraints (**x**,** y **axes)**
- **Need to solve for** 4 **intrinsic parameters (ignore skew and distortion) and** $6M$ **extrinsic parameters (intrinsic parameters stay the same)**
- $2NM$ **must be ≥ to** $6M+4$ **to solve for these parameters (or equivalently** $(N\text{-}3)M \geq 2$**)**
- **If** $N = 5$ **then we need only** $M = 1$ **image!?**

Camera calibration - 2D calibration chessboard



- Only 4 points are needed to express that a planar perspective view can do
- A homography can yield at most 8 paras from 4 (x, y) pairs
- Per chessboard view, then, the equations can give us only 4 corners of information $(N\text{-}3)M \geq 2 \rightarrow (4\text{-}3)M \geq 2 \rightarrow M \geq 2$

- N corners and M images of the chessboard (in different positions)
- M images of the chessboard provide $2NM$ constraints (x, y axes)
- Need to solve for 4 intrinsic parameters (ignore skew and distortion) and $6M$ extrinsic parameters (intrinsic parameters stay the same)
- $2NM$ must be ≥ to $6M+4$ to solve for these parameters (or equivalently $(N\text{-}3)M \geq 2$)
- If $N = 5$ then we need only $M = 1$ image!?

in practice we need more images of larger chessboard to achieve better stability!