MAEG 5720: Computer Vision in Practice

Lecture 11:

Camera Model and Calibration

Dr. Terry Chang

2021-2022

Semester 1

香港中文大學
The Chinese University of Hong Kong

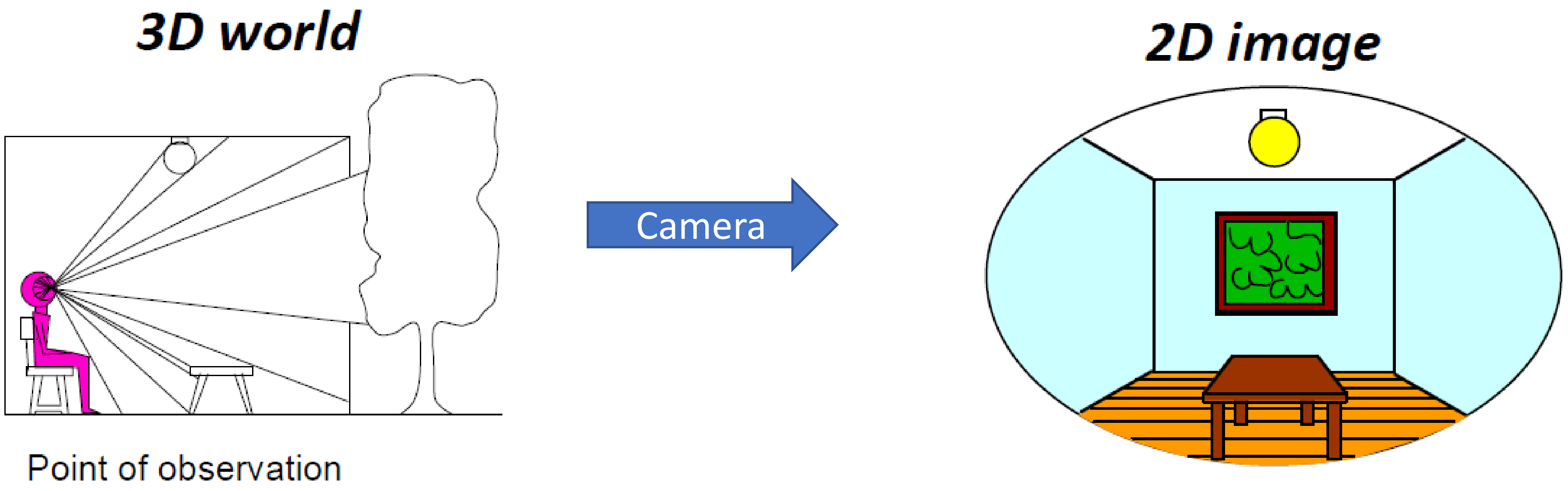
Department of Mechanical and Automation Engineering
機械與自動化工程學系

Content

- Camera models
  - Finite cameras
  - The projective cameras
  - The camera anatomy
- Camera Calibration
  - Direct Linear Transform
  - Zhang's methods

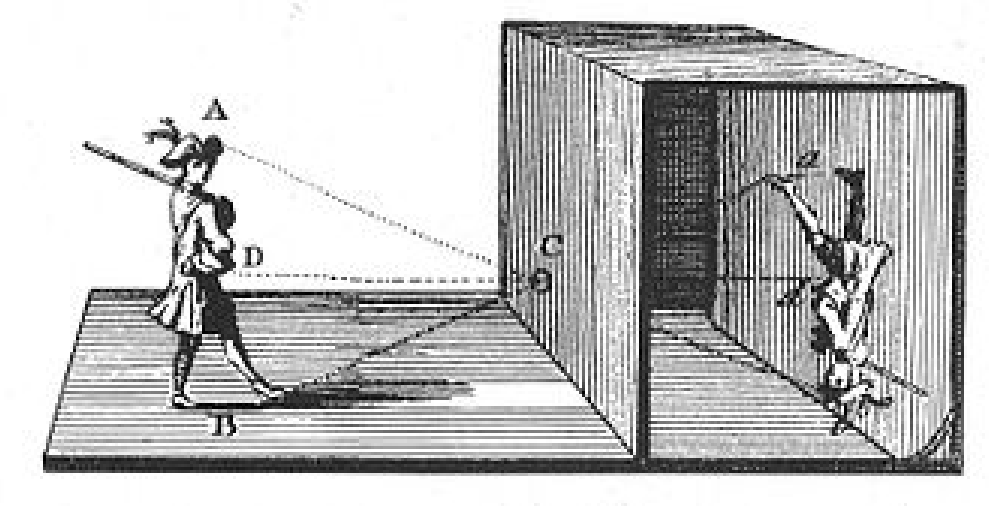
Camera Models

- A camera is a mapping between the 3D world (object space) and a 2D image through the central projection.



Finite Cameras

- We start with the most specialized and simplest camera model, which is the basic pinhole camera.
- Then progressively generalize this model through a series of gradations.



Pinhole Camera Model

- The centre of project is ***camera centre***, also known as the ***optical centre***.
- The line from the camera centre perpendicular to the image plane is called the ***principal axis*** or ***principal ray*** of the camera
- The point where the principal axis meets the image plane is called ***principal point***
- The plane through the camera centre parallel to the image plane is called ***principal plane***

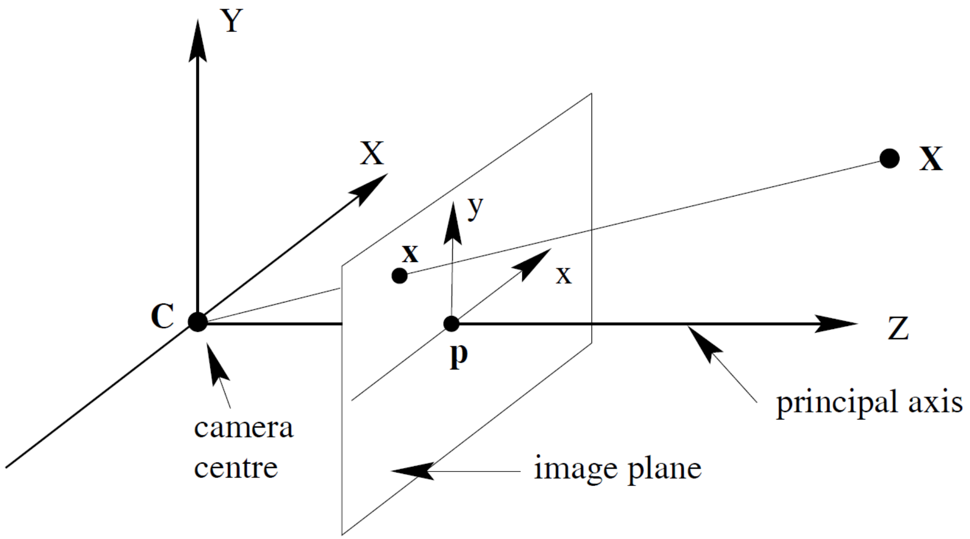


Image Credit: R. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision”

Camera Models – Pinhole Camera

- *By Similar Triangle*

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

- One can quickly computes the point

$$(X, Y, Z)^{\mathrm{T}} \rightarrow \left(\frac{fX}{Z}, \frac{fY}{Z}, f\right)^{\mathrm{T}}$$

- Describes the ***central projection*** mapping from ***world*** to ***image coordinates***

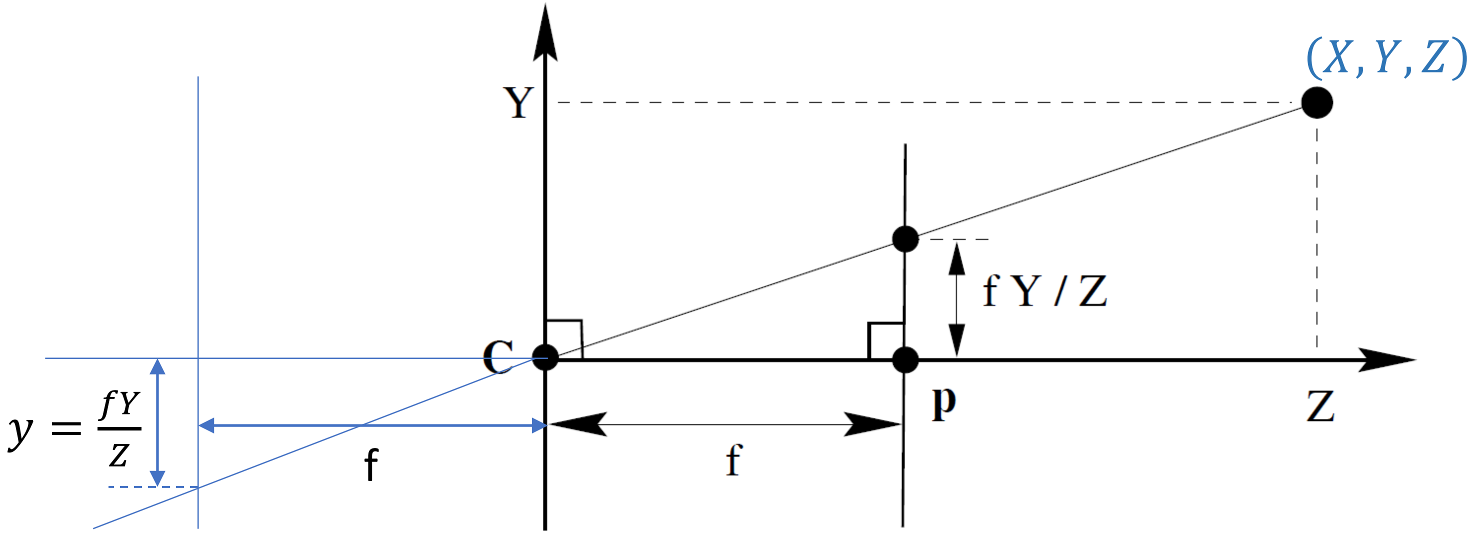


Image Credit: R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision"

Central projection using homogeneous coordinates

- Using homogenous coordinate representation, the central projection is very simply expressed as a linear mapping as follow:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

The matrix can be written as $diag(f, f, 1)[\mathrm{I} \mid \mathbf{0}]$ where

- $diag(f, f, 1)$ is diagonal matrix and

- $[\mathrm{I} \mid \mathbf{0}]$ represents a matrix divided up into a 3×3 identity matrix plus a column(Zero) vector

Central projection using homogeneous coordinates

- The notation
  - X for the world point represented by a homogenous 4-vector $(\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1)^{\mathrm{T}}$
  - x for the image point represented by a homogenous 3-vector $(\mathrm{x}, \mathrm{y}, 1)^{\mathrm{T}}$
  - P for camera projection matrix represented by a 3×4 matrix.

- The mapping becomes

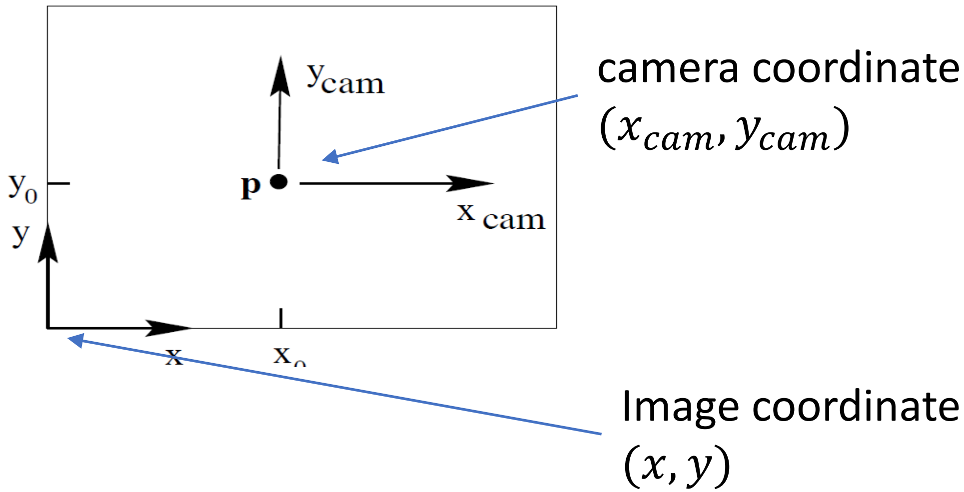
$$\mathrm{x} = \mathrm{PX}$$

- Which defines the camera matrix for pinhole model of central project as

$$\mathrm{P} = diag(f, f, 1)\ [\mathrm{I} \mid \mathbf{0}]$$

The Principal point offset

- The pinhole camera model assumed that the origin of coordinates in the image plane is at the principal point. In practice, it may not be.



- In general, the mapping is

$$(X, Y, Z)^{\mathrm{T}} \rightarrow \left(\frac{f\mathrm{X}}{\mathrm{Z}} + p_x, \frac{f\mathrm{Y}}{\mathrm{Z}} + p_y\right)^{\mathrm{T}}$$

where $\left(p_x, p_y\right)^{\mathrm{T}}$ are principal point

The Principal point offset

- In homogenous coordinates, the mapping is expressed as follow:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

- Writing the ***camera calibration matrix*** K is defined as follow

$$\mathrm{K} = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}$$

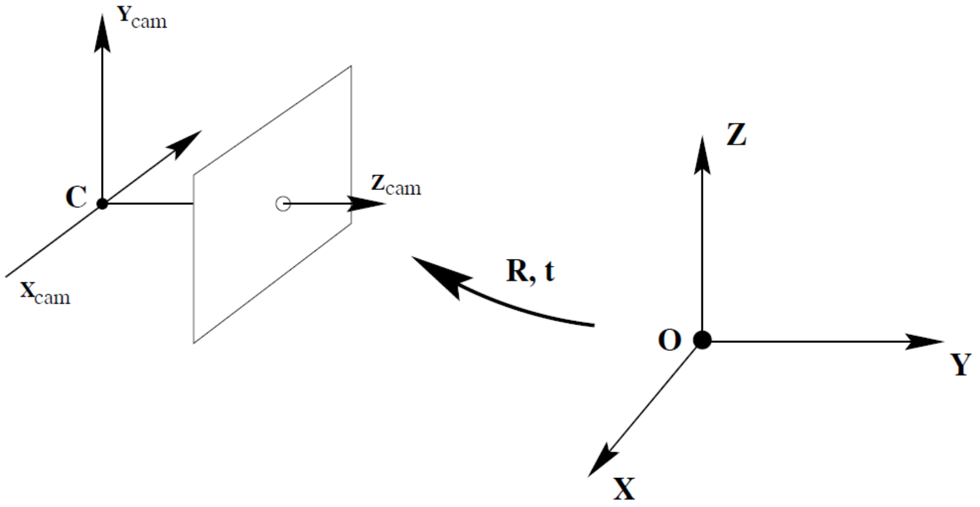
- Assuming the ***camera*** is located at the ***origin*** of a Euclidean coordinate system and ***principal axis*** of camera points ***straight down the Z-axis***

$$\mathrm{x} = \mathrm{K}\,[\mathrm{I} \mid \mathbf{0}]\mathbf{X}_{\mathbf{cam}}$$

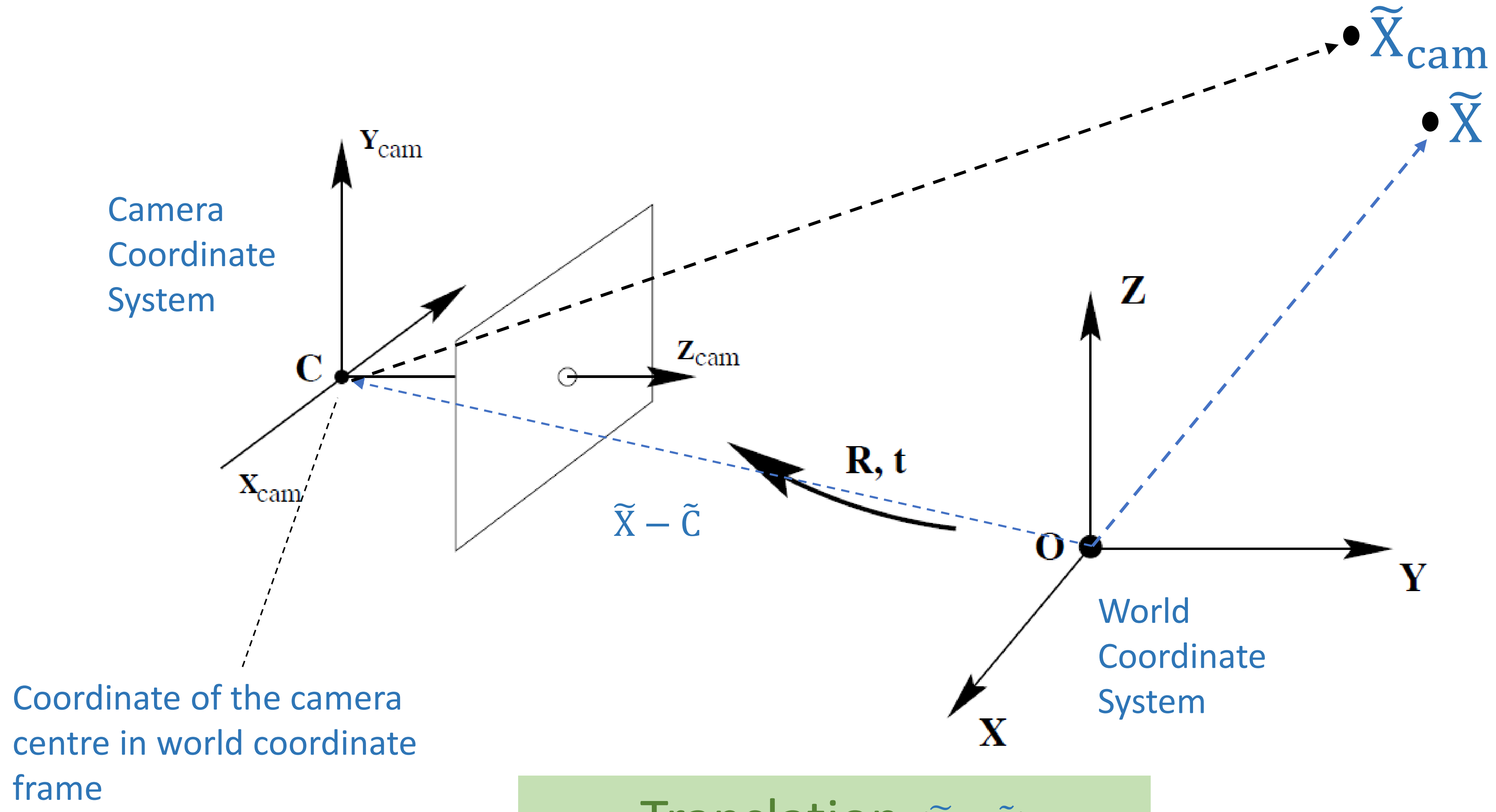
$\mathbf{X}_{\mathbf{cam}}$ is called ***camera coordinate frame***

Camera Rotation and Translation

- In general, points in space will be expressed in different Euclidean coordinate frame, known as the ***world coordinate frame***.
- The two coordinate frames are related via a ***rotation*** and a ***translation***.
- Let $\widetilde{\mathrm{X}}$ is an inhomogeneous 3-vector of a point in ***world coordinate frame*** and $\widetilde{\mathrm{X}}_{\mathrm{cam}}$ represent the same point in the ***camera coordinate frame***,

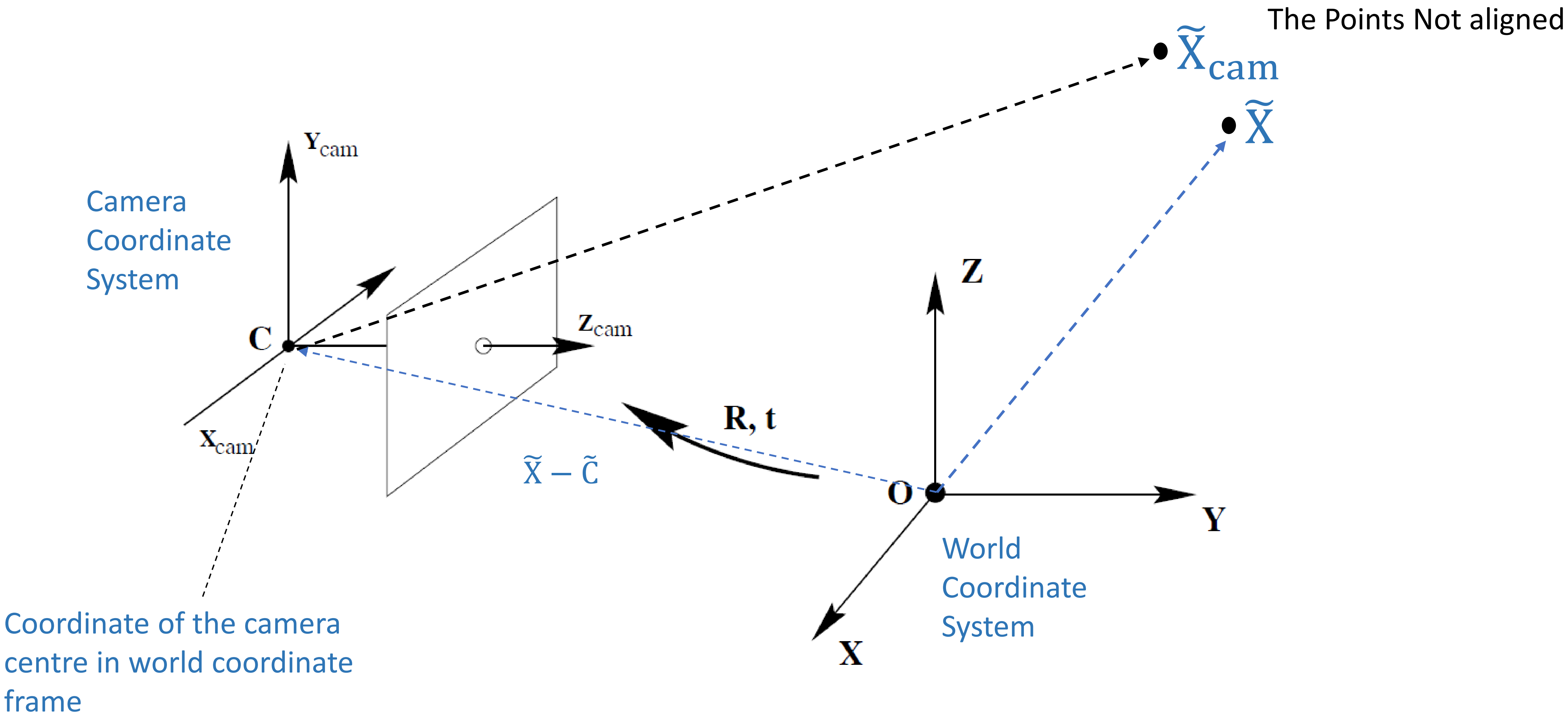


Camera Rotation and Translation

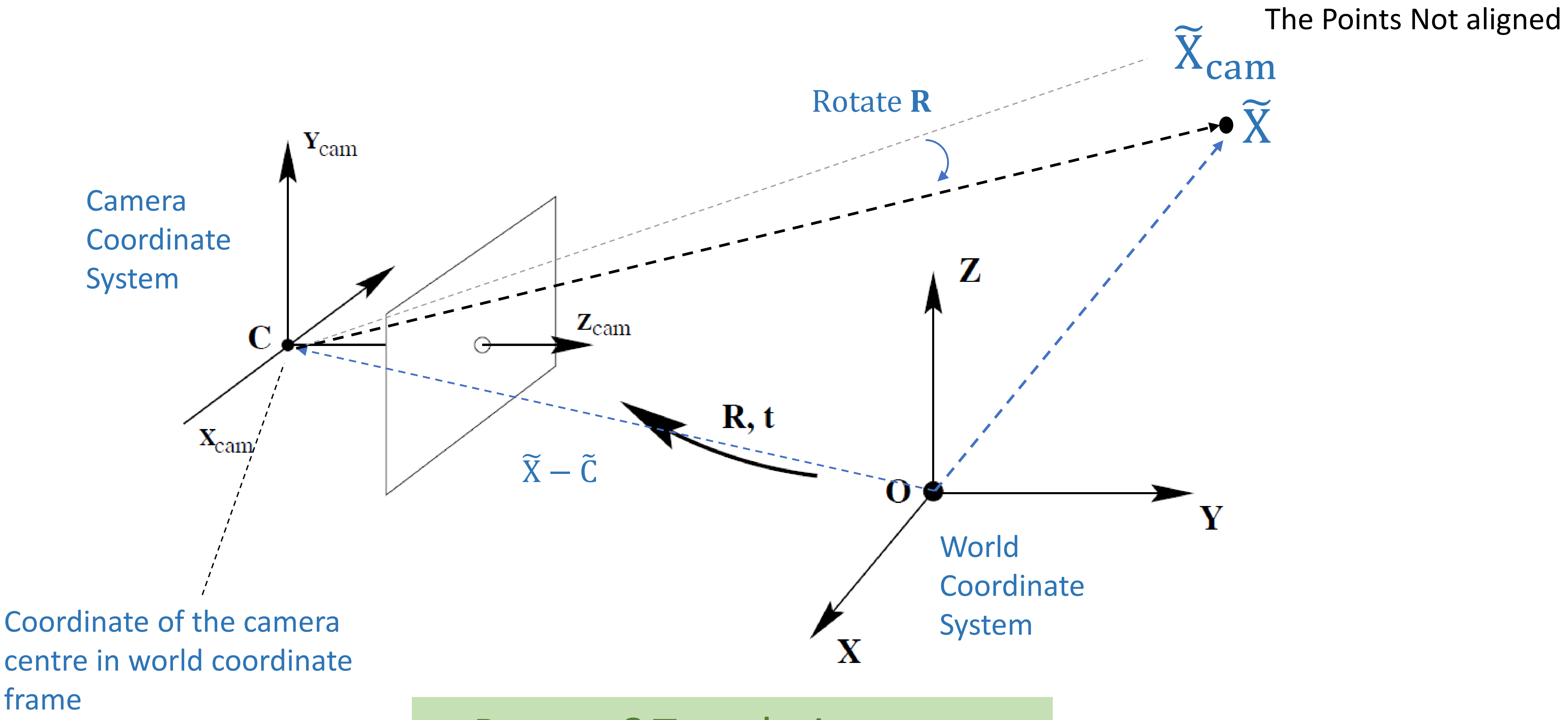


Translation $(\tilde{X} - \tilde{C})$

Camera Rotation and Translation



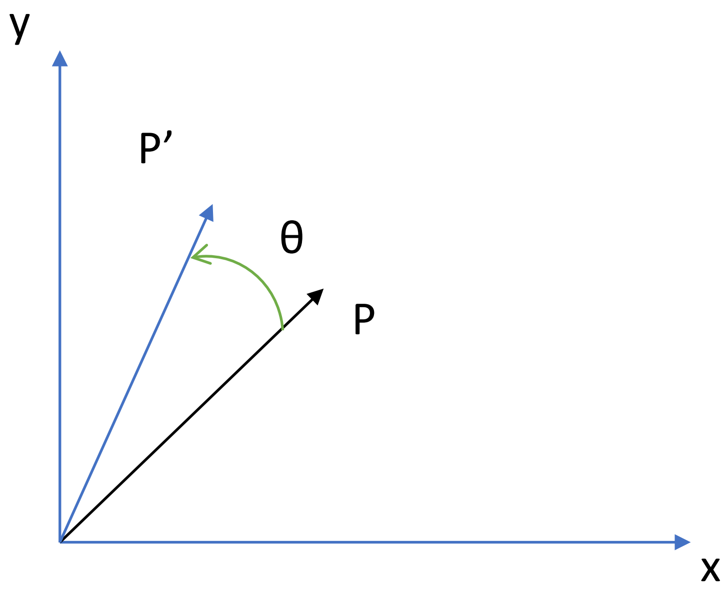
Camera Rotation and Translation

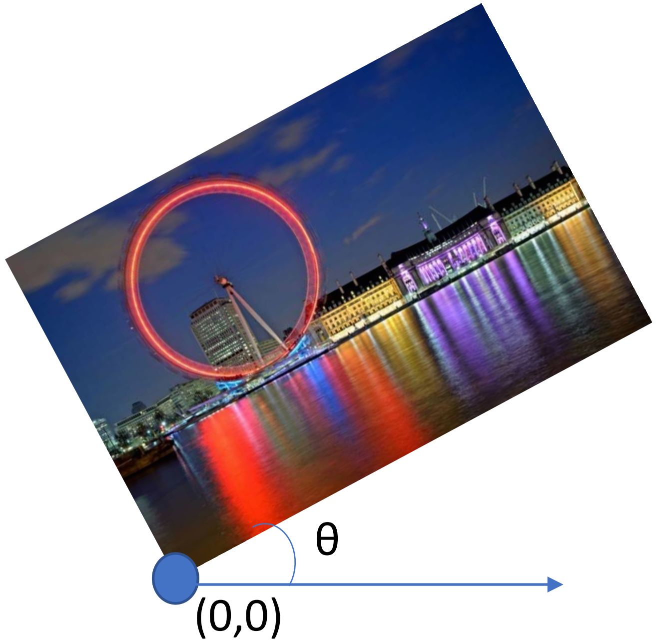


Rotate &Translation $R(\tilde{X} - \tilde{C})$

Recall – Rotation in 2D

- $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
- $R = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$





Degree of freedom = 1

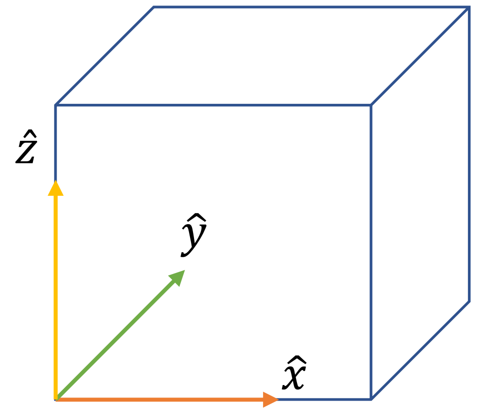
3D Rotation Matrix

In 3D, we add a 3rd dimension to the rotation matrix

$$R = [\hat{x} \quad \hat{y} \quad \hat{z}] \in \mathbb{SO}(3)$$

Each column is a 3D unit vector

$$\hat{x}, \hat{y}, \hat{z} \in \mathbb{R}^3$$
$$||\hat{x}|| = ||\hat{y}|| = ||\hat{z}|| = 1$$

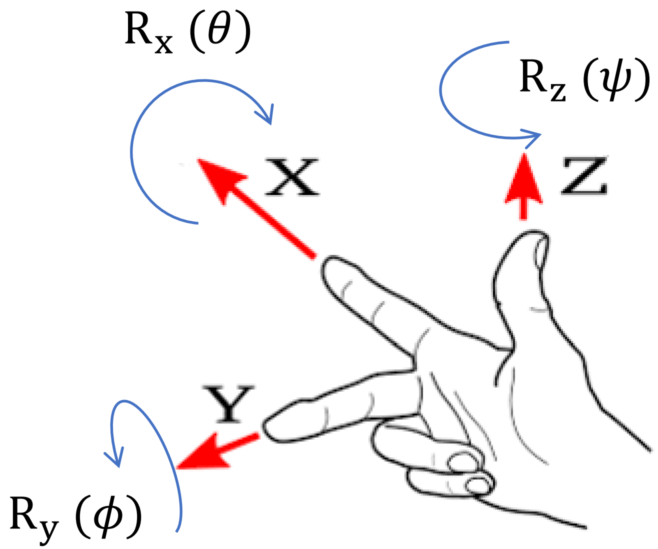


All columns are orthonormal

Right Hand Rule

- X-axis is forward from the index finger
- Y-axis is left from the middle finger
- Z-axis is up from the thumb

- Roll θ about x-axis
- Pitch ϕ about y-axis
- Yield ψ about z-axis

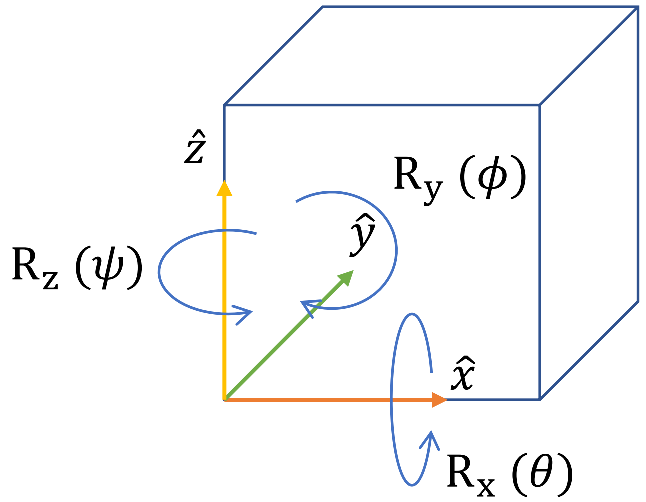


Single Axis Rotations

- $R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$

- $R_y(\phi) = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix}$

- $R_z(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$



Combinations of single axis rotations

- A rotation matrix can be formed by 3 sequential rotation about the primary axes represented by

$$\mathrm{R}\,(\theta, \phi, \psi)$$

- The sequence can be arbitrary chosen but not the repeated axis in succession since

$$\mathrm{R_x}\,(\theta)\mathrm{R_x}\,(\theta) = \mathrm{R_x}\,(2\theta)$$

In total, there are $3 \times 2 \times 2 = 12$ combinations

Combination of Elementary Rotations

Euler Angles:

- Same Axis twice $(3 \times 2 \times 1 = 6 \text{ combinations})$

XYX	XZX
YXY	YZY
ZXZ	ZYZ

Cardan Angles:

- All three axes $(3 \times 2 \times 1 = 6 \text{ combinations})$

XYZ	XZY
YXZ	YZX
ZXY	ZYX

Roll, Pitch and Yaw to Rotation Matrix

$$
\begin{aligned}
&R_X R_Y R_Z \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos B & 0 & \sin B \\ \sin A \sin B & \cos A & -\sin A \cos B \\ -\cos A \sin B & \sin A & \cos A \cos B \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos B \cos C & -\cos B \sin C & \sin B \\ \sin A \sin B \cos C + \cos A \sin C & -\sin A \sin B \sin C + \cos A \cos C & -\sin A \cos B \\ -\cos A \sin B \cos C + \sin A \sin C & \cos A \sin B \sin C + \sin A \cos C & \cos A \cos B \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&R_Y R_X R_Z \\
&= \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos B & \sin B \sin A & \sin B \cos A \\ 0 & \cos A & -\sin A \\ -\sin B & \cos B \sin A & \cos B \cos A \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos B \cos C + \sin B \sin A \sin C & -\cos B \sin C + \sin B \sin A \cos C & \sin B \cos A \\ \cos A \sin C & \cos A \cos C & -\sin A \\ -\sin B \cos C + \cos B \sin A \sin C & \sin B \sin C + \cos B \sin A \cos C & \cos B \cos A \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&R_Z R_X R_Y \\
&= \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \\
&= \begin{pmatrix} \cos C & -\sin C \cos A & \sin C \sin A \\ \sin C & \cos C \cos A & -\cos C \sin A \\ 0 & \sin A & \cos A \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \\
&= \begin{pmatrix} \cos C \cos B - \sin C \sin A \sin B & -\sin C \cos A & \cos C \sin B + \sin C \sin A \cos B \\ \sin C \cos B + \cos C \sin A \sin B & \cos C \cos A & \sin C \sin B - \cos C \sin A \cos B \\ -\cos A \sin B & \sin A & \cos A \cos B \end{pmatrix}
\end{aligned}
$$

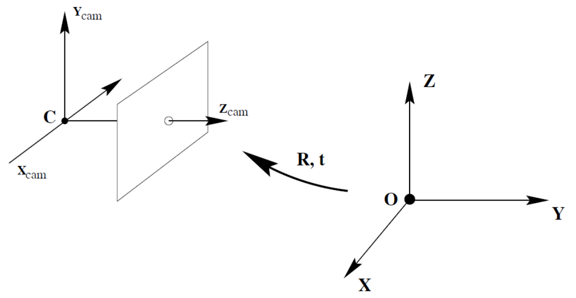
$$
\begin{aligned}
&R_X R_Z R_Y \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \\
&= \begin{pmatrix} \cos C & -\sin C & 0 \\ \cos A \sin C & \cos A \cos C & -\sin A \\ \sin A \sin C & \sin A \cos C & \cos A \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \\
&= \begin{pmatrix} \cos C \cos B & -\sin C & \cos C \sin B \\ \cos A \sin C \cos B + \sin A \sin B & \cos A \cos C & \cos A \sin C \sin B - \sin A \cos B \\ \sin A \sin C \cos B - \cos A \sin B & \sin A \cos C & \sin A \sin C \sin B + \cos A \cos B \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&R_Y R_Z R_X \\
&= \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \\
&= \begin{pmatrix} \cos B \cos C & -\cos B \sin C & \sin B \\ \sin C & \cos C & 0 \\ -\sin B \cos C & \sin B \sin C & \cos B \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \\
&= \begin{pmatrix} \cos B \cos C & -\cos B \sin C \cos A + \sin B \sin A & \cos B \sin C \sin A + \sin B \cos A \\ \sin C & \cos C \cos A & -\cos C \sin A \\ -\sin B \cos C & \sin B \sin C \cos A + \cos B \sin A & -\sin B \sin C \sin A + \cos B \cos A \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&R_Z R_Y R_X \\
&= \begin{pmatrix} \cos C & -\sin C & 0 \\ \sin C & \cos C & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos B & 0 & \sin B \\ 0 & 1 & 0 \\ -\sin B & 0 & \cos B \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \\
&= \begin{pmatrix} \cos C \cos B & -\sin C & \cos C \sin B \\ \sin C \cos B & \cos C & \sin C \sin B \\ -\sin B & 0 & \cos B \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos A & -\sin A \\ 0 & \sin A & \cos A \end{pmatrix} \\
&= \begin{pmatrix} \cos C \cos B & -\sin C \cos A + \cos C \sin B \sin A & \sin C \sin A + \cos C \sin B \cos A \\ \sin C \cos B & \cos C \cos A + \sin C \sin B \sin A & -\cos C \sin A + \sin C \sin B \cos A \\ -\sin B & \cos B \sin A & \cos B \cos A \end{pmatrix}
\end{aligned}
$$

Camera Rotation and Translation

- We can write



$$\tilde{X}_{cam} = R(\tilde{\mathbf{X}} - \tilde{\mathbf{C}})$$

  - Where
    - $\tilde{C}$ represents the coordinate of the camera in the world coordinate frame
    - R is a 3×3 rotation matrix representing the orientation of the camera coordinate frame.

- In homogenous coordinate, the equation becomes

$$\mathbf{X}_{cam} = \begin{bmatrix} R & -R\tilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \mathbf{X}$$

General Mapping of Pinhole Camera

- Since

$$\mathbf{x} = \mathrm{K}\,[\mathrm{I} \mid \mathbf{0}]\mathbf{X}_{\mathbf{cam}} \quad \text{and} \quad \mathbf{X}_{\mathrm{cam}} = \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathrm{C}} \\ 0 & 1 \end{bmatrix} \mathbf{X}$$

- We have

$$\mathbf{x} = \mathrm{K} \begin{bmatrix} \mathrm{R} & -\mathrm{R}\tilde{\mathrm{C}} \\ 0 & 1 \end{bmatrix} \mathbf{X}$$

- This becomes the general mapping of a pinhole camera

$$\mathbf{x} = \mathrm{K}\,\mathrm{R}\left[\mathrm{I} \mid -\tilde{\mathrm{C}}\right]\mathbf{X}$$

General Mapping of a Pinhole camera

- The general mapping of a pinhole camera

$$\mathbf{x} = \mathrm{K}\,\mathrm{R}[\mathrm{I} \mid -\tilde{\mathrm{C}}]\mathbf{X}$$

- The camera matrix $\mathrm{P} = \mathrm{K}\,\mathrm{R}[\mathrm{I} \mid -\tilde{\mathrm{C}}]$ has 9 degrees of freedom
  - 3 for K (the elements f, p_x, p_y) called internal or intrinsic parameters
  - 3 for R and 3 for $\tilde{\mathrm{C}}$ called external or extrinsic parameters
- It is convenient to express the camera centre as

$$\begin{aligned}\widetilde{\boldsymbol{X}}_{\mathrm{cam}} &= \mathrm{R}(\widetilde{\mathbf{X}} - \tilde{\mathbf{C}}) \\ &= \mathrm{R}\widetilde{\mathbf{X}} + \mathbf{t} \quad \text{where } \mathbf{t} = -\mathrm{R}\tilde{\mathbf{C}}\end{aligned}$$

- The camera matrix is simply

$$\mathrm{P} = \mathrm{K}\,\mathrm{R}[\mathrm{I} \mid \mathbf{t}]$$

From pinhole to CCD cameras

- The pin hole camera model assumes image coordinates having equal scales in both axial direction
- CCD cameras has non-square pixels, unequal scale factors in each direction is required
- Let m_x, m_y be pixel per unit distance in the x and y direction. We have to multiple extra factor $diag(m_x, m_y, 1)$ to the transform
- The general form of calibration matrix of CCD

$$K = \begin{bmatrix} m_x f & & p_x \\ & m_x f & p_y \\ & & 1 \end{bmatrix}$$

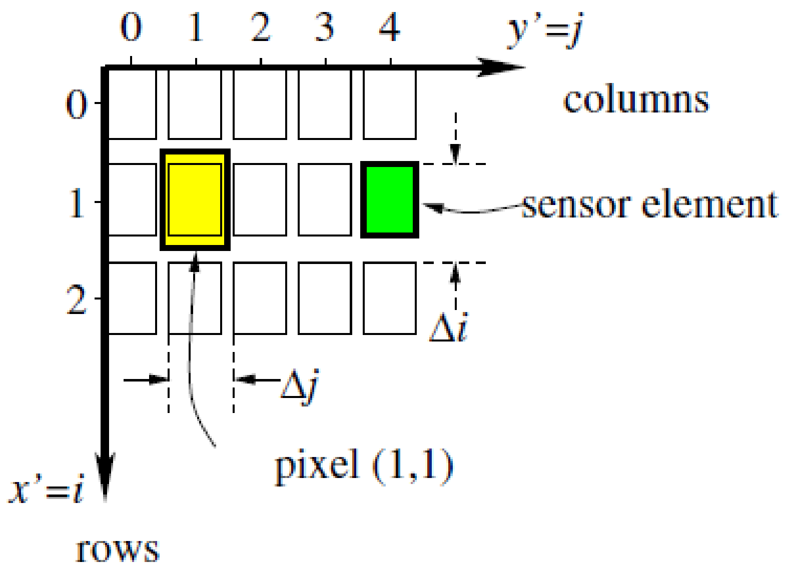


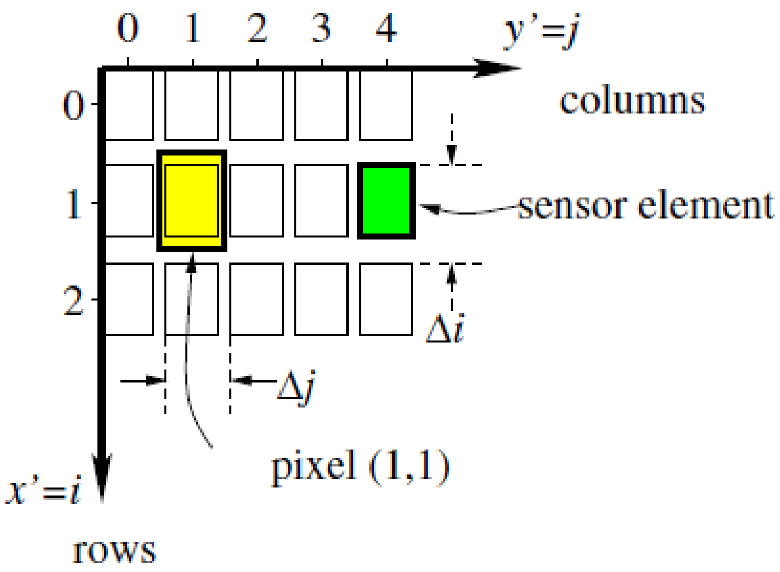
Image Courtesy: Forstner

From pinhole to CCD cameras (alternative representation

- Correction of the scale factor by

- Let $\alpha_x = f m_x$, $\alpha_y = f m_y$ be the focal length in x and y direction

$$K = \begin{bmatrix} \alpha_x & & p_x \\ & \alpha_y & p_y \\ & & 1 \end{bmatrix}$$



- A CCD camera thus has 10 degree of freedom.

Alternative representation of the scale

Alternative Representation:

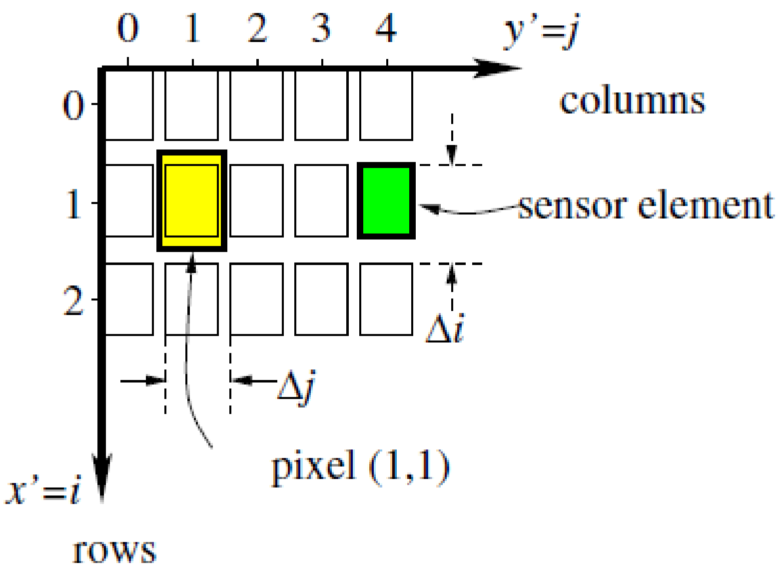
- Correction of scale factor by

$$1 + m = \Delta j / \Delta i$$

- The calibration matrix bcomes

$$\mathrm{K} = \begin{bmatrix} f & & p_x \\ & f(1+m) & p_y \\ & & 1 \end{bmatrix}$$

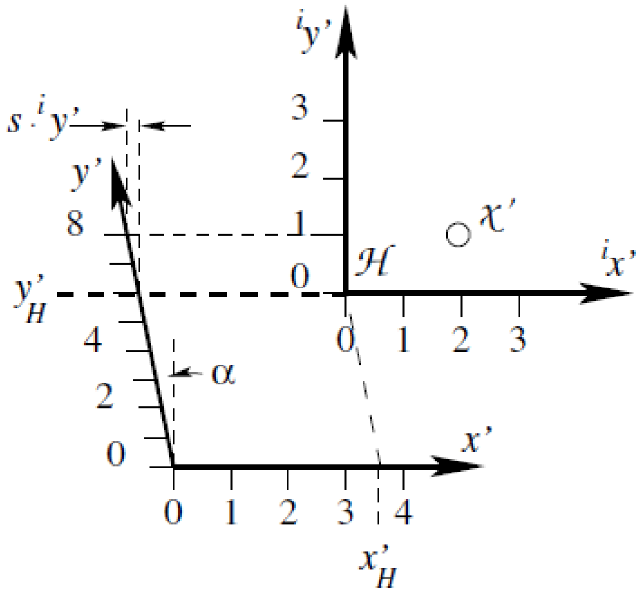
- A CCD camera thus has 10 degree of freedom.



Adding Skew to the calibration matrix

- For added generality, we consider the calibration matrix of the form

$$K = \begin{bmatrix} \alpha_x & s & p_x \\ & \alpha_y & p_y \\ & & 1 \end{bmatrix}$$



$${}^sx = {}^ix + s\ {}^iy + x_H$$
$${}^sy = {}^iy + m\ {}^iy + y_H$$

where $s = tan\alpha$

- The parameter s is referred to the skew parameters
- The camera with calibration matrix K above is called a ***finite projective camera*** which has ***11 degree of freedom***

Image Courtesy: Forstner

The Finite Projective Camera

- A camera below is called a finite projective camera

$$\mathbf{x} = \mathrm{P}\mathbf{X} \qquad \text{where } \mathrm{P} = \mathrm{KR}\left[\mathrm{I} \mid -\tilde{\mathbf{C}}\right]$$

- The projective matrix P is a **3×4 matrix** defined by to an ***arbitrary scale*** with ***11 degrees of freedom.***
- We can write $\mathrm{P} = [\mathbf{p}_1 \mathbf{p}_2\ \mathbf{p}_3\ \mathbf{p}_4]$ and let $\mathrm{M} = [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]$ be the left 3×3 submatrix of P, then we have:

$$\mathrm{P} = \mathrm{M}[\mathrm{I} \mid \mathrm{M}^{-1}\mathbf{p}_4] = \mathrm{KR}\left[\mathrm{I} \mid -\tilde{\mathbf{C}}\right]$$

- The set of camera matrices of finite projective cameras is identical with the soft of homogeneous **3×4 matrix** for which the left hand 3×3 submatrix is ***non-singular***

Projective Finite Camera - The Camera Anatomy

The Camera Centre:

- The matrix P is *rank 3* but it has *4 columns*, therefore it has ***1-dimensional right null-space***. Suppose it is generated by the 4-vector $\mathbf{C}$, i.e.

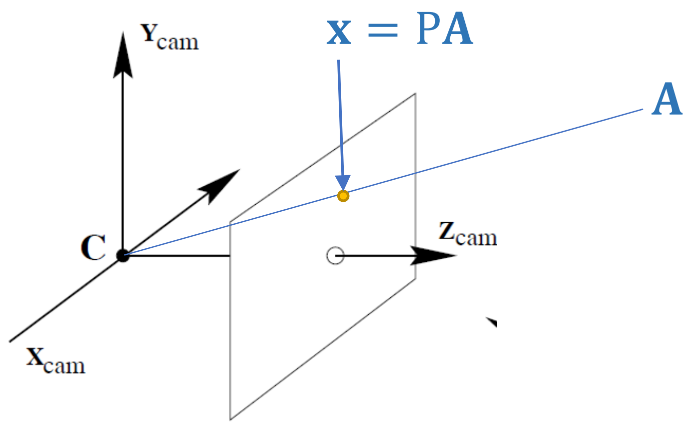
$$\mathrm{PC} = 0$$

- $\mathbf{C}$ is the homogenous 4-vector representing the ***camera centre***
- PC is the image point $(0,0,0)^{\mathrm{T}}$ is ***undefined***.

Projective Finite Camera - The Camera Centre

- **Proof:**
- Consider the line containing $\mathbf{C}$ and any point $\mathbf{A}$ in 3-space. Points on this line may be represented:

$$\mathbf{X}(\lambda) = \lambda\mathbf{A} + (1 - \lambda)\mathbf{C}$$



- Under the mapping $\mathbf{x} = \mathrm{P}\mathbf{X}$, we have

$$\mathbf{x} = \mathrm{P}\mathbf{X}(\lambda) = \lambda\mathrm{P}\mathbf{A} + \cancel{(1 - \lambda)\mathbf{PC}} = \lambda\mathrm{P}\mathbf{A}$$

- Therefore $\mathrm{PC} = 0$ and all points on the line are mapped to the same point $\mathrm{P}\mathbf{A}$, the line must through the camera centre $\mathbf{C}$

Projective Finite Camera - The Camera Anatomy

Column vectors:

- The columns of the projective camera are 3-vectors which has a geometric meaning as particular image points.

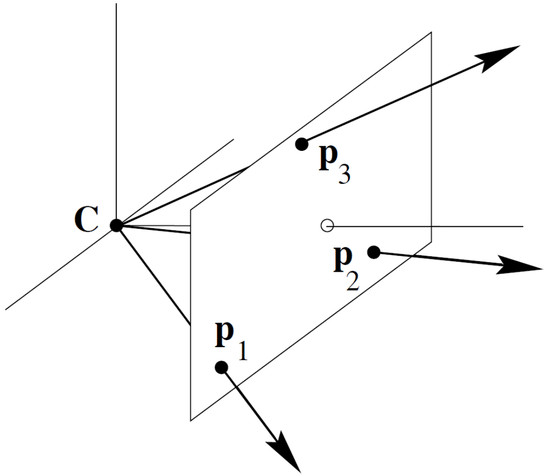
- With the notation that the columns of P are $\mathbf{p}_i, i = 1, \ldots, 4$. (i.e.) $\mathrm{P} = [\mathbf{p}_1 | \, \mathbf{p}_2 | \, \mathbf{p}_3 | \, \mathbf{p}_4]$.

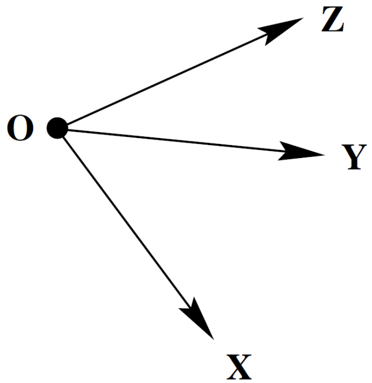
- Then $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ are the vanishing point of the world coordinate X, Y, Z axes respectively

For example:

x-axis has direction $\mathbf{D} = (1,0,0,0)^{\mathrm{T}}$

$$\mathbf{p}_1 = \mathrm{PD} = [\mathbf{p}_1 | \, \mathbf{p}_2 | \, \mathbf{p}_3 | \, \mathbf{p}_4] \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$





Projective Finite Camera - The Camera Anatomy

Row Vectors:

- The ***rows*** of the projective camera are 4-vectors which may be interpreted geometrically as ***particular world planes***

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{41} \\ p_{21} & p_{22} & p_{23} & p_{42} \\ p_{31} & p_{23} & p_{33} & p_{43} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^{1\mathrm{T}} \\ \mathbf{p}^{2\mathrm{T}} \\ \mathbf{p}^{3\mathrm{T}} \end{bmatrix}$$

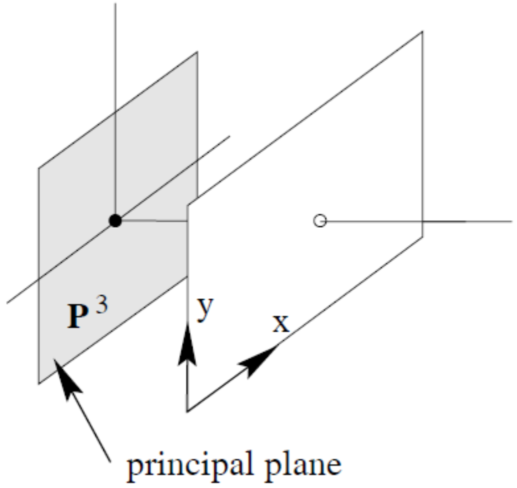
Projective Finite Camera - The Camera Anatomy

The principal plane:

- The plane through the ***camera centre*** *parallel* to the image plane
- It consists the set of points $\mathbf{X}$ which are imaged on ***the line at infinity*** explicitly $\mathrm{P}\mathbf{X} = (x, y, 0)^{\mathrm{T}}$
- $\mathbf{p}^{3\mathrm{T}}$ represents the principal plane of the camera

- **Proof:**

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{bmatrix} \mathbf{p}^{1\mathrm{T}} \\ \mathbf{p}^{2\mathrm{T}} \\ \mathbf{p}^{3\mathrm{T}} \end{bmatrix} \mathbf{X}, \text{ therefore } \mathbf{p}^{3\mathrm{T}}\mathbf{X} = \mathbf{0}$$



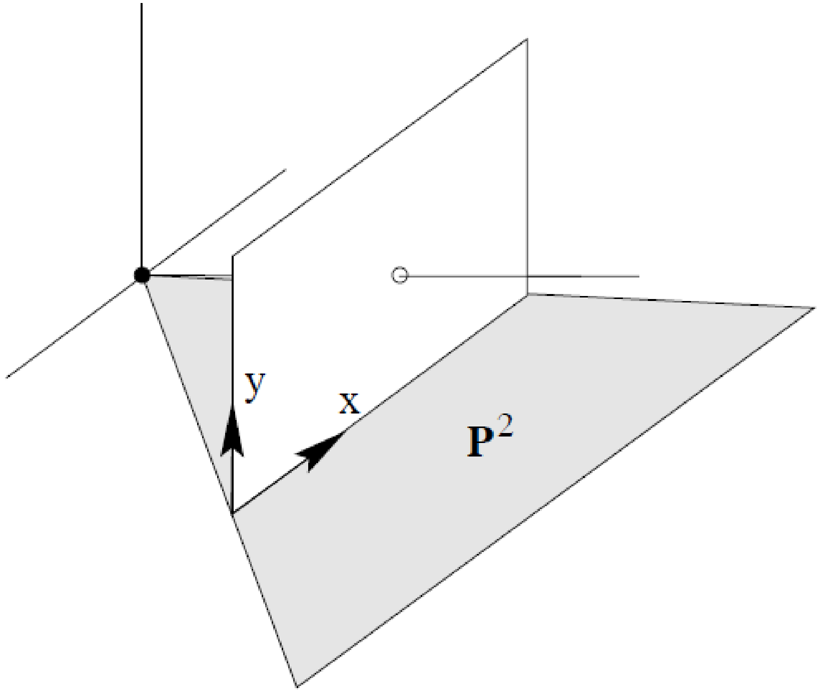
- Since $\mathrm{P}\mathbf{C} = 0 \Rightarrow \mathbf{C}$ lies on the principal plane.

Projective Finite Camera - The Camera Anatomy

Axis planes

- Consider the set of points on $\mathbf{p}^1$ satisfies $\mathbf{p}^{1\mathrm{T}}\mathbf{X}=0$ and imaged at $\mathrm{P}\mathbf{X} = (0, y, w)^{\mathbf{T}}$, which are points on the y-axis.

$$\text{i.e. } \mathbf{x} = \begin{bmatrix} \mathbf{p}^{1\mathrm{T}} \\ \mathbf{p}^{2\mathrm{T}} \\ \mathbf{p}^{3\mathrm{T}} \end{bmatrix} \mathbf{X} \qquad \mathbf{p}^{1\mathrm{T}}\mathbf{X}=0 \Rightarrow \mathbf{x} = \begin{pmatrix} 0 \\ y \\ w \end{pmatrix}$$



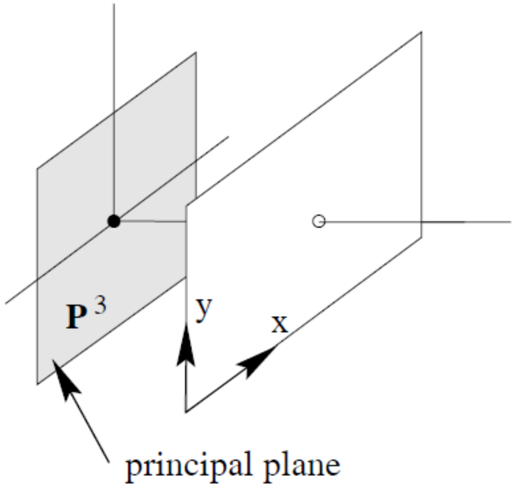
- It follows $\mathbf{p}^{1\mathrm{T}}\mathbf{X} = 0$ and $\mathrm{P}\mathbf{C} = 0$ and so $\mathbf{C}$ also lies on $\mathbf{p}^1$
- Similarly for the case of $\mathbf{p}^2$ imaged at $\mathrm{P}\mathbf{X} = (x, 0, w)^{\mathbf{T}}$

- The ***Camera Centre $\mathbf{C}$ lies on the intersection all three pl***

Projective Finite Camera - The Camera Anatomy

The principal point:

- The principal axis is the line passing through the camera centre $\mathbf{C}$, with direction perpendicular to the principal plane .
- The principal axis intercepts the image plane at the principal point.
- A plane $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^{\mathrm{T}}$ has normal vector $(\pi_1, \pi_2, \pi_3)^{\mathrm{T}}$ which may be represented by a point $(\pi_1, \pi_2, \pi_3, 0)^{\mathrm{T}}$ on the plane at infinity.



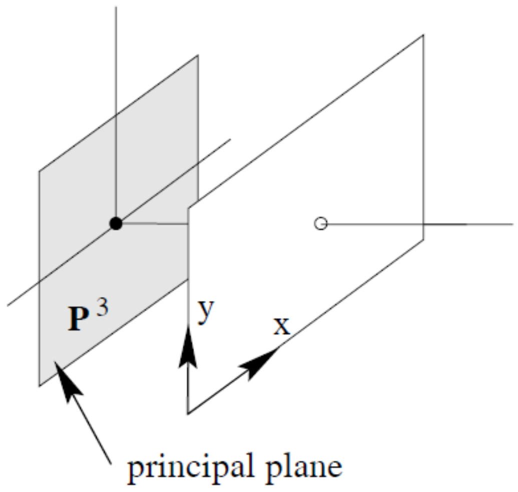
Projective Finite Camera - The Camera Anatomy

The principal point:

- In the case of $\mathbf{p}^3$, the point at plane at infinity is $\hat{\mathbf{p}}^3 = (p_{11}, p_{12}, p_{13}, 0)^\mathrm{T}$
- Projecting $\hat{\mathbf{p}}^3$ using the camera matrix P gives the principal point of the camera,

$$\mathbf{x}_0 = \mathrm{P}\hat{\mathbf{p}}^3$$

- As only the left 3×3 part of $\mathrm{P} = [\mathrm{M} \mid \mathbf{p}_4]$ is involved,



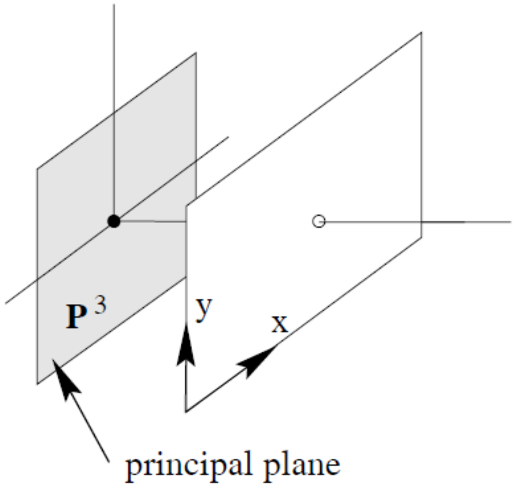
- The principal point is computed as

$$\mathbf{x}_0 = \mathrm{M}\mathrm{m}^3 \quad \text{where } \mathrm{m}^{3\mathrm{T}} \text{ is third row of } \mathrm{M}$$

Projective Finite Camera - The Camera Anatomy

The principal axis vector:

- Although any point $\mathbf{X}$ not on the principal plane may be mapped to an image point according to $\mathrm{x} = \mathrm{P}\mathbf{X}$
- In reality only half the points in space, in front of the camera, may be seen in an image.
- Let $\mathrm{P} = [\mathrm{M} \mid \mathbf{p}_4]$, it is shown the vector $\mathbf{m}^3$ points in the direction of principal axis
- However P is only defined update to sign, we need to determine if $\mathbf{m}^3$ or $-\mathbf{m}^3$ points in the positive direction.



Projective Finite Camera - The Camera Anatomy

The principal axis vector:

- The equation for projection of a 3D point to image is given by

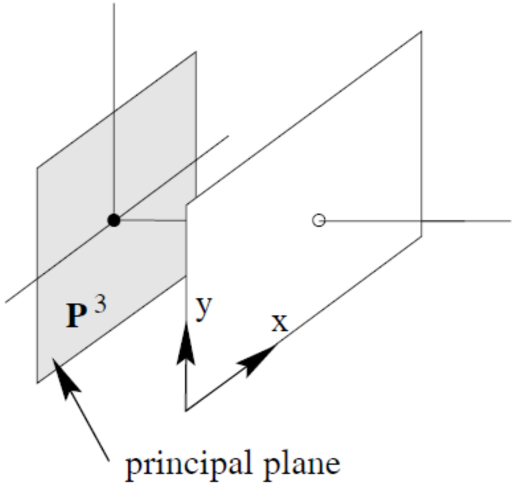
  $\mathbf{x} = \mathrm{P}_{\mathrm{cam}}\mathbf{X}_{\mathrm{cam}} = \mathrm{K}[\mathrm{I}\ |0]\mathbf{X}_{\mathrm{cam}}$ where $\mathbf{X}_{\mathrm{cam}}$ is 3D points in camera coordinate.

- The vector $\mathbf{v} = \det(\mathrm{M})\mathbf{m}^3 = (0,0,1)^{\mathrm{T}}$ points to the front of camera irrespective of the scaling of $\mathrm{P}_{\mathrm{cam}}$

- In world coordinate, we have

  $\mathrm{P} = k\mathrm{K}\left[\mathrm{R}\ |-\mathrm{R}\tilde{\mathbf{C}}\right] = [\mathrm{M}\ |\ \mathbf{p}_4]$, where $\mathrm{M}{=}k\mathrm{KR}$
- Since $\det(\mathrm{R}) > 0$, $\mathbf{v} = \det(\mathrm{M})\ \mathbf{m}^3$ is also unaffected by scale.

- **In summary, $\mathbf{v} = \mathbf{det}(\mathbf{M})\ \mathbf{m}^3$ is a vector in the direction of the principal axis, directed toward the front of camera**



Summary of the Properties of a Projective Camera

Camera centre. The camera centre is the 1-dimensional right null-space $\mathbf{C}$ of $\mathtt{P}$, i.e. $\mathtt{P}\mathbf{C} = \mathbf{0}$.

- ◇ **Finite camera** ($\mathtt{M}$ is not singular) $\mathbf{C} = \begin{pmatrix} -\mathtt{M}^{-1}\mathbf{p}_4 \\ 1 \end{pmatrix}$
- ◇ **Camera at infinity** ($\mathtt{M}$ is singular) $\mathbf{C} = \begin{pmatrix} \mathbf{d} \\ 0 \end{pmatrix}$ where $\mathbf{d}$ is the null 3-vector of $\mathtt{M}$, i.e. $\mathtt{M}\mathbf{d} = \mathbf{0}$.

Column points. For $i = 1, \ldots, 3$, the column vectors $\mathbf{p}_i$ are vanishing points in the image corresponding to the X, Y and Z axes respectively. Column $\mathbf{p}_4$ is the image of the coordinate origin.

Principal plane. The principal plane of the camera is $\mathbf{P}^3$, the last row of $\mathtt{P}$.

Axis planes. The planes $\mathbf{P}^1$ and $\mathbf{P}^2$ (the first and second rows of $\mathtt{P}$) represent planes in space through the camera centre, corresponding to points that map to the image lines $x = 0$ and $y = 0$ respectively.

Principal point. The image point $\mathbf{x}_0 = \mathtt{M}\mathbf{m}^3$ is the principal point of the camera, where $\mathbf{m}^{3\mathsf{T}}$ is the third row of $\mathtt{M}$.

Principal ray. The principal ray (axis) of the camera is the ray passing through the camera centre $\mathbf{C}$ with direction vector $\mathbf{m}^{3\mathsf{T}}$. The principal axis vector $\mathbf{v} = \det(\mathtt{M})\mathbf{m}^3$ is directed towards the front of the camera.

Action of a Projective Camera on Points

Forward Projection:

- A general projective camera maps a point in space $\mathbf{X}$ to an image point through $\mathbf{x} = \mathrm{P}\mathbf{X}$.
- Points $\mathrm{D} = (\mathrm{d}^{\mathrm{T}}, 0)^{\mathrm{T}}$ on the plane at infinity represent ***vanishing points***. Such points map to

$$\mathbf{x} = \mathrm{P}\mathbf{D} = [\mathrm{M} \mid \mathbf{p}_4]\mathbf{D} = \mathrm{M}\mathbf{d}$$

- x is ***only affected*** by M, the first 3×3 submatrix of P

Action of a Projective Camera on Points

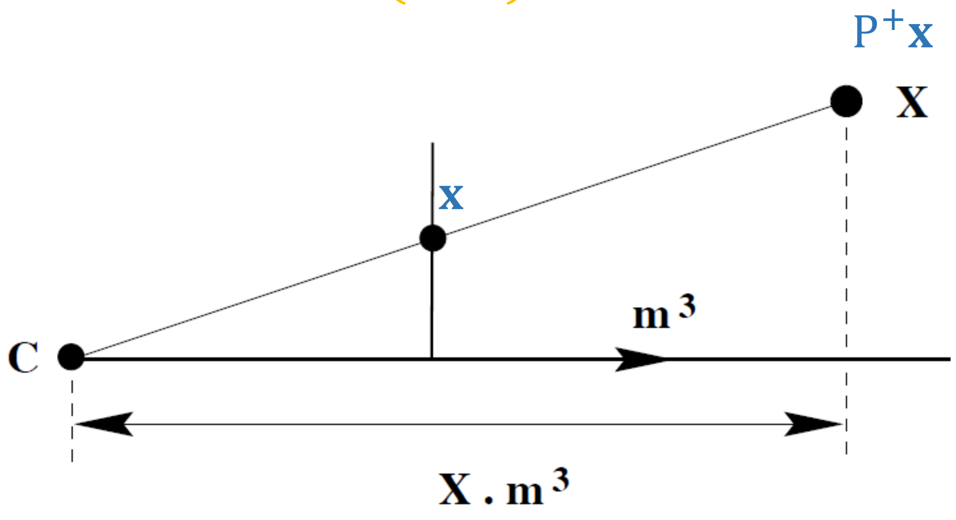
Back-projection of points to rays:

- Given a point $\mathbf{x}$ in an image, we wish to determine the set of points in space map to this point.
- We know two points on the ray.
  - The *camera centre* $\mathbf{C}$ where $\mathrm{P}\mathbf{C} = \mathbf{0}$
  - The point $\mathrm{P}^{+}\mathbf{x}$ where P^{+} is *pseudo-inverse* of P where $\mathrm{P}^{+} = \mathrm{P}^{\mathrm{T}}(\mathrm{P}\mathrm{P}^{\mathrm{T}})^{-1}$
- The point $\mathrm{P}^{+}\mathbf{x}$ lies on the ray because

$$\mathrm{P}(\mathrm{P}^{+}\mathbf{x}) = \mathrm{I}\mathbf{x} = \mathbf{x}$$

- Hence the ray is a line

$$\boxed{\mathbf{X}(\lambda) = \mathrm{P}^{+}\mathbf{x} + \lambda\mathbf{C}}$$



Action of a Projective Camera on Points

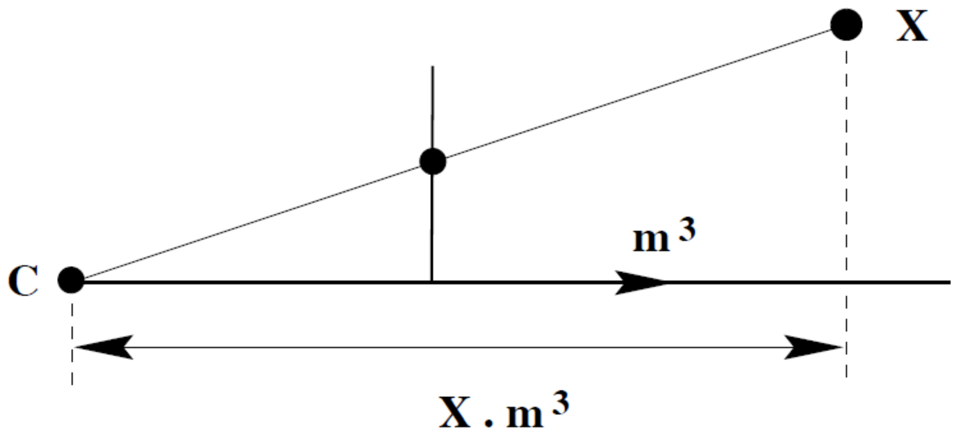
- In the case of finite camera $\mathbf{M}^{-1}$, we can write $\mathrm{P} = [\mathrm{M} \mid \mathbf{p_4}]$
  - The ***camera centre*** is $\tilde{\mathbf{C}} = -\mathbf{M}^{-1}\mathbf{p_4}$
  - An image point $\mathbf{x}$ back-projects to a ray intersecting the plane at infinity π_∞ at the ***ideal point***

$$\mathbf{D} = \left(\left(\mathbf{M}^{-1}\mathbf{x}\right)^{\mathbf{T}}, \mathbf{0} \right)^{\mathbf{T}}$$

- The line joining two points

$$\mathrm{X}(\mu) = \mu \begin{pmatrix} \mathbf{M}^{-1}\mathbf{x} \\ 0 \end{pmatrix} + \begin{pmatrix} -\mathbf{M}^{-1}\mathbf{p_4} \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -\mathbf{M}^{-1}(\mu\mathbf{x} - \mathbf{p_4}) \\ 1 \end{pmatrix}$$



Action of a Projective Camera on Points

Depth of points:

- Consider a camera matrix $\mathrm{P} = [\mathrm{M} \mid \mathbf{p_4}]$, projecting a point $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1)^{\mathrm{T}} = (\tilde{\mathbf{X}}^{\mathrm{T}}, 1)^{\mathrm{T}}$ to the image point $\mathbf{x} = w(x, y, 1)^{\mathrm{T}} = \mathrm{P}\mathbf{X}$
- We have

$$\mathbf{x} = \begin{pmatrix} wx \\ wy \\ w \end{pmatrix} = \begin{bmatrix} \mathbf{p}^{1\mathrm{T}} \\ \mathbf{p}^{2\mathrm{T}} \\ \mathbf{p}^{3\mathrm{T}} \end{bmatrix} \mathbf{X}$$

- **hence** $w = \mathbf{p}^{3\mathrm{T}}\mathbf{X} = \mathbf{p}^{3\mathrm{T}}(\mathbf{X} - \mathbf{C})$ since P**C**=0

Inhomogeneous coordinates

- $\mathrm{w} = \mathbf{p}^{3\mathrm{T}}(\mathbf{X} - \mathbf{C}) = \mathrm{m}^{3\mathrm{T}}(\tilde{\mathbf{X}} - \tilde{\mathbf{C}})$ where m^3 is the principal ray direction

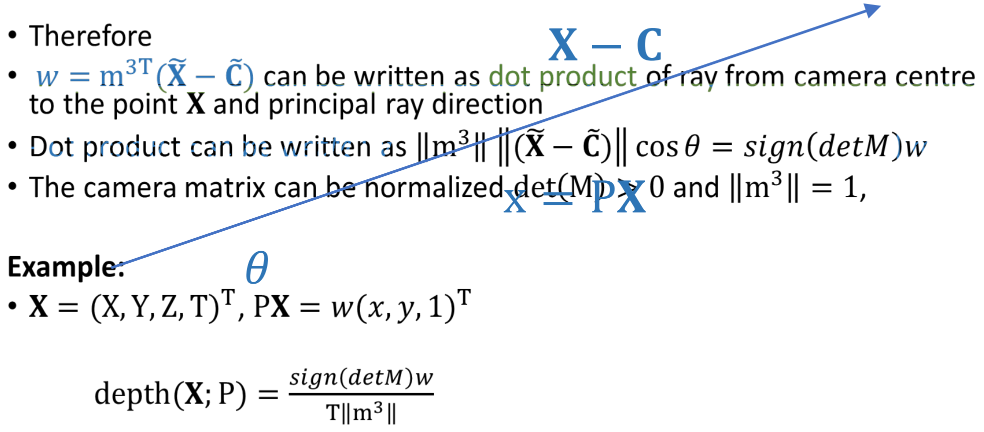
Action of a Projective Camera on Points

- Therefore
- $w = \mathrm{m}^{3\mathrm{T}}(\widetilde{\mathbf{X}} - \widetilde{\mathbf{C}})$ can be written as dot product of ray from camera centre to the point $\mathbf{X}$ and principal ray direction
- Dot product can be written as $\|\mathrm{m}^3\| \, \|(\widetilde{\mathbf{X}} - \widetilde{\mathbf{C}})\| \cos\theta = sign(detM)w$
- The camera matrix can be normalized $\det(\mathrm{M}) > 0$ and $\|\mathrm{m}^3\| = 1$,

Example:

- $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \mathrm{T})^{\mathrm{T}}$, $\mathrm{P}\mathbf{X} = w(x, y, 1)^{\mathrm{T}}$

$$\mathrm{depth}(\mathbf{X}; \mathrm{P}) = \frac{sign(detM)w}{\mathrm{T}\|\mathrm{m}^3\|}$$



Decomposition of the Camera Matrix

Decomposition of the camera matrix:

Let P be a camera representing a general projective camera.

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{41} \\ p_{21} & p_{22} & p_{23} & p_{42} \\ p_{31} & p_{23} & p_{33} & p_{43} \end{bmatrix}$$

We wish to find ***camera centre, orientation of the camera*** and the ***internal parameters*** of the camera from P

Decomposition of the Camera Matrix

Finding the camera centre:

- The camera centre $\mathbf{C}$ is the point for which P$\mathbf{C}$=0.

Numerically:

- it is the right null-vector obtained by SVD of P

$\mathrm{SVD(P)} = \mathrm{U\Sigma V^T}$ where $\mathbf{C}$ is last column of V corresponding to the least singular value σ

Algebraically:

$\mathbf{C}$ can be obtained by the null-space of P

$$\mathbf{X} = \det([\mathbf{p_2}, \mathbf{p_3}, \mathbf{p_4}]) \qquad \mathbf{Y} = \det([\mathbf{p_1}, \mathbf{p_3}, \mathbf{p_4}])$$
$$\mathbf{Z} = \det([\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_4}]) \qquad \mathbf{T} = \det([\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}])$$

Decomposition of the Camera Matrix

Finding the camera orientation and internal parameters

- In the case of finite camera, we have

$$\mathrm{P} = \left[\mathrm{M} \mid -\mathrm{M}\tilde{\mathbf{C}}\right] = \mathrm{K}\left[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}\right]$$

- Recall that $\mathrm{K} = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix}$ which is an *upper-triangular matrix*
- R is the rotation matrix which is ***orthogonal matrix***
- K and R can be found easily by ***RQ-Decomposition*** of M as $\mathrm{M} = \mathrm{KR}$
- The ambiguity in the decomposition is removed by requiring that K have *positive diagonal entries*

QR Decomposition

Finding the camera orientation and internal parameters:

- The matrix K has the form:

$$\mathrm{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Where

- α_x is the scale factor in the x-coordinate direction
- α_y is the scale factor in the y-coordinate direction
- s is the skew
- $(x_0, y_0)^T$ are the coordinates of the principal point
- The aspect ratio is α_x/α_y

QR Decomposition

- Given a matrix A, we can decompose it into

$$A = QR$$

Where Q is an orthogonal matrix (i.e. $Q^T Q = I$) and R is an upper triangular matrix. We can solve by Gram-Schmidt process.

$$A = [a_1 \quad | \quad a_2 \quad | \quad \dots \quad | \quad a_n]$$

QR Decomposition

By Gram-Schmidt

$$u_1 = a_1, \qquad q_1 = \frac{u_1}{\|u_1\|},$$

$$u_2 = a_2 - (a_2 \cdot q_1)q_1, \qquad q_2 = \frac{u_2}{\|u_2\|},$$

$$u_{k+1} = a_{k+1} - (a_{k+1} \cdot q_1)q_1 - \cdots - ((a_{k+1} \cdot q_k)q_k, \qquad q_{k+1} = \frac{u_{k+1}}{\|k+1\|},$$

- $A = [a_1 \quad | \quad a_2 \quad | \quad \ldots \quad | \quad a_n]$

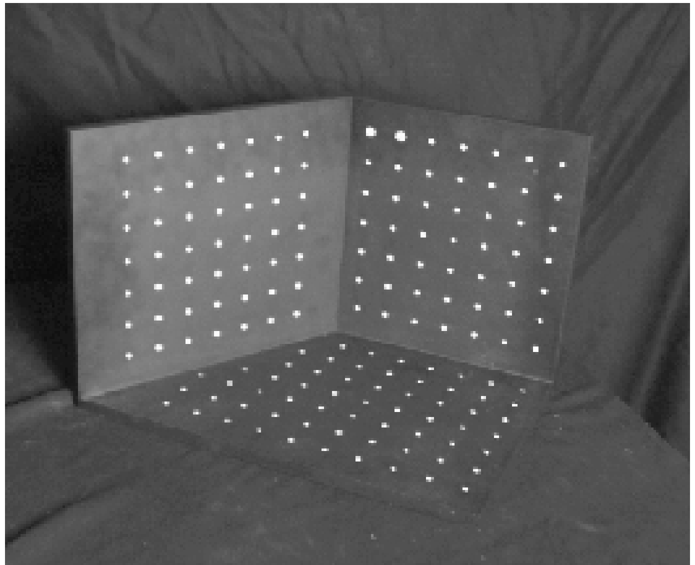
$$= [q_1 \quad | \quad q_2 \quad | \quad \ldots \quad | \quad q_n] \begin{bmatrix} a_1 \cdot q_1 & a_2 \cdot q_1 & \ldots & a_n \cdot q_1 \\ 0 & a_2 \cdot q_2 & \ldots & a_n \cdot q_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \cdot q_n \end{bmatrix} = QR$$

Camera Calibration

- Mapping from World Coord to image

$$\mathbf{x} = \mathrm{P}\mathbf{X} = \mathrm{K}\left[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}\right]\mathbf{X}$$

- Goal: To estimate the intrinsic and extrinsic parameters of the camera
- Given: Known 3D points
- Observation: corresponding 2d points



Camera Calibration

We have

3 rotations

$$\mathbf{x} = \mathrm{P}\mathbf{X} = \mathrm{K}[\mathrm{R} \mid -\mathrm{R}\tilde{\mathbf{C}}]\mathbf{X}$$

5 intrinsic paras

3 translations

- Total: 11 parameters, 6 extrinsic, 5 intrinsic

- How many points to we need?
  - Six points, each point gives two equations

Direct Linear Transform (DLT)

Mapping from World Coordinate to image:

$$\mathbf{x} = \mathrm{P}\mathbf{X}$$

- where

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{41} \\ p_{21} & p_{22} & p_{23} & p_{42} \\ p_{31} & p_{23} & p_{33} & p_{43} \end{bmatrix}$$

Direct Linear Transform

For every data point i, we have

$$\mathbf{x}_{\boldsymbol{i}} = \begin{bmatrix} u_i \\ \boldsymbol{v_i} \\ \boldsymbol{w_i} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

- Collinear equations

$$x_i = \frac{u_i}{w_i} = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

$$y_i = \frac{v_i}{w_i} = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

Direct Linear Transform

$$x_i = \frac{u_i}{w_i} = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}} \qquad y_i = \frac{v_i}{w_i} = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

- Rearranging the equation, we have

$$x_i\,(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) - (p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}) = 0$$
$$y_i\,(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) - (p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}) = 0$$

- Rewrite as

$$\begin{array}{llll} -(p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}) & & +x_i\,(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) & =0 \\ & -(p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}) & +x_i\,(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) & =0 \end{array}$$

Direct Linear Transform

- For every point *i*, we have

$$\begin{pmatrix} -X_i, & -Y_i, & -Z_i, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_iZ_i, & x_i \end{pmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} = 0$$

Direct Linear Transform

$$\begin{pmatrix} 0, & 0, & 0, & 0, & -X_i, & -Y_i, & -Z_i, & -1, & y_iX_i, & y_iY_i, & y_iZ_i, & y_i \end{pmatrix} \begin{bmatrix} {\color{red}p_{11}} \\ {\color{red}p_{12}} \\ {\color{red}p_{13}} \\ {\color{red}p_{14}} \\ {\color{blue}p_{21}} \\ {\color{blue}p_{22}} \\ {\color{blue}p_{23}} \\ {\color{blue}p_{24}} \\ {\color{green}p_{31}} \\ {\color{green}p_{32}} \\ {\color{green}p_{33}} \\ {\color{green}p_{34}} \end{bmatrix} = 0$$

Combining the equations

2Xn equations

$$\begin{bmatrix} -X_i, & -Y_i, & -Z_i, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_iZ_i, & x_i \\ 0, & 0, & 0, & 0, & -X_i, & -Y_i, & -Z_i, & -1, & y_iX_i, & y_iY_i, & y_iZ_i, & y_i \\ \vdots & & & & & & & & & & & \vdots \\ \vdots & & & & & & & & & & & \vdots \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} = 0$$

Rewrite as:

$$A_{2n\times 12} P_{12\times 1} = 0$$

Singular Value Decomposition (SVD)

Given the linear system

$$AP = 0$$

By SVD,

$$A_{2n\times 12} = U_{2n\times 12}\,\Sigma_{12\times 12}\,V^T{}_{12\times 12} = \sum_{i=1}^{12} u_i \sigma_i v_i^T$$

For A has rank 11 (for n>6), the system has a nontrivial solution p which is proportional to the column of V corresponding to the smaller singular value σ which is σ_{12}

Singular Value Decomposition (SVD)

- Therefore the estimated *p* is

$$\hat{p} = v_{12} = \begin{bmatrix} \widehat{p_1} & \widehat{p_2} & \widehat{p_3} & \widehat{p_4} \\ \widehat{p_5} & \widehat{p_6} & \widehat{p_7} & \widehat{p_8} \\ \widehat{p_9} & \widehat{p_{10}} & \widehat{p_{11}} & \widehat{p_{12}} \end{bmatrix}$$

Note for the DLT solution of Camera Calibration

- A is of rank 11, we need
  - number of points $n > 6$
  - No error measurement
- Pay attention to the sample points
  - No solution if all points X_i are on a plane
- A= $\begin{bmatrix} -X_i, & -Y_i, & -Z_i, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_iZ_i, & x_i \\ 0, & 0, & 0, & 0, & -X_i, & -Y_i, & -Z_i, & -1, & y_iX_i, & y_iY_i, & y_iZ_i, & y_i \\ \vdots & & & & & & & & & & & \vdots \\ \vdots & & & & & & & & & & & \vdots \end{bmatrix}$

- If all X_i on a plane (say Z=0), we have
- A= $\begin{bmatrix} -X_i, & -Y_i, & 0, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & 0, & x_i \\ 0, & 0, & 0, & 0, & -X_i, & -Y_i, & 0, & -1, & y_iX_i, & y_iY_i, & 0, & y_i \\ \vdots & & & & & & & & & & & \vdots \\ \vdots & & & & & & & & & & & \vdots \end{bmatrix}$ reduced rank and no – solution!

Decomposition of P

- Now we have

$$\hat{P} = \begin{bmatrix} \widehat{p_1} & \widehat{p_2} & \widehat{p_3} & \widehat{p_4} \\ \widehat{p_5} & \widehat{p_6} & \widehat{p_7} & \widehat{p_8} \\ \widehat{p_9} & \widehat{p_{10}} & \widehat{p_{11}} & \widehat{p_{12}} \end{bmatrix}$$

- But we are interested in projection

$$\widehat{\mathrm{P}} = \widehat{\mathrm{K}}\widehat{\mathrm{R}}\left[\mathrm{I}_3 | -\widehat{\mathbf{C}}\right]$$

We need to compute

$\widehat{\mathrm{K}}$, $\widehat{\mathrm{R}}$, $\widehat{\mathbf{C}}$ from $\widehat{\mathrm{p}}$

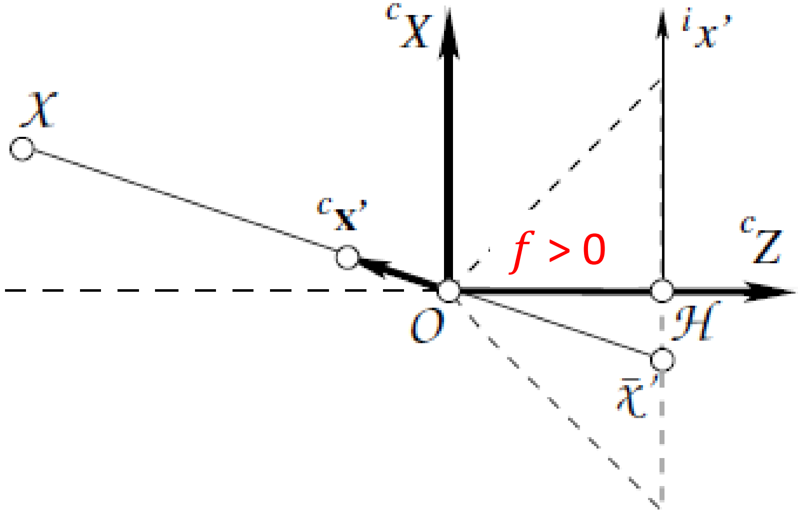
Decomposition of P

The intrinsic parameters $\mathrm{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$

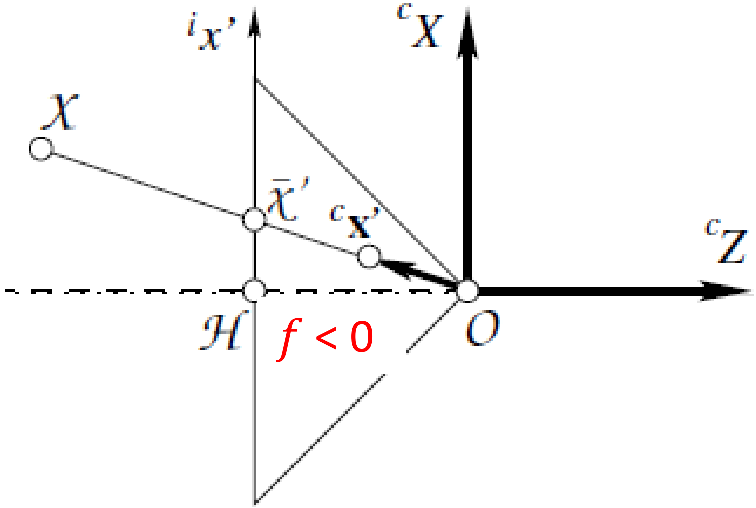
- If $\widehat{\mathrm{K}}$ diagonal element is position, α_x and is α_y positive and also f.
- To get negative camera constant f, apply another rotation

$$R = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Image plane in the front



f > 0, image behind the origin
Image upside down



f < 0, image in front of the origin, Image orientation maintained

Image Courtesy: Forstner

Summary of DLT steps

1. Vectorize the P: $\boldsymbol{p} = (p_k)$
2. Construct the Linear System

$$\begin{bmatrix} -X_i, & -Y_i, & -Z_i, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_iZ_i, & x_i \\ 0, & 0, & 0, & 0, & -X_i, & -Y_i, & -Z_i, & -1, & y_iX_i, & y_iY_i, & y_iZ_i, & y_i \\ \vdots & & & & & & & & & & & \vdots \\ \vdots & & & & & & & & & & & \vdots \end{bmatrix} \boldsymbol{p} = 0$$

3. Apply SVD on A and select the last column of V as the solution to compute a solution $\widehat{\mathrm{P}}$
4. To calculate the extrinsic and intrinsic parameters, decompose $\widehat{\mathrm{P}}$

$$\widehat{\mathrm{P}} = [\widehat{\mathrm{P}}_1 | \quad \widehat{\mathrm{P}}_2] = \widehat{\mathrm{K}}\widehat{\mathrm{R}}[\mathrm{I}_3 | -\hat{\mathrm{C}}]$$

$$\widehat{\mathbf{X}_\mathrm{o}} = -\left(\widehat{\mathrm{P}}_1\right)^{-1}\widehat{\mathrm{P}}_2$$

$\widehat{\mathrm{P}}_1 = \widehat{\mathrm{K}}\widehat{\mathrm{R}}$ computed by ***QR decomposition***

Summary of DLT

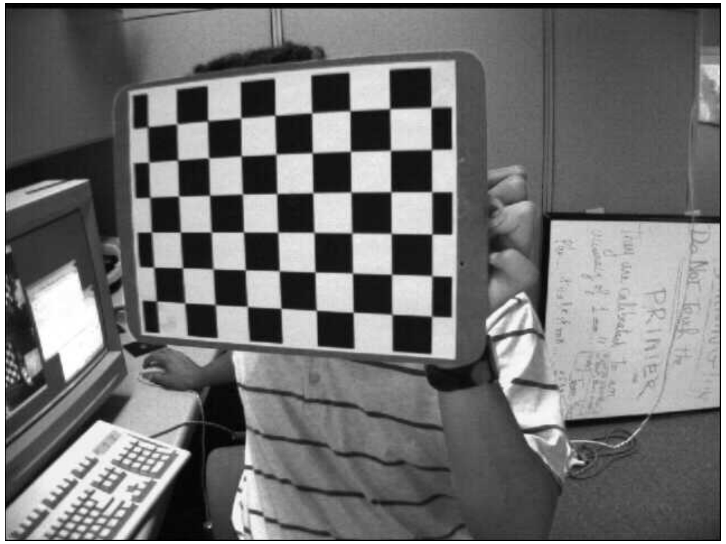
- For un-calibrated camera, we need at least 6 points

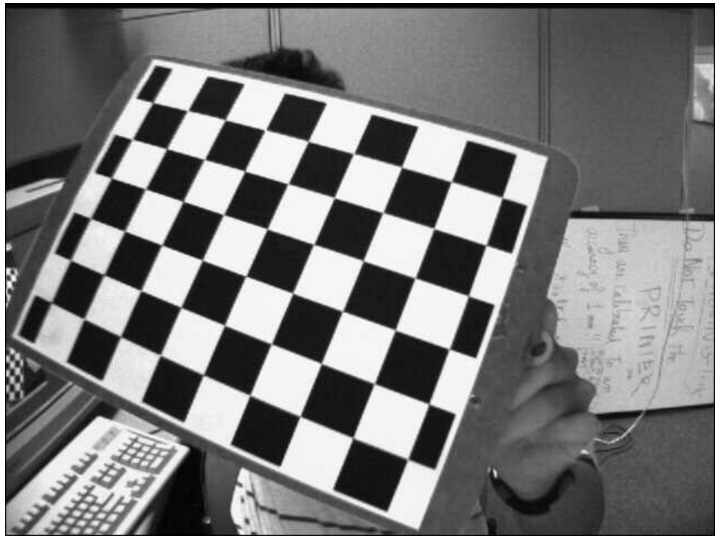
- Advantages:
  - Very simple to formulate and solve

- Disadvantages
  - You need to 3D co-ordinates of the control points.
  - Solution non-stable if control points lie on a plane.
  - Doesn’t model radial distortion.

Multi-Plane Camera Calibration(Zhang 2000)

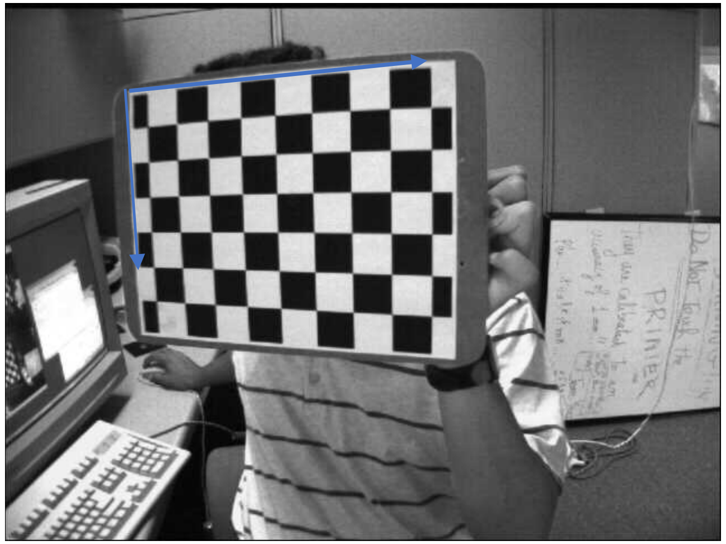
- Use a 2D known pattern (checkerboard) for calibration

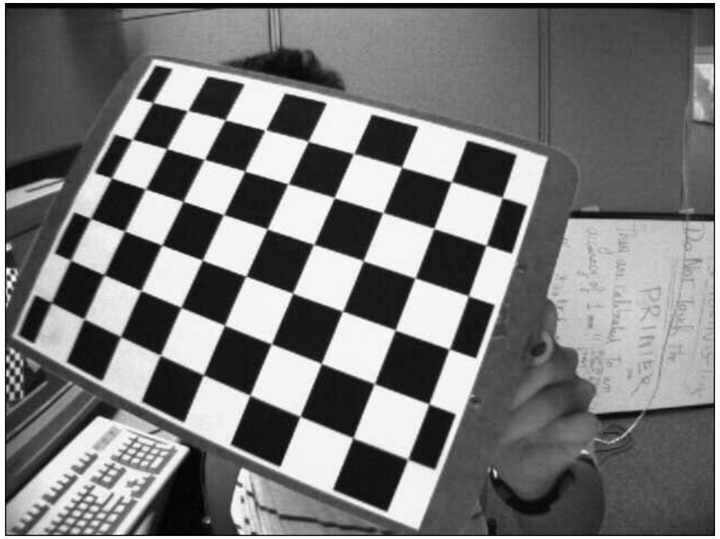




Checkerboard Calibration

- Set the world coordinate on the checkerboard with X-Y plane (Z=0) lies on the checkerboard





The mapping Function

- As all points lie on the checkerboard and therefore have Z=0
- The mapping function

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathrm{K} \begin{bmatrix} r_{11} & r_{12} & \boxed{r_{13}} & t_X \\ r_{21} & r_{22} & \boxed{r_{23}} & t_Y \\ r_{31} & r_{32} & \boxed{r_{33}} & t_Z \end{bmatrix} \begin{bmatrix} X \\ Y \\ \boxed{Z} \\ 1 \end{bmatrix}$$

0

- We can take away 3rd Column of the extrinsic matrix, and the mapping function becomes

Simplified Mapping Function

- The Z-Coordinate of all points on checkerboard is zero
- The 3rd column of the extrinsic matrix is deleted

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

We can rewrite as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathrm{K}_{3\times3}[r_1, \quad r_2, \quad t]_{3\times3} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \mathrm{H} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Where H is a homography between the transformation

$$\mathrm{H} = [h_1, \quad h_2, \quad h_3] = \mathrm{K}\,[r_1, \quad r_2, \quad t]$$

Setting up the Equations

- For all the point X_i, we have

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_1, & h_2, & h_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad i = 1 \ldots n \text{ for } n \text{ points}$$

- We perform similar steps in DLT with 3 columns taken away

$$\begin{bmatrix} -X_i, & -Y_i, & -Z_i, & -1, & 0, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_iZ_i, & x_i \\ 0, & 0, & 0, & 0, & -X_i, & -Y_i, & -Z_i, & -1, & y_iX_i, & y_iY_i, & y_iZ_i, & y_i \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = 0$$

Setting up Equations

- Solving the system of linear equations will give you H
- H has DoF of 8, each point gives you two equations, therefore you need at least 4 points

$$\begin{bmatrix} -X_i, & -Y_i, & -1, & 0, & 0, & 0, & x_iX_i, & x_iY_i, & x_i \\ 0, & 0, & 0, & -X_i, & -Y_i, & -1, & y_iX_i, & y_iY_i, & y_i \\ \vdots & & & & & & & & \vdots \\ & & & & & & & & \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = [0]$$

A 2nx9 **h** 9x1 **0** 2n

How to solve?

SVD

Estimation of Calibration matrix from H

- $\mathrm{H} = [h_1, \quad h_2, \quad h_3] = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix}$

- How can we can K and R and X_0? Can we use QR decomposition again?
- No!

$$\begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \text{ not a rotation matrix anymore.}$$

Estimation of Calibration matrix from H

$$\mathrm{H} = [h_1, \quad h_2, \quad h_3] = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix}$$

As

$$\mathrm{H} = [h_1, \quad h_2, \quad h_3] = \mathrm{K}\,[r_1, \quad r_2, \quad t]$$

we have

$$r_1 = \mathrm{K}^{-1} h_1 \text{ ----------------------eq(1)}$$
$$r_2 = \mathrm{K}^{-1} h_2 \text{ ----------------------eq(2)}$$

Other Constraints

Any rotation matrix is an orthonormal matrix

$r_1 = \mathrm{K}^{-1} h_1$ ---------------------eq(1)
$r_2 = \mathrm{K}^{-1} h_2$ ---------------------eq(2)

i.e $$r_i \cdot r_j = \begin{cases} 0 & when\ i \neq j \\ 1 & when\ i = j \end{cases}$$

$$r_1^T r_2 = {h_1}^T K^{-T} K^{-1} h_2 = 0$$ --eq(3)

$$r_1^T r_1 = r_2^T r_2 = {h_1}^T K^{-T} K^{-1} h_1 - {h_2}^T K^{-T} K^{-1} h_2 = 0$$ -------eq(4)

Other Constraints

Let B be a symmetric and positive definite matrix such that

$$\mathrm{B} = \mathrm{K}^{-\mathrm{T}}\mathrm{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{11} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{23} & B_{33} \end{bmatrix} \text{where } B_{13} = B_{31}, B_{23} = B_{32}$$

Since B is symmetric matrix, B is defined by 6D vector

$$\mathrm{B} = \left[B_{11}, B_{12,} B_{13}, B_{22,} B_{23}, B_{33}\right]^{T}$$

Other Constraints

Let B be a symmetric and positive definite matrix such that

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

We have

$$h_1{}^T\,\mathrm{B}h_2 = 0 \text{ - (eq. 5)}$$

$$h_1{}^T\mathrm{B}h_1 - h_2{}^T\mathrm{B}h_2 = 0 \text{ - (eq. 6)}$$

Note: One image gives two linear equation in elements of B, how many images we need?

Ans: 3 images

Solving K from B

- $\begin{bmatrix} B_{11} & B_{12} & B_{11} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} c & s & x_H \\ 0 & c(1+m) & y_H \\ 0 & 0 & 1 \end{bmatrix}^{-T} \begin{bmatrix} c & s & x_H \\ 0 & c(1+m) & y_H \\ 0 & 0 & 1 \end{bmatrix}^{-1}$

$$c = \sqrt{\frac{\lambda}{B_{11}}}$$

$$s = -\frac{B_{12}c^3(1+m)}{\lambda}$$

$$\lambda = B_{33} - \frac{{B_{33}}^2 + u_y(B_{12}B_{13} - B_{11}B_{23})}{B_{11}}$$

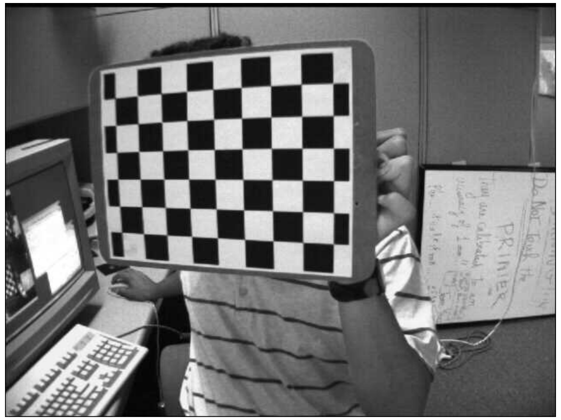
$$m = \frac{1}{c}\sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - {B_{12}}^2} - 1}$$

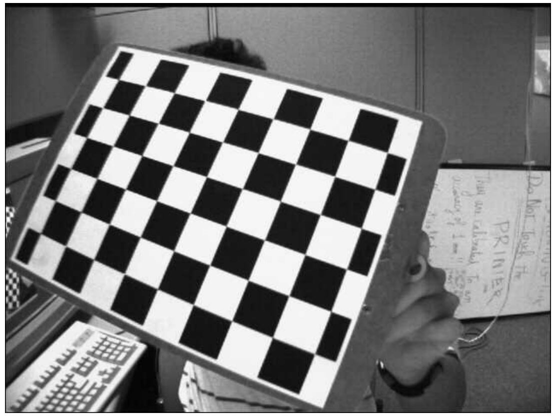
$$u_y = \frac{B_{12}B_{13} - B_{13}B_{23}}{B_{11}B_{22} - {B_{12}}^2}$$

$$u_x = \frac{s u_y}{c} - B_{13}\frac{c^2}{\lambda}$$

What do we need?

- Each Plane gives ***2*** equations
- Since B has ***6 degree of freedom***, we nee at least ***3 different views*** of a plane





- We need at least ***4 points*** per plane to compute the Homography
- Solve B and then you can compute K by Cholesky decomposition

Solving for Extrinsic Parameters

- Once K is known, the extrinsic parameters

$$r_1 = \lambda K^{-1} h_1$$
$$r_2 = \lambda K^{-1} h_2$$

where

$$\lambda = \frac{1}{\|K^{-1} h_1\|}$$

r_3 can be retrieved by cross-production of

$$r_3 = r_1 \times r_2$$

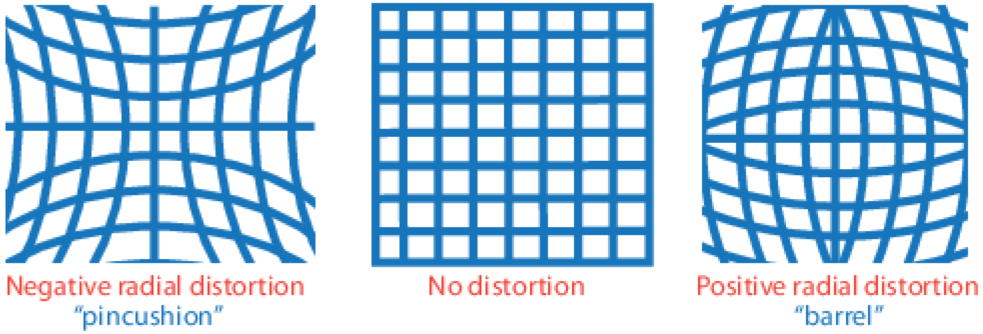
Translation t can be found by

$$t = \lambda K^{-1} h_3$$

Non-linear Error

- Non-linear Effect
  - Radial Distortion
  - Tangential Distortion

- Radial Distortion
  - $x_{distorted} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$
  - $y_{distorted} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$



- Tangential Distortion
  - $x_{distorted} = x + [2p_1 xy + p_2(r^2 + 2x^2)]$
  - $y_{distorted} = y + [p_1(r^2 + 2x^2) + 2p_2 xy]$

Error Minimization

- Lens distortion can be calculated by minimizing a non-linear error function

- $\min_{(\mathrm{K},\mathbf{q},R_{i,}\mathbf{t}_i)} \sum_i \sum_j \left\| \mathbf{X}_{\boldsymbol{ij}} - \hat{\mathrm{x}}(\mathrm{K},\mathbf{q},R_{i,}\mathbf{t}_i,\mathbf{X}_{\boldsymbol{ij}}) \right\|^2$

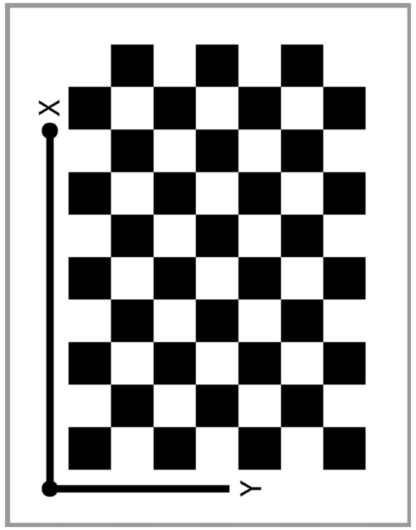
- Solved by using Levenberg-Marquardt
  - [Levenberg–Marquardt algorithm - Wikipedia]

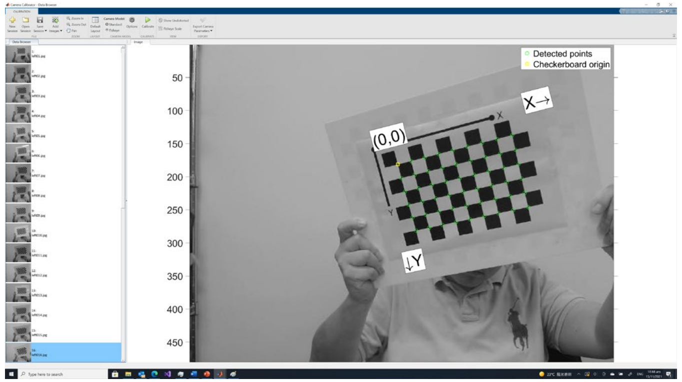
Summary of Steps for Checkboard calibration

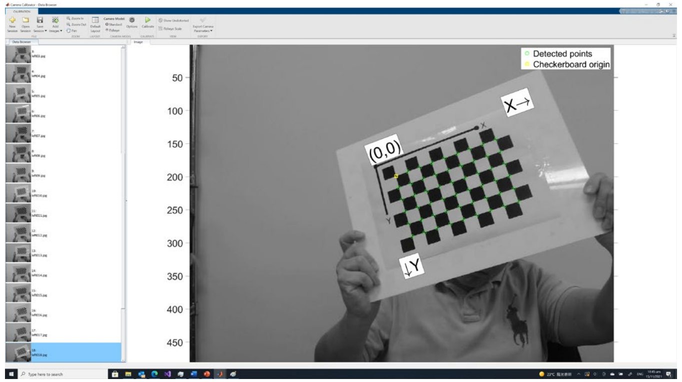
1. Print a ***pattern*** and attach it to a ***planar*** surface
2. Take a few images of the model plane under ***different orientations*** by moving either the plane or the camera
3. Detect the feature points in the images
4. Estimate the ***five intrinsic parameters*** and all the extrinsic parameters using the closed-form solution
5. Estimate the coefficient of the ***radial distortion*** by solving the ***linear least-squares***
6. Refine all parameters, including lens distortion parameters by minimizing

MATLAB Camera Calibration Toolbox

- MATLAB provides the toolbox for camera calibration
- You will need a checkerboard. Type "open checkerboardPattern.pdf"
- Measure the checkerboard sizes and take 15+ pictures of the checkerboard pattern with different angles.



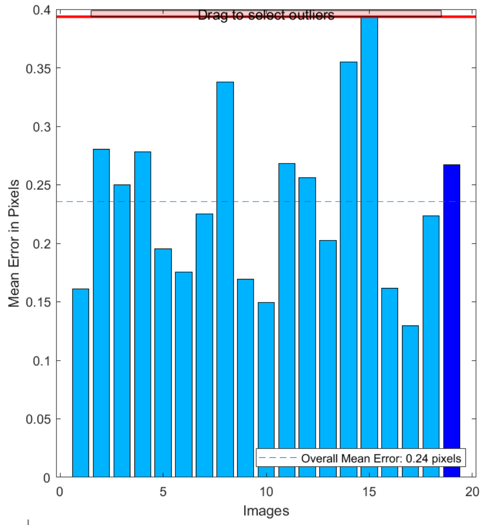


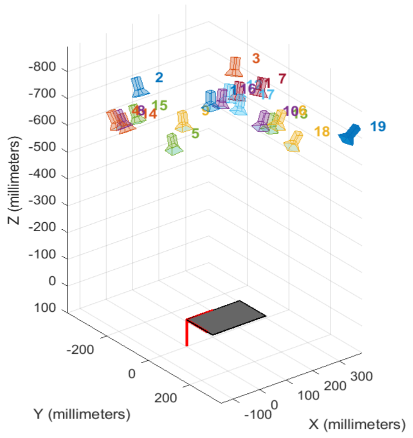


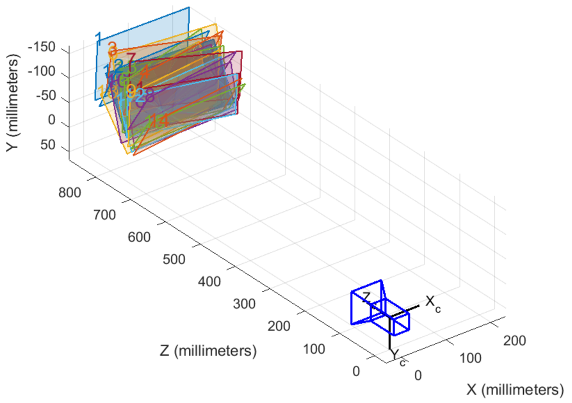
https://www.mathworks.com/help/vision/ug/using-the-single-camera-calibrator-app.html

MATLAB Camera Calibration Toolbox

- Open the camera calibrator in MATLAB
- Add all the images and calibrate



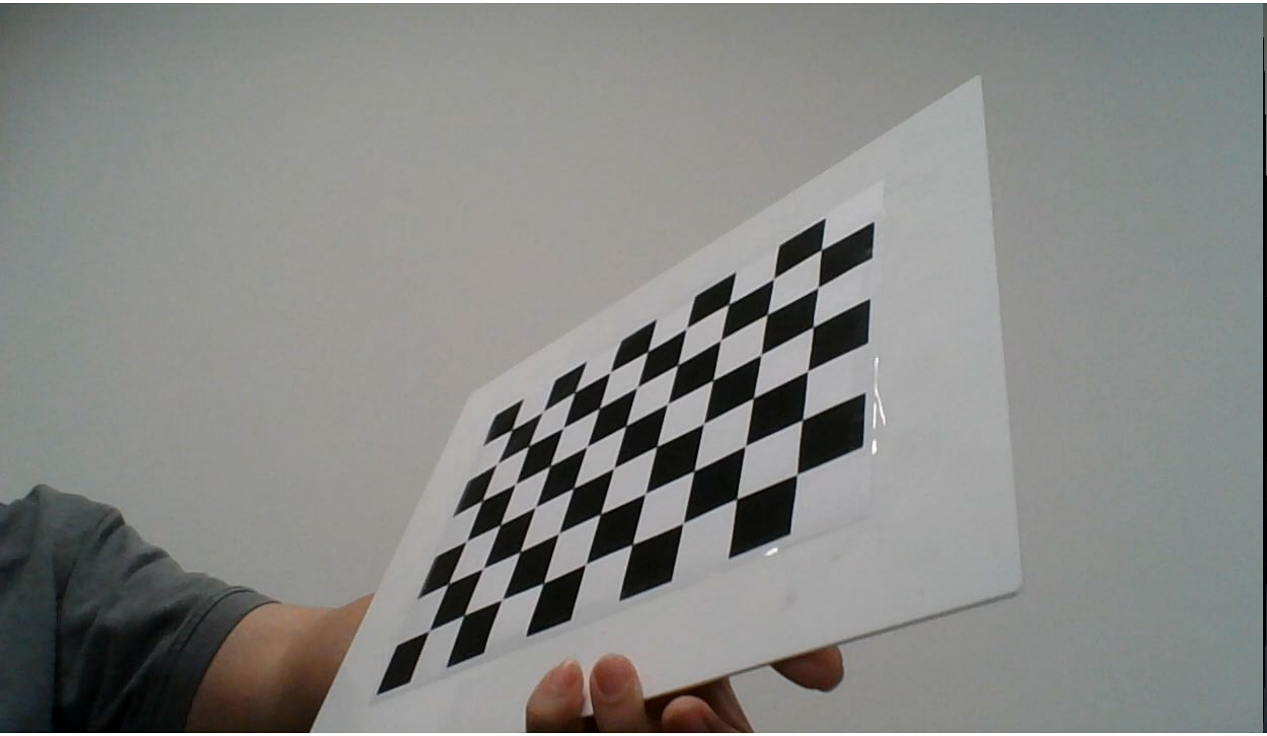




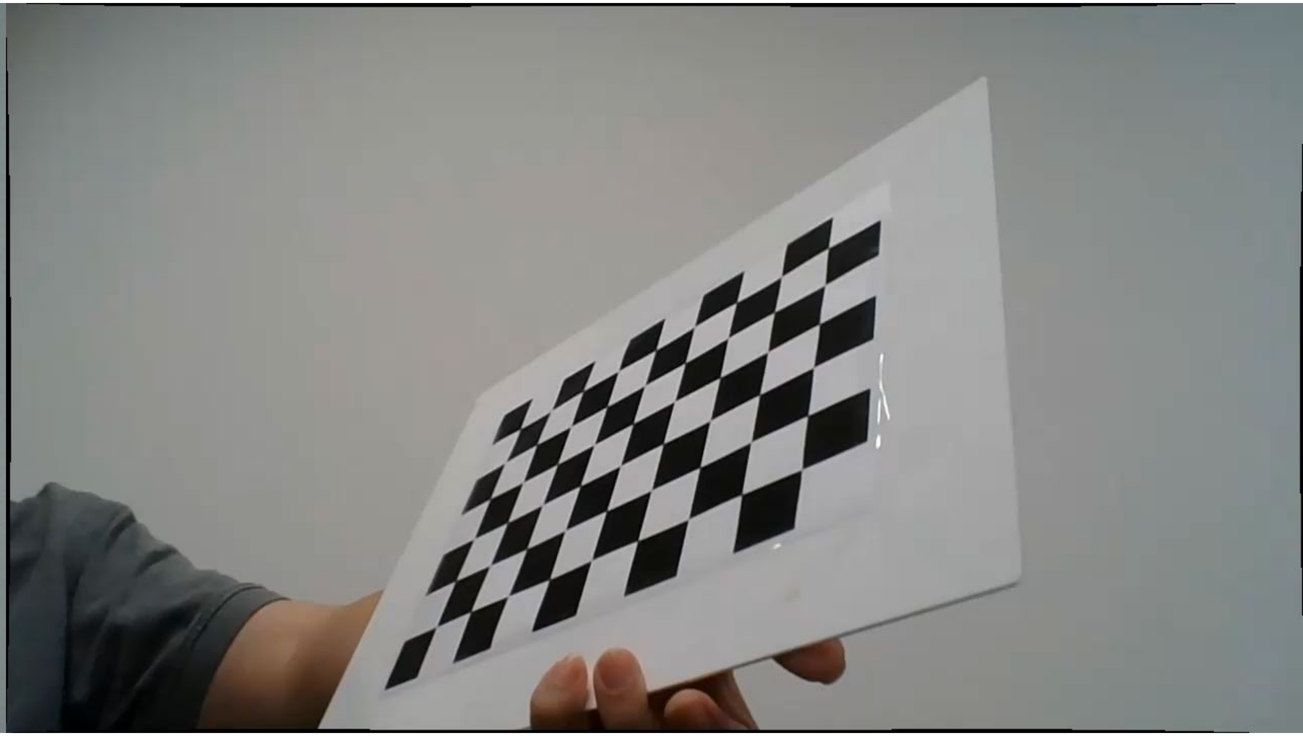
FocalLength	[650.0602,650.1585]
PrincipalPoint	[305.5123,253.3729]
ImageSize	[480,640]
RadialDistortion	[-0.0018,0.0326]
TangentialDistortion	[0,0]
Skew	0
IntrinsicMatrix	[650.0602,0,0;0,650.1585,0;305.5123,253.3729,1]

https://www.mathworks.com/help/vision/ug/using-the-single-camera-calibrator-app.html

Applications: Calibrate distorted video

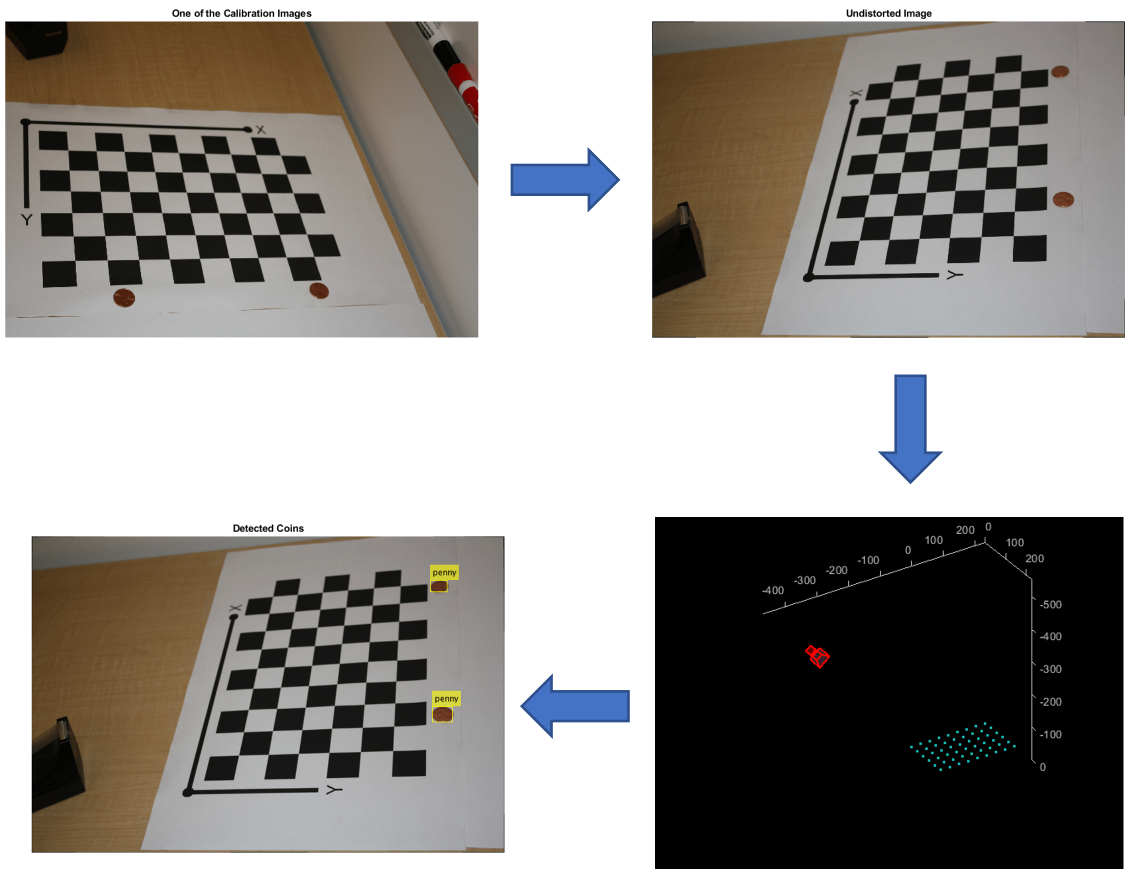


Original Video



Undistorted Video

Applications: Measurement of Planar Objects



1. Calibrate the camera
2. Take the images and undistort them.
3. Detect the coins (blob analysis or Hough transform
4. Estimate the camera extrinsic parameters R, T using "extrinsic" function by MATLAB
5. Get world coordinate of the coins

[Measuring Planar Objects with a Calibrated Camera - MATLAB & Simulink (mathworks.com)]

Reading

- Wolfgang Forstner, Bernhard P. Wrobel (2016), Photogrammetric Computer Vision, Statistic, Geometry, Orientation and Reconstruction, Springer. Chapter 12

- Zhang, Z. (2000). A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330–1334.