# Computer Vision (CS763)

**Teaching cameras to "see"**

# Camera Extrinsics and Intrinsics

**Arjun Jain**

# Mapping as a Two Step Process

1. Projection of the affine camera

$$^{s}\mathbf{x} = \mathsf{P}\mathbf{X}$$

2. Consideration of non-linear effects

$$^{a}\mathbf{x} = {}^{a}\mathsf{H}_{s}(\boldsymbol{x})\, {}^{s}\mathbf{x}$$

**Individual mapping for each point!**

# What To Do If We Want To Get Information About The Scene?

# Inversion of the Mapping

- **Goal:** map from ${}^{a}\mathbf{x}$ back to $\mathbf{X}$
- 1st step: ${}^{a}\mathbf{x} \rightarrow {}^{s}\mathbf{x}$
- 2nd step: ${}^{s}\mathbf{x} \rightarrow \mathbf{X}$

# Inversion of the Mapping

- Goal: map from ${}^{a}\mathbf{x}$ back to $\mathbf{X}$
- **1st step:** ${}^{a}\mathbf{x} \rightarrow {}^{s}\mathbf{x}$
- 2nd step: ${}^{s}\mathbf{x} \rightarrow \mathbf{X}$

$${}^{a}\mathbf{x} \rightarrow {}^{s}\mathbf{x}$$

- The general nature of ${}^{a}\mathrm{H}_{s}(\boldsymbol{x})$ in ${}^{a}\mathbf{x} = {}^{a}\mathrm{H}_{s}(\boldsymbol{x})\, {}^{s}\mathbf{x}$ requires an iterative solution

**depends on the coordinate of the point to transform**

# Inversion Step 1: ${}^{a}\mathbf{x} \rightarrow {}^{s}\mathbf{x}$

- Iteration due to unknown $\boldsymbol{x}$ in ${}^{a}\mathsf{H}_{s}(\boldsymbol{x})$
- Start with ${}^{a}\mathbf{x}$ as the initial guess

$$\mathbf{x}^{(1)} = [\,{}^{a}\mathsf{H}_{s}(\,{}^{a}\mathbf{x})]^{-1}\;{}^{a}\mathbf{x}$$

often w.r.t. the principal point

- and iterate

$$\mathbf{x}^{(\nu+1)} = \left[\,{}^{a}\mathsf{H}_{s}(\mathbf{x}^{(\nu)})\right]^{-1}\;{}^{a}\mathbf{x}$$

- As ${}^{a}\mathbf{x}$ is often a good initial guess, this procedure converges quickly

# Inversion Step 2: ${}^{s}\mathbf{x} \rightarrow \mathbf{X}$

- The next step is the **inversion of the projective mapping**
- We cannot reconstruct the 3D point but the ray though the 3D point
- With the known matrix P, we can write

$$
\begin{aligned}
\lambda \mathbf{x} &= \mathsf{P}\mathbf{X} = \mathsf{K}R[I_3| - \boldsymbol{X}_O]\mathbf{X} \\
&= [\mathsf{K}R| - \mathsf{K}R\boldsymbol{X}_O] \begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix} \\
&= \mathsf{K}R\boldsymbol{X} - \mathsf{K}R\boldsymbol{X}_O
\end{aligned}
$$

factor resulting from the H.C.

# Inversion Step 2: ${}^{s}\mathbf{x} \rightarrow \mathbf{X}$

- Starting from $\lambda \mathbf{x} = \mathsf{K}R\boldsymbol{X} - \mathsf{K}R\boldsymbol{X}_O$
- we obtain

$$\begin{aligned} \boldsymbol{X} &= (\mathsf{K}R)^{-1}\mathsf{K}R\boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x} \\ &= \boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x} \end{aligned}$$

- The term $\lambda(\mathsf{K}R)^{-1}\mathbf{x}$ describes the direction of the ray from the camera origin $\boldsymbol{X}_O$ to the 3D point $\boldsymbol{X}$

# Determining Depth

- Ideally, we if we now have 2 cameras and two projections of the the same 3D point, we can obtain two such rays and simply find the intersection

$$\begin{aligned} \boldsymbol{X} &= (\mathsf{K}R)^{-1}\mathsf{K}R\boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x} \\ &= \boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x} \end{aligned}$$

- Why will this not work? What to do? (We will see this in detail when doing Stereo)

# Classification of Cameras

extrinsic parameters

intrinsic parameters

| | | | | | |
|---|---|---|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | $c$ | $x_H, y_H$ | $m, s$ | $q_1, q_2, \ldots$ |

# Classification of Cameras

| extrinsic parameters | |
|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | |
| normalized | |

**Example: pinhole camera for which the principal point is the origin of the image coordinate system, the x- and y-axis of the image coordinate system is aligned with the x-/y-axis of the world c.s. and the distance between the origin and the image plane is 1**

# Classification of Cameras

extrinsic parameters

| | | |
|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | |
| normalized | | |
| unit camera | | |

**Example: pinhole camera for which the principal point (x, y) is the origin of the image coordinate system and the distance between the origin and the image plane is 1**

# Classification of Cameras

| extrinsic parameters | | intrinsic parameters | |
|---|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | $c$ | |
| normalized | | | |
| unit camera | | | |
| ideal camera | | | |

**Example: pinhole camera for which the x/y coordinate of the principal point is the origin of the image coordinate system**

# Classification of Cameras

| extrinsic parameters | | intrinsic parameters | | |
|---|---|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | $c$ | $x_H, y_H$ | |
| normalized | | | | |
| unit camera | | | | |
| ideal camera | | | | |
| Euclidian camera | | | | |

**Example: pinhole camera using a Euclidian sensor in the image plane**

# Classification of Cameras

| extrinsic parameters | | intrinsic parameters | | | |
|---|---|---|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | $c$ | $x_H, y_H$ | $m, s$ | |
| normalized | | | | | |
| unit camera | | | | | |
| ideal camera | | | | | |
| Euclidian camera | | | | | |
| affine camera | | | | | |

**Example: camera that preserves straight lines**

# Classification of Cameras

| extrinsic parameters | | intrinsic parameters | | | |
|---|---|---|---|---|---|
| $\boldsymbol{X}_0$ $(X, Y, Z)$ | $R$ $(\omega, \phi, \kappa)$ | $c$ | $x_H, y_H$ | $m, s$ | $q_1, q_2, \ldots$ |
| normalized | | | | | |
| unit camera | | | | | |
| ideal camera | | | | | |
| Euclidian camera | | | | | |
| affine camera | | | | | |
| general camera | | | | | |

**Example: camera with non-linear distortions**

# Calibration Matrices

| camera | calibration matrix | #parameters |
|---|---|---|
| unit | ${}^0\mathsf{K} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | 6 (6+0) |
| ideal | ${}^k\mathsf{K} = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | 7 (6+1) |
| Euclidian | ${}^p\mathsf{K} = \begin{bmatrix} c & 0 & x_H \\ 0 & c & y_H \\ 0 & 0 & 1 \end{bmatrix}$ | 9 (6+3) |
| affine | $\mathsf{K} = \begin{bmatrix} c & cs & x_H \\ 0 & c(1+m) & y_H \\ 0 & 0 & 1 \end{bmatrix}$ | 11 (6+5) |
| general | ${}^a\mathsf{K} = \begin{bmatrix} c & cs & x_H + \Delta x \\ 0 & c(1+m) & y_H + \Delta y \\ 0 & 0 & 1 \end{bmatrix}$ | 11+N |

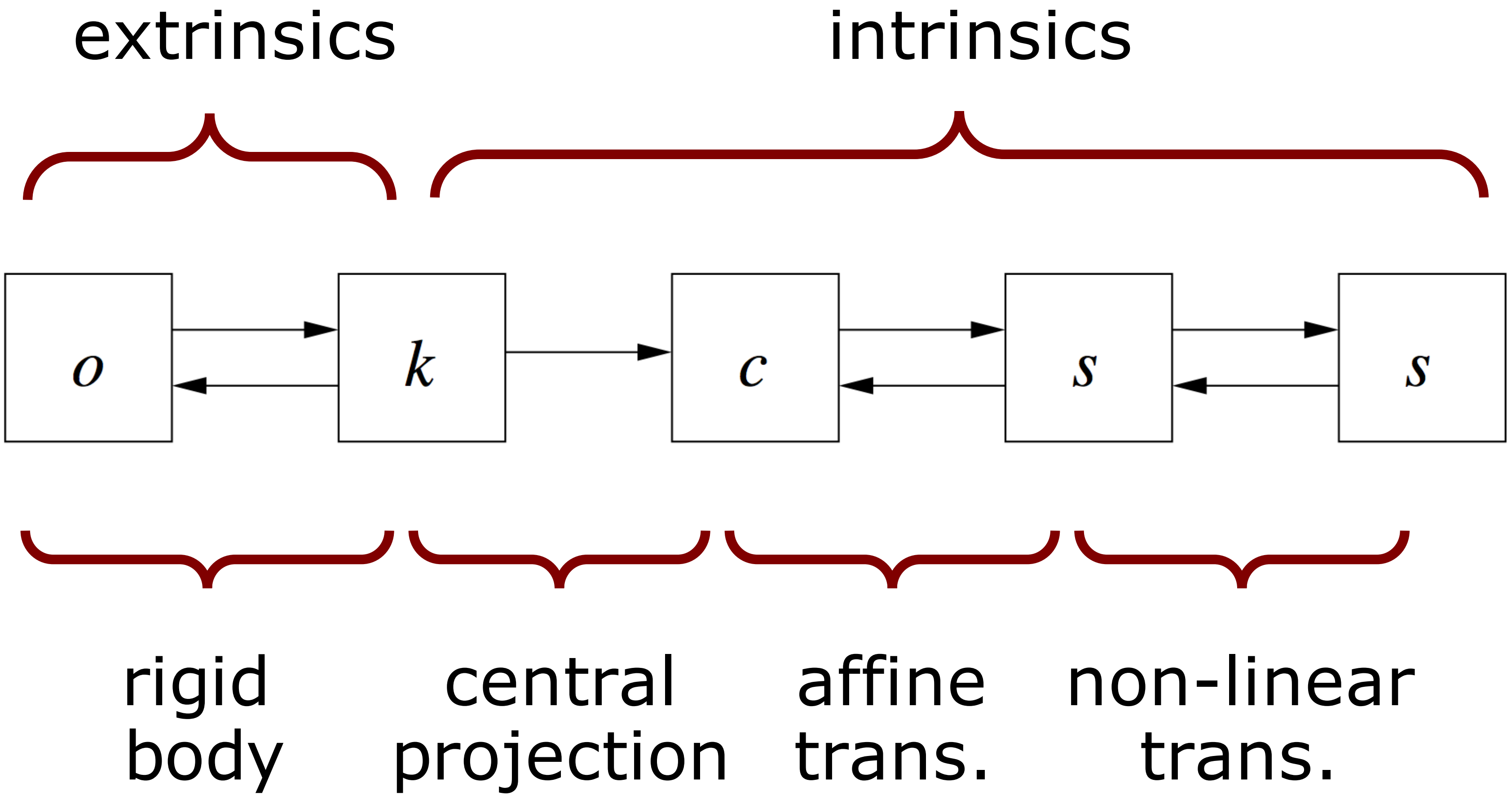
# Calibrated Camera

- If the intrinsics are **unknown**, we call the camera **uncalibrated**
- If the intrinsics are **known**, we call the camera **calibrated**
- If the intrinsics are known and do not change, the camera is called **metric camera**
- The process of obtaining the intrinsics is called **camera calibration**

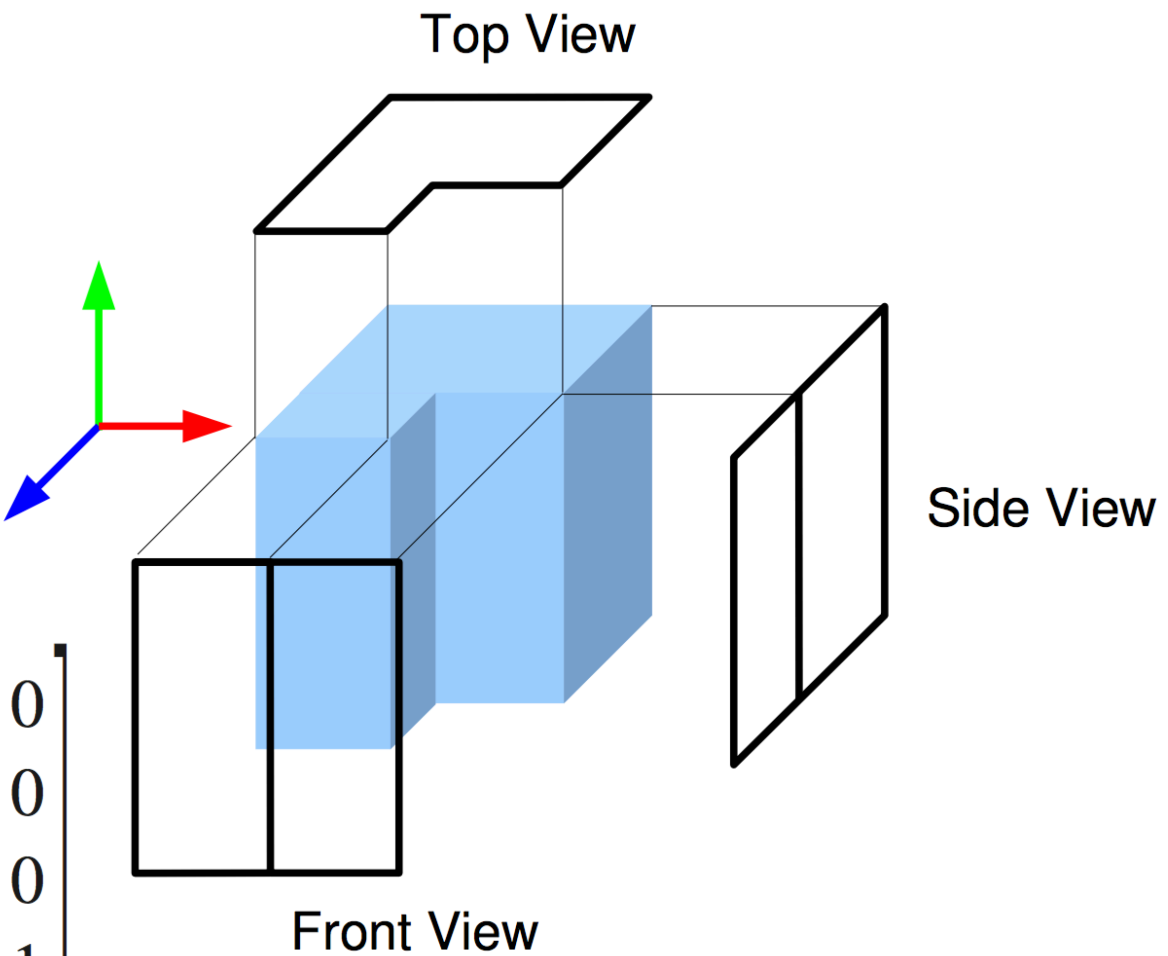
# Summary

- We described the mapping from the world c.s. to the image c.s.
- **Extrinsics** = world to camera c.s.
- **Intrinsics** = camera to sensor c.s.
- **DLT** = Direct linear transform
- Non-linear errors
- Inversion of the mapping process

# Summary of the Mapping



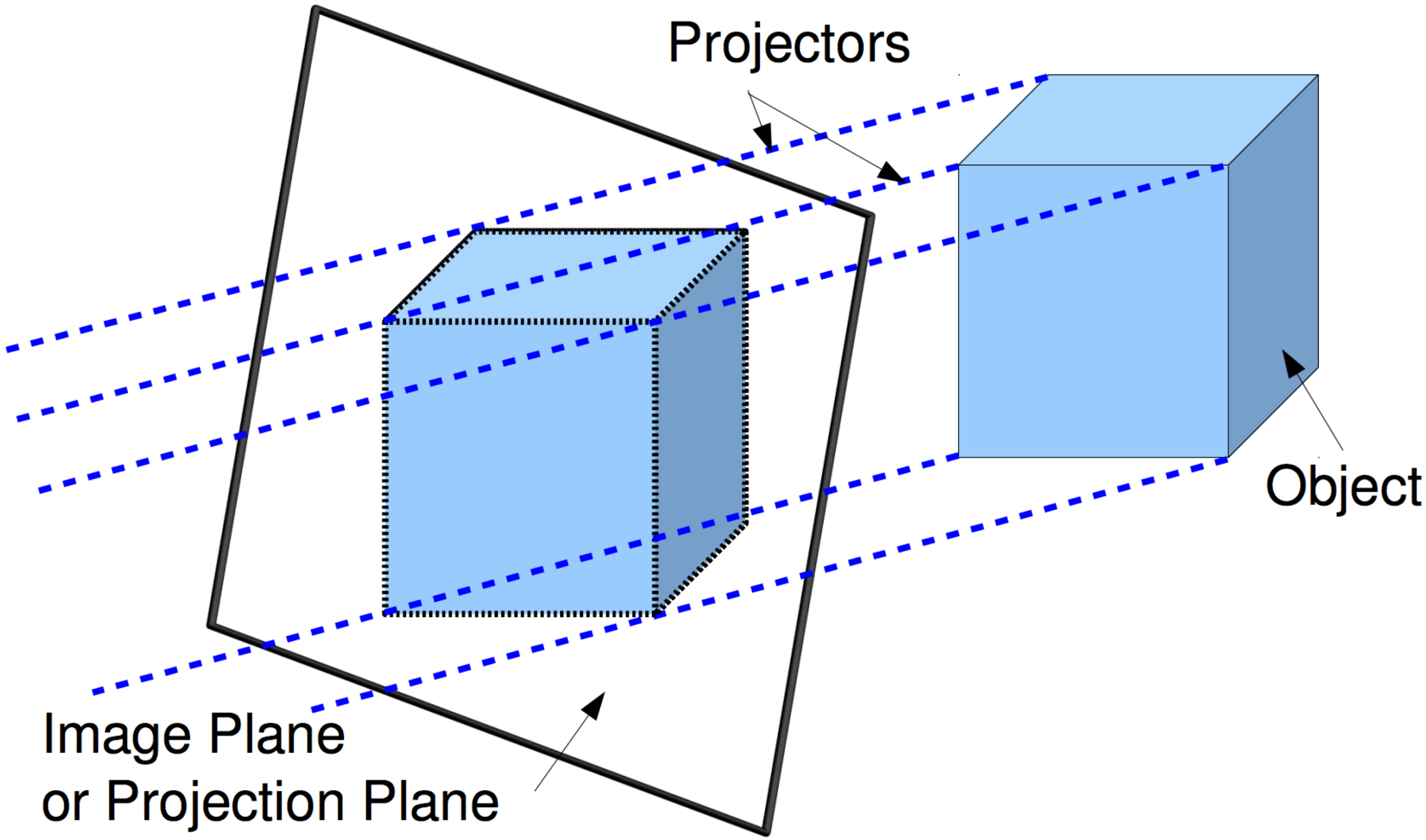
# Orthographic Camera Model




$$P(z=0)=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

https://www.cse.iitb.ac.in/~paragc/teaching/2017/cs475/lectures/06_viewing_1.pdf

# Orthographic Camera Model




https://www.cse.iitb.ac.in/~paragc/teaching/2017/cs475/lectures/06_viewing_1.pdf

# Orthographic Camera Model

- Therefore, the projection defined by the orthographic camera is $(X, Y, Z)^T \longmapsto (X, Y)$

# Weak-perspective Camera Model

- If the relative variation in depth (Z coordinate) of different points in the scene is negligible compared to the average depth Z′ of the scene, then the projection can be defined as weak perspective:

$$X \longmapsto \frac{fX}{Z' + \delta} = \frac{fX}{Z'}\left(\frac{1}{1 + \frac{\delta}{Z'}}\right) \approx \frac{fX}{Z'}\left(1 - \frac{\delta}{Z'} + \left(\frac{\delta}{Z'}\right)^2\right) \approx \frac{fX}{Z'}$$

$$Y \longmapsto \frac{fY}{Z' + \delta} = \frac{fY}{Z'}\left(\frac{1}{1 + \frac{\delta}{Z'}}\right) \approx \frac{fY}{Z'}\left(1 - \frac{\delta}{Z'} + \left(\frac{\delta}{Z'}\right)^2\right) \approx \frac{fY}{Z'}$$

# Weak-perspective Camera Model

- Weak perspective projection is Orthographic projection followed by uniform scaling by a factor

- That is, $(X, Y, Z)^T \longmapsto \left(X \frac{f}{z'}, Y \frac{f}{z'}\right)$

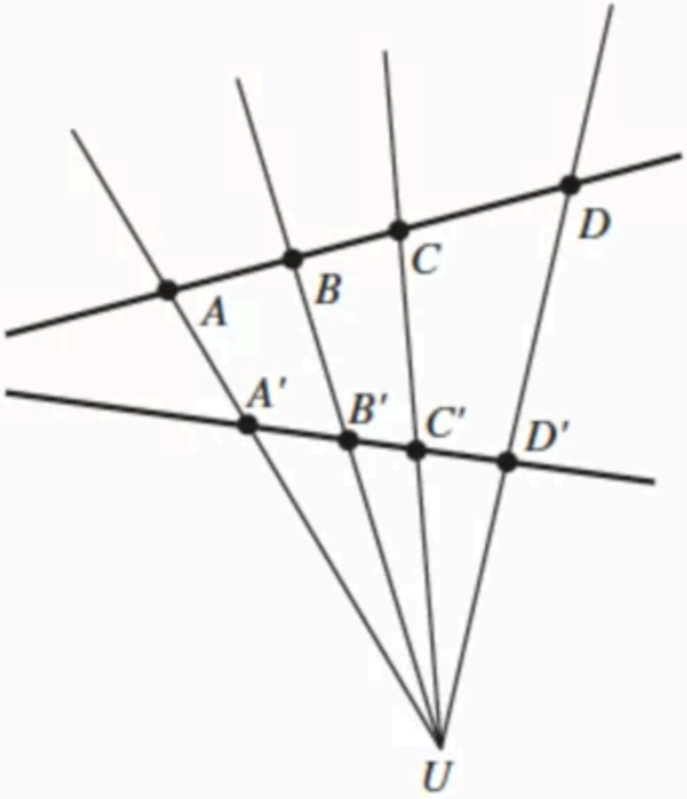
# Projective Invariant: Cross Ratio

# Use it for Single View Metrology

https://www.coursera.org/learn/robotics-perception/lecture/AV44m/cross-ratios-and-single-view-metrology

# What is Cross Ratio?

collinear




Given four points $A, B, C, D$, we define the cross-ratio of their distances as

$$CR(A,B,C,D) = \frac{AC}{AD} : \frac{BC}{BD}.$$

$CR(A,B,C,D)$ remains invariant under projective transformations

$$\frac{AC}{AD} : \frac{BC}{BD} = \frac{A'C'}{A'D'} : \frac{B'C'}{B'D'}$$

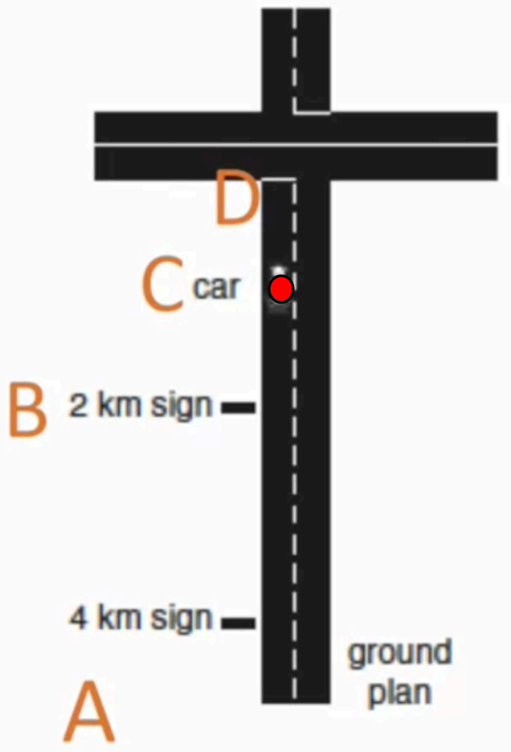
https://www.cut-the-knot.org/pythagoras/Cross-Ratio.shtml

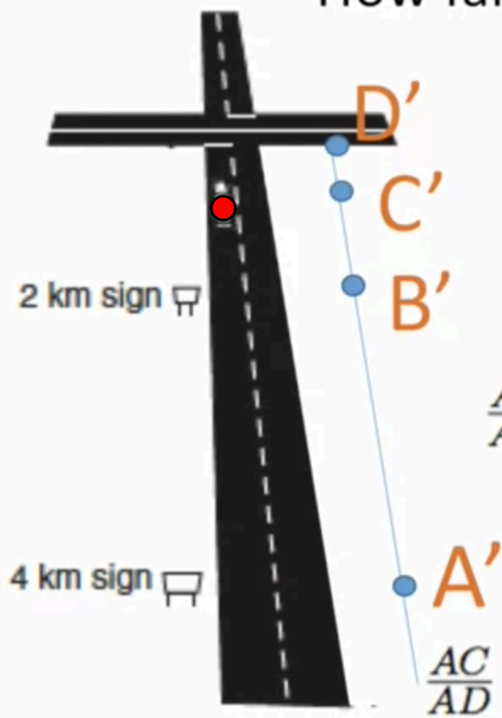
# What is Cross Ratio?

- Cross-ratios are invariant under projective transformations, i.e. the cross ratio of four collinear in 3D = cross ratio of their projections (i.e., images) onto a 2D image plane!

- Recall – lengths, areas, ratios of lengths, ratios of areas are not preserved under projective transformations (btw ratios of lengths and areas are preserved under affine transformation).

# Use of Cross Ratio?




How far is car from intersection:

CD=?

A’D’=300 pixels

A’C’=275pixels

B’C’=50 pixels

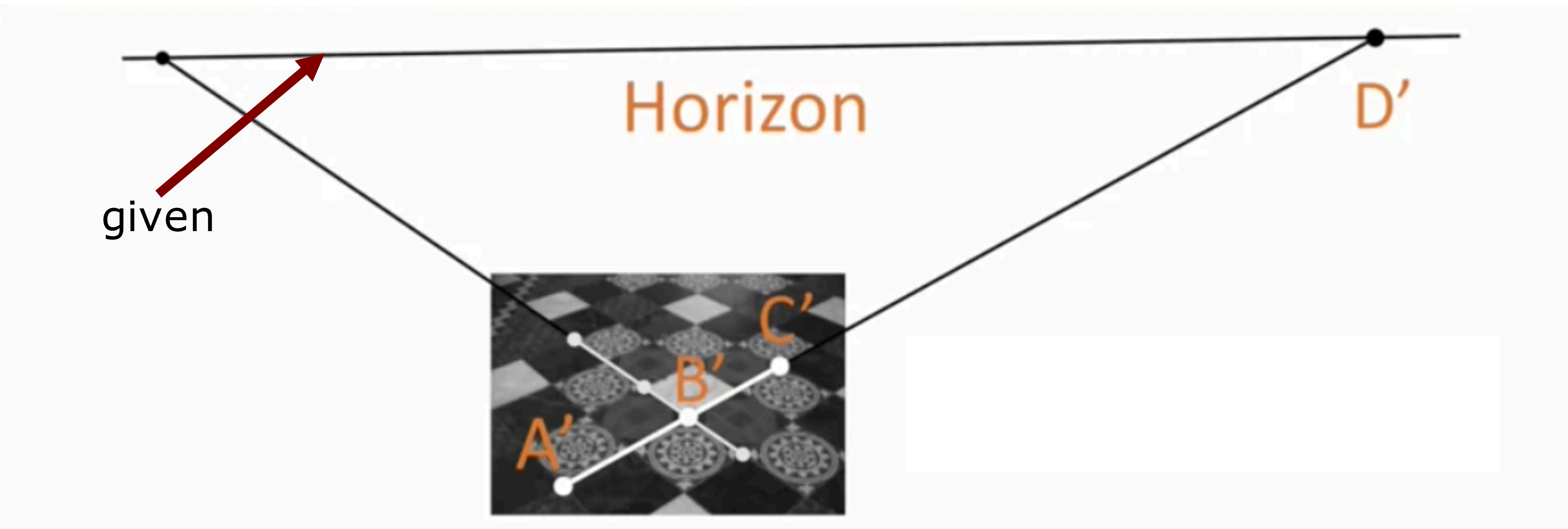
B’D’=75 pixels

$$\frac{A'C'}{A'D'} : \frac{B'C'}{B'D'} = \frac{275}{300} : \frac{50}{75} = 1.375$$

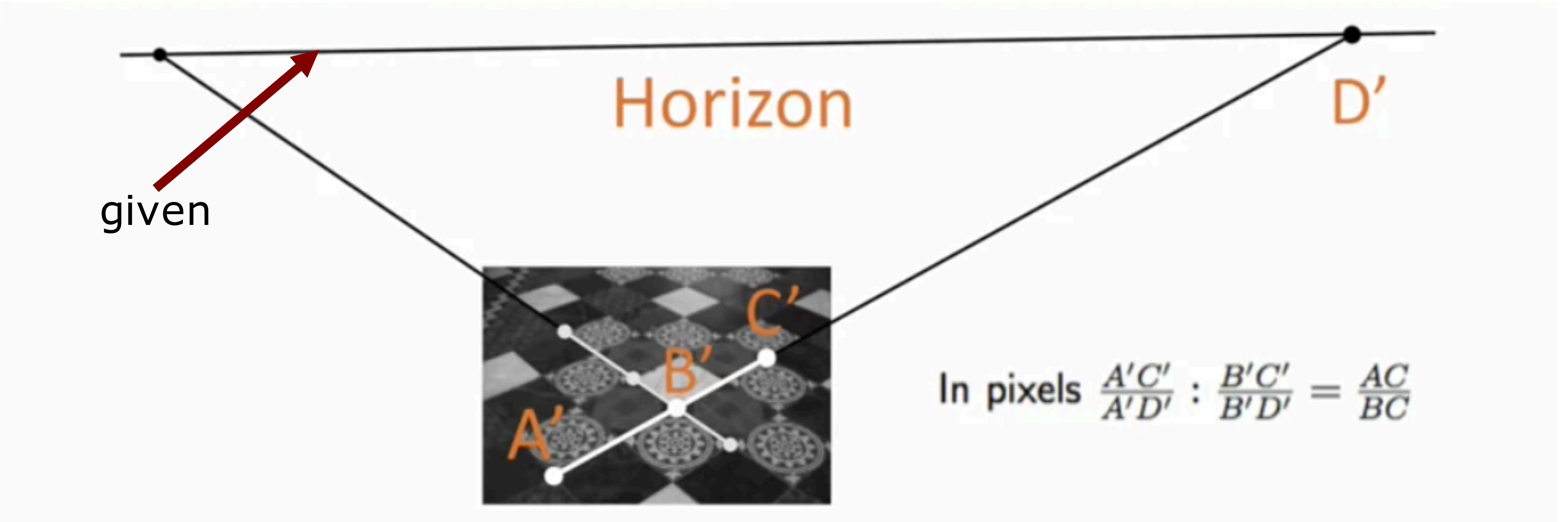
$$\frac{AC}{AD} : \frac{BC}{BD} = \frac{4-CD}{4} : \frac{2-CD}{2} = 1.375$$

$$CD = 0.857km$$

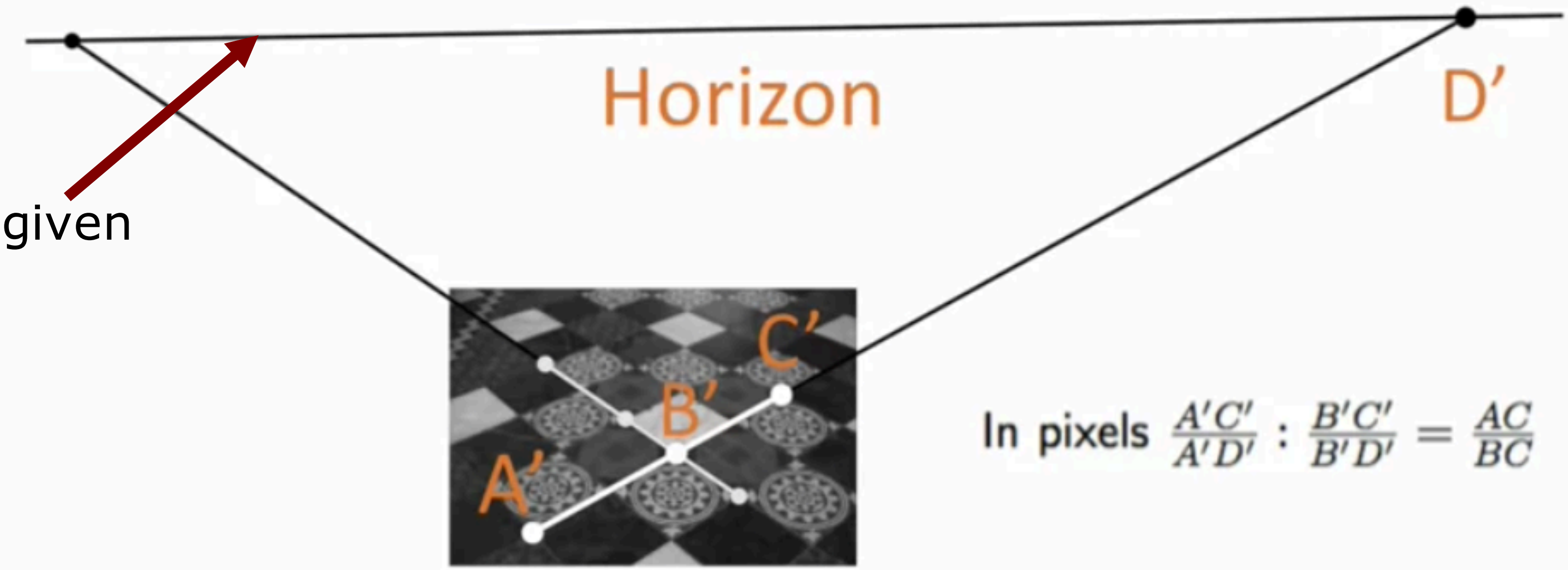
# What happens to the cross-ratio if one of the points is at ∞?



# What happens to the cross-ratio if one of the points is at ∞?



# What happens to the cross-ratio if one of the points is at ∞?

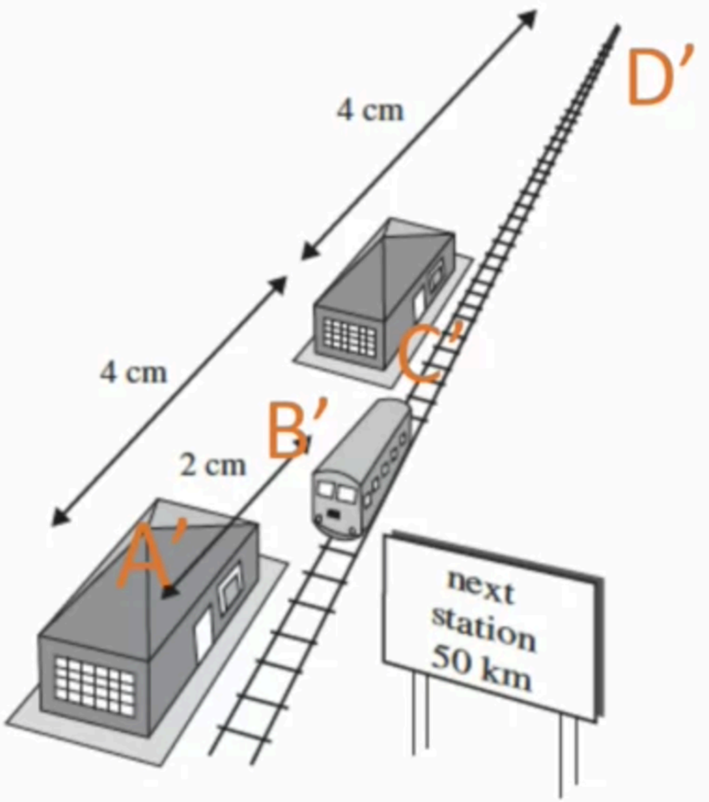



In pixels $\frac{A'C'}{A'D'} : \frac{B'C'}{B'D'} = \frac{AC}{BC}$

When a point $D$ is at infinity, the cross-ratio becomes a ratio !

$$\frac{AC}{AD} : \frac{BC}{BD} = \frac{AC}{BC}$$

# Lets use cross-ratio when one point is at ∞




How far away is the train from the next station? Or BC=?

$$\frac{A'C'}{A'D'} : \frac{B'C'}{B'D'} = \frac{AC}{BC}$$

$$\frac{4}{8} : \frac{2}{6} = \frac{50}{BC}$$

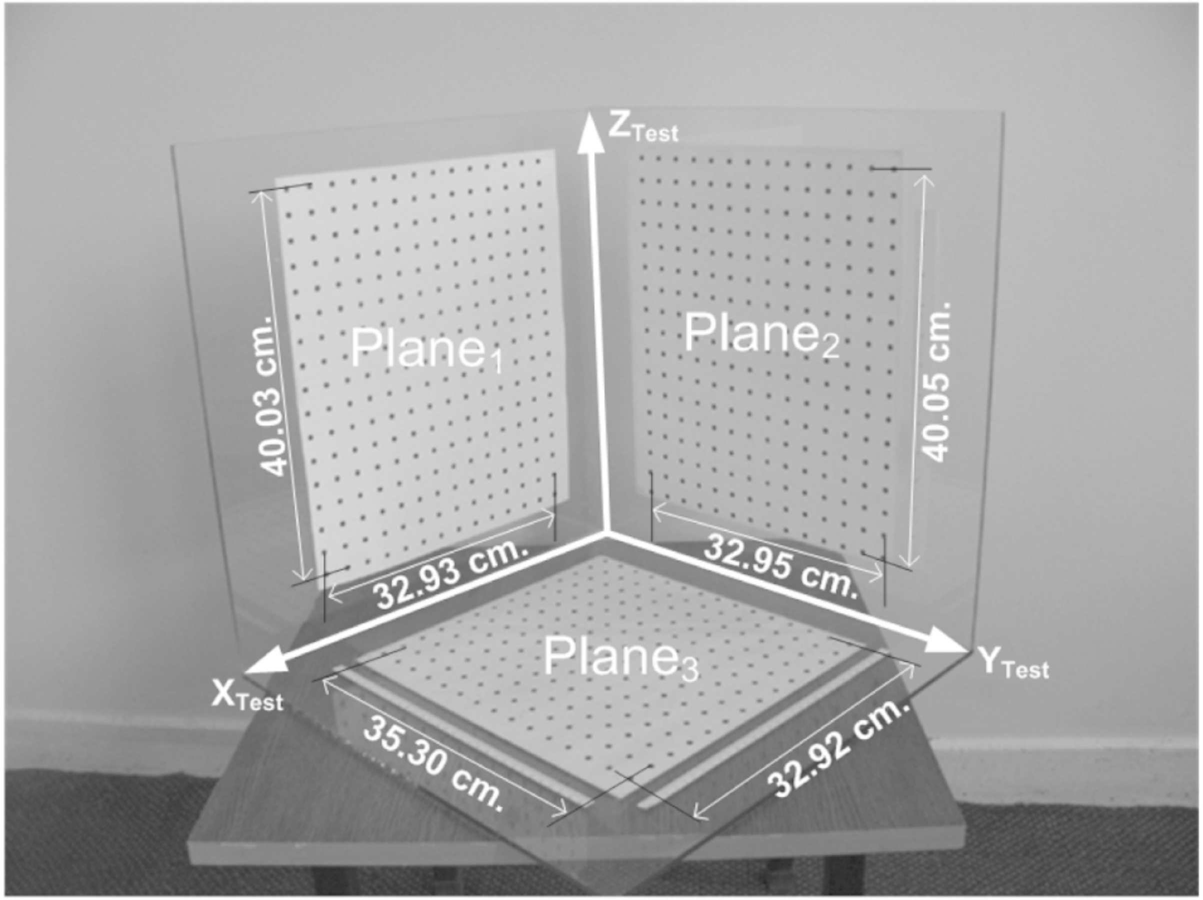
$$\frac{4}{8} : \frac{2}{6} = \frac{50}{BC} \text{ and } BC = 33.33 \text{ km}.$$

Brannan et al. Geometry

# Camera Calibration: Direct Linear Transform and Zhang’s Method

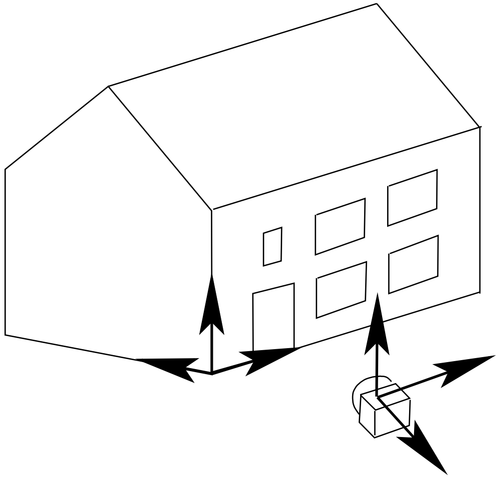
# Estimating Camera Parameters Given the Geometry

known control points in 3D



# Possible Applications

- Camera calibration (obtain intrinsics)
- Estimating the pose of a camera given knowledge about the 3D scene
- Estimating the object location given the known pose of a camera



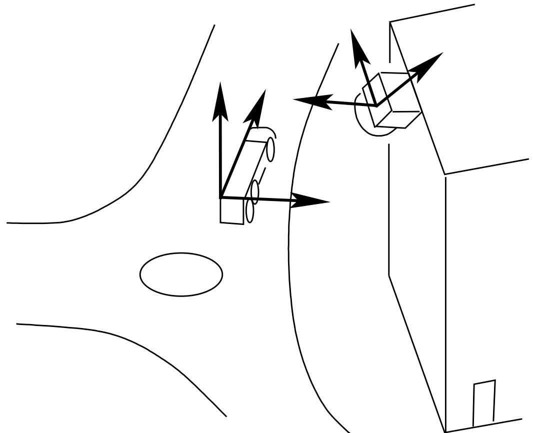


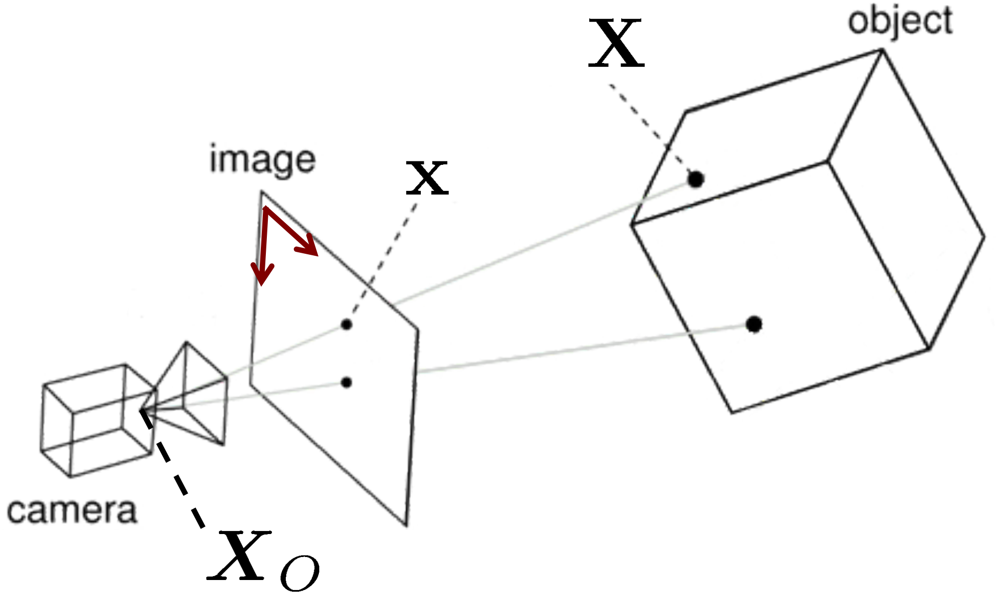
Image courtesy: Förstner

# Estimate Ex- and Intrinsics

- **Wanted:** Extrinsic and intrinsic parameters of a camera
- **Given:** Coordinates of object points (control points)
- **Observed:** Coordinates $(x, y)$ of those object points in an image

# Mapping

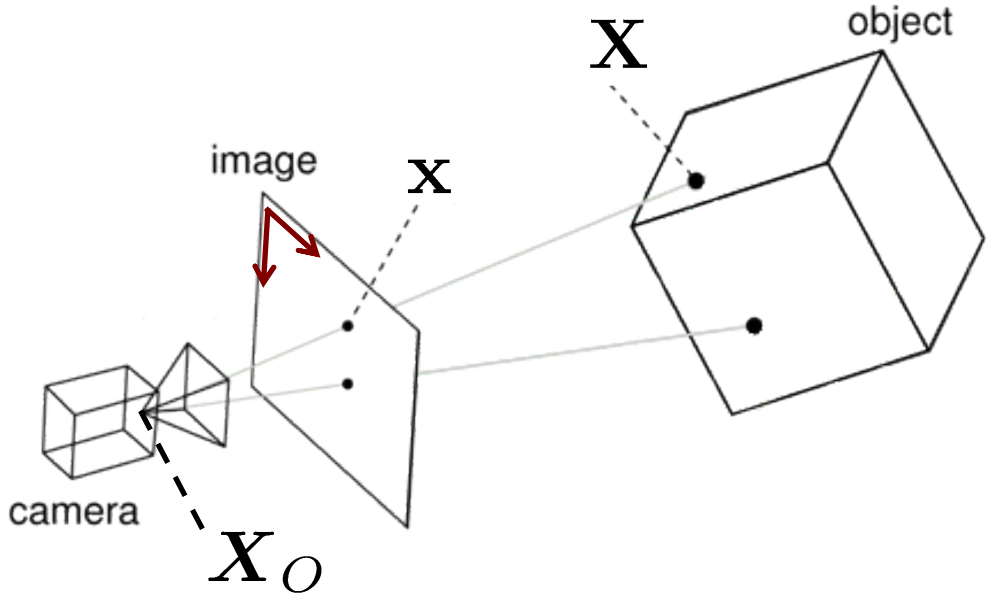
Direct linear transform (DLT) maps any object point $\mathbf{X}$ to the image point $\mathbf{x}$



# Mapping

Direct linear transform (DLT) maps any object point $\mathbf{X}$ to the image point $\mathbf{x}$

$$\begin{aligned} \mathbf{x} &= \mathsf{K}R[I_3| - \boldsymbol{X}_O]\mathbf{X} \\ &= \mathsf{P}\ \mathbf{X} \end{aligned}$$



# Mapping

Direct linear transform (DLT) maps any object point $\mathbf{X}$ to the image point $\mathbf{x}$

$$\underset{3\times 1}{\mathbf{x}} = \underset{3\times 3}{\mathsf{K}} \underset{3\times 3}{R} \underbrace{[\underset{3\times 3}{I_3} \mid -\underset{3\times 1}{\boldsymbol{X}_O}]}_{3\times 4} \underset{4\times 1}{\mathbf{X}}$$

$$= \underset{3\times 4}{\mathsf{P}} \underset{4\times 1}{\mathbf{X}}$$

# Camera Parameters

$$\mathbf{x} = \mathsf{K}R[I_3| - \boldsymbol{X}_O]\mathbf{X} = \mathsf{P}\ \mathbf{X}$$

- **Intrinsic Parameters (I.P.)**
  - Intrinsic parameters of the camera
  - Given through $\mathsf{K}$
- **Extrinsic Parameters (E.P.)**
  - Extrinsic parameters of the camera
  - Given through $\boldsymbol{X}_O$ and $R$
- Projection matrix $\mathsf{P} = \mathsf{K}R[I_3| - \boldsymbol{X}_O]$ contains both I.P. and O.P.

# Direct Linear Transform (DLT)

Compute the **11 intrinsic and extrinsic parameters (I.P. and E.P.)**

$$\mathbf{x} = \mathsf{K} R [I_3 | - \boldsymbol{X}_O] \mathbf{X}$$

observed image point

**c, s, m, $x_H$, $y_H$**

**3 rotations**

**3 translations**

control point coordinates (given)

# How Many Points Are Needed?

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

Each point gives **???** observation equations

# How Many Points Are Needed?

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \mathsf{P} \begin{bmatrix} U \\ V \\ W \\ T \end{bmatrix}$$

Each point gives **???** observation equations

# How Many Points Are Needed?

$$\begin{bmatrix} u/w \\ v/w \\ 1 \end{bmatrix} = \mathrm{P} \begin{bmatrix} U/T \\ V/T \\ W/T \\ 1 \end{bmatrix}$$

Each point gives **???** observation equations

# How Many Points Are Needed?

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathsf{P} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Each point gives **two** observation equations, one for each image coordinate

$$\begin{aligned} x &= \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \\ y &= \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \end{aligned}$$

# Spatial Resection vs. DLT

- **Calibrated camera**
  - 6 unknowns
  - We need at least **3 points**
  - Problem solved by **spatial resection**
- **Uncalibrated camera**
  - 11 unknowns
  - We need at least **6 points**
  - Assuming the model of an **affine camera**
  - Problem solved by **DLT**

# DLT: Direct Linear Transform

## Computing the Orientation of an Uncalibrated Camera Using ≥ 6 Known Points

# DLT: Problem Specification

- Task: Estimate the 11 elements of $\mathsf{P}$
- Given:
  - 3D coordinates $\mathbf{X}_i$ of $I \geq 6$ object points
  - Observed image coordinates $\mathbf{x}_i$ of an uncalibrated camera with the mapping

$$\mathbf{x}_i = \mathsf{P}\,\mathbf{X}_i \qquad i = 1, \ldots, I$$

# Rearrange the DLT Equation

$$\mathbf{x}_i = \underset{3\times 4}{\mathsf{P}}\mathbf{X}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \mathbf{X}_i$$

# Rearrange the DLT Equation

$$\mathbf{x}_i = \underset{3\times 4}{\mathsf{P}}\,\mathbf{X}_i = \begin{bmatrix} \boxed{p_{11} \quad p_{12} \quad p_{13} \quad p_{14}} \\ \boxed{p_{21} \quad p_{22} \quad p_{23} \quad p_{24}} \\ \boxed{p_{31} \quad p_{32} \quad p_{33} \quad p_{34}} \end{bmatrix} \mathbf{X}_i$$

$$= \begin{bmatrix} \boxed{\mathbf{A}^\top} \\ \boxed{\mathbf{B}^\top} \\ \boxed{\mathbf{C}^\top} \end{bmatrix} \mathbf{X}_i$$

# Rearrange the DLT Equation

$$\mathbf{x}_i = \underset{3\times 4}{\mathsf{P}} \mathbf{X}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \mathbf{X}_i$$

$$= \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{B}^\top \\ \mathbf{C}^\top \end{bmatrix} \mathbf{X}_i$$

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{A}^\top \mathbf{X}_i \\ \mathbf{B}^\top \mathbf{X}_i \\ \mathbf{C}^\top \mathbf{X}_i \end{bmatrix}$$

# Rearrange the DLT Equation

$$\mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{A}^\top \mathbf{X}_i \\ \mathbf{B}^\top \mathbf{X}_i \\ \mathbf{C}^\top \mathbf{X}_i \end{bmatrix}$$

$$\Rightarrow$$

$$x_i = \frac{u_i}{w_i} = \frac{\mathbf{A}^\top \mathbf{X}_i}{\mathbf{C}^\top \mathbf{X}_i} \qquad y_i = \frac{v_i}{w_i} = \frac{\mathbf{B}^\top \mathbf{X}_i}{\mathbf{C}^\top \mathbf{X}_i}$$

# Rearrange the DLT Equation

$$x_i = \frac{u_i}{w_i} = \frac{\mathbf{A}^\mathsf{T}\mathbf{X}_i}{\mathbf{C}^\mathsf{T}\mathbf{X}_i} \quad \Rightarrow \quad x_i\,\mathbf{C}^\mathsf{T}\mathbf{X}_i - \mathbf{A}^\mathsf{T}\mathbf{X}_i = 0$$

$$y_i = \frac{v_i}{w_i} = \frac{\mathbf{B}^\mathsf{T}\mathbf{X}_i}{\mathbf{C}^\mathsf{T}\mathbf{X}_i} \quad \Rightarrow \quad y_i\,\mathbf{C}^\mathsf{T}\mathbf{X}_i - \mathbf{B}^\mathsf{T}\mathbf{X}_i = 0$$

Leads to an system of equation, which is **linear in the parameters $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$**

$$\begin{aligned} -\mathbf{X}_i^\mathsf{T}\mathbf{A} \qquad\qquad\qquad & +x_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0 \\ -\mathbf{X}_i^\mathsf{T}\mathbf{B} \quad & +y_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0 \end{aligned}$$

# Estimating the Elements of P

- Collect the elements of P within a vector $\boldsymbol{p}$

$$\boldsymbol{p} = (p_k) = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \end{bmatrix} = \mathrm{vec}(\mathsf{P}^{\mathsf{T}})$$

**rows of P as column-vectors, one below the other (12x1)**

# Estimating the Elements of P

- Rewrite
$$\begin{array}{rrr} -\mathbf{X}_i^\mathsf{T}\mathbf{A} & & +x_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0 \\ & -\mathbf{X}_i^\mathsf{T}\mathbf{B} & +y_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0 \end{array}$$
- as
$$\boldsymbol{a}_{x_i}^\mathsf{T}\boldsymbol{p} = 0$$
$$\boldsymbol{a}_{y_i}^\mathsf{T}\boldsymbol{p} = 0$$
- with

$$\begin{aligned} \boldsymbol{p} &= (p_k) = \mathrm{vec}(\mathsf{P}^\mathsf{T}) \\ \boldsymbol{a}_{x_i}^\mathsf{T} &= (-\mathbf{X}_i^\mathsf{T}, \mathbf{0}^\mathsf{T}, x_i\mathbf{X}_i^\mathsf{T}) \\ &= (-X_i, -Y_i, -Z_i, -1, 0, 0, 0, 0, x_iX_i, x_iY_i, x_iZ_i, x_i) \\ \boldsymbol{a}_{y_i}^\mathsf{T} &= (\mathbf{0}^\mathsf{T}, -\mathbf{X}_i^\mathsf{T}, y_i\mathbf{X}_i^\mathsf{T}) \\ &= (0, 0, 0, 0, -X_i, -Y_i, -Z_i, -1, y_iX_i, y_iY_i, y_iZ_i, y_i) \end{aligned}$$

# Verifying Correctness

$$\boldsymbol{a}_{x_i}^{\mathsf{T}}\boldsymbol{p} = (-X_i,\ -Y_i,\ -Z_i,\ -1, 0,\ 0,\ 0,\ 0, x_i X_i,\ x_i Y_i,\ x_i Z_i,\ x_i) \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}$$

# Verifying Correctness

$$\boldsymbol{a}_{x_i}^{\top}\boldsymbol{p} \quad = \quad (-X_i,\, -Y_i,\, -Z_i,\, -1, 0,\, 0,\, 0,\, 0, x_iX_i,\, x_iY_i,\, x_iZ_i,\, x_i) \begin{bmatrix} \boxed{\begin{matrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \end{matrix}} \\ \boxed{\begin{matrix} p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \end{matrix}} \\ \boxed{\begin{matrix} p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{matrix}} \end{bmatrix} \begin{matrix} \boldsymbol{A} \\ \\ \boldsymbol{B} \\ \\ \boldsymbol{C} \end{matrix}$$

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix}$$

# Verifying Correctness

$$
\boldsymbol{a}_{x_i}^{\mathsf{T}} \boldsymbol{p} = \left( \boxed{-X_i,\ -Y_i,\ -Z_i,\ -1}, \boxed{0,\ 0,\ 0,\ 0}, \boxed{x_i X_i,\ x_i Y_i,\ x_i Z_i,\ x_i} \right) \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} \begin{array}{l} \boldsymbol{A} \\ \\ \\ \\ \boldsymbol{B} \\ \\ \\ \\ \boldsymbol{C} \\ \\ \\ \\ \end{array}
$$

$$
= \left( -\mathbf{X}_i^{\mathsf{T}},\ \mathbf{0},\ x_i \mathbf{X}_i^{\mathsf{T}} \right) \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix}
$$

# Verifying Correctness

$$
\begin{aligned}
\boldsymbol{a}_{x_i}^{\mathsf{T}} \boldsymbol{p} &= (-X_i,\ -Y_i,\ -Z_i,\ -1, 0,\ 0,\ 0,\ 0, x_i X_i,\ x_i Y_i,\ x_i Z_i,\ x_i) \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} \\
&= ( \quad -\mathbf{X}_i^{\mathsf{T}}, \quad \mathbf{0}, \quad x_i \mathbf{X}_i^{\mathsf{T}} \quad ) \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix} \\
&= -\mathbf{X}_i^{\mathsf{T}} \mathbf{A} + x_i\, \mathbf{X}_i^{\mathsf{T}} \mathbf{C}
\end{aligned}
$$

# Verifying Correctness

$$
\begin{aligned}
\boldsymbol{a}_{y_i}^{\top}\boldsymbol{p} &= (0,\ 0,\ 0,\ 0,\ -X_i,\ -Y_i,\ -Z_i,\ -1,\ y_iX_i,\ y_iY_i,\ y_iZ_i,\ y_i) \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} \\
&= (\quad \mathbf{0}, \quad -\mathbf{X}_i^{\top}, \quad y_i\mathbf{X}_i^{\top} \quad) \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix} \\
&= \quad -\mathbf{X}_i^{\top}\mathbf{B} \quad + y_i\,\mathbf{X}_i^{\top}\mathbf{C}
\end{aligned}
$$

# Estimating the Elements of P

- For each point, we have

$$\boldsymbol{a}_{x_i}^{\mathsf{T}} \boldsymbol{p} = 0$$

$$\boldsymbol{a}_{y_i}^{\mathsf{T}} \boldsymbol{p} = 0$$

- Collecting everything together

$$\begin{bmatrix} \boldsymbol{a}_{x_1}^{\mathsf{T}} \\ \boldsymbol{a}_{y_1}^{\mathsf{T}} \\ \cdots \\ \boldsymbol{a}_{x_i}^{\mathsf{T}} \\ \boldsymbol{a}_{y_i}^{\mathsf{T}} \\ \cdots \\ \boldsymbol{a}_{x_I}^{\mathsf{T}} \\ \boldsymbol{a}_{y_I}^{\mathsf{T}} \end{bmatrix} \boldsymbol{p} = \underset{2I \times 12}{\mathrm{M}} \underset{12 \times 1}{\boldsymbol{p}} \stackrel{!}{=} 0$$

**Zero vector of size 2*I***

# Estimating the elements of P

- In case of redundant observations, we get contradictions

$$\mathsf{M}\ \boldsymbol{p} = \boldsymbol{w}$$

- Find $\boldsymbol{p}$ such that it minimizes

$$\begin{aligned} \Omega &= \boldsymbol{w}^{\mathsf{T}}\boldsymbol{w} \\ \Rightarrow \quad \widehat{\boldsymbol{p}} &= \arg\min_{\boldsymbol{p}}\ \boldsymbol{w}^{\mathsf{T}}\boldsymbol{w} \\ &= \arg\min_{\boldsymbol{p}}\ \boldsymbol{p}^{\mathsf{T}}\mathsf{M}^{\mathsf{T}}\mathsf{M}\boldsymbol{p} \end{aligned}$$

with $||P||_2 = \sum_{ij} p_{ij}^2 = \|\boldsymbol{p}\| = 1$

# Comments

$$\widehat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \ \boldsymbol{w}^{\mathsf{T}} \boldsymbol{w}$$

- A Least-Squares problem
- An algebraic minimization problem
  - Contradictions just algebraic, not directly related to geometric entities
  - Due to homogeneity scale of $\boldsymbol{p}$ is arbitrary, so the one of $\boldsymbol{w}$, is too
  - Require $\|\boldsymbol{p}\| = 1$

# Solving the Linear System

- Solving a system of linear equations of the form $A\,\boldsymbol{x} = \mathbf{0}$ is equivalent to finding the null space of A
- Thus, we can apply the SVD to solve M $\boldsymbol{p} \stackrel{!}{=} \mathbf{0}$

# Solution

$$\mathrm{M}\ \boldsymbol{p} \overset{!}{=} \mathbf{0}$$

- Singular value decomposition (SVD)

$$\underset{2I\times 12}{\mathrm{M}} = \underset{2I\times 12}{U}\ \underset{12\times 12}{\mathrm{S}}\ \underset{12\times 12}{V^{\mathsf{T}}} = \sum_{i=1}^{12} s_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathsf{T}}$$

with properties $U^{\mathsf{T}} U = I_{12}$ , $V^{\mathsf{T}} V = I_{12}$
and $s_1 \geq s_2 \geq \ldots \geq s_{12}$

http://16720.courses.cs.cmu.edu/lec/svd.pdf
https://www.cse.iitb.ac.in/~ajitvr/CS763_Spring2017/SVD.pdf

# Solution

- Linear system $\mathsf{M}\,\boldsymbol{p} = \boldsymbol{w} \stackrel{!}{=} \mathbf{0}$
- Minimize $\boldsymbol{\Omega} = \boldsymbol{w}^\mathsf{T}\boldsymbol{w}$
- using SVD $\mathsf{M} = U\,\mathsf{S}\,V^\mathsf{T}$
- Applying SVD leads to:

$$
\begin{aligned}
\Omega &= \boldsymbol{p}^\mathsf{T}\mathsf{M}^\mathsf{T}\mathsf{M}\boldsymbol{p} \\
&= \boldsymbol{p}^\mathsf{T} V\mathsf{S}U^\mathsf{T}U\mathsf{S}V^\mathsf{T}\boldsymbol{p} \\
&= \boldsymbol{p}^\mathsf{T} V\mathsf{S}^2 V^\mathsf{T}\boldsymbol{p} \\
&= \boldsymbol{p}^\mathsf{T}\left(\sum_{i=1}^{12} s_i^2 \boldsymbol{v}_i \boldsymbol{v}_i^\mathsf{T}\right)\boldsymbol{p}
\end{aligned}
$$

# Solution

$$\Omega = \boldsymbol{p}^{\top} \left( \sum_{i=1}^{12} s_i^2 \boldsymbol{v}_i \boldsymbol{v}_i^{\top} \right) \boldsymbol{p}$$

- Due to orthogonality of *V*

$$\boldsymbol{v}_i^{\top} \boldsymbol{v}_j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

- If we choose $\boldsymbol{p} = \boldsymbol{v}_i$

$$\Omega = \boldsymbol{v}_i^{\top} (s_i^2 \boldsymbol{v}_i \boldsymbol{v}_i^{\top}) \boldsymbol{v}_i = s_i^2 \, \boldsymbol{v}_i^{\top} \boldsymbol{v}_i \boldsymbol{v}_i^{\top} \boldsymbol{v}_i = s_i^2$$

- Choosing $\boldsymbol{p} = \boldsymbol{v}_{12}$ (the singular vector belonging to the smallest singular value $s_{12}$ ) minimizes $\Omega$

# Solution

- Estimate of $\boldsymbol{p}$ is given by

$$\widehat{\boldsymbol{p}} = \begin{bmatrix} \widehat{\mathbf{A}} \\ \widehat{\mathbf{B}} \\ \widehat{\mathbf{C}} \end{bmatrix} = \boldsymbol{v}_{12}$$

- and leads to the estimated projection matrix

$$\widehat{\mathsf{P}} = \begin{bmatrix} \widehat{\mathbf{A}}^{\mathsf{T}} \\ \widehat{\mathbf{B}}^{\mathsf{T}} \\ \widehat{\mathbf{C}}^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \widehat{p}_1 & \widehat{p}_2 & \widehat{p}_3 & \widehat{p}_4 \\ \widehat{p}_5 & \widehat{p}_6 & \widehat{p}_7 & \widehat{p}_8 \\ \widehat{p}_9 & \widehat{p}_{10} & \widehat{p}_{11} & \widehat{p}_{12} \end{bmatrix}$$

# Solution

- M is of rank 11, if
  - Number of points ≥6
  - Assumption: no gross errors
- Enhancement: **Gold Standard Algorithm**
  - Start with $\widehat{\mathrm{P}}$ as initial estimate
  - Refine iteratively $\boldsymbol{min}_{\widehat{\boldsymbol{P}}} \sum_{\boldsymbol{i}} \boldsymbol{d}(\boldsymbol{x}_{\boldsymbol{i}}, \widehat{\boldsymbol{P}}\boldsymbol{X}_{\boldsymbol{i}})$

# Critical Surfaces

- M is of rank 11, if
  - Number of points ≥6
  - Assumption: no gross errors
- **No solution**, if all points $\mathbf{X}_i$ are located on a **plane**

$$
\mathsf{M} = \begin{bmatrix} \cdots \\ \boldsymbol{a}_{x_i}^{\mathsf{T}} \\ \boldsymbol{a}_{y_i}^{\mathsf{T}} \\ \cdots \end{bmatrix}
$$

**e.g., assume all** $Z_i=0$

$$
= \begin{bmatrix} & & & & & \cdots & & & & & & \\ -X_i & -Y_i & -Z_i & -1 & 0 & 0 & 0 & 0 & x_i X_i & x_i Y_i & x_i Z_i & x_i \\ 0 & 0 & 0 & 0 & -X_i & -Y_i & -Z_i & -1 & y_i X_i & y_i Y_i & y_i Z_i & y_i \\ & & & & & \cdots & & & & & & \end{bmatrix}
$$

# Critical Surfaces

- M is of rank 11, if
  - Number of points ≥6
  - Assumption: no gross errors
- **No solution**, if all points $\mathbf{X}_i$ are located on a **plane**

$$
\begin{aligned}
\mathsf{M} &= \begin{bmatrix} \cdots \\ \boldsymbol{a}_{x_i}^{\mathsf{T}} \\ \boldsymbol{a}_{y_i}^{\mathsf{T}} \\ \cdots \end{bmatrix} \\
&= \begin{bmatrix} & & & & & \cdots & & & & & & \\ -X_i & -Y_i & \boxed{0} & -1 & 0 & 0 & \boxed{0} & 0 & x_iX_i & x_iY_i & \boxed{0} & x_i \\ 0 & 0 & \boxed{0} & 0 & -X_i & -Y_i & \boxed{0} & -1 & y_iX_i & y_iY_i & \boxed{0} & y_i \\ & & & & & \cdots & & & & & & \end{bmatrix}
\end{aligned}
$$

# Critical Surfaces

- M is of rank 11, if
  - Number of points ≥6
  - Assumption: no gross errors
- **No solution**, if
  - All points $\mathbf{X}_i$ are located on a **plane**
  - (All points $\mathbf{X}_i$ and projection center $\boldsymbol{X}_o$ are located on a twisted cubic curve)

# Decomposition of P

- We have $\widehat{\mathrm{P}}$ , how to obtain $\widehat{\mathrm{K}}, \widehat{R}, \widehat{\boldsymbol{X}_O}$?

# Decomposition of P

- We have $\widehat{\mathsf{P}}$, how to obtain $\widehat{\mathsf{K}}, \widehat{R}, \widehat{\boldsymbol{X}}_O$?
- Structure of the projection matrix

$$\widehat{\mathsf{P}} = \widehat{\mathsf{K}}\widehat{R}\,[\boldsymbol{I}_3 | -\widehat{\boldsymbol{X}}_O] = [\underbrace{\widehat{\mathsf{H}}_\infty}_{3\times 3} | \underbrace{\widehat{\mathbf{h}}}_{3\times 1}]$$

- with

$$\widehat{\mathsf{H}}_\infty = \widehat{\mathsf{K}}\widehat{R} \qquad \widehat{\mathbf{h}} = -\widehat{\mathsf{K}}\widehat{R}\widehat{\mathbf{X}}_O$$

# Decomposition of P

$$\widehat{\mathsf{H}}_\infty = \widehat{\mathsf{K}}\widehat{R} \qquad \widehat{\mathbf{h}} = -\widehat{\mathsf{K}}\widehat{R}\mathbf{X}_O$$

- We get the projection center

$$\widehat{\boldsymbol{X}_O} = -\widehat{\mathsf{H}}_\infty^{-1}\,\widehat{\mathbf{h}}$$

- QR decomposition of $\widehat{\mathsf{H}}_\infty^{-1}$ yields rotation and calibration matrix

$$\widehat{\mathsf{H}}_\infty^{-1} = (\widehat{\mathsf{K}}\;\widehat{R})^{-1} = \widehat{R}^{-1}\;\widehat{\mathsf{K}}^{-1} = \widehat{R}^{\mathsf{T}}\;\widehat{\mathsf{K}}^{-1}$$

- Due to homogeneity normalize

$$\widehat{\widehat{K}} = \frac{1}{\widehat{\mathsf{K}}_{33}}\widehat{\mathsf{K}}$$

# Decomposition of P

- Decomposition $\widehat{\mathsf{H}}_\infty^{-1} = \widehat{R}^{\mathsf{T}}\,\widehat{\mathsf{K}}^{-1}$ results in $\widehat{\mathsf{K}}$ with **positive** diagonal elements
- To get **negative camera constant** $\widehat{c}$, choose

$$\widehat{\mathsf{K}} \leftarrow \widehat{\mathsf{K}} R(z,\pi) \qquad \widehat{R} \leftarrow R(z,\pi)\widehat{R}$$

  - Using

$$R(z,\pi) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

decomposition still holds $\widehat{\mathsf{H}}_\infty = \widehat{\mathsf{K}} R(z,\pi)\; R(z,\pi)\widehat{R} = \widehat{\mathsf{K}}\widehat{R}$

# DLT in a Nutshell

1. Vectorize $\mathsf{P}$ : $\boldsymbol{p} = (p_k) = \mathrm{vec}(\mathsf{P}^{\mathsf{T}})$

# DLT in a Nutshell

1. Vectorize $\mathsf{P}$ : $\boldsymbol{p} = (p_k) = \mathrm{vec}(\mathsf{P}^\mathsf{T})$
2. Build the $\mathsf{M}$ for the linear system

$$\mathsf{M} = \begin{bmatrix} \boldsymbol{a}_{x_1}^\mathsf{T} \\ \boldsymbol{a}_{y_1}^\mathsf{T} \\ \cdots \\ \boldsymbol{a}_{x_I}^\mathsf{T} \\ \boldsymbol{a}_{y_I}^\mathsf{T} \end{bmatrix} \qquad \underbrace{\mathsf{M}}_{2I\mathrm{x}12} \; \underbrace{\boldsymbol{p}}_{12\mathrm{x}1} \stackrel{!}{=} 0$$

with

$$\begin{aligned} \boldsymbol{a}_{x_i}^\mathsf{T} &= (-X_i, -Y_i, -Z_i, -1, 0, 0, 0, 0, x_i X_i, x_i Y_i, x_i Z_i, x_i) \\ \boldsymbol{a}_{y_i}^\mathsf{T} &= (0, 0, 0, 0, -X_i, -Y_i, -Z_i, -1, y_i X_i, y_i Y_i, y_i Z_i, y_i) \end{aligned}$$

# DLT in a Nutshell

3. Solve by SVD $\mathsf{M} = U\,S\,V^{\mathsf{T}}$
   Solution is last column of $V$

$$\widehat{\boldsymbol{p}} = \begin{bmatrix} \widehat{\mathbf{A}} \\ \widehat{\mathbf{B}} \\ \widehat{\mathbf{C}} \end{bmatrix} = \boldsymbol{v}_{12} \;\Rightarrow\; \widehat{\mathsf{P}} = \begin{bmatrix} \widehat{\mathbf{A}}^{\mathsf{T}} \\ \widehat{\mathbf{B}}^{\mathsf{T}} \\ \widehat{\mathbf{C}}^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \widehat{p}_1 & \widehat{p}_2 & \widehat{p}_3 & \widehat{p}_4 \\ \widehat{p}_5 & \widehat{p}_6 & \widehat{p}_7 & \widehat{p}_8 \\ \widehat{p}_9 & \widehat{p}_{10} & \widehat{p}_{11} & \widehat{p}_{12} \end{bmatrix}$$

# DLT in a Nutshell

3. Solve by SVD $\mathsf{M} = U\,S\,V^{\mathsf{T}}$
   Solution is last column of $V$

$$\widehat{\boldsymbol{p}} = \begin{bmatrix} \widehat{\mathbf{A}} \\ \widehat{\mathbf{B}} \\ \widehat{\mathbf{C}} \end{bmatrix} = \boldsymbol{v}_{12} \Rightarrow \widehat{\mathsf{P}} = \begin{bmatrix} \widehat{\mathbf{A}}^{\mathsf{T}} \\ \widehat{\mathbf{B}}^{\mathsf{T}} \\ \widehat{\mathbf{C}}^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \widehat{p}_1 & \widehat{p}_2 & \widehat{p}_3 & \widehat{p}_4 \\ \widehat{p}_5 & \widehat{p}_6 & \widehat{p}_7 & \widehat{p}_8 \\ \widehat{p}_9 & \widehat{p}_{10} & \widehat{p}_{11} & \widehat{p}_{12} \end{bmatrix}$$

4. If individual parameters are needed

$$\widehat{\mathsf{P}} = \widehat{\mathsf{K}}\widehat{R}\,[I_3 | -\widehat{\boldsymbol{X}}_O] = [\underset{3\times 3}{\widehat{\mathsf{H}}_\infty} | \underset{3\times 1}{\widehat{\mathbf{h}}}]$$

$$\widehat{\mathsf{H}}_\infty = \widehat{\mathsf{K}}\widehat{R} \qquad \widehat{\mathbf{h}} = -\widehat{\mathsf{K}}\widehat{R}\mathbf{X}_O$$

$$\widehat{\boldsymbol{X}}_O = -\widehat{\mathsf{H}}_\infty^{-1}\,\widehat{\mathbf{h}} \qquad \widehat{\mathsf{H}}_\infty^{-1} = \widehat{R}^{\mathsf{T}}\,\widehat{\mathsf{K}}^{-1} \qquad \widehat{\widehat{K}} = \frac{1}{\widehat{\mathsf{K}}_{33}}\widehat{\mathsf{K}}$$

# Discussion DLT

- We realize $\mathsf{P} \leftrightarrow (\mathsf{K}, R, \boldsymbol{X}_O)$ both ways
- We are free to choose sign of $c$
- Solution is instable if the control points lie approximately on a plane

# DLT Summary

- We can estimate the camera parameters given control points
- **Uncalibrated camera**
  - **DLT**
  - Using **≥6 points**
  - Direct solution

# Slide Information

- The slides have been created by Cyrill Stachniss (cyrill.stachniss@igg.uni-bonn.de) as part of the photogrammetry and robotics courses.
- A lot of material from Ajit Rajwade's CS763 course
- Thanks to Parag for some slides
- **I tried to acknowledge all people from whom I used images or videos. In case I made a mistake or missed someone, please let me know.**
- The photogrammetry material heavily relies on the very well written lecture notes by Wolfgang Förstner and the Photogrammetric Computer Vision book by Förstner & Wrobel.
- Parts of the robotics material stems from the great Probabilistic Robotics book by Thrun, Burgard and Fox.
- If you are a university lecturer, feel free to use the course material. If you adapt the course material, please make sure that you keep the acknowledgements to others and please acknowledge me as well. To satisfy my own curiosity, please send me email notice if you use my slides.

Arjun Jain, ajain@cse.iitb.ac.in