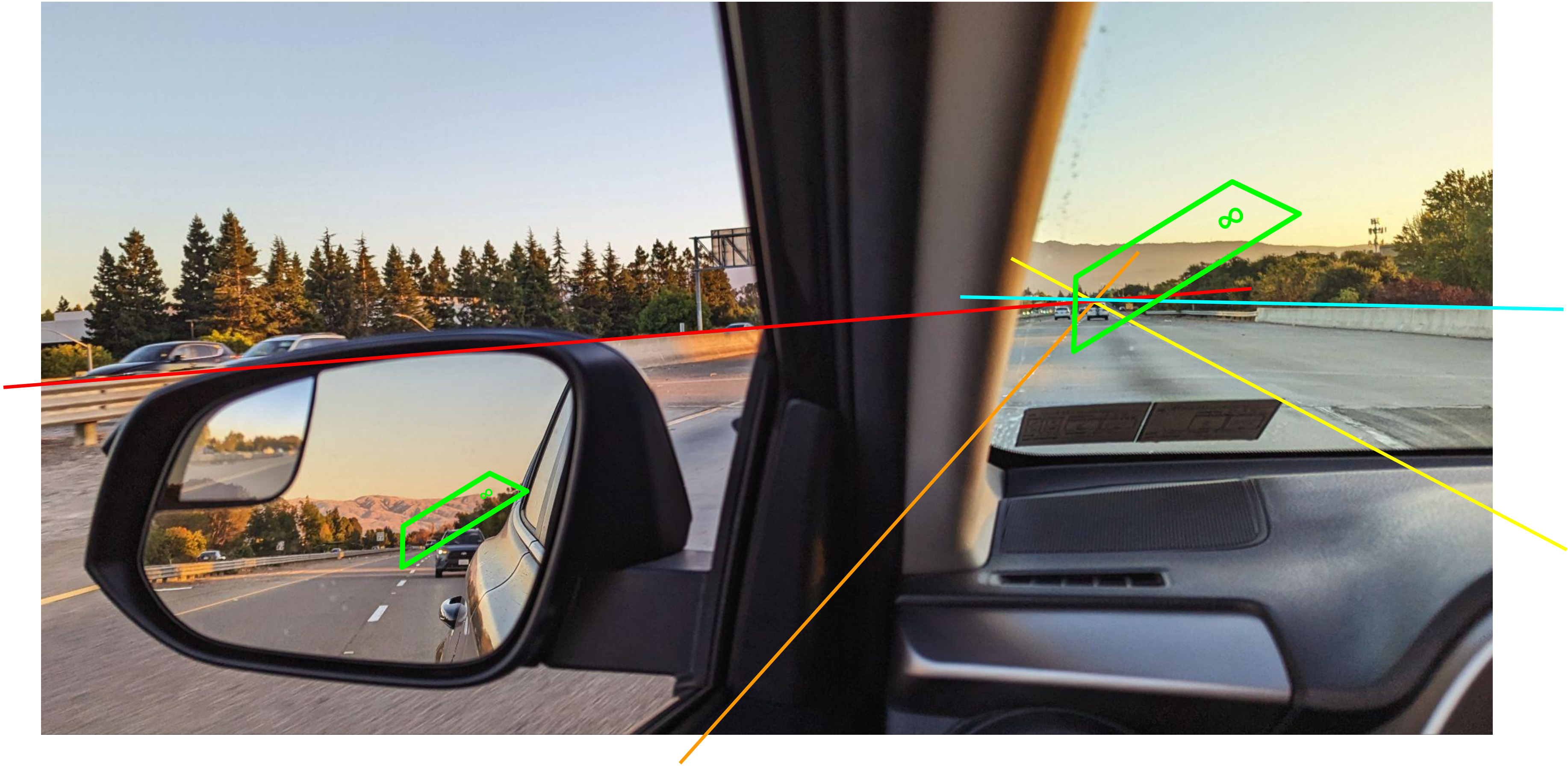
Traditional Computer Vision

jopark



Agenda

1. Camera
   a. Camera Matrix
   b. Camera Calibration
2. Questions in Geometry Vision
   a. Pose Estimation
   b. Triangulation
   c. (Epipolar Geometry and) Reconstruction
3. Photogrammetry
   a. Stereo Vision
   b. N-view Reconstruction

Camera Matrix

Camera Matrix

A mapping from 3D world to 2D image

$$\underset{\substack{\text{2D image}\\\text{point}}}{\boldsymbol{x}} = \underset{\substack{\text{camera}\\\text{matrix}}}{\mathbf{P}}\,\underset{\substack{\text{3D world}\\\text{point}}}{\mathbf{X}}$$

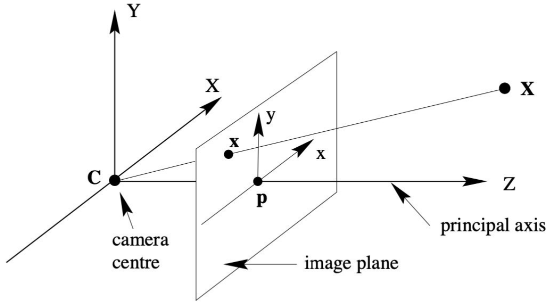
Camera Matrix

Generic camera

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous image 3 x 1	Camera matrix 3 x 4	homogeneous world point 4 x 1

Camera Matrix



Pinhole camera

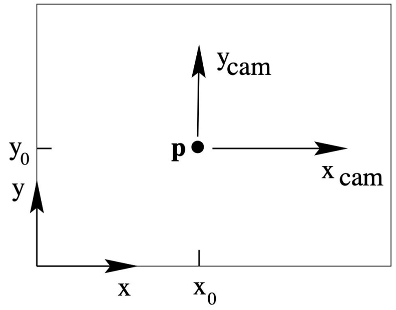
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous image 3 x 1

Camera matrix 3 x 4

homogeneous world point 4 x 1

Camera Matrix



Accounts for image origin

$$
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
$$

homogeneous image 3 x 1	Camera matrix 3 x 4	homogeneous world point 4 x 1

Camera Matrix

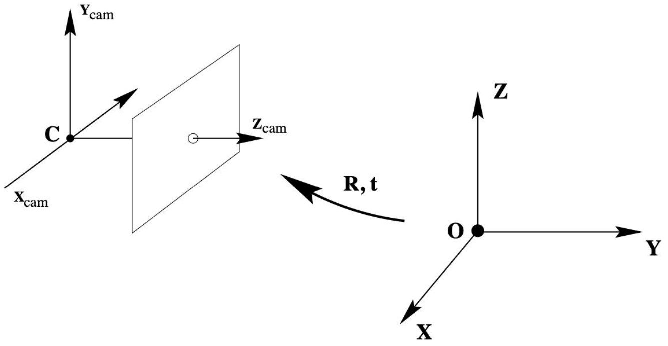
Decomposed P

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous image 3 x 1	(3 x 3)	(3 x 4)	homogeneous world point 4 x 1

$$\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}]$$

Camera Matrix



Accounts for world origin

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \end{array}\right] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous image 3 x 1

intrinsic parameters

extrinsic parameters

homogeneous world point 4 x 1

$$\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}]$$

Camera Calibration

Camera Calibration

Conic is a curve described by a second-degree equation in the plane

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$ // *inhomogeneous*

$$ax_1{}^2 + bx_1x_2 + cx_2{}^2 + dx_1x_3 + ex_2x_3 + fx_3{}^2 = 0$$ // *homogeneous*

$$\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = 0 \quad \mathtt{C} = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}$$ // *homogeneous, matrix*

Camera Calibration

Camera params and Image of Absolute Conic (IAC)

- Mapping from plane at infinity (Π_∞) to image plane: $\mathbf{H = KR}$
- Conic maps: $\mathbf{C} \rightarrow \mathbf{H^{-T}CH^{-1}}$
- Absolute Conic maps: $\mathbf{I} \rightarrow \mathbf{(KR)^{-T}I(KR)^{-1} = (KK^T)^{-1} = w}$

Find IAC, Dual IAC, then K

- IAC has no real point
- Algebraic tool that constrains image components such as VP, VL, etc.
- Cholesky decomposition on DIAC: $\mathbf{w^* = w^{-1} = KK^T}$

Questions in Geometry Vision

Questions in Geometry Vision

1. **Pose Estimation**
   3D-2D correspondences, 3D Structure ⇒ Motion
2. **Triangulation**
   2D-2D correspondences, Motion ⇒ 3D Structure
3. **Reconstruction**
   2D-2D correspondences ⇒ 3D Structure, Motion

Pose Estimation

Pose Estimation

3D-2D correspondences, 3D Structure ⇒ Motion

$$\{\mathbf{X}_i, \boldsymbol{x}_i\} \longrightarrow \mathbf{P}$$

X in {world}
x in {image}

Pose Estimation

Generic projection, Inhomogeneous coordinates, N points

$$\begin{bmatrix} \boldsymbol{X}_1^\top & \mathbf{0} & -x'\boldsymbol{X}_1^\top \\ \mathbf{0} & \boldsymbol{X}_1^\top & -y'\boldsymbol{X}_1^\top \\ \vdots & \vdots & \vdots \\ \boldsymbol{X}_N^\top & \mathbf{0} & -x'\boldsymbol{X}_N^\top \\ \mathbf{0} & \boldsymbol{X}_N^\top & -y'\boldsymbol{X}_N^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \boldsymbol{p}_3 \end{bmatrix} = \mathbf{0}$$

SVD

Pose Estimation

Camera center, Intrinsics, Rotation

$$\mathbf{P} \Rightarrow \begin{matrix}\mathbf{K}[\mathbf{R}| - \mathbf{R}\mathbf{c}] \\ \textit{or} \\ [\mathbf{M}| - \mathbf{M}\mathbf{c}]\end{matrix}$$

$$\mathbf{P}\mathbf{c} = \mathbf{0} \qquad \mathbf{M} = \mathbf{K}\mathbf{R}$$

SVD **RQ (ambiguity)**

Triangulation

Triangulation

2D-2D correspondences, Motion ⇒ 3D Structure

$$\{\boldsymbol{x}_i, \boldsymbol{x}'_i\}$$

$$\mathbf{P}, \mathbf{P}'$$

$$\mathbf{X}$$

Triangulation

Noisy measurements, Similarity ⇒ **DLT (Direct Linear Transform)**

$$\mathbf{x} = \alpha \mathbf{P} \boldsymbol{X}$$

$$\mathbf{x} \times \mathbf{P} \boldsymbol{X} = \mathbf{0}$$

$$\begin{bmatrix} y\boldsymbol{p}_3^\top - \boldsymbol{p}_2^\top \\ \boldsymbol{p}_1^\top - x\boldsymbol{p}_3^\top \\ y'\boldsymbol{p}_3'^\top - \boldsymbol{p}_2'^\top \\ \boldsymbol{p}_1'^\top - x'\boldsymbol{p}_3'^\top \end{bmatrix} \boldsymbol{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{A} \boldsymbol{X} = \mathbf{0}$$

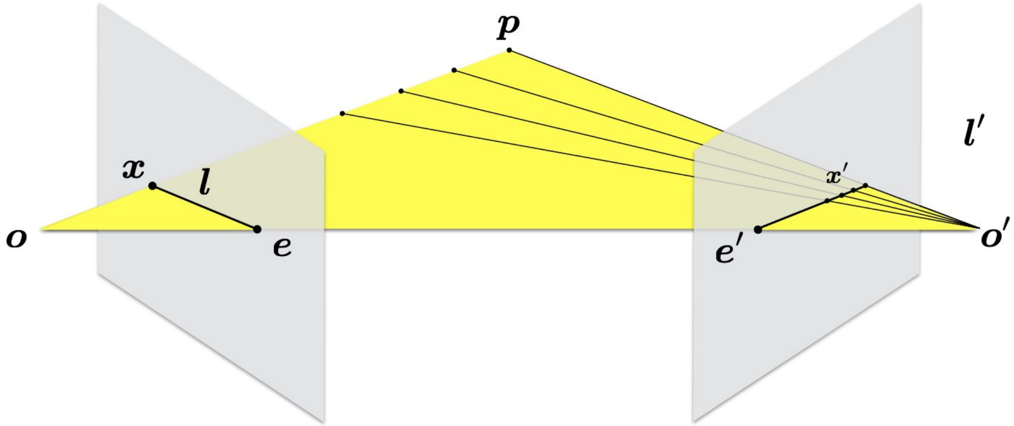
SVD

Epipolar Geometry, E, F

Epipolar Geometry

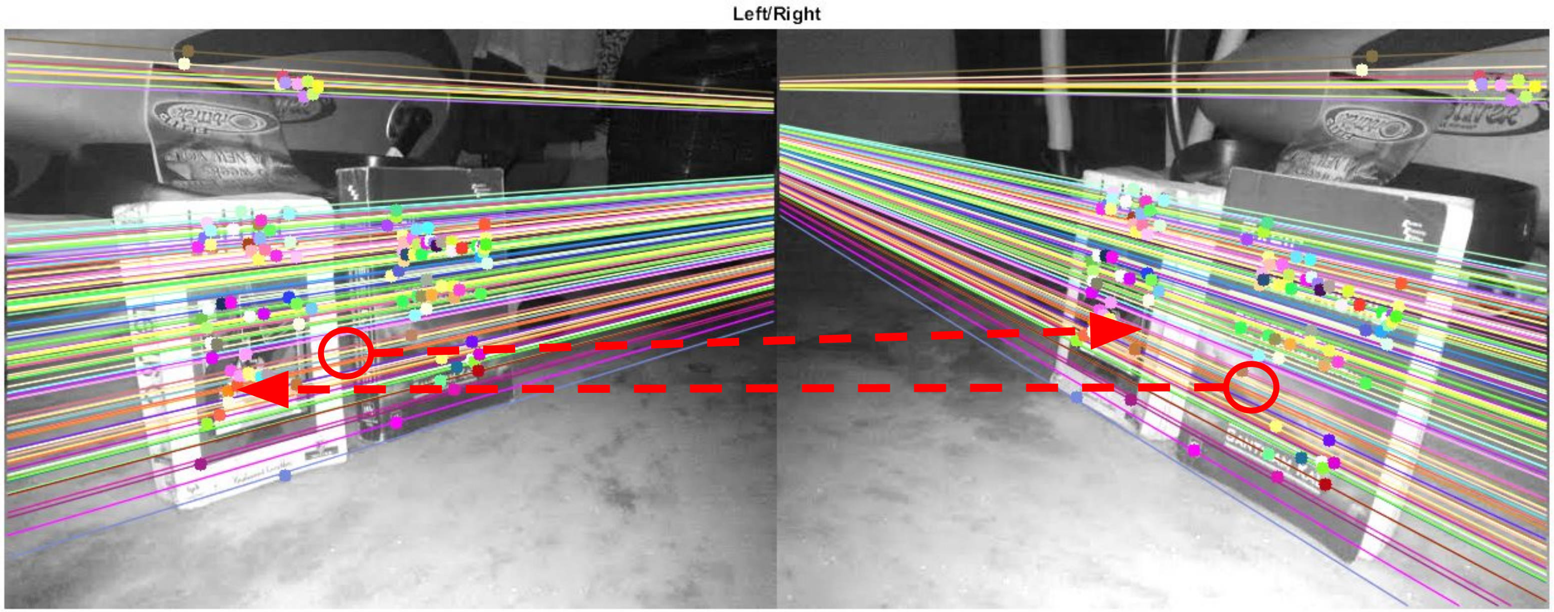
For point correspondences {x, x’},

1. **Epipole**: Intersection of the baseline and image planes (e, e’)
2. **Epipolar line**: Intersection of epipolar plane and image planes (l, l’)
3. **Epipolar plane**: Plane containing the baseline



Epipolar Geometry

Point matching via epipolar line



Essential Matrix

Algebraic representation of epipolar geometry (3x3 matrix)

$$\boldsymbol{x}'^{\top}\mathbf{E}\boldsymbol{x} = 0$$ *// calibrated camera points*

and

$$\mathbf{E}\boldsymbol{x} = \boldsymbol{l}'$$

$$\mathbf{E} = \mathbf{R}[\mathbf{t}_{\times}]$$

Fundamental Matrix

Generalization of E for **uncalibrated** cameras

$$\boldsymbol{x}'^{\top}\mathbf{F}\boldsymbol{x} = 0$$ *// plain image points*

and

$$\mathbf{F}\boldsymbol{x} = \boldsymbol{l}'$$

$$\mathbf{F} = \mathbf{K}'^{-\top}\mathbf{E}\mathbf{K}^{-1}$$

Fundamental Matrix

8-point algorithm

$$\begin{bmatrix} x'_m & y'_m & 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} = 0$$

$$\Downarrow$$

$$\begin{bmatrix} x_1 x'_1 & x_1 y'_1 & x_1 & y_1 x'_1 & y_1 y'_1 & y_1 & x'_1 & y'_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_M x'_M & x_M y'_M & x_M & y_M x'_M & y_M y'_M & y_M & x'_M & y'_M & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = \mathbf{0}$$

SVD

Fundamental Matrix

Solve for e, e’

$$\mathbf{F}\boldsymbol{e} = \mathbf{0}$$

SVD

Reconstruction

Reconstruction

2D-2D correspondences ⇒ 3D Structure, Motion

$$\{\boldsymbol{x}_i, \boldsymbol{x}'_i\} \longrightarrow \begin{array}{c} \mathbf{X} \\ \mathbf{P}, \mathbf{P}' \end{array}$$

Reconstruction

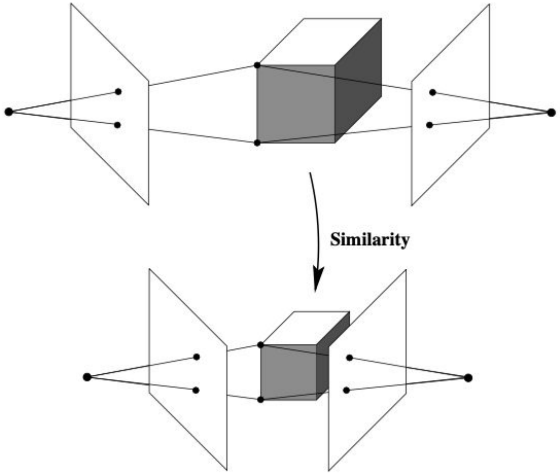
Steps

1. Compute F from {x, x'}   *// 8-point algorithm (Fundamental Matrix)*
2. Compute P, P' from F
3. Compute X_i from P, P' and {x_i, x'_i}   *// DLT (Triangulation)*

$$\mathbf{P} = [\mathbf{I}|\mathbf{0}] \quad \textit{and} \quad \mathbf{P}' = [[\boldsymbol{e}_\times]\mathbf{F}|\boldsymbol{e}']$$

Reconstruction

Ambiguity in reconstruction

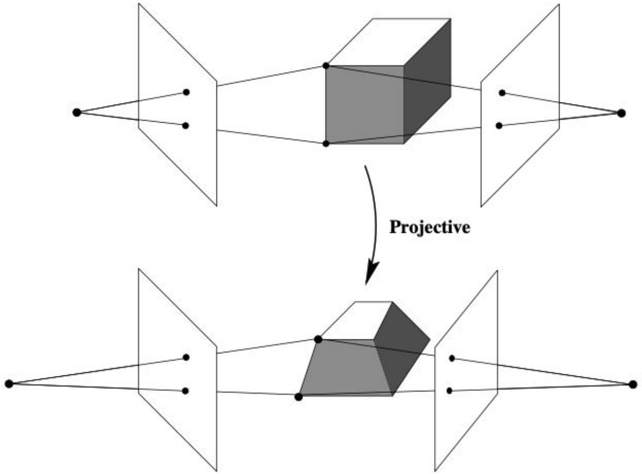


Calibrated Camera
reconstructs **up to Similarity $\mathbf{H}_S$**

$$(\mathrm{P}_1, \mathrm{P}'_1, \{\mathbf{X}_{1i}\})$$

$$\mathrm{P}_2 = \mathrm{P}_1\mathrm{H}^{-1}$$
$$\mathrm{P}'_2 = \mathrm{P}'_1\mathrm{H}^{-1}$$
$$\mathbf{X}_{2i} = \mathrm{H}\mathbf{X}_{1i}$$

$$(\mathrm{P}_2, \mathrm{P}'_2, \{\mathbf{X}_{2i}\})$$

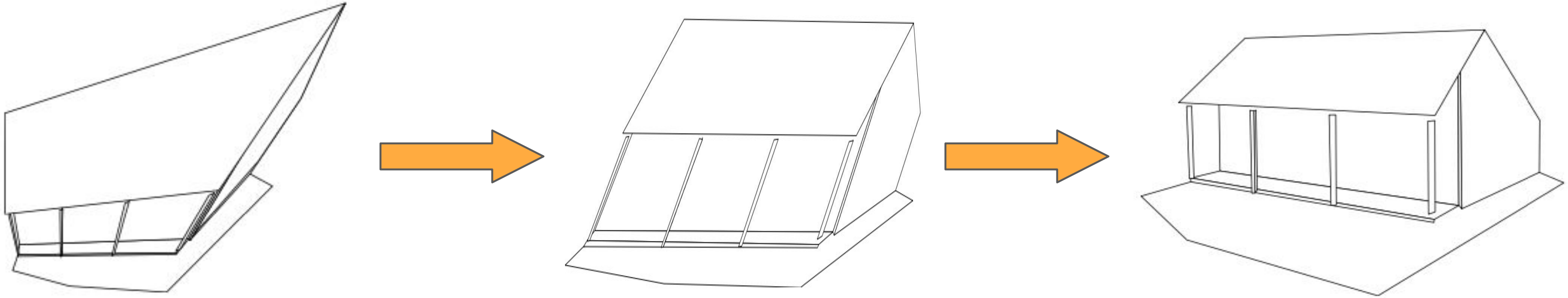


Uncalibrated Camera
reconstructs **up to Projectivity $\mathbf{H}_P$**

Reconstruction

Recover **Projective** → **Affine** → **Euclidean** reconstructions

(... below images are still 2D **projections**)



Use parallelism, orthogonality, known constraints in K, etc.
to find **invariants** of each transform class

Reconstruction

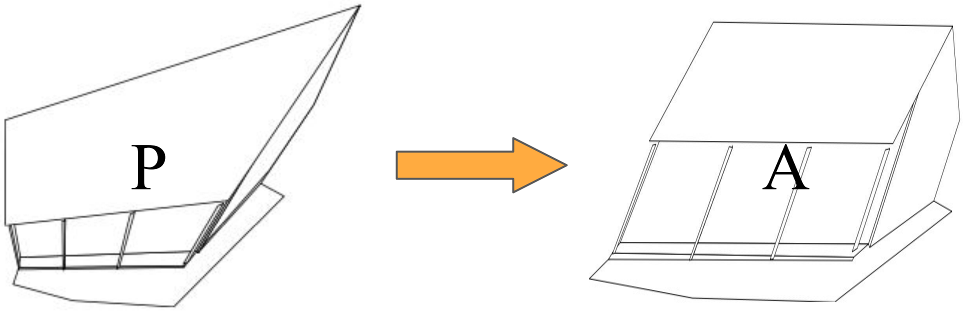
Recover **Projective** → **Affine** reconstructions

Canonical expressions of

- X_∞: (x, y, z, 0)
- $\boldsymbol{\pi_\infty}$**: (0, 0, 0, 1)**

Find imaged π_∞ from parallelism, VPs

$\mathbf{H_P}$**: Imaged** $\boldsymbol{\pi_\infty}$ **→ Canonical** $\boldsymbol{\pi_\infty}$



$$(\mathtt{P}_1, \mathtt{P}'_1, \{\mathbf{X}_{1i}\})$$

$$\begin{aligned} \mathtt{P}_2 &= \mathtt{P}_1\mathtt{H}^{-1} \\ \mathtt{P}'_2 &= \mathtt{P}'_1\mathtt{H}^{-1} \\ \mathbf{X}_{2i} &= \mathtt{H}\mathbf{X}_{1i} \end{aligned}$$

$$(\mathtt{P}_2, \mathtt{P}'_2, \{\mathbf{X}_{2i}\})$$

Reconstruction

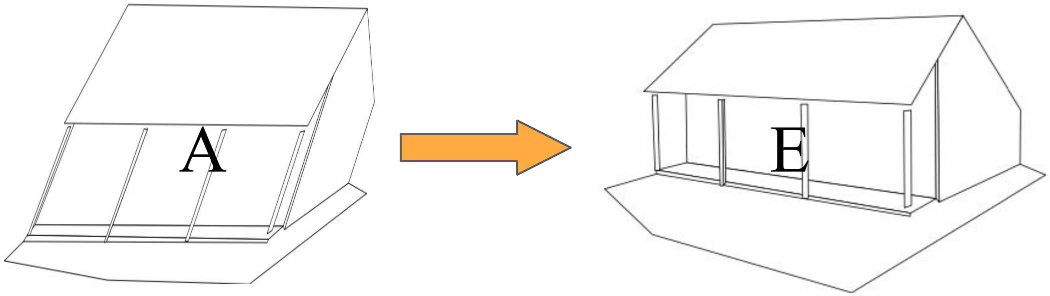
Recover **Affine** → **Euclidean** reconstructions

Canonical expressions of Absolute Conic

$\boldsymbol{\Omega_\infty}$**:** $\mathbf{x_1^{\,2} + x_2^{\,2} + x_3^{\,3} = 0}$

Find imaged Ω_∞ from orthogonality, VP, VL, K

$\mathbf{H_A}$**: Imaged** $\boldsymbol{\Omega_\infty}$ **→ Canonical** $\boldsymbol{\Omega_\infty}$



$$(\mathtt{P}_1, \mathtt{P}'_1, \{\mathbf{X}_{1i}\})$$

$$\begin{aligned} \mathtt{P}_2 &= \mathtt{P}_1\mathtt{H}^{-1} \\ \mathtt{P}'_2 &= \mathtt{P}'_1\mathtt{H}^{-1} \\ \mathbf{X}_{2i} &= \mathtt{H}\mathbf{X}_{1i} \end{aligned}$$

$$(\mathtt{P}_2, \mathtt{P}'_2, \{\mathbf{X}_{2i}\})$$

Photogrammetry

Photogrammetry

“**Photogrammetry** is the science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images and patterns of electromagnetic radiant imagery and other phenomena.”

Photogrammetry

Examples

- Extraction of **3D measurements** from 2D data
- Extraction of accurate color ranges and values representing such quantities as **albedo, specular reflection, metallicity, or ambient occlusion** from photographs

Stereo Vision

Stereo Vision

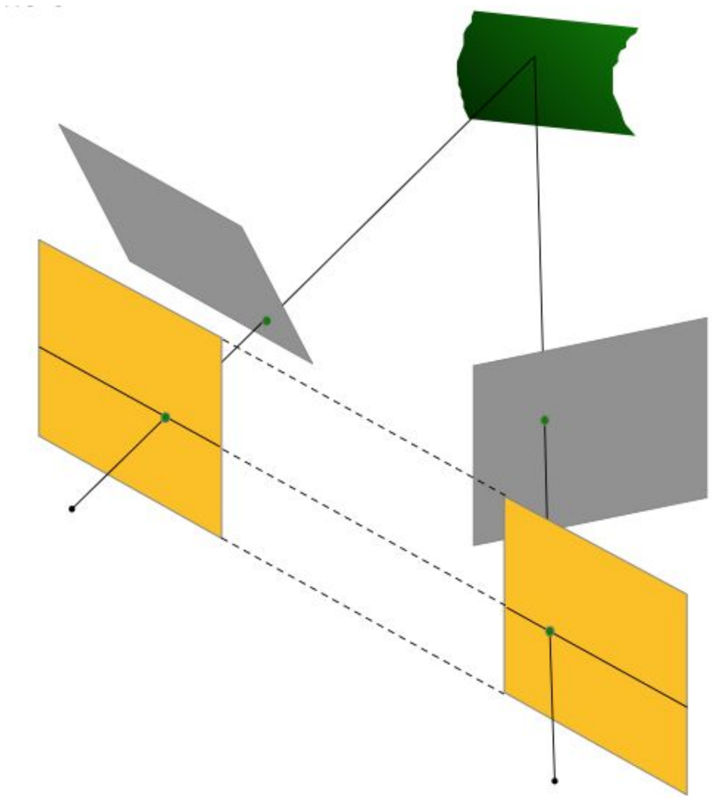
Steps

1. Rectify images
2. For each pixel
   a. Find epipolar line
   b. Find the best match along the line
   c. Compute depth from disparity

Stereo Vision

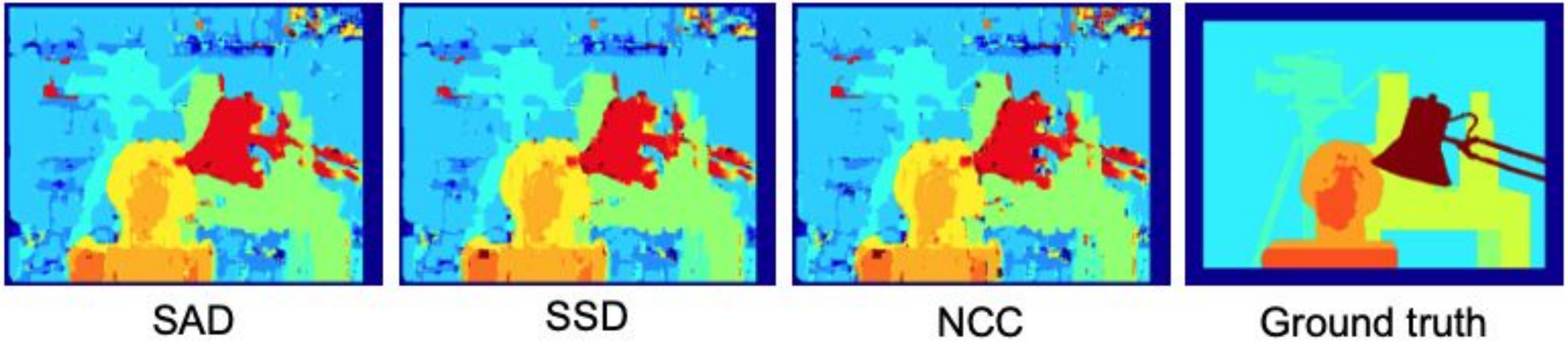
Stereo rectification

- Parallel image planes
- Horizontal epipolar lines
- Same y-coords
- Two 3x3 homographies H_1, H_2



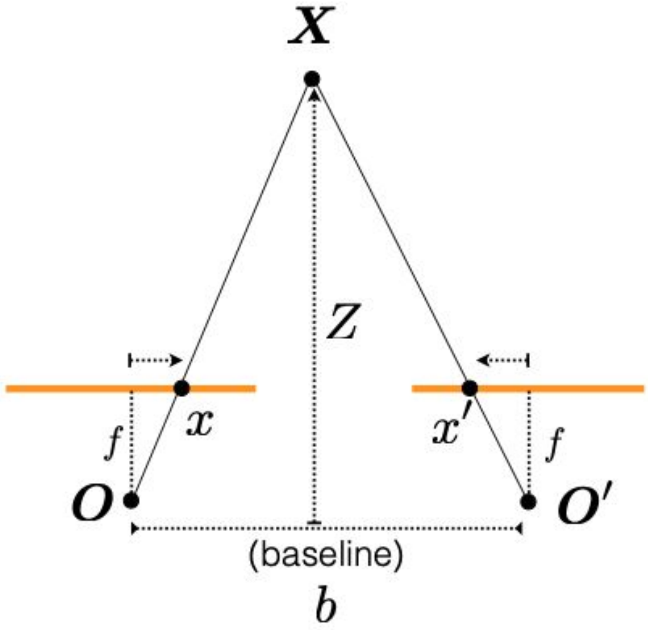
Stereo Vision

Find the best match



Stereo Vision

Disparity



Assumes f (K), b (R, T) are already known

Left : $$\frac{X}{Z} = \frac{x}{f}$$

Right : $$\frac{b - X}{Z} = \frac{x'}{f}$$

Disparity : $$\frac{bf}{Z}$$

Stereo Vision

Improve depth estimation

- Match uniqueness, Points ordering, Depth smoothness

N-view Reconstruction

N-view Reconstruction

Cameras calibrated or not?

Constraints on K? K identical for all cameras?

External information provided?

...

$$p_{ij} = M_i P_j$$

// Different notations...

N-view Reconstruction

Projective ambiguity in reconstruction

Recall **IAC, DIAC** for a single camera calibration… we find **w**, then **K**

In N-view, **avoid explicit calibration of K**

N-view Reconstruction

2. Autocalibration

- Recover **metric** reconstruction from **projective** reconstruction
- Projective $\{M_i\}$ and $\{P_j\}$ $\Rightarrow$ Metric $\{M_i\}$, it's decomposition $\{K_i, R_i, T_i\}$, and $\{P_j\}$

N-view Reconstruction

2. Autocalibration

$$M'_i = M_i Q \qquad P'_j = Q^{-1} P_j \qquad \text{, where } Q = \begin{bmatrix} K_1 & 0 \\ v^T & 1 \end{bmatrix} = \begin{bmatrix} Q_3 & 0 \\ & 1 \end{bmatrix}$$ ***// 8 unknowns***

DIAC

$$\omega_i^* = M_i Q_3 Q_3^{\mathrm{T}} M_i^{\mathrm{T}}$$

$$\omega^* = K_i K_i^{\mathrm{T}}$$

// 8+5m unknowns

Need some constraints. Let's apply constraints on K ...

N-view Reconstruction

2. Autocalibration

$$K = \begin{bmatrix} \alpha_x & s & x_o \\ 0 & \alpha_y & y_o \\ 0 & 0 & 1 \end{bmatrix} \Longrightarrow \omega^* = KK^T = \begin{bmatrix} \alpha_x^2 + s^2 + x_o^2 & s\alpha_y + x_o y_o & x_o \\ s\alpha_y + x_o y_o & \alpha_y^2 + y_o^2 & y_o \\ x_o & y_o & 1 \end{bmatrix} \quad \omega^* = MQ_3Q_3^T M^T$$

Principal point known : $(M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{13} = 0 \quad (M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{23} = 0$

Zero-skew : $(M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{12} = 0$ ***or*** $(M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{13}\,(M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{23}$

Aspect Ratio 1.0 : $(M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{11} = (M_i Q_3 Q_3{}^{\mathrm{T}} M_i^{\mathrm{T}})_{22}$

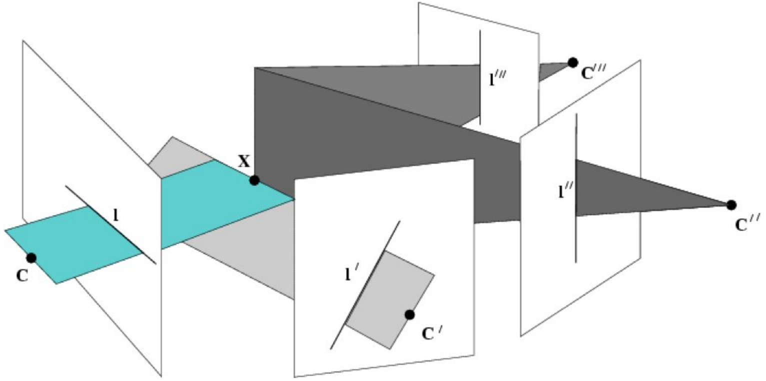
4*m constraints (4 per img) → Can be solved in Q_3 if **m >= 3**

What if **K is constant** for all cameras?

N-view Reconstruction

1. Projective reconstruction

- Projective relationship: Trifocal/Quadrifocal tensors, ...
- Refinement (Bundle adjustment): Reproj err, Nonlinear LS, LM/Newton, …
- Sparseness, Hierarchical approach (2/3-view relations as sub problems), ...



$$Q^{ijkl}\varepsilon_{ii'i''}\mathbf{x}^1_{i'}\varepsilon_{jj'j''}\mathbf{x}^2_{j'}\varepsilon_{kk'k''}\mathbf{x}^3_{k'}\varepsilon_{ll'l''}\mathbf{x}^4_{l'} = 0$$

$$l_p l'_q l''_r l'''_s Q^{pqrs} = 0$$

Thank you!