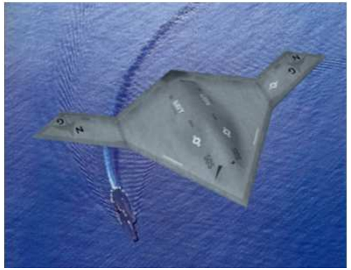
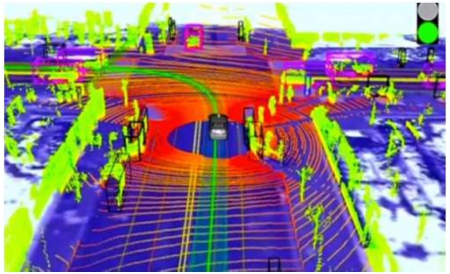
Intelligent Image and Graphics Processing

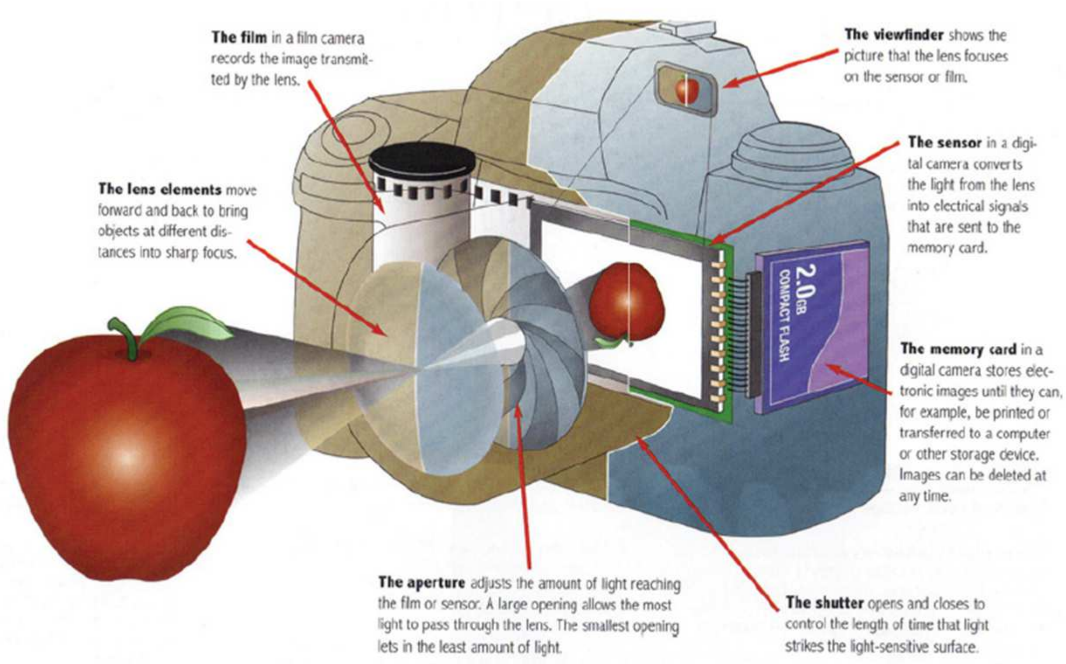
智能图像图形处理

布树辉 bushuhui@nwpu.edu.cn
http://www.adv-ci.com





3D Reconstruction



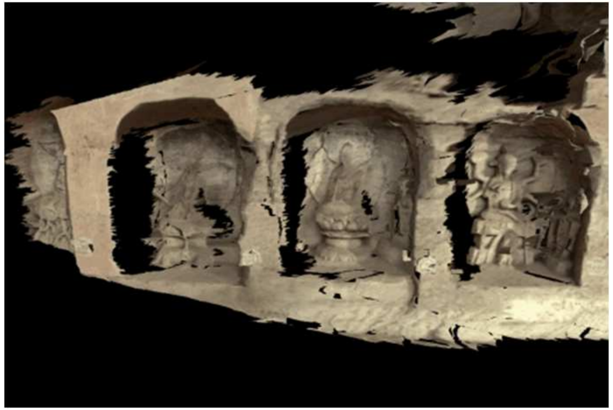
- Recover the 3D world
- Inverse problem

Structure from Motion (Batch processing)

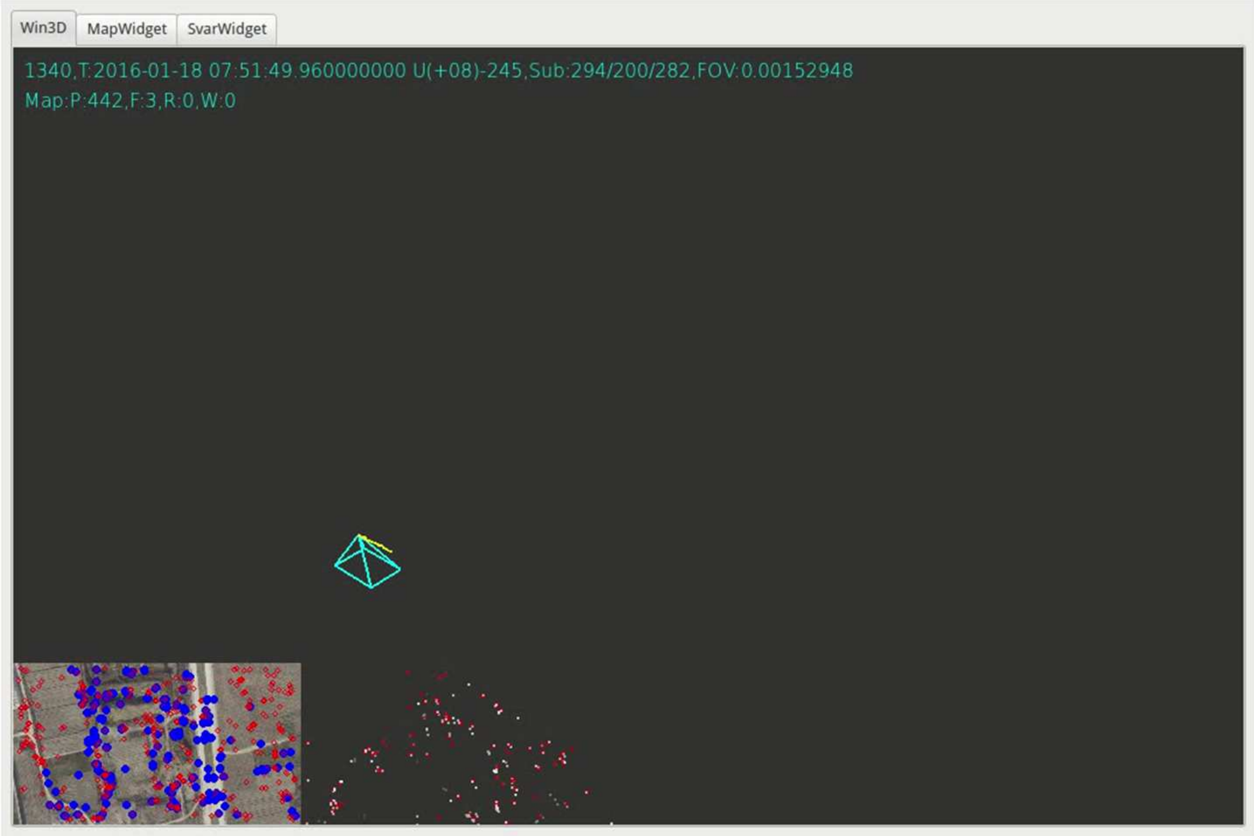
Tracking

Depth
Reconver

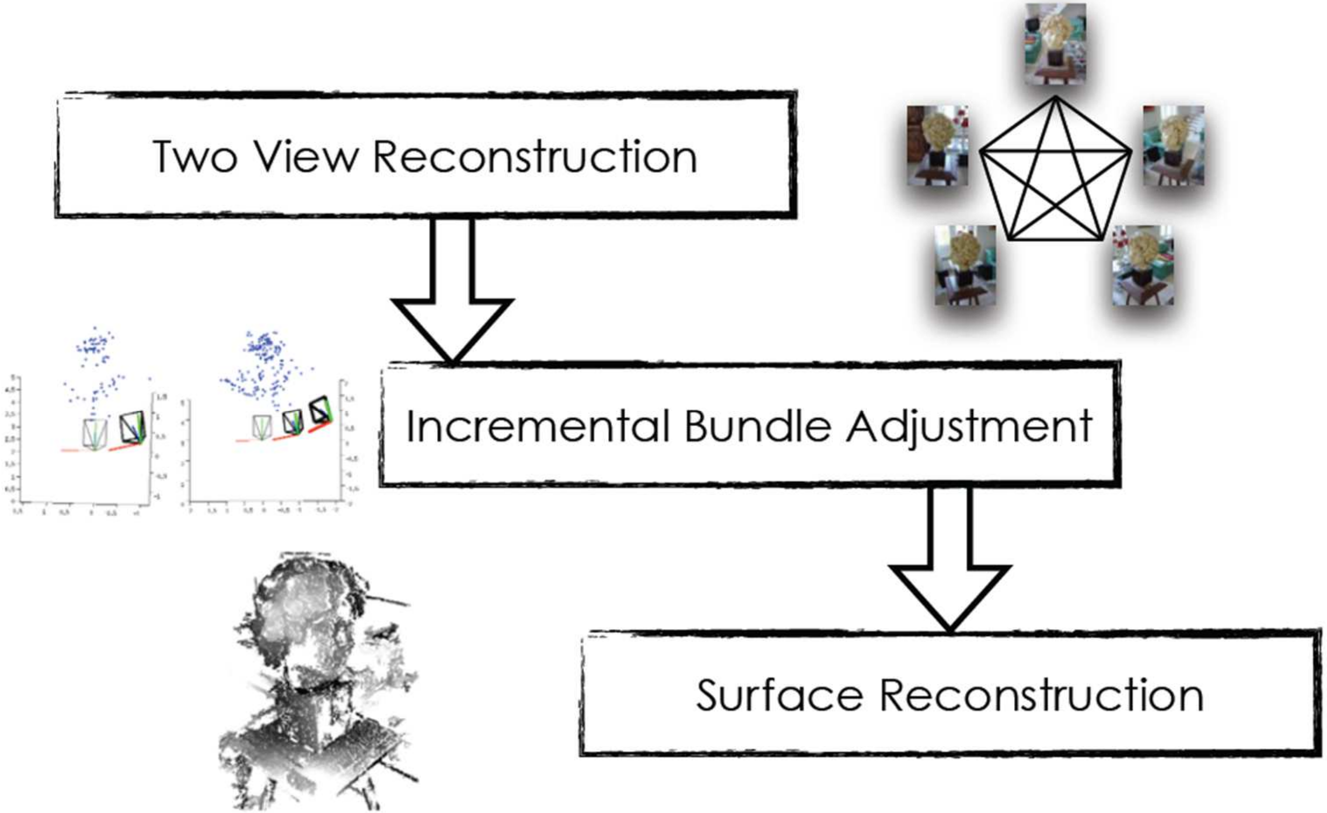
3D
Reconstruction



Visual SLAM (Real-time Processing)



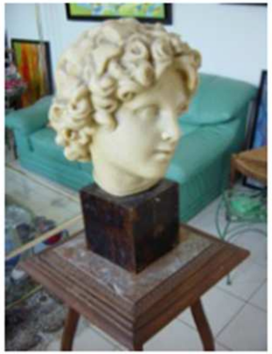
Pipeline


Two View Reconstruction
Incremental Bundle Adjustment
Surface Reconstruction

Pipeline

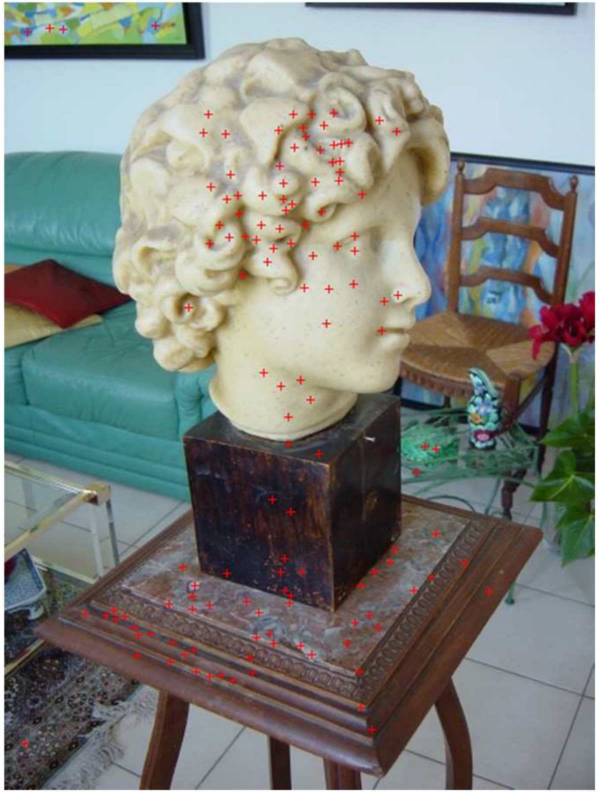
- Input
  - Images
  - Intrinsics
- Output
  - 3D model and camera poses
- Optional
  - Some cases of unknown intrinsics are also discussed

Two View Reconstruction

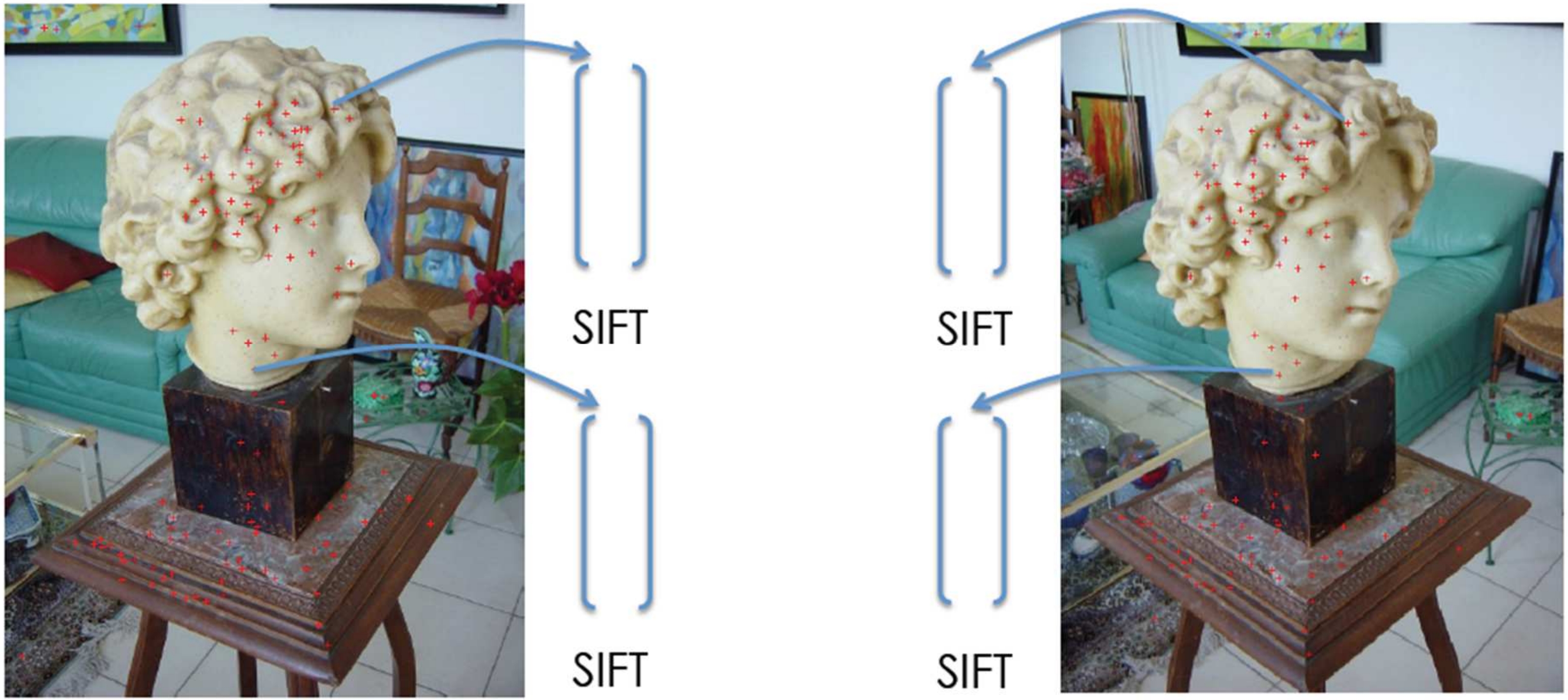


$\Rightarrow$ **Match** $\Rightarrow$ $\mathbf{F}$ $\Rightarrow$ $\mathbf{E}$ $\Rightarrow$ $\mathbf{R}, \mathbf{t}$ $\Rightarrow$ $\mathbf{X}$ $\Rightarrow$ **Bundle Adjustment**

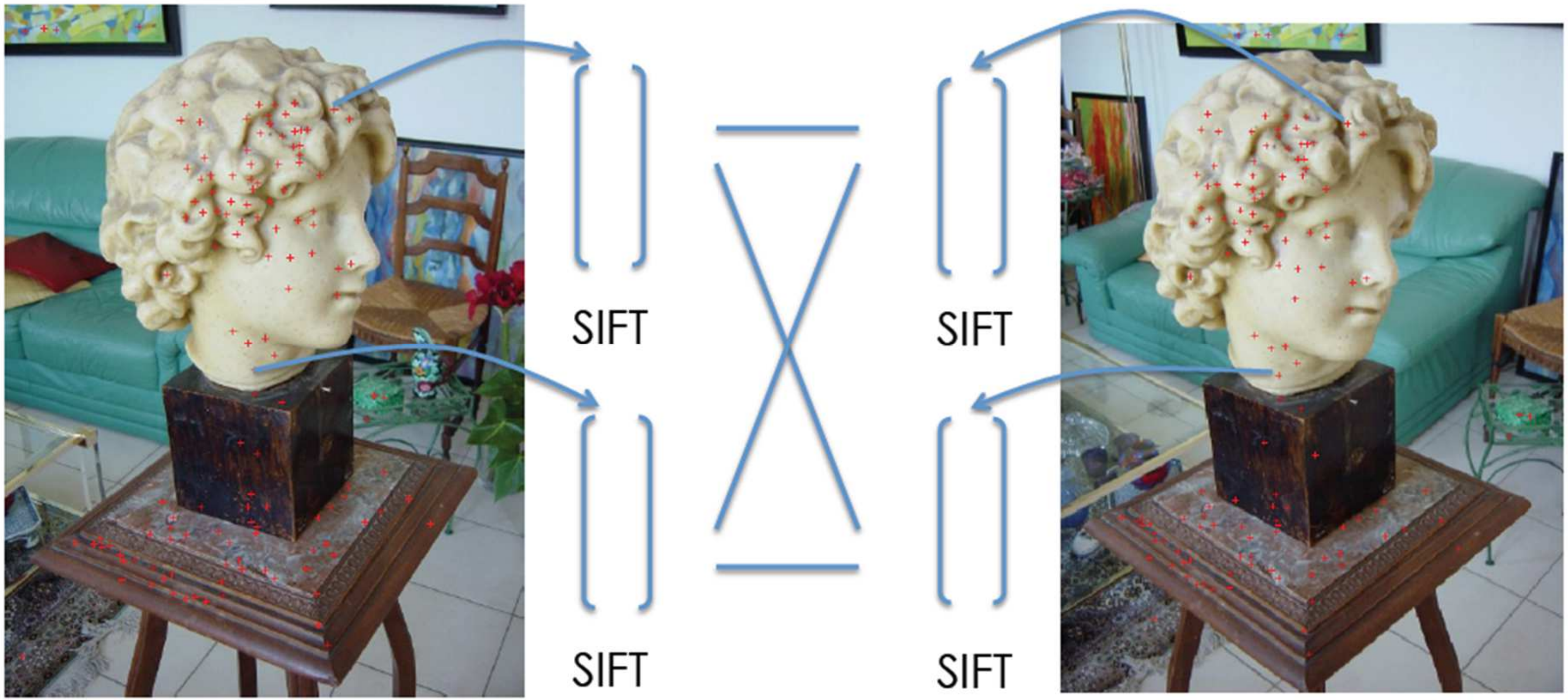
Keypoints Detection



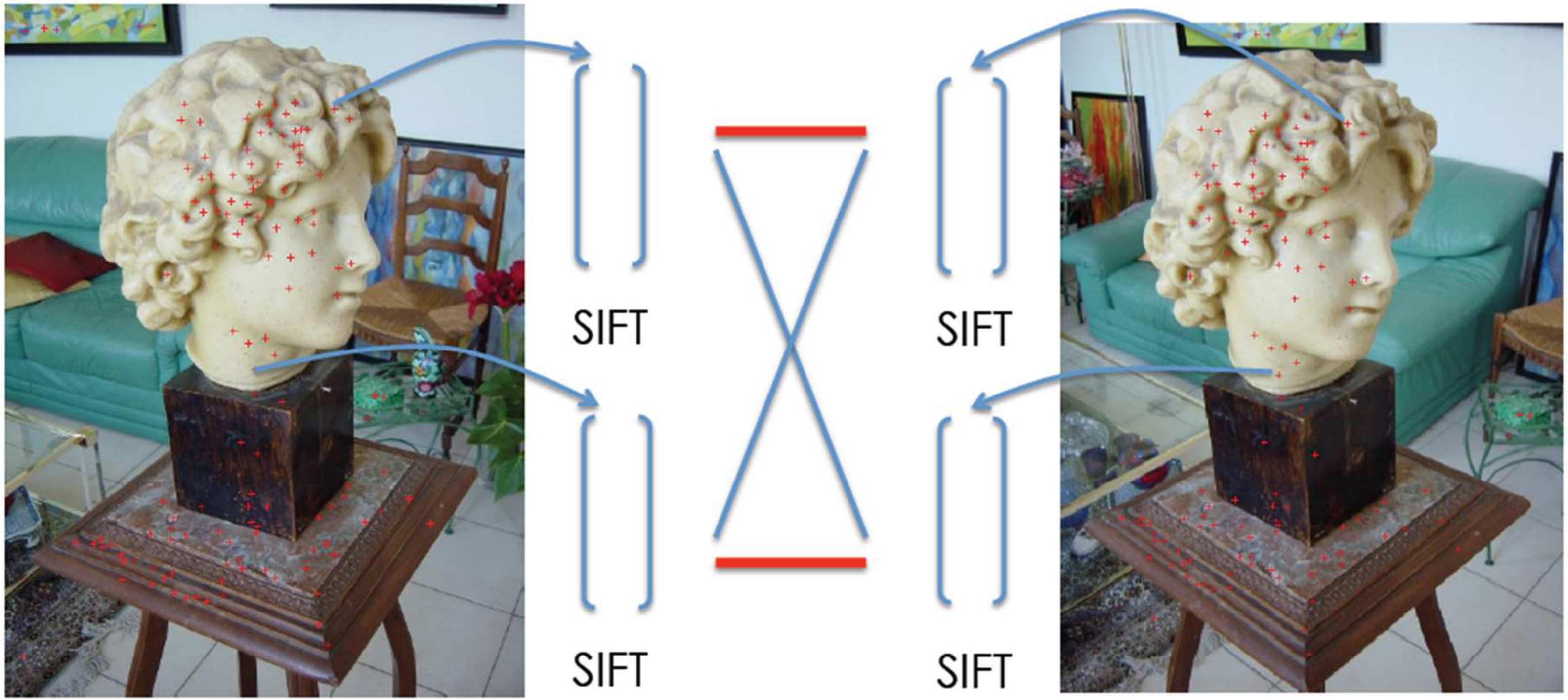
Descriptor of Each Point



Match Points



Match Points



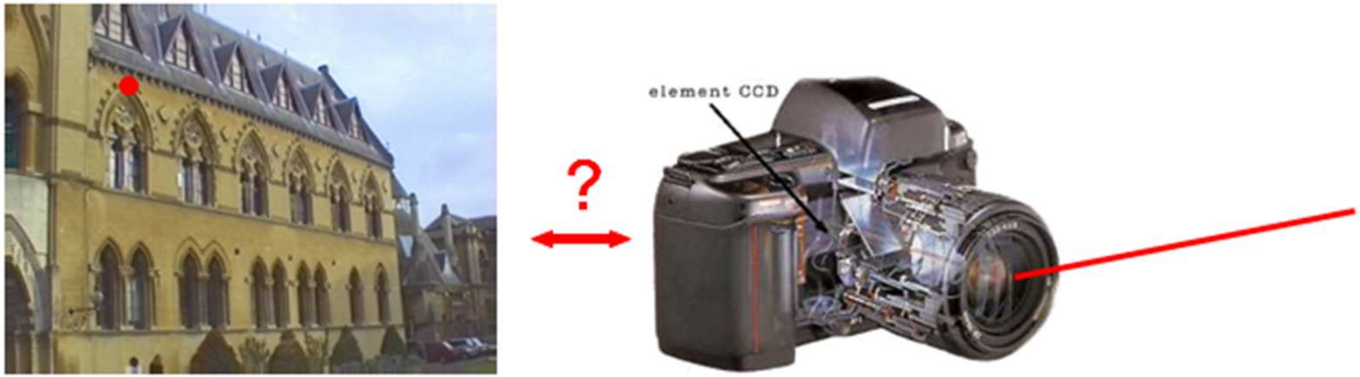
Match Points

- Need fast nearest neighbor searching
  - Curse of Dimensionality
  - Flann http://www.cs.ubc.ca/research/flann/
- Ratio test for robustness
  - The nearest neighbor should be much better than the other candidates
- Consistency check
  - Whether the correspondences have consistent structure and motion
  - Make use of two-view relation

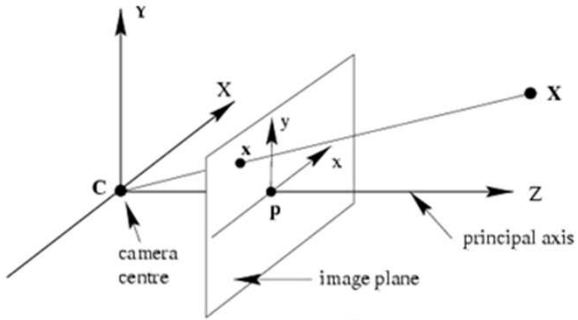
What is the relation between the correspondences?

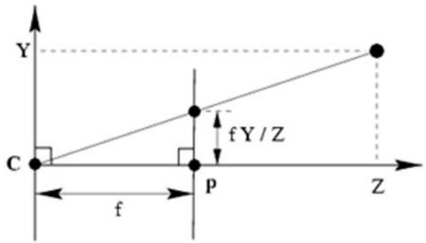
Camera Model

- Relation between pixels and rays in space
  - Pinhole Perspective Projection



Pinhole Camera Model





non-homogeneous coordinates

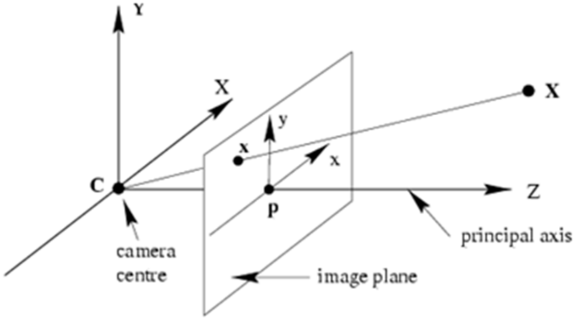
$$(X, Y, Z)^T \mapsto (fX / Z, fY / Z)^T$$

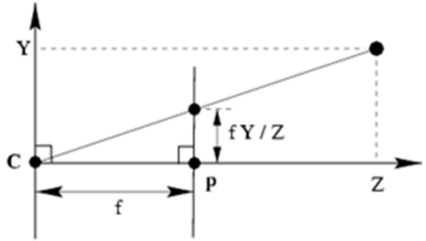
homogeneous coordinates

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

linear projection in homogeneous coordinates!

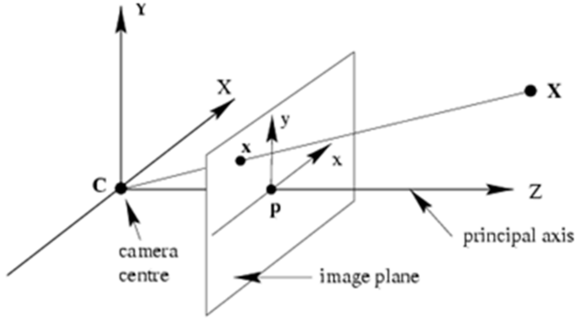
Pinhole Camera Model

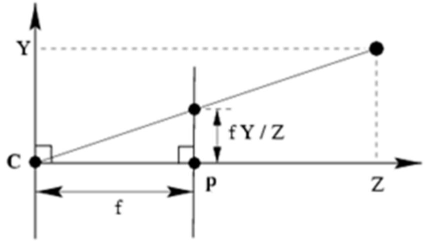




$$\begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & 0 \\ & 1 & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Pinhole Camera Model



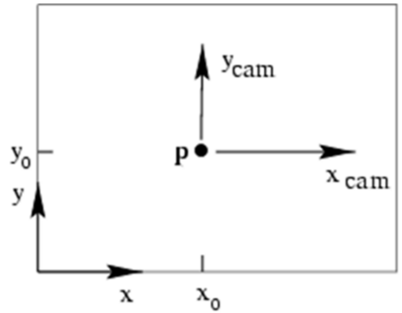


$$\mathrm{x} = \mathrm{PX}$$

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

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

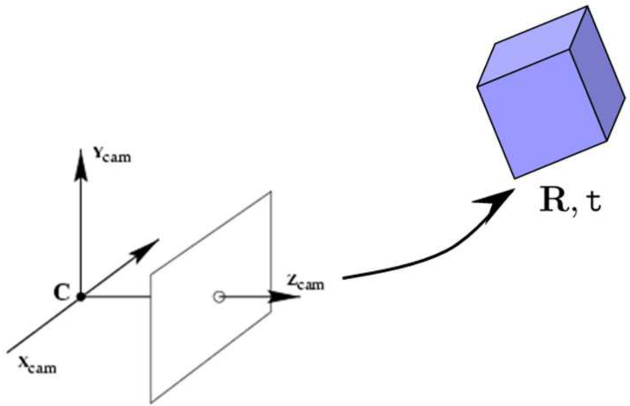
Principal Point Offset



$$x = K\left[I \mid 0\right]X_{cam}$$

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

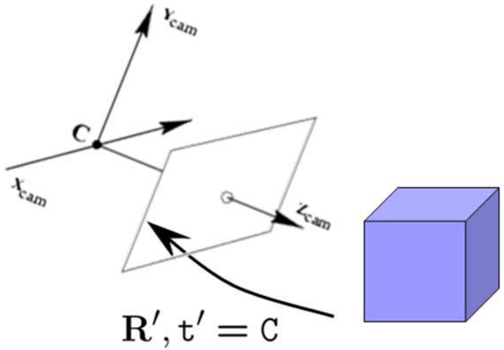
Object Motion



$$x = \mathbf{K}\begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} X_{cam} \qquad X_{cam} = \begin{bmatrix} \mathbf{R} & t \\ 0 & 1 \end{bmatrix} X_{obj}$$

$$\begin{aligned} x &= \mathbf{K}\begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & t \\ 0 & 1 \end{bmatrix} X_{obj} \\ &= \mathbf{K}\begin{bmatrix} \mathbf{R} & t \end{bmatrix} X_{obj} \end{aligned}$$

Camera Motion

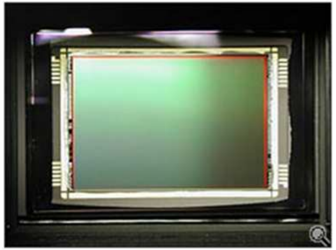


$$\mathtt{X}_{\mathtt{obj}} = \begin{bmatrix} \mathbf{R}' & \mathtt{t}' \\ 0 & 1 \end{bmatrix} \mathtt{X}_{\mathtt{cam}}$$

$$\mathtt{X}_{\mathtt{cam}} = \begin{bmatrix} \mathbf{R}'^{\top} & -\mathbf{R}'^{\top}\mathtt{t}' \\ 0 & 1 \end{bmatrix} \mathtt{X}_{\mathtt{obj}}$$

$$\mathtt{x} = \mathbf{K} \begin{bmatrix} \mathbf{R}'^{\top} & -\mathbf{R}'^{\top}\mathtt{t}' \end{bmatrix} \mathtt{X}_{\mathtt{obj}}$$

Real Camera



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

General Projective Camera

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

$$\mathrm{P} = \mathrm{KR}\left[\mathrm{I} \mid \mathrm{t}\right]$$ 11 dof (5+3+3)

$$\mathrm{P} = \mathrm{K}\underbrace{\left[\mathrm{R} \mid \mathrm{t}\right]}$$

intrinsic camera parameters

extrinsic camera parameters

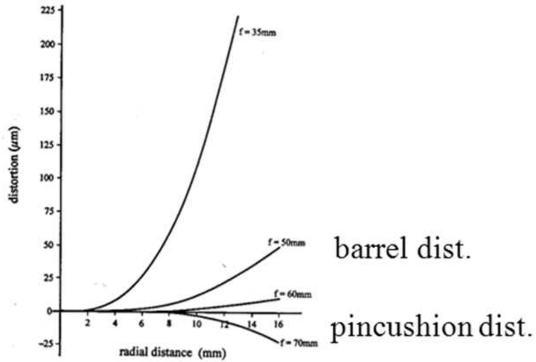
Radial Distortion

- Due to spherical lenses (cheap)
- Model:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} R\left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathrm{t} \\ 0_3^\top & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \right)$$

$$R(x, y) = (1 + K_1(x^2 + y^2) + K_2(x^2 + y^2)^2 + \ldots) \begin{bmatrix} x \\ y \end{bmatrix}$$

straight lines are not straight anymore



Radial Distortion Example

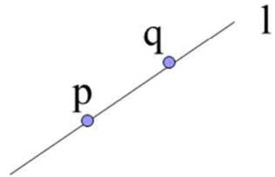
Homography

- A line **l** is represented by the homogeneous 3-vector

$$l = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$$

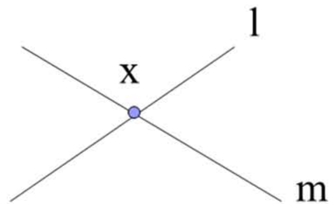
for the line $l_1x+l_2y+l_3=0$. Only the ratio of the homogeneous line coordinates is significant.

- point on line: $\mathbf{l}.\mathbf{x} = 0$ or $\mathbf{l}\mathbf{x} = 0$ or $\mathbf{x}\mathbf{l} = 0$



- two points define a line: $\mathbf{l} = \mathbf{p} \times \mathbf{q}$

- two lines define a point: $\mathbf{x} = \mathbf{l} \times \mathbf{m}$



Homography – Matrix Notation for Vector Product

The vector product $\mathbf{v} \times \mathbf{x}$ can be represented as a matrix multiplication

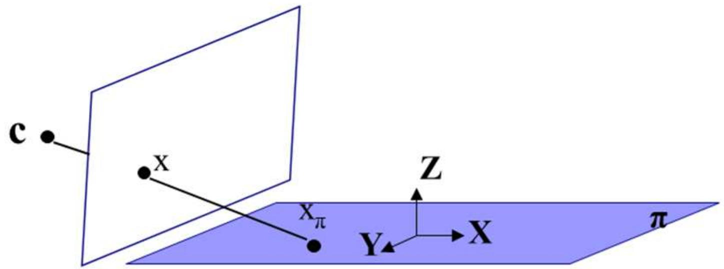
$$\mathbf{v} \times \mathbf{x} = [\mathbf{v}]_{\times} \mathbf{x}$$

where

$$[v]_{\times} = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}$$

- $[\mathbf{v}]_{\times}$ is a 3×3 skew-symmetric matrix of rank 2.
- $\mathbf{v}$ is the null-vector of $[\mathbf{v}]_{\times}$, since $\mathbf{v} \times \mathbf{v} = [\mathbf{v}]_{\times}\mathbf{v} = \mathbf{0}$.

Homography – Plane Projective Transformations



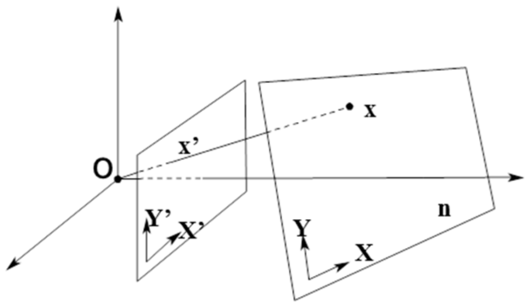
Choose the world coordinate system such that the plane of the points has zero Z coordinate. Then the 3 ×4 matrix P reduces to

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{21} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

which is a 3×3 matrix representing a general plane to plane projective transformation.

Homography – Plane Projective Transformations

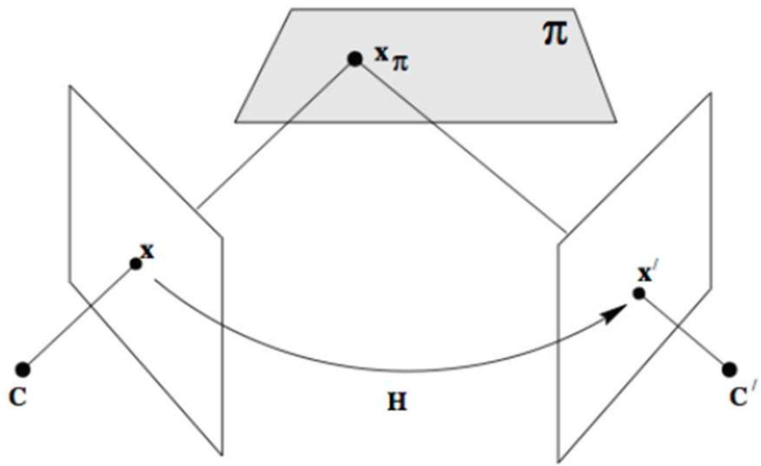
$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



or $\mathbf{x} = \mathrm{H}\mathbf{x}$, where H is a 3 × 3 non-singular homogeneous matrix.

- This is the most general transformation between the world and image plane under imaging by a perspective camera.
- It is often only the 3 × 3 form of the matrix that is important in establishing properties of this transformation.
- A projective transformation is also called a “homography” and a “collineation”.
- H has 8 degrees of freedom.

Homography



$$\mathbf{x}' = \mathbf{H}_\pi \mathbf{x}$$

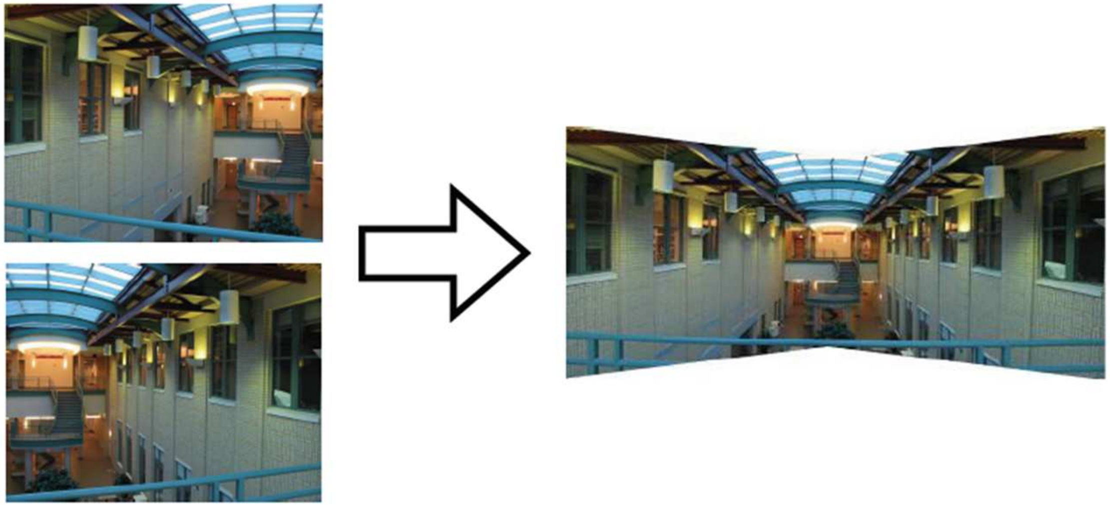
Homography

- 8 degree of freedom
- Can represent camera rotation
  - Make panorama images

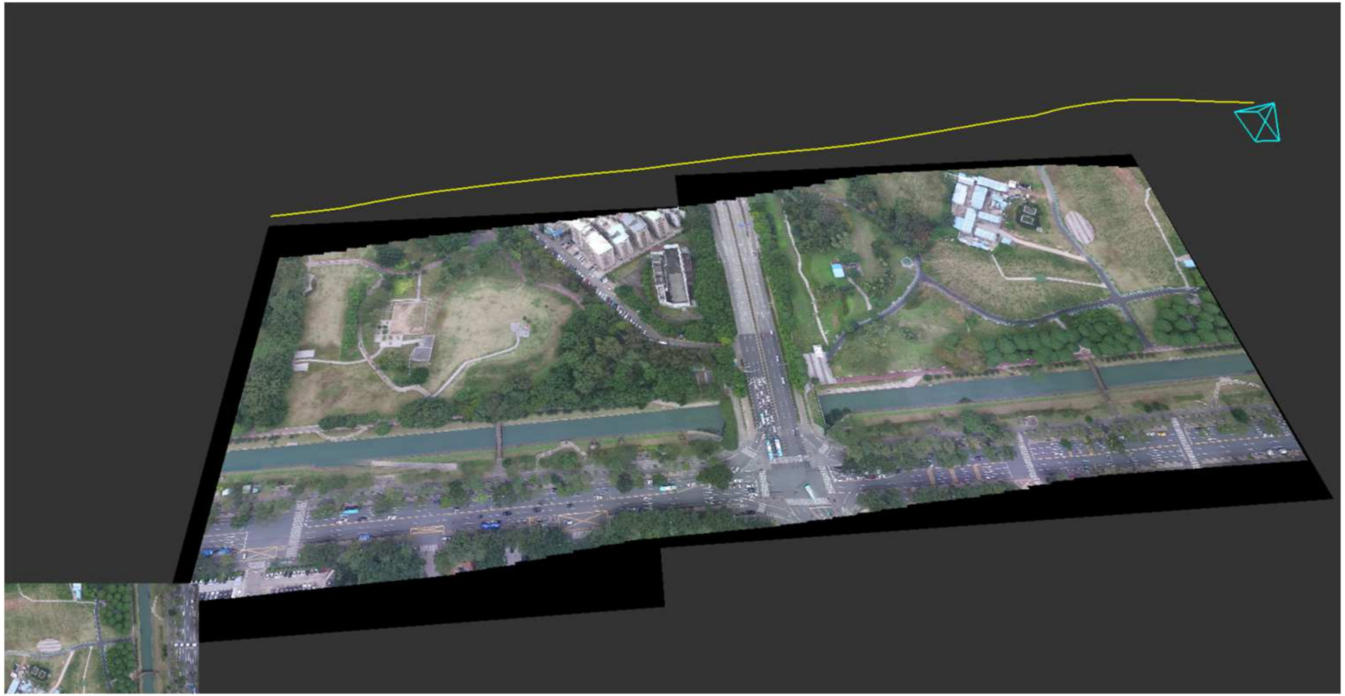
Homography

- 8 degree of freedom
- Can represent camera rotation
  - Make panorama images

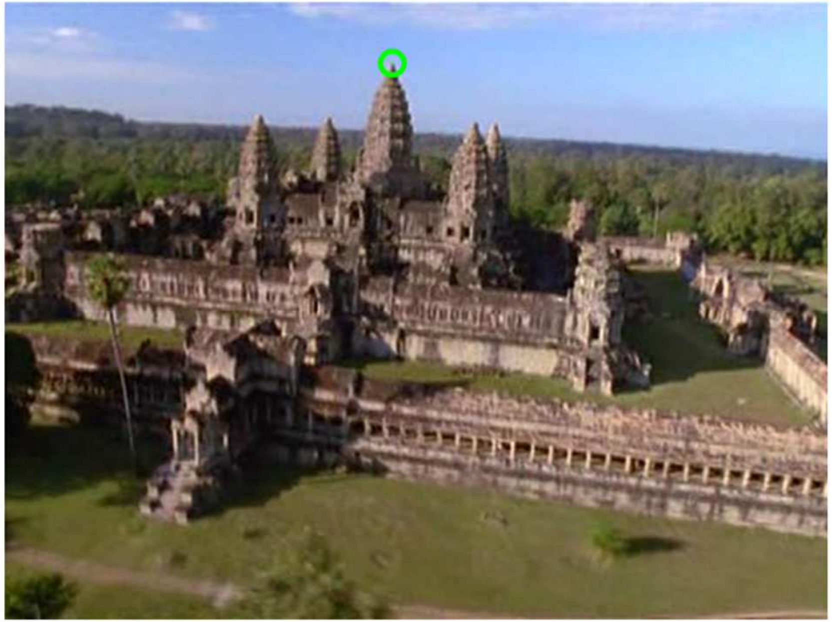


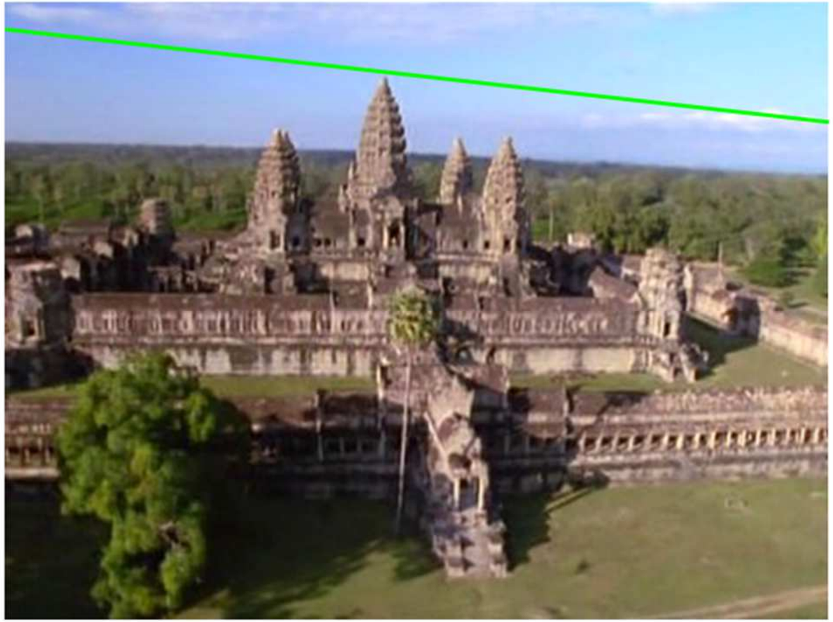
Homography

Limitation: Only homography can not handle translation!

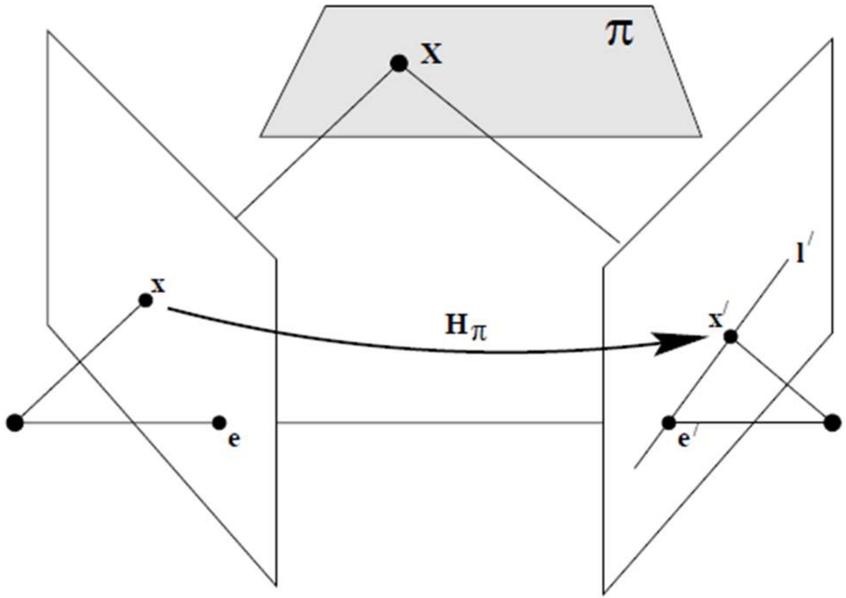


Epipolar Geometry





Fundamental Matrix



$$\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$$

$$\mathbf{x}' = \mathtt{H}_\pi \mathbf{x}$$

$$\mathbf{l}' = [\mathbf{e}']_\times \mathtt{H}_\pi \mathbf{x} = \mathtt{F}\mathbf{x}$$

x' lies on the l'

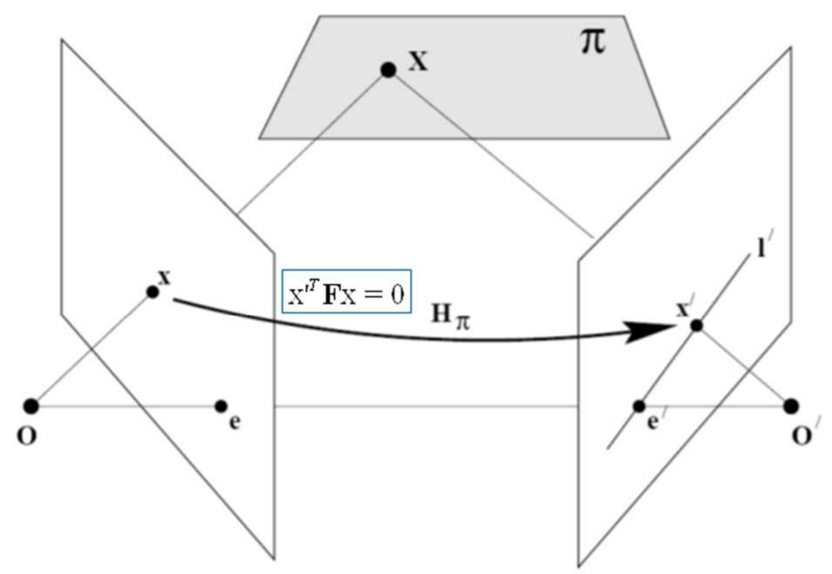
$$0 = \mathbf{x}'^\top \mathbf{l}' = \mathbf{x}'^\top \mathtt{F}\mathbf{x}$$

Features of Fundamental Matrix

- Depends only on the relative pose and internal parameters

$$F = K_2^{-T}[t]_\times R K_1^{-1}$$

- F is a 3 × 3 rank 2 homogeneous matrix
- F**e** = **0**
- It has 7 degrees of freedom
- Compute from 7 image point correspondences



Fundamental Matrix – Eight-point Algorithm

- Given a correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$
- Assume

$$\mathbf{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

$$\mathbf{f} = \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{21} & f_{22} & f_{23} & f_{31} & f_{32} & f_{33} \end{bmatrix}^{\mathsf{T}}$$

- We can get

$$\mathbf{x}'^{\mathsf{T}} \mathbf{F} \mathbf{x} = 0$$

$$\Downarrow$$

$$\begin{bmatrix} x'x & x'y & x' & y'x & y'y & y' & x & y & 1 \end{bmatrix} \mathbf{f} = 0$$

Fundamental Matrix – Eight-point Algorithm

- Given 8 correspondences

$$\begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ x'_2x_2 & x'_2y_2 & x'_2 & y'_2x_2 & y'_2y_2 & y'_2 & x_2 & y_2 & 1 \\ x'_3x_3 & x'_3y_3 & x'_3 & y'_3x_3 & y'_3y_3 & y'_3 & x_3 & y_3 & 1 \\ x'_4x_4 & x'_4y_4 & x'_4 & y'_4x_4 & y'_4y_4 & y'_4 & x_4 & y_4 & 1 \\ x'_5x_5 & x'_5y_5 & x'_5 & y'_5x_5 & y'_5y_5 & y'_5 & x_5 & y_5 & 1 \\ x'_6x_6 & x'_6y_6 & x'_6 & y'_6x_6 & y'_6y_6 & y'_6 & x_6 & y_6 & 1 \\ x'_7x_7 & x'_7y_7 & x'_7 & y'_7x_7 & y'_7y_7 & y'_7 & x_7 & y_7 & 1 \\ x'_8x_8 & x'_8y_8 & x'_8 & y'_8x_8 & y'_8y_8 & y'_8 & x_8 & y_8 & 1 \end{bmatrix} \mathbf{f} = 0 \Rightarrow \mathbf{Af} = 0$$

- Nontrivial solution
  - $\mathbf{f}$ is in null space of A

Fundamental Matrix – Singular Value Decomposition

Any m x n real matrix

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{T}}$$

$\mathbf{U}$ left-singular vectors

$\boldsymbol{\Sigma}$ diagonal values are singular values

$\mathbf{V}$ right-singular vectors

Properties

$$\mathbf{U}\mathbf{U}^{\mathsf{T}} = \mathbf{I} \quad \mathbf{V}\mathbf{V}^{\mathsf{T}} = \mathbf{I}$$

$\boldsymbol{\Sigma}$
1. Uniquely determined by **A**
2. Entries are nonnegative

Fundamental Matrix – Eight-Point Algorithm

- In presence of multiple correspondences

$$\min \|\mathbf{Af}\| \quad \text{subject to} \quad \|\mathbf{f}\| = 1$$

SVD!

The solution is the singular vector of $\mathbf{A}$ corresponding to the least singular value

- Rank constraint

$$\mathbf{F} \rightarrow \mathbf{F}' \qquad \det \mathbf{F}' = 0$$

- Minimize Frobenius norm

$$\min_{\mathbf{F}'} \|\mathbf{F} - \mathbf{F}'\|_{\mathsf{F}} \quad \text{subject to} \quad \det \mathbf{F}' = 0$$

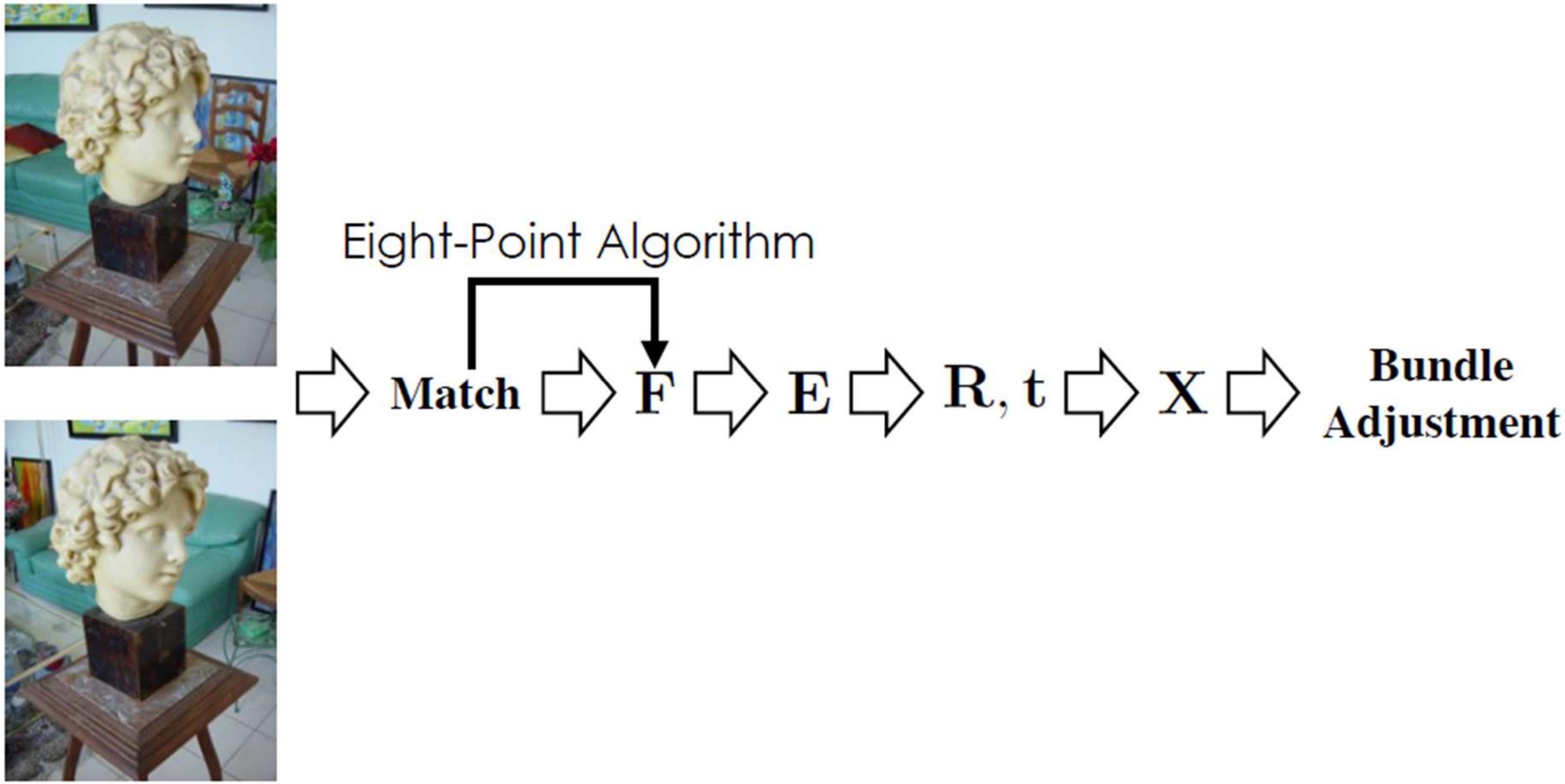
SVD!

$$\mathbf{F} = \mathbf{U}\,\mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)\mathbf{V}^{\mathsf{T}} \Rightarrow \mathbf{F}' = \mathbf{U}\,\mathrm{diag}(\sigma_1, \sigma_2, 0)\mathbf{V}^{\mathsf{T}}$$

Fundamental Matrix – RANSAC Estimation

- For many times
  - Pick 8 points
  - Compute a solution for $\mathbf{F}$ using these 8 points
  - Count number of inliers that with geometric error close to 0
- Pick the one with the largest number of inliers
- Only the inliers are kept as correspondences

Two View Reconstruction



Essential Matrix

- A relation of camera coordinates

$$\mathbf{x}_{\text{cam}} = \mathbf{K}^{-1}\mathbf{x} \quad \mathbf{x}'_{\text{cam}} = \mathbf{K}'^{-1}\mathbf{x}' \quad \Rightarrow \quad \tilde{\mathbf{x}}'^{\mathsf{T}}\mathbf{E}\mathbf{x}' = 0$$

- Related to fundamental matrix

$$\mathbf{E} = \mathbf{K}'^{\mathsf{T}}\mathbf{F}\mathbf{K}$$

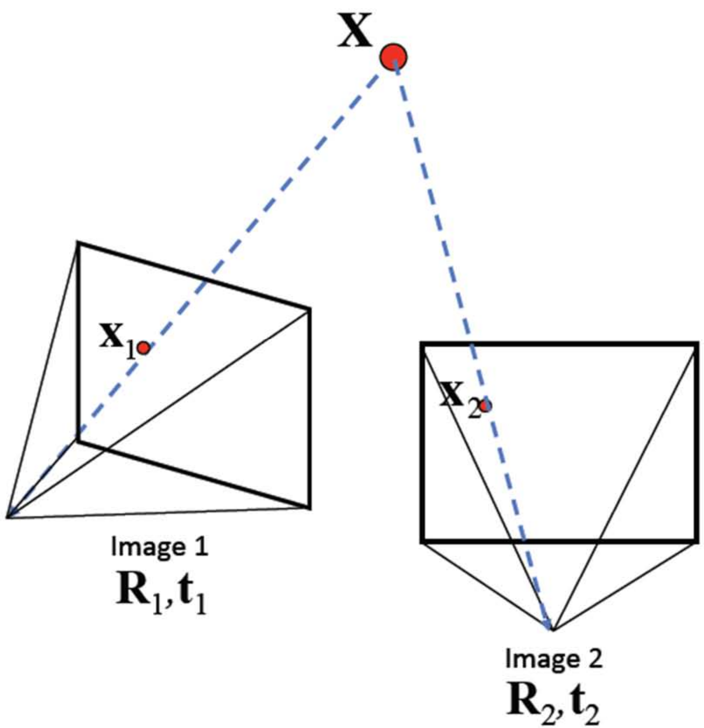
- Can get from camera extrinsics

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

- Properties

1) $\det \mathbf{E} = 0$
2) Two identical nonzero singular values
3) Five degree of freedom

Essential Matrix



$$\mathbf{x}'^{\mathsf{T}}\mathbf{F}\mathbf{x} = 0$$

$$\Downarrow$$

$$\mathbf{E} = \mathbf{K}'^{\mathsf{T}}\mathbf{F}\mathbf{K}$$

$$\Downarrow$$

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

$$\Downarrow$$

$$\mathbf{R}, \mathbf{t}$$

Essential Matrix

Result 9.19. For a given essential matrix

$$\mathbf{E} = \mathbf{U}\mathrm{diag}(1,1,0)\mathbf{V}^T,$$

and the first camera matrix $\mathbf{P}_1 = [\mathbf{I} | \mathbf{0}]$, there are four possible choices for the second camera matrix $\mathbf{P}_2$:

$$\mathbf{P}_2 = [\mathbf{U}\mathbf{W}\mathbf{V}^T | +\mathbf{u}_3]$$

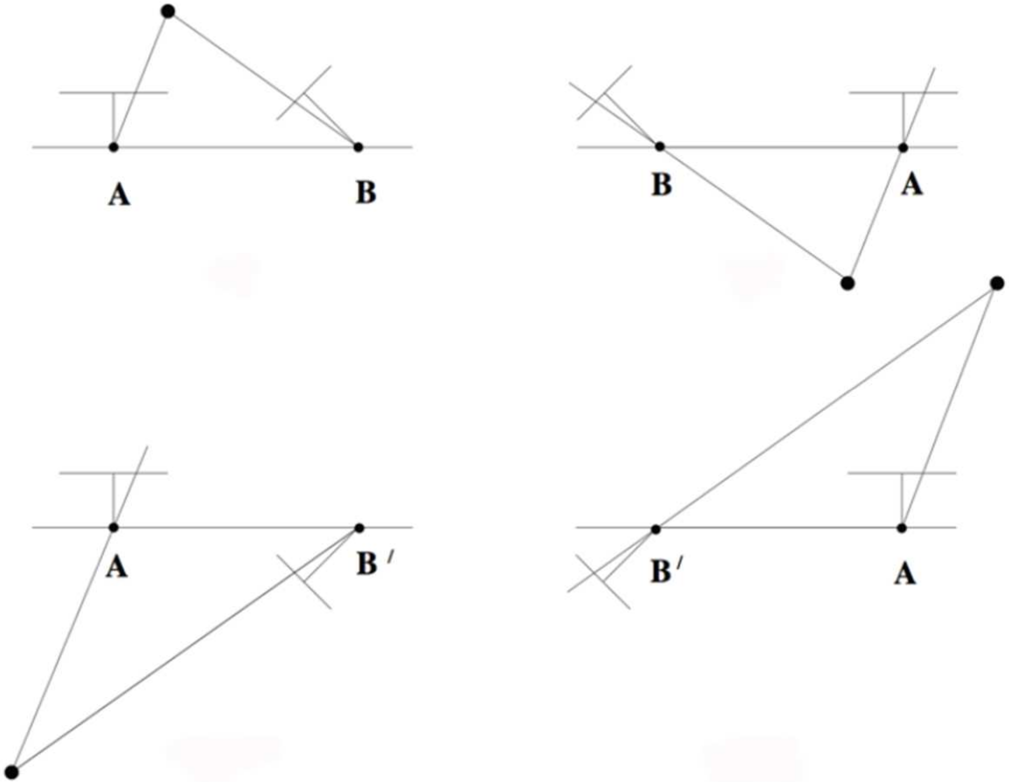
$$\mathbf{P}_2 = [\mathbf{U}\mathbf{W}\mathbf{V}^T | -\mathbf{u}_3]$$

$$\mathbf{P}_2 = [\mathbf{U}\mathbf{W}^T\mathbf{V}^T | +\mathbf{u}_3]$$

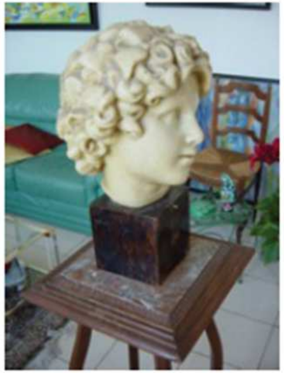
$$\mathbf{P}_2 = [\mathbf{U}\mathbf{W}^T\mathbf{V}^T | -\mathbf{u}_3]$$

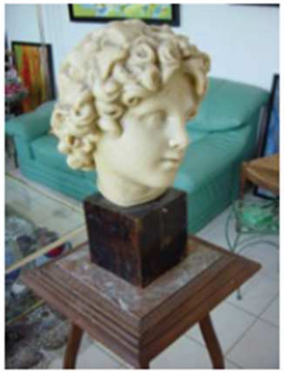
$$\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

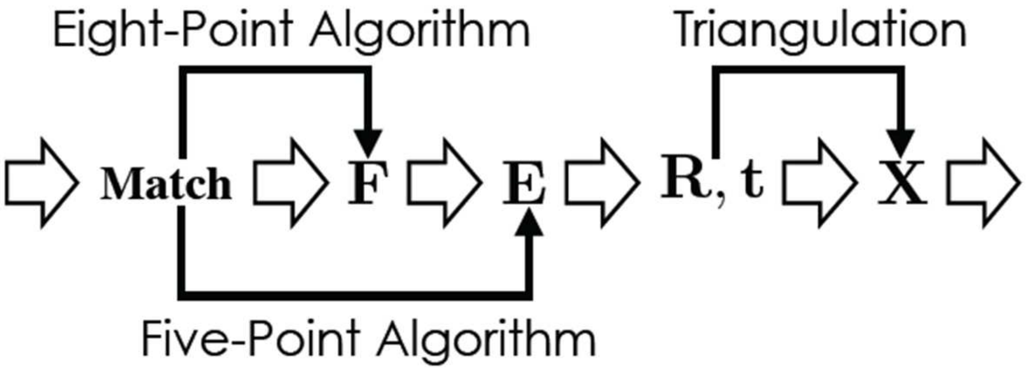
Essential Matrix – Four Possible Solution

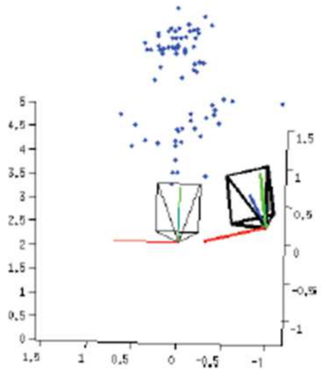


Two View Reconstruction

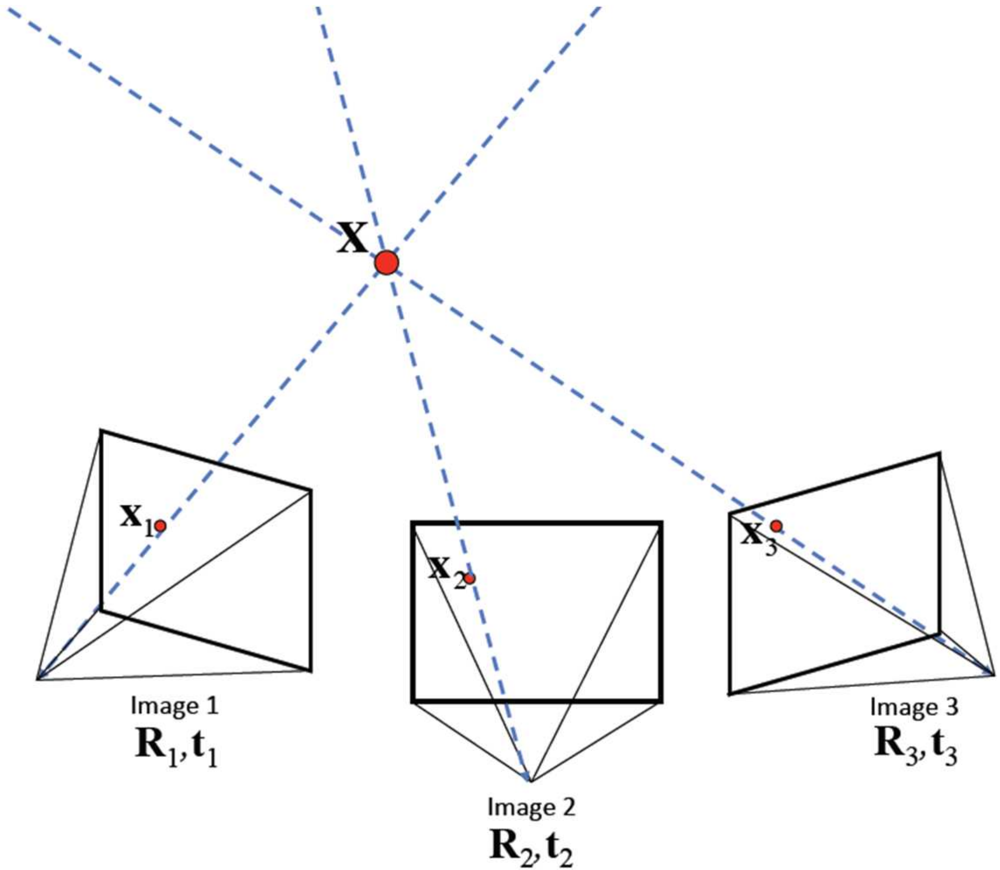








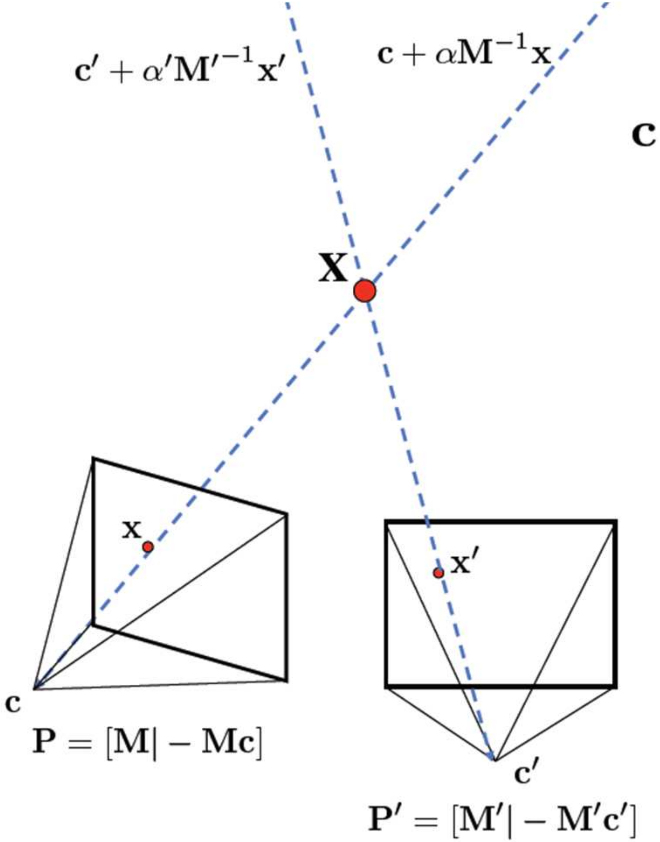
Triangulation



Triangulation

- Mid-point algorithm [Hartley and Sturm 1997]
  - \+ Easy to implement
  - \+ Easy to generalize to multiple views
  - \- Not "optimal"
- Optimal method
  - \+ Minimize reprojection error
  - \- Not big improvement in practice
  - \- Hard to generalize to multiple views

Mid-point Algorithm



$$\mathbf{c} + \alpha\mathbf{M}^{-1}\mathbf{x} = \mathbf{c}' + \alpha'\mathbf{M}'^{-1}\mathbf{x}'$$

2 unknowns + 3 equations

$\Downarrow$

Minimize the squared distance between two rays

$\Downarrow$

$$\mathbf{X} = (\mathbf{c} + \alpha\mathbf{M}^{-1}\mathbf{x} + \mathbf{c}' + \alpha'\mathbf{M}'^{-1}\mathbf{x}')/2$$

View Direction

- Given camera extrinsic $\mathbf{R}, \mathbf{t}$

$$\mathbf{x}_{\text{cam}} = \mathbf{R}\mathbf{x}_{\text{world}} + \mathbf{t} \quad \Longrightarrow \quad \mathbf{x}_{\text{world}} = \mathbf{R}^{\mathsf{T}}(\mathbf{x}_{\text{cam}} - \mathbf{t})$$

- Camera center

$$\mathbf{x}_{\text{cam}} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{c} = \mathbf{x}_{\text{world}} = \mathbf{R}^{\mathsf{T}}\mathbf{0} - \mathbf{R}^{\mathsf{T}}\mathbf{t} = -\mathbf{R}^{\mathsf{T}}\mathbf{t}$$

- View direction

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \mathbf{0} \quad \Longrightarrow \quad \left(\mathbf{R}^{\mathsf{T}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \mathbf{R}^{\mathsf{T}}\mathbf{t} \right) - \mathbf{c} = \mathbf{R}(3,:)^{\mathsf{T}}$$

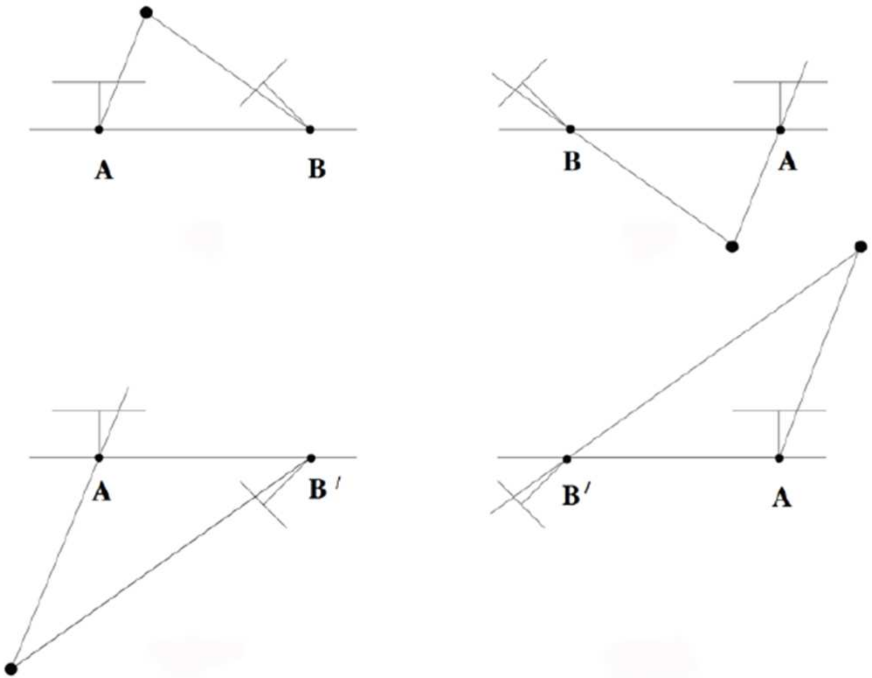
Front of the Camera

- A point: $\mathbf{x}_{\text{world}}$
- Direction: $\mathbf{x}_{\text{world}} - \mathbf{c}$
- Angle: $\mathbf{a}^{\mathsf{T}}\mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos\theta$
- What to test?

$$(\mathbf{x}_{\text{world}} - \mathbf{c})^{\mathsf{T}}\mathbf{R}(3,:) > 0$$

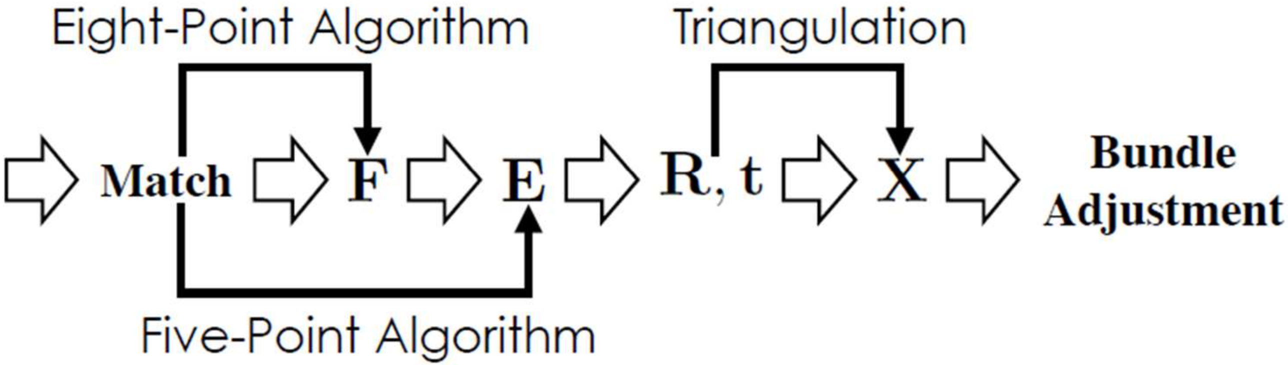
Field of view

Pick the Solution

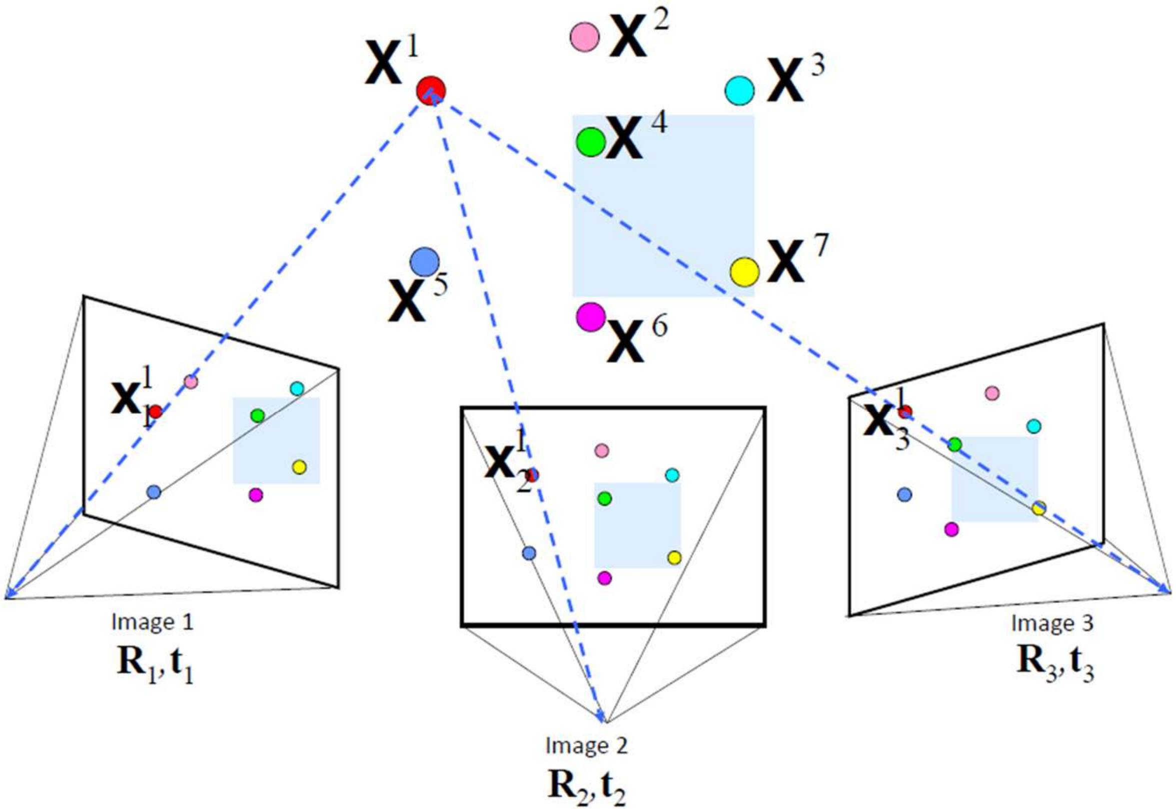


- With maximal number of points in front of both cameras.
- "Cheiralilty Constraints"

Two View Reconstruction



Bundle Adjustment



Bundle Adjustment

	Point 1	Point 2	Point 3
Image 1	$\mathbf{x}_1^1 = \mathbf{K}\left[\mathbf{R}_1 \middle| \mathbf{t}_1\right]\mathbf{X}^1$	$\mathbf{x}_1^2 = \mathbf{K}\left[\mathbf{R}_1 \middle| \mathbf{t}_1\right]\mathbf{X}^2$	
Image 2	$\mathbf{x}_2^1 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^1$	$\mathbf{x}_2^2 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^2$	$\mathbf{x}_2^3 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^3$
Image 3	$\mathbf{x}_3^1 = \mathbf{K}\left[\mathbf{R}_3 \middle| \mathbf{t}_3\right]\mathbf{X}^1$		$\mathbf{x}_3^3 = \mathbf{K}\left[\mathbf{R}_3 \middle| \mathbf{t}_3\right]\mathbf{X}^3$

Bundle Adjustment

	Point 1	Point 2	Point 3
Image 1	$\mathbf{x}_1^1 = \mathbf{K}\left[\mathbf{R}_1 \middle| \mathbf{t}_1\right]\mathbf{X}^1$	$\mathbf{x}_1^2 = \mathbf{K}\left[\mathbf{R}_1 \middle| \mathbf{t}_1\right]\mathbf{X}^2$	
Image 2	$\mathbf{x}_2^1 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^1$	$\mathbf{x}_2^2 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^2$	$\mathbf{x}_2^3 = \mathbf{K}\left[\mathbf{R}_2 \middle| \mathbf{t}_2\right]\mathbf{X}^3$
Image 3	$\mathbf{x}_3^1 = \mathbf{K}\left[\mathbf{R}_3 \middle| \mathbf{t}_3\right]\mathbf{X}^1$		$\mathbf{x}_3^3 = \mathbf{K}\left[\mathbf{R}_3 \middle| \mathbf{t}_3\right]\mathbf{X}^3$

Bundle Adjustment

- Input: Observed 2D image position

$$\begin{array}{ccc} \tilde{\mathbf{x}}_1^1 & \tilde{\mathbf{x}}_1^2 & \\ \tilde{\mathbf{x}}_2^1 & \tilde{\mathbf{x}}_2^2 & \tilde{\mathbf{x}}_2^3 \\ \tilde{\mathbf{x}}_3^1 & & \tilde{\mathbf{x}}_3^3 \end{array}$$

- Output:

Unknown Camera Parameters (with some guess)

$$[\mathbf{R}_1|\mathbf{t}_1],[\mathbf{R}_2|\mathbf{t}_2],[\mathbf{R}_3|\mathbf{t}_3]$$

Unknown Point 3D coordinate (with some guess)

$$\mathbf{X}^1,\mathbf{X}^2,\mathbf{X}^3,\cdots$$

Bundle Adjustment

A valid solution $\left[\mathbf{R}_1|\mathbf{t}_1\right], \left[\mathbf{R}_2|\mathbf{t}_2\right], \left[\mathbf{R}_3|\mathbf{t}_3\right]$ and $\mathbf{X}^1, \mathbf{X}^2, \mathbf{X}^3, \cdots$ must let

$$\text{Re-projection} \left\{ \begin{array}{lll} \mathbf{x}_1^1 = \mathbf{K}\left[\mathbf{R}_1|\mathbf{t}_1\right]\mathbf{X}^1 & \mathbf{x}_1^2 = \mathbf{K}\left[\mathbf{R}_1|\mathbf{t}_1\right]\mathbf{X}^2 & \\ \mathbf{x}_2^1 = \mathbf{K}\left[\mathbf{R}_2|\mathbf{t}_2\right]\mathbf{X}^1 & \mathbf{x}_2^2 = \mathbf{K}\left[\mathbf{R}_2|\mathbf{t}_2\right]\mathbf{X}^2 & \mathbf{x}_2^3 = \mathbf{K}\left[\mathbf{R}_2|\mathbf{t}_2\right]\mathbf{X}^3 \\ \mathbf{x}_3^1 = \mathbf{K}\left[\mathbf{R}_3|\mathbf{t}_3\right]\mathbf{X}^1 & & \mathbf{x}_3^3 = \mathbf{K}\left[\mathbf{R}_3|\mathbf{t}_3\right]\mathbf{X}^3 \end{array} \right.$$

$$=$$

$$\text{Observation} \left\{ \begin{array}{lll} \tilde{\mathbf{x}}_1^1 & \tilde{\mathbf{x}}_1^2 & \\ \tilde{\mathbf{x}}_2^1 & \tilde{\mathbf{x}}_2^2 & \tilde{\mathbf{x}}_2^3 \\ \tilde{\mathbf{x}}_3^1 & & \tilde{\mathbf{x}}_3^3 \end{array} \right.$$

Bundle Adjustment

A valid solution $\left[\mathbf{R}_1|\mathbf{t}_1\right], \left[\mathbf{R}_2|\mathbf{t}_2\right], \left[\mathbf{R}_3|\mathbf{t}_3\right]$ and $\mathbf{X}^1, \mathbf{X}^2, \mathbf{X}^3, \cdots$ must let the Re-projection close to the Observation, i.e. to minimize the reprojection error

$$\min \sum_i \sum_j \left(\tilde{\mathbf{x}}_i^j - \mathbf{K}\left[\mathbf{R}_i|\mathbf{t}_i\right]\mathbf{X}^j\right)^2$$

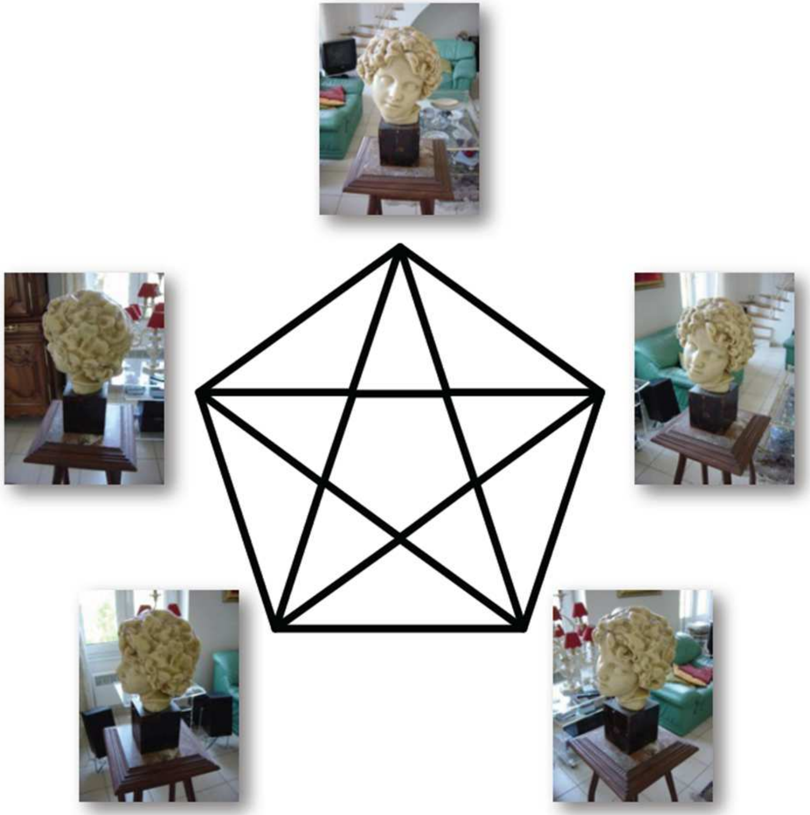
Bundle Adjustment

Refine a visual reconstruction to produce **jointly optimal** 3D structure and viewing parameter (camera pose and/or calibration) estimates [Triggs 00]

Optimality → Cost Function → Error Model

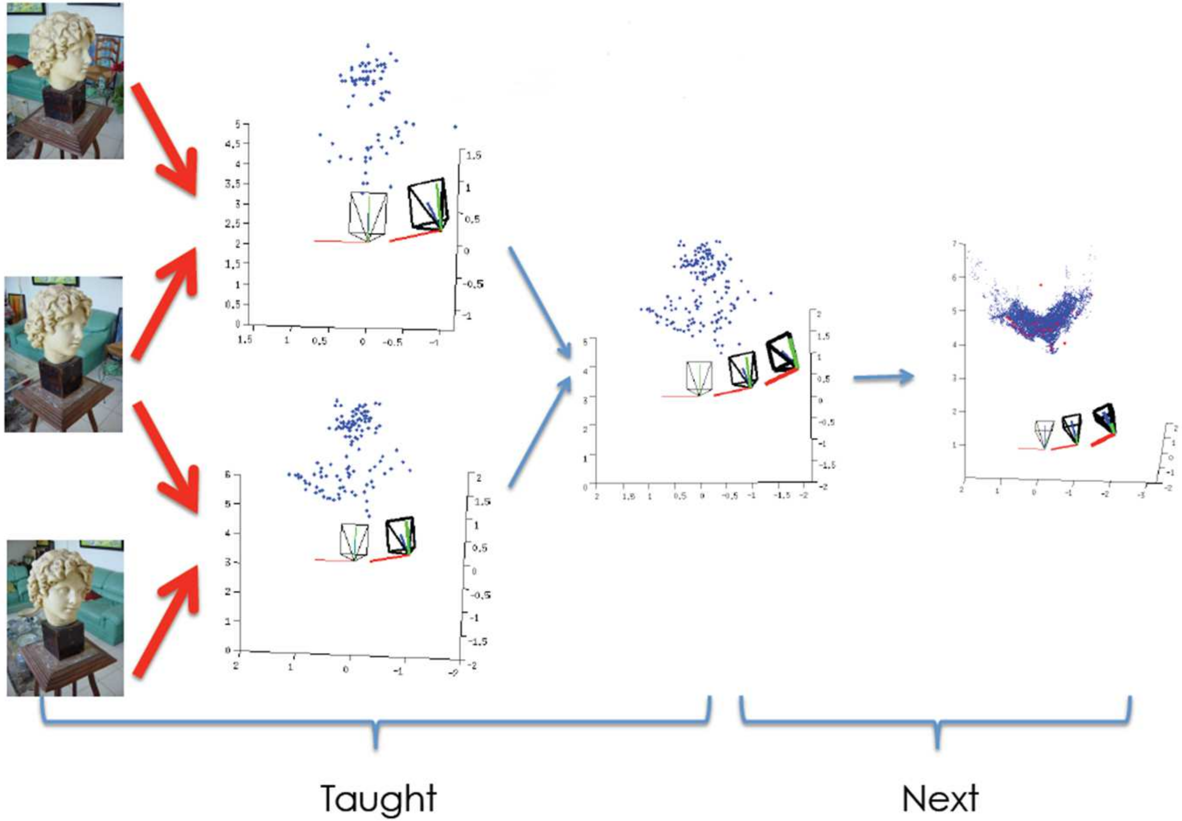
Which Two Views?

Select First Two View

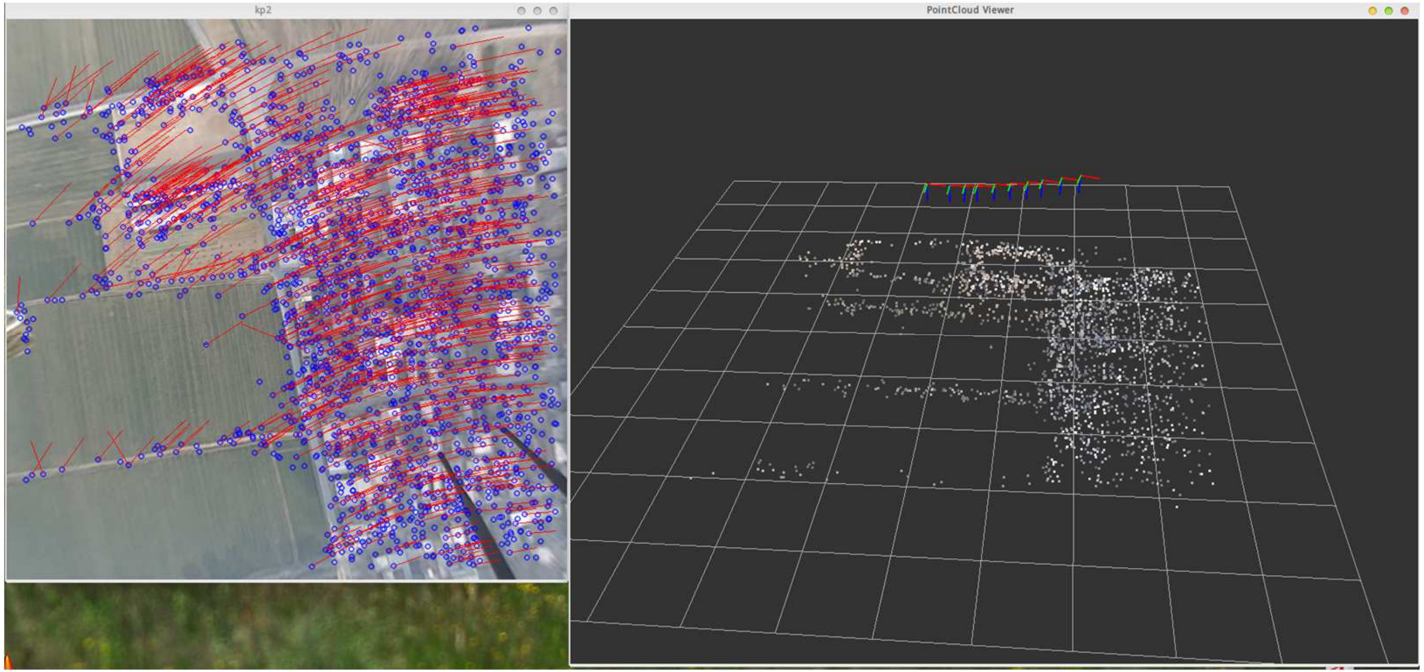


- Select two views that can be reliably reconstructed
- When? How?

Pipeline



Demo



http://www.adv-ci.com/blog/source/sequence-sfm/

How to Efficiently Learn

1. Read example program > 3 times
   Find good examples at *github*
2. STL, OpenCV, OpenGL, Eigen3, …
3. Practice key components: SE3, RANSAC, ICP, Bundle Adjustment
4. Global -> Detail
5. Read the bible <<Multi-view Geometry>>