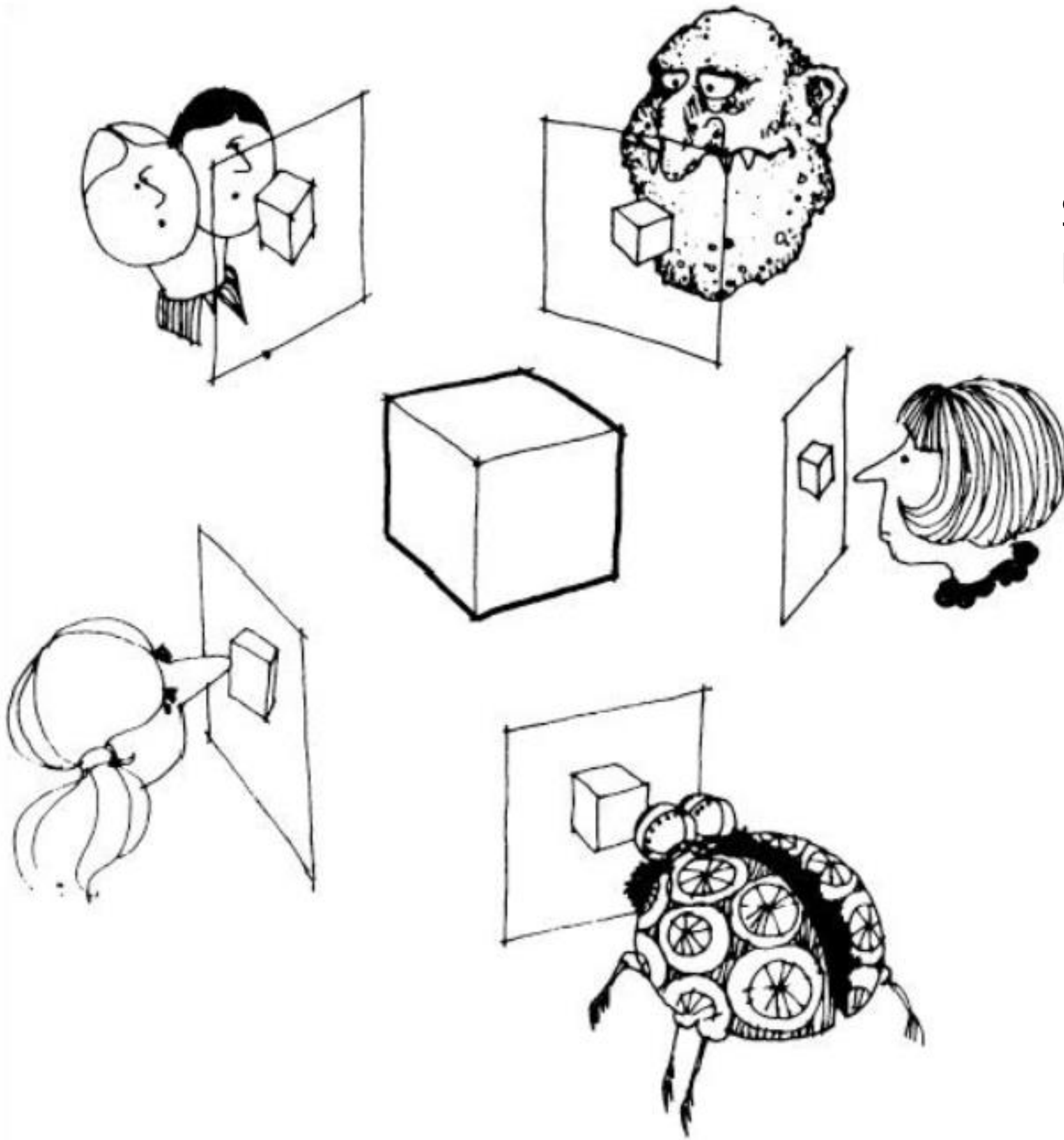


Multi-View Geometry:

Structure from Motion









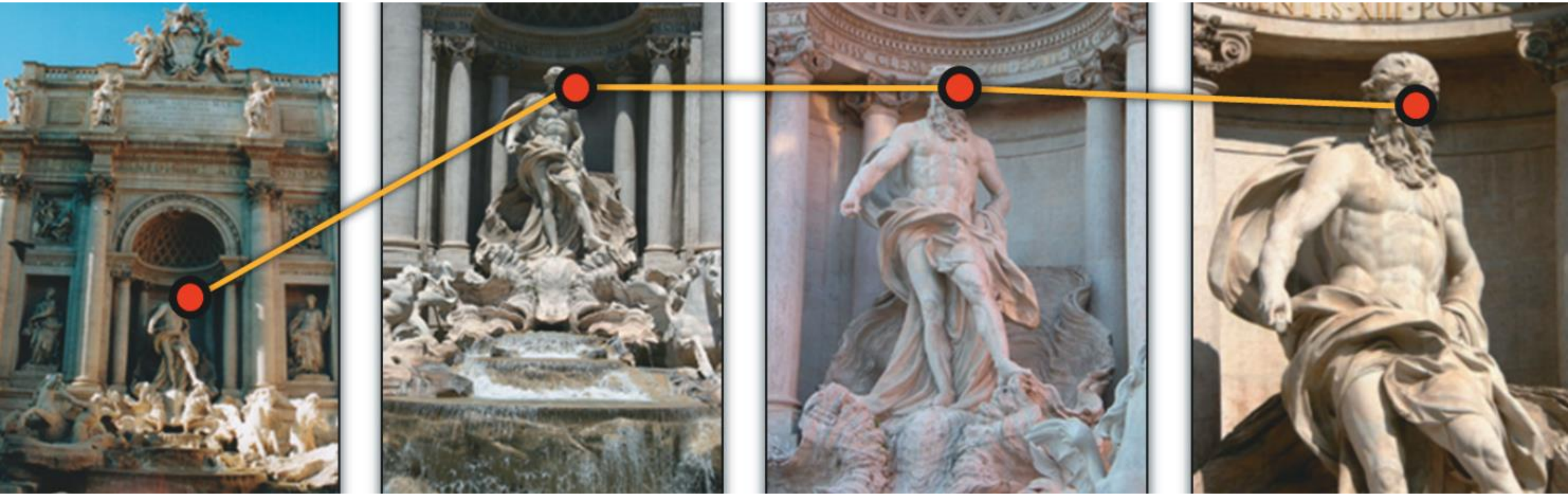




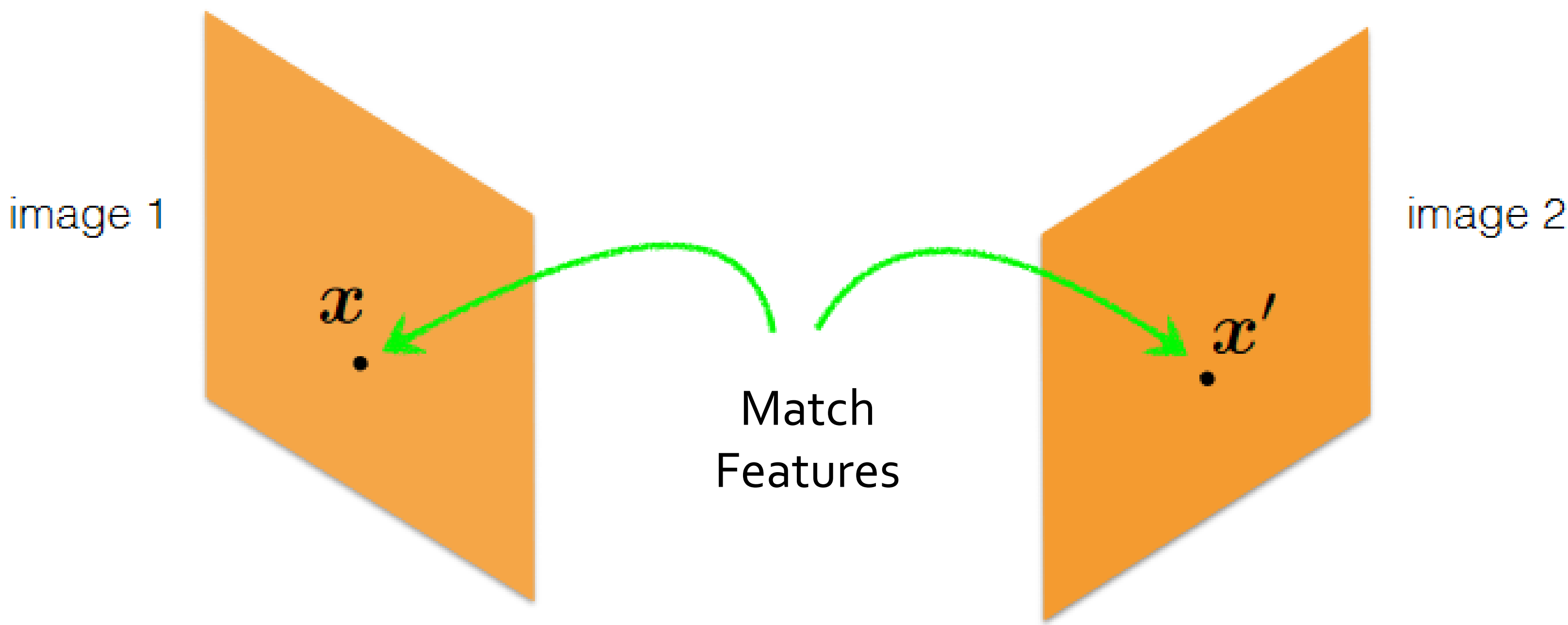
M·AGRIPPA·L·F·COS·TERTIVM·FECIT

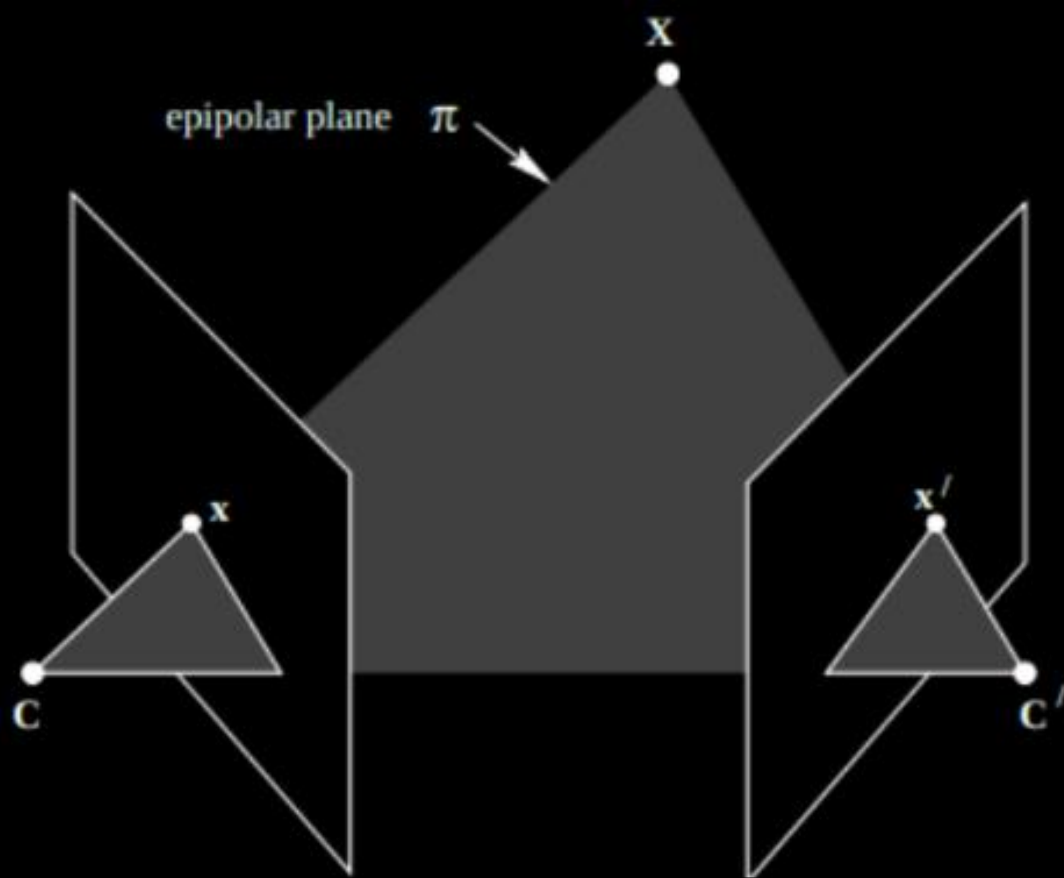
Before you Build, Some Questions:

- How does your camera work?
    - Pinhole
    - Camera Intrinsics
- How do you determine that two images are similar?
    - Feature Extraction (SIFT and friends)
- How do you know the camera positions?
    - Epipolar Geometry

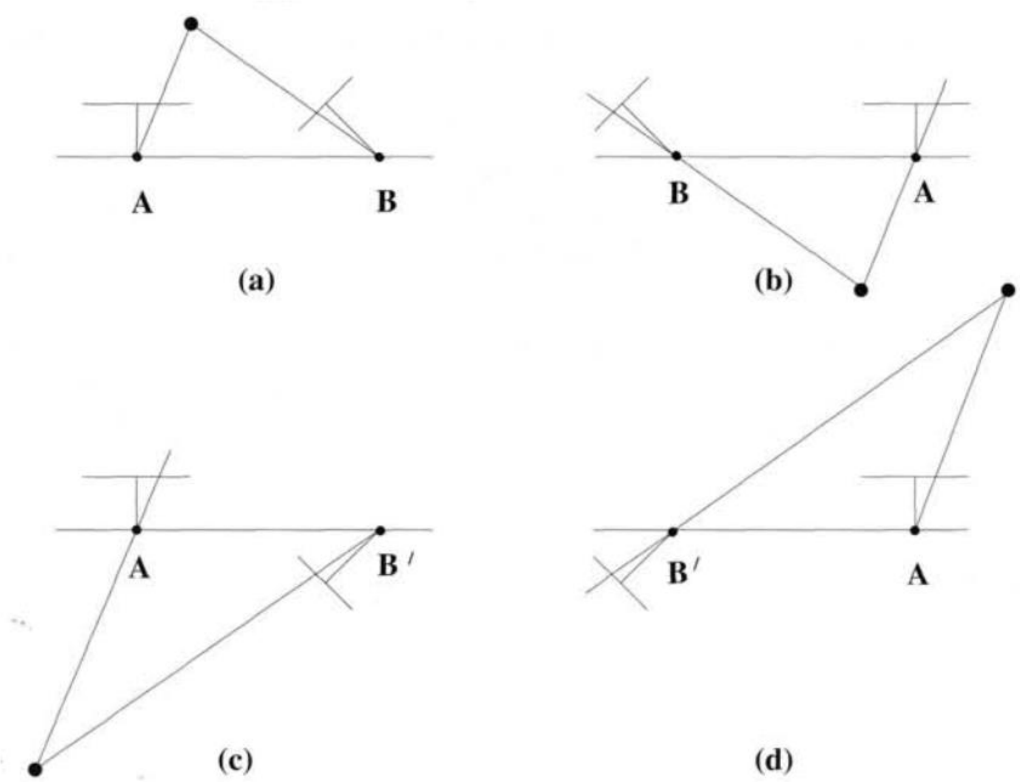


So far, 2D-2D Visual Odometry

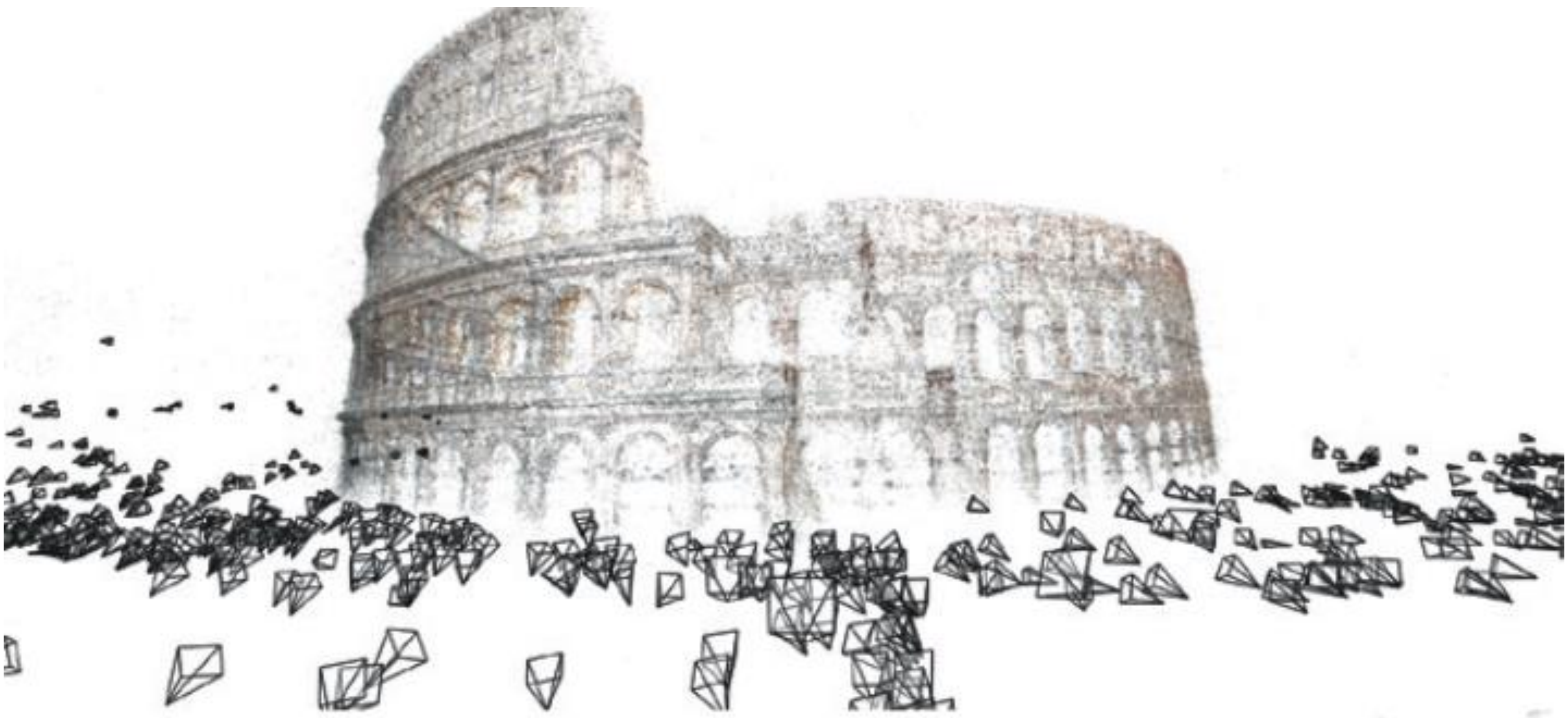




$$x'^T F x = 0$$



The four possible solutions for calibrated reconstruction from E. Between the left and right sides there is a baseline reversal. Between the top and bottom rows camera B rotates 180° about the baseline. Note, only in (a) is the reconstructed point in front of both cameras.

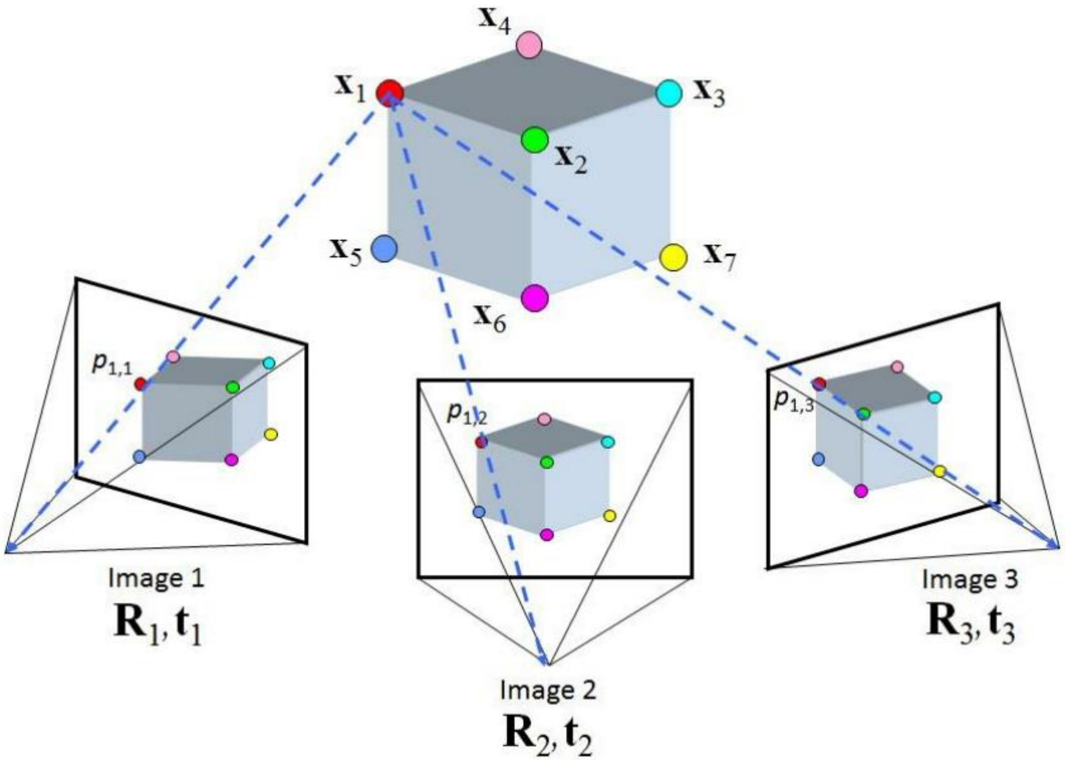


Building Rome in a Day

https://grail.cs.washington.edu/rome/

Watermarks are annoying

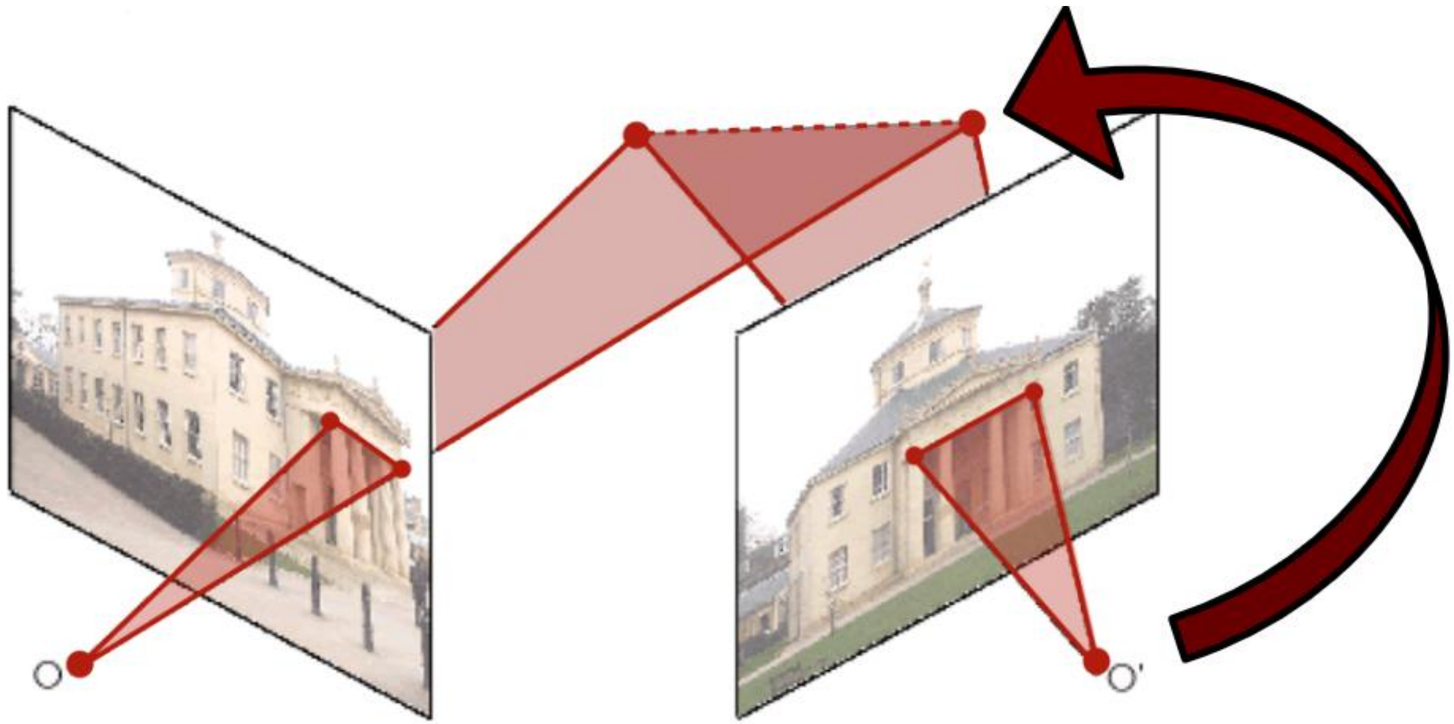
Structure from Motion



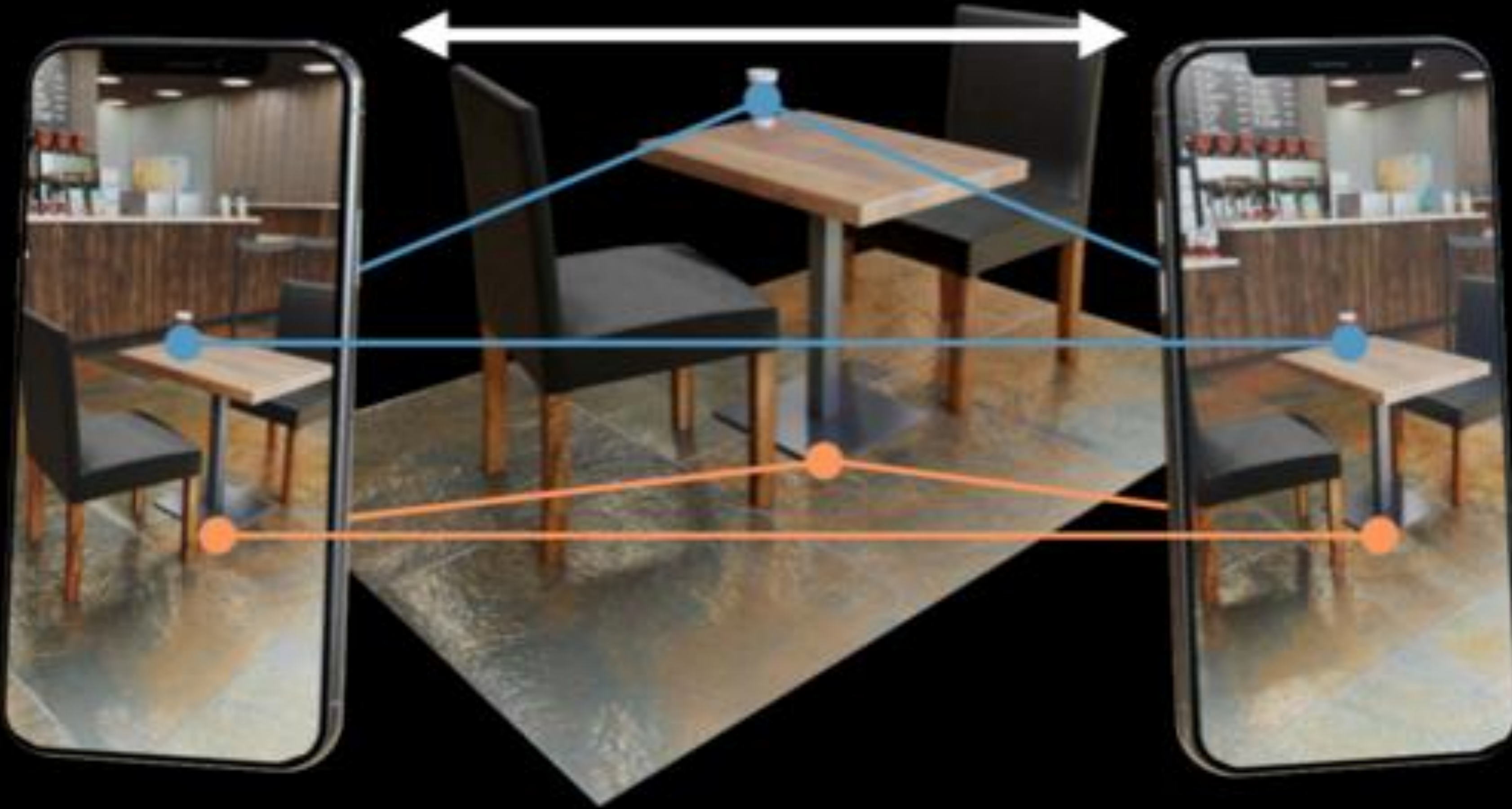
Up Next:

Stereo Camera and Depth

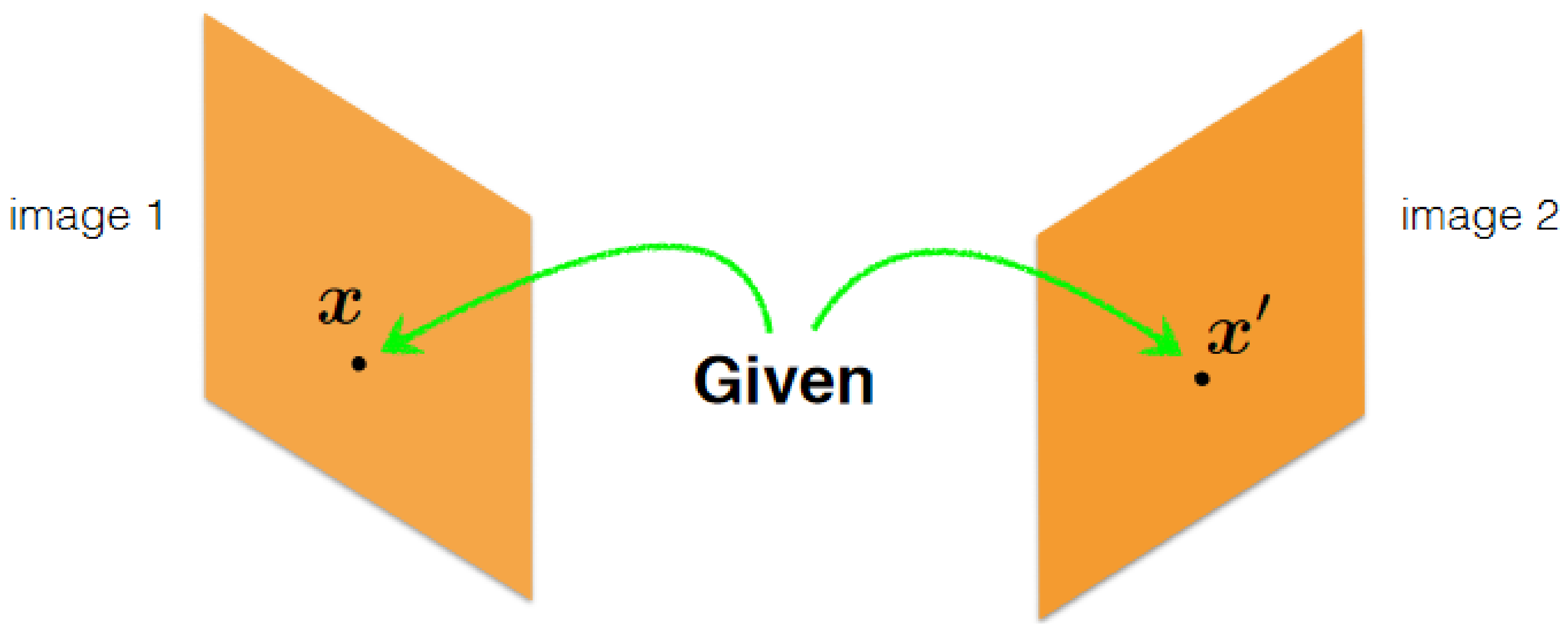
Triangulation



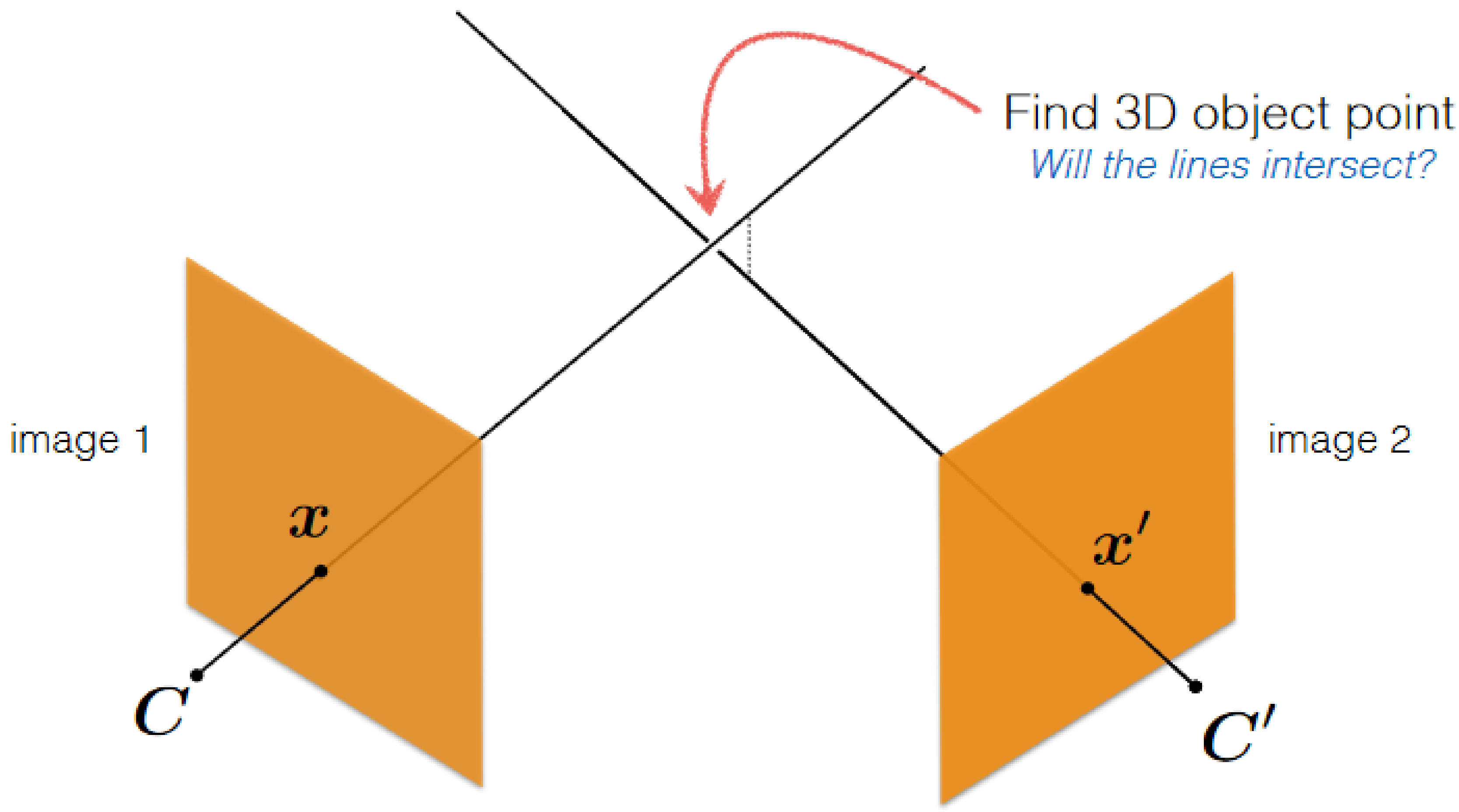
Given the relative orientation of two images, compute the points in 3D



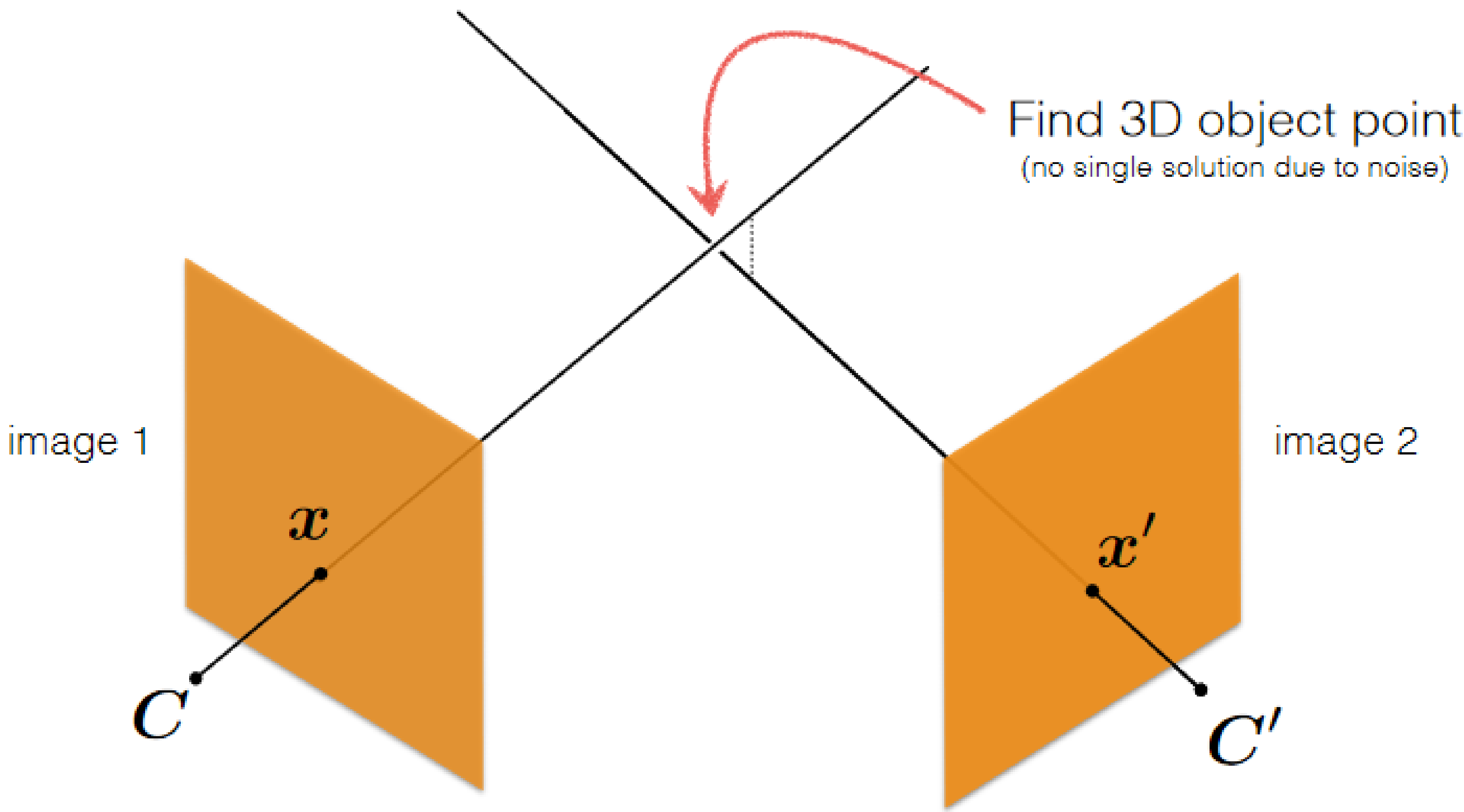
Triangulation



Triangulation



Triangulation



Triangulation

Given a set of (noisy) matched points

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

and camera matrices

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

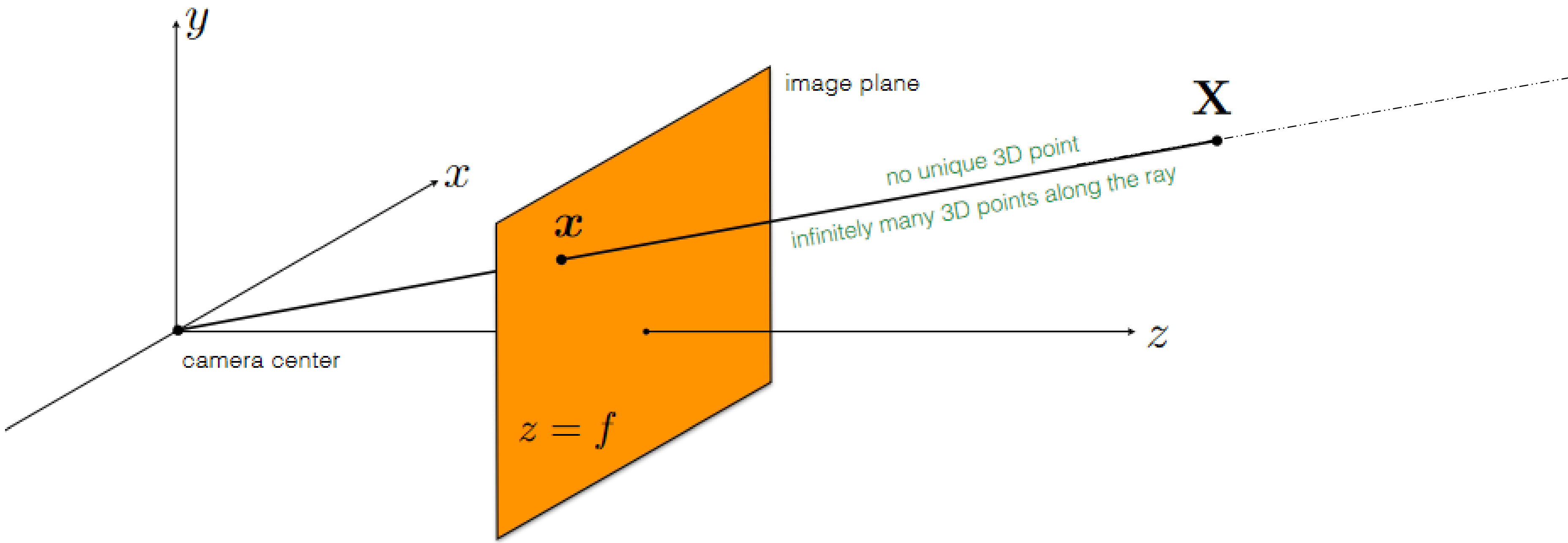
Estimate the 3D point

$$\mathbf{X}$$

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

known known

*Can we compute **X** from a single correspondence **x**?*



y
x
image plane
$\mathbf{X}$
no unique 3D point
infinitely many 3D points along the ray
$\boldsymbol{x}$
camera center
z
$z = f$

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

known known

*Can we compute **X** from two correspondences **x** and **x'**?*

yes if perfect measurements

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

known known

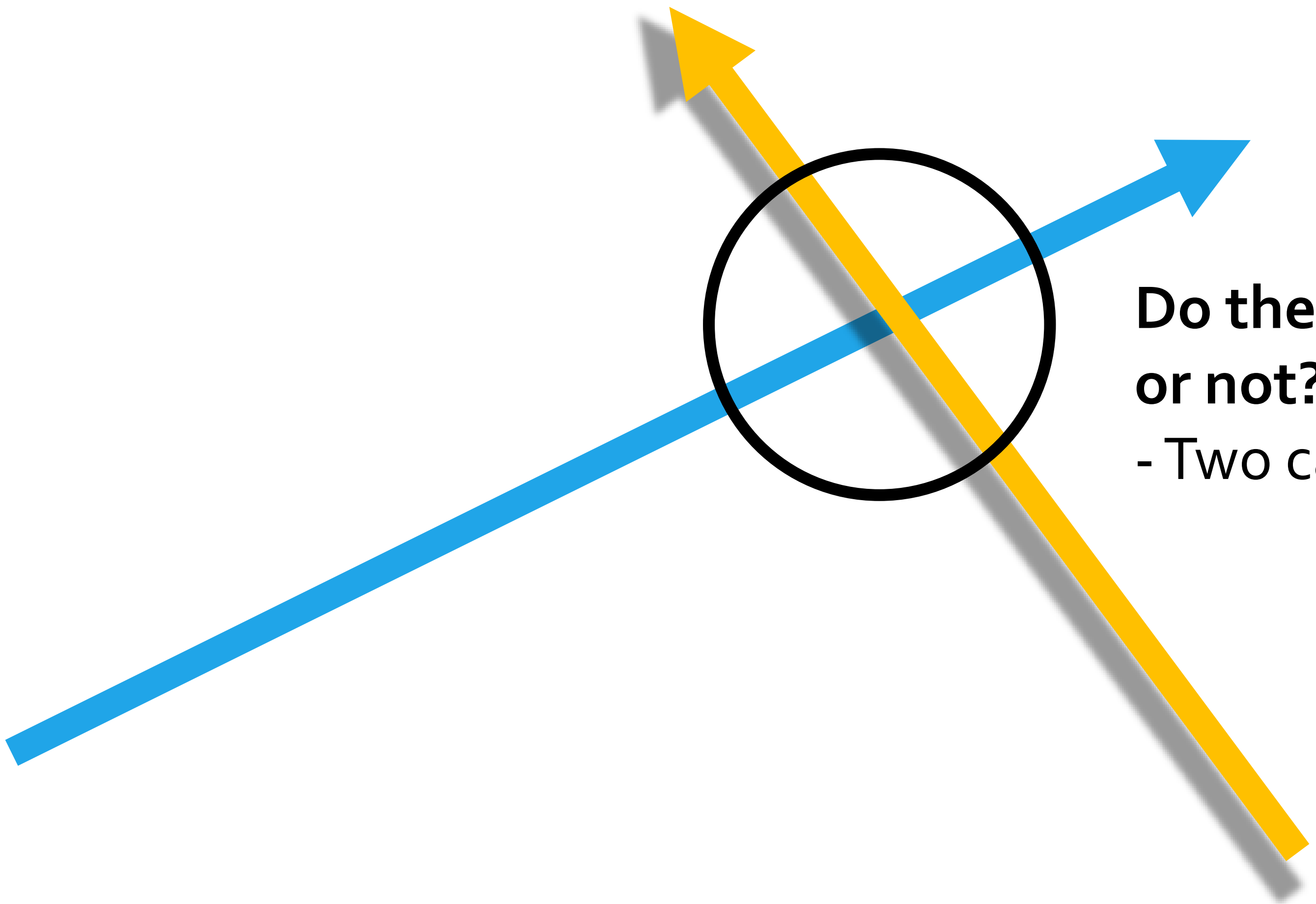
Can we compute ***X*** *from* *<u>two</u>* *correspondences* ***x*** *and* ***x'****?*

yes if perfect measurements

There will not be a point that satisfies both constraints because the measurements are usually noisy

$$\mathbf{x}' = \mathbf{P}'\boldsymbol{X} \qquad \mathbf{x} = \mathbf{P}\boldsymbol{X}$$

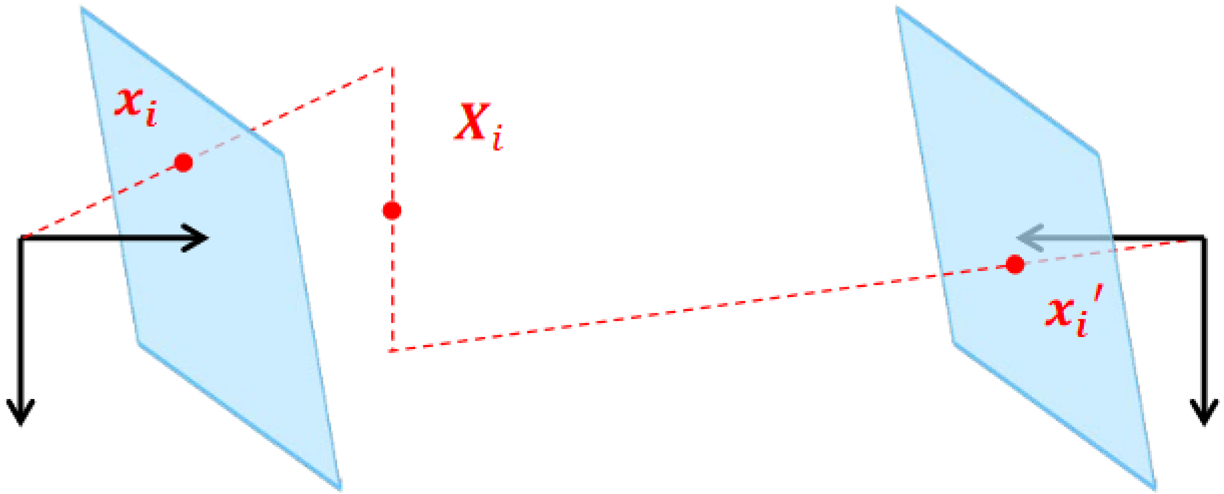
Need to find the **best fit**

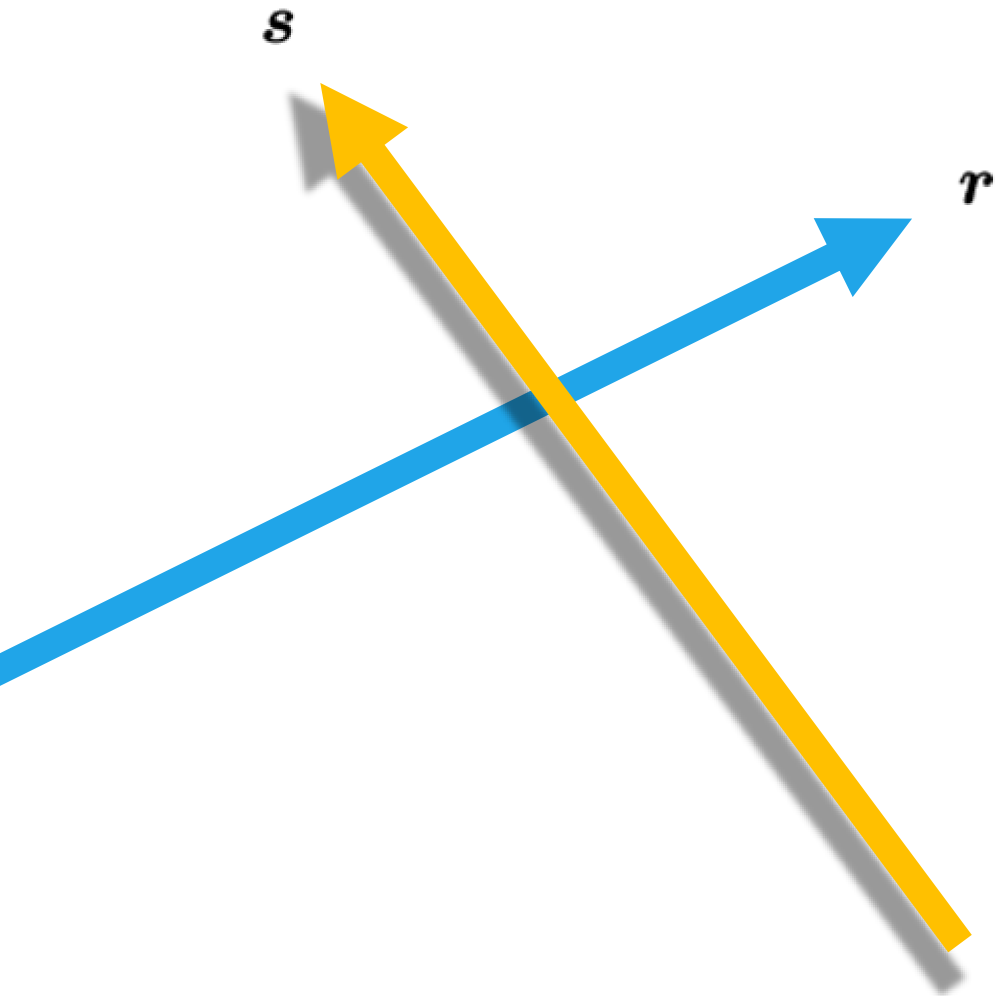


Do they intersect or not?

- Two cases

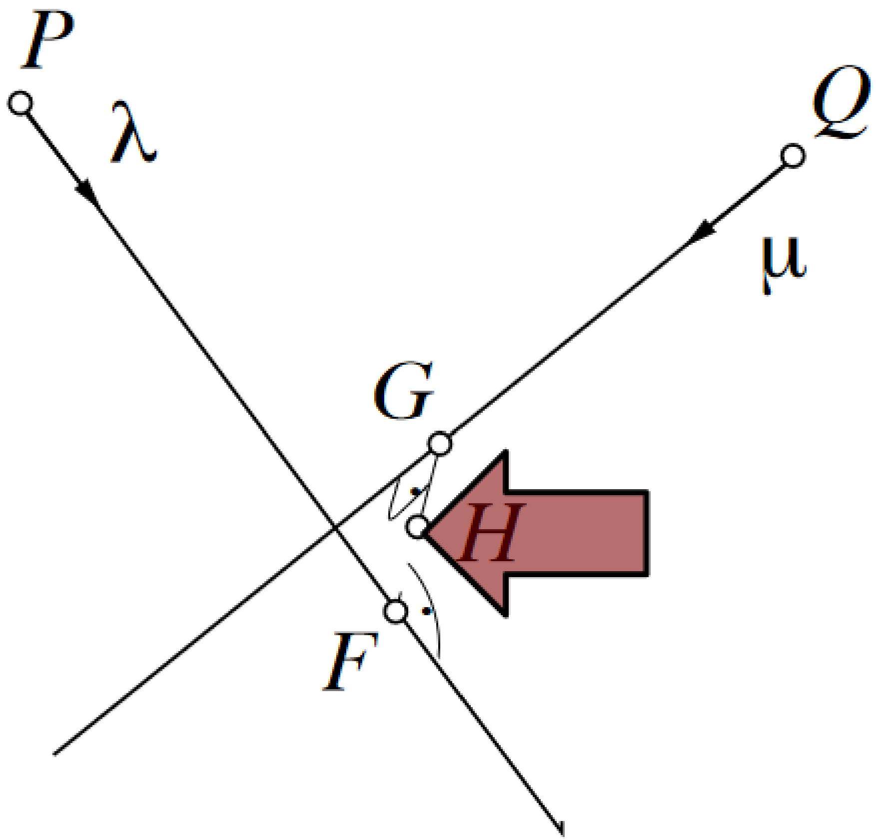
Minimize 3D Error





Directions

$$\boldsymbol{r} = K_1^{-1}\boldsymbol{x_1}$$
$$\boldsymbol{s} = R_2^1 K_2^{-1}\boldsymbol{x_2}$$



Line Equations

$$\boldsymbol{f} = \boldsymbol{P} + \lambda \boldsymbol{r}$$
$$\boldsymbol{g} = \boldsymbol{Q} + \mu \boldsymbol{s}$$

The shortest line between these two is perpendicular

For a point f and g, the line connecting the two is

$$(\boldsymbol{f} - \boldsymbol{g})$$

Since this line is perpendicular to our rays

$$(\boldsymbol{f} - \boldsymbol{g}) \cdot \boldsymbol{r} = 0 \qquad (\boldsymbol{f} - \boldsymbol{g}) \cdot \boldsymbol{s} = 0$$

$$(\boldsymbol{f} - \boldsymbol{g}) \cdot \boldsymbol{r} = 0 \qquad (\boldsymbol{f} - \boldsymbol{g}) \cdot \boldsymbol{s} = 0$$

Then,

$$\begin{aligned}(\boldsymbol{q} + \lambda \boldsymbol{s} - \boldsymbol{p} - \mu \boldsymbol{r}) \cdot \boldsymbol{s} &= 0 \\ (\boldsymbol{q} + \lambda \boldsymbol{s} - \boldsymbol{p} - \mu \boldsymbol{r}) \cdot \boldsymbol{r} &= 0\end{aligned}$$

Two Equations, Two Unknowns, We can solve it!

Now, take the midpoint of this line.

Getting 3D Point – More Methods

Minimal 3D error – Choose $\boldsymbol{X}_i$ to be the mid-point between back projected image points

Minimal algebraic error – Combine the two perspective models to get a homogeneous system of linear equations, then determine $\boldsymbol{X}_i$ by SVD

Minimal reprojection error – Determine the epipolar plane (and points $\widehat{\boldsymbol{u}}_i$ and $\widehat{\boldsymbol{u}}_i'$) that minimize the reprojection error by minimizing a 6^{th} order polynomial

Revisting Epipolar and Triangulation

Pure rotation ambiguity

- Triangulation will fail
- Baseline should be high
  - If too high, feature matching difficult

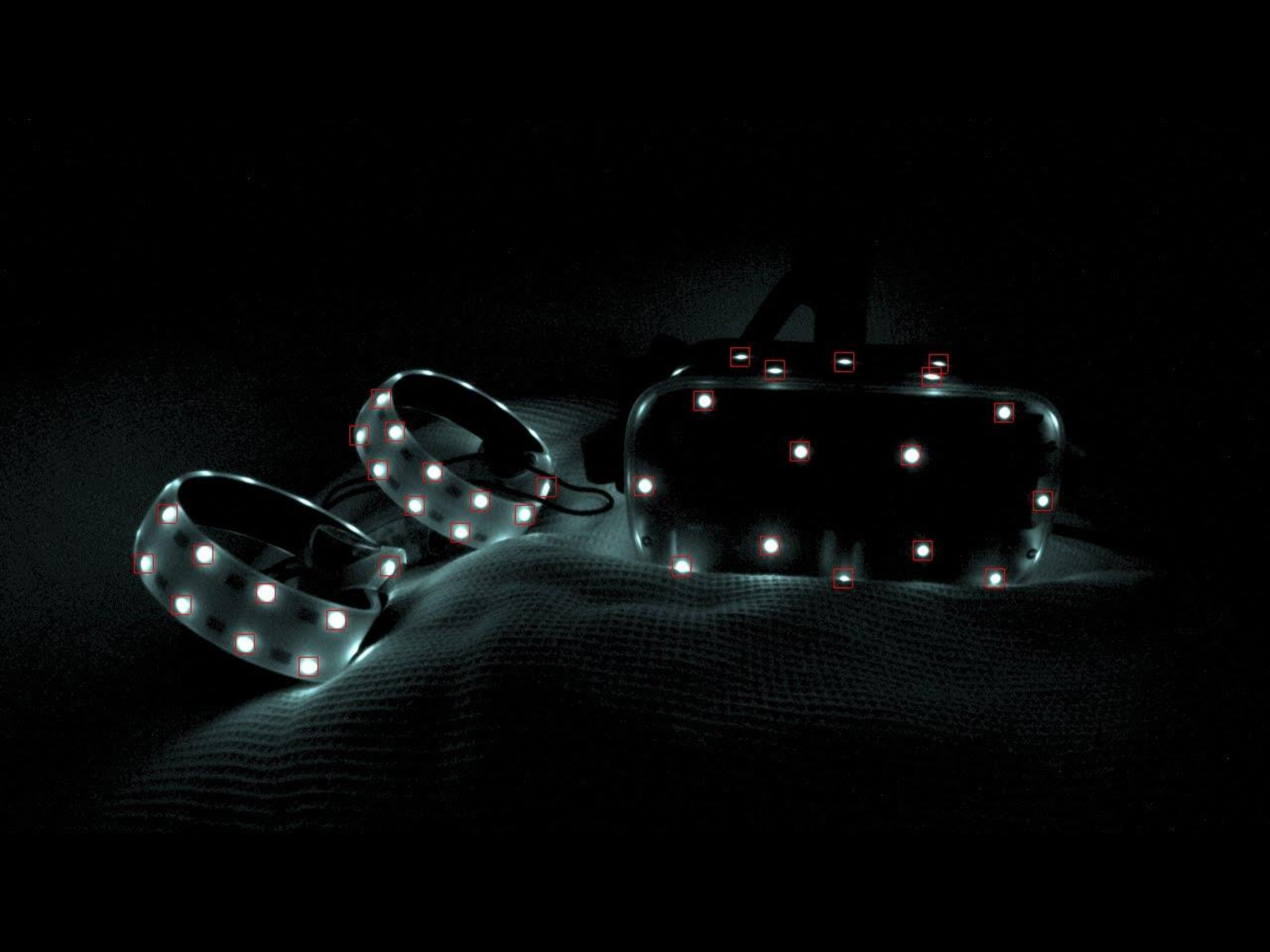
Scale Ambiguity

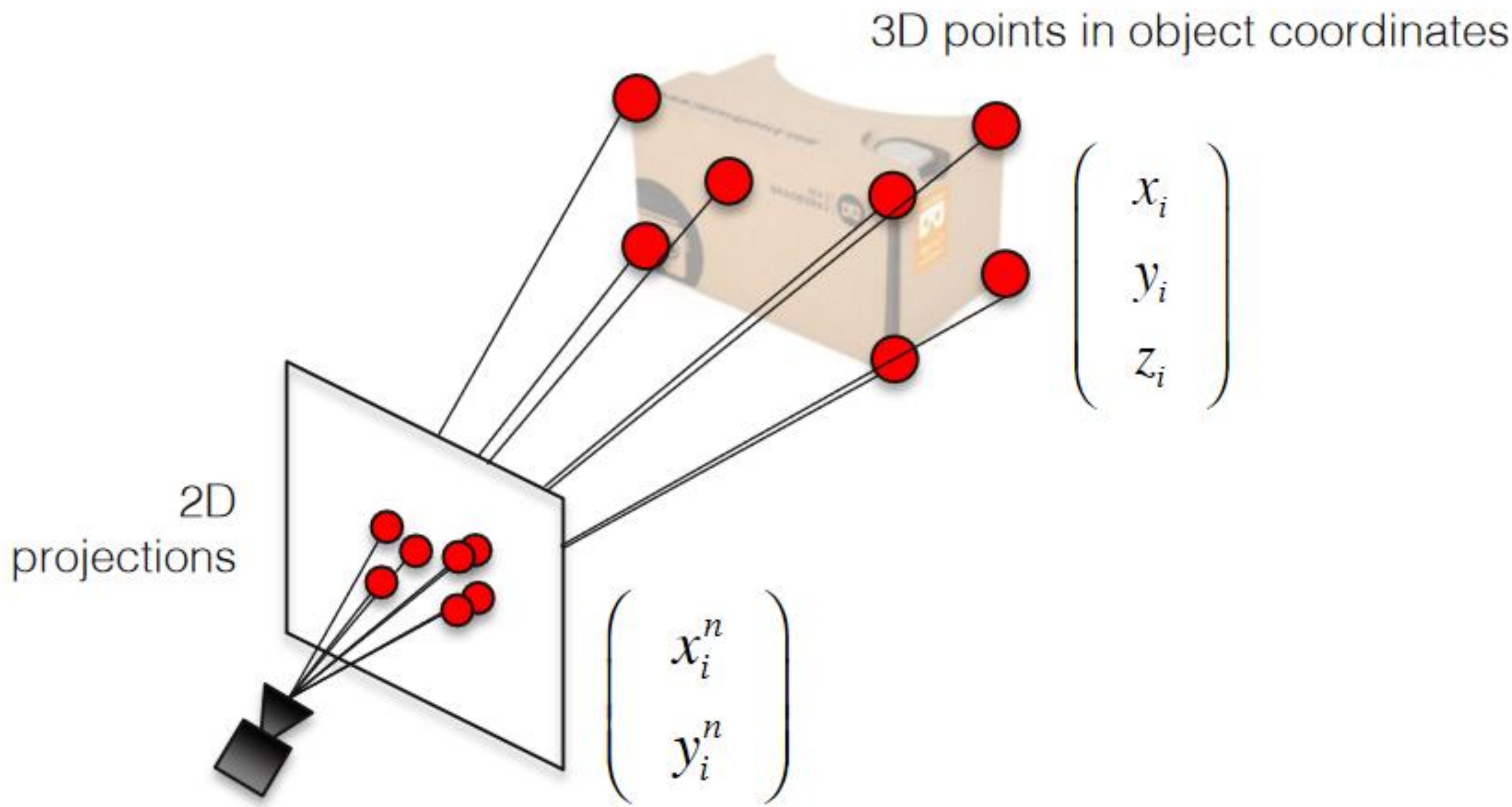
- 1cm
- 1m
- 1km
- We need some ground truth

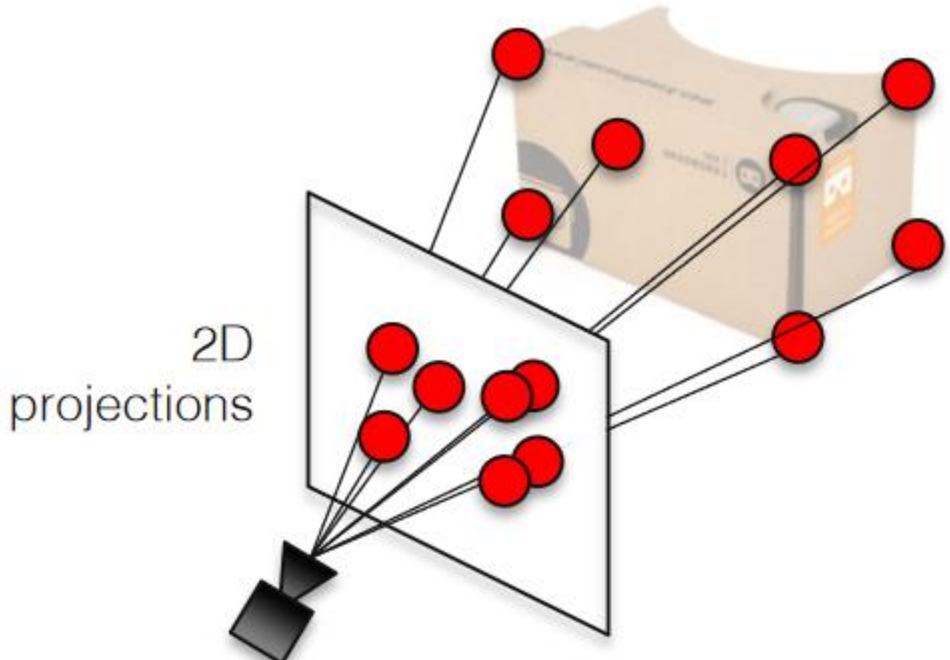
Moving on, now we look at 3D-2D

The Goal:

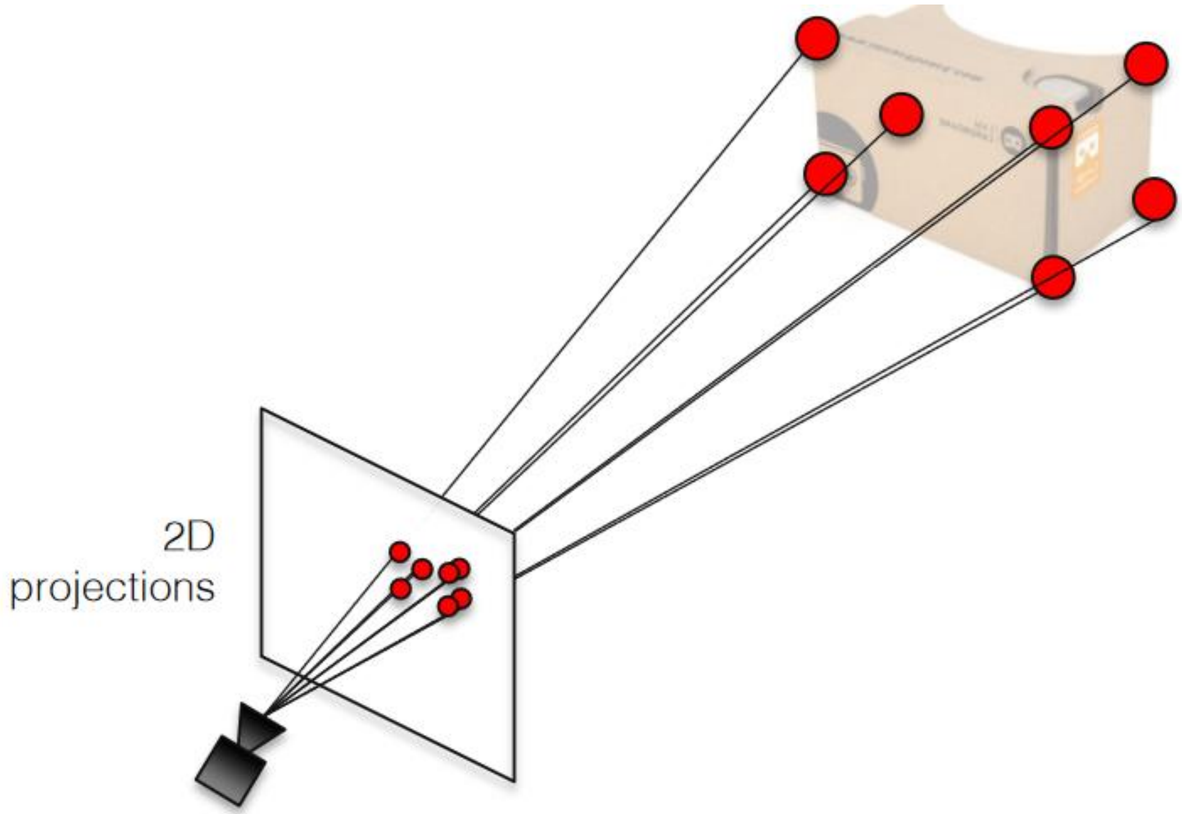
Estimate pose of the camera from 3D coordinates of object and 2D image coordinates.




3D points in object coordinates
$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$
2D projections
$\begin{pmatrix} x_i^n \\ y_i^n \end{pmatrix}$

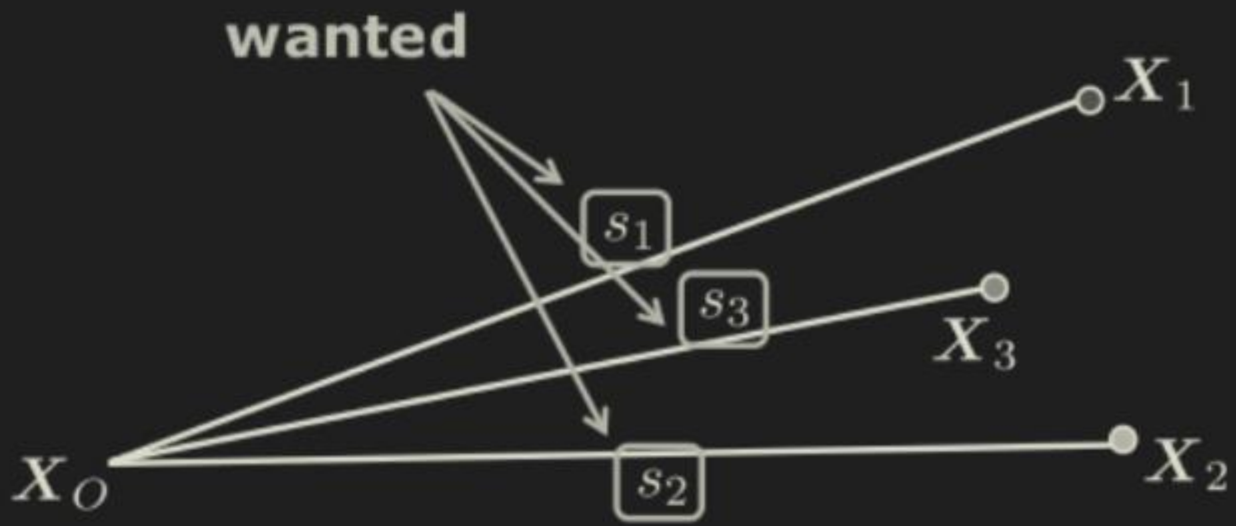


when object is closer,
projection is bigger



when object is father,
projection is smaller

Grunert's Solution



Get length of rays,

orientation

In triangle $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2$

$$s_1^2 + s_2^2 - 2\underbrace{\boxed{s_1}\boxed{s_2}}_{\text{wanted}}\cos\boxed{\gamma} = \underbrace{\boxed{c^2}}_{\text{known}}$$

In the end, you'll get a 4^{th} Degree Polynomial which can be solved.

But, multiple solutions are possible, so consider a 4^{th} point to confirm the right solution.

(1)

(2)

(3)

Colmap Video

2D-2D

Eight Point Algorithm
Triangulation
Stereo

2D-3D

DLT

PnP

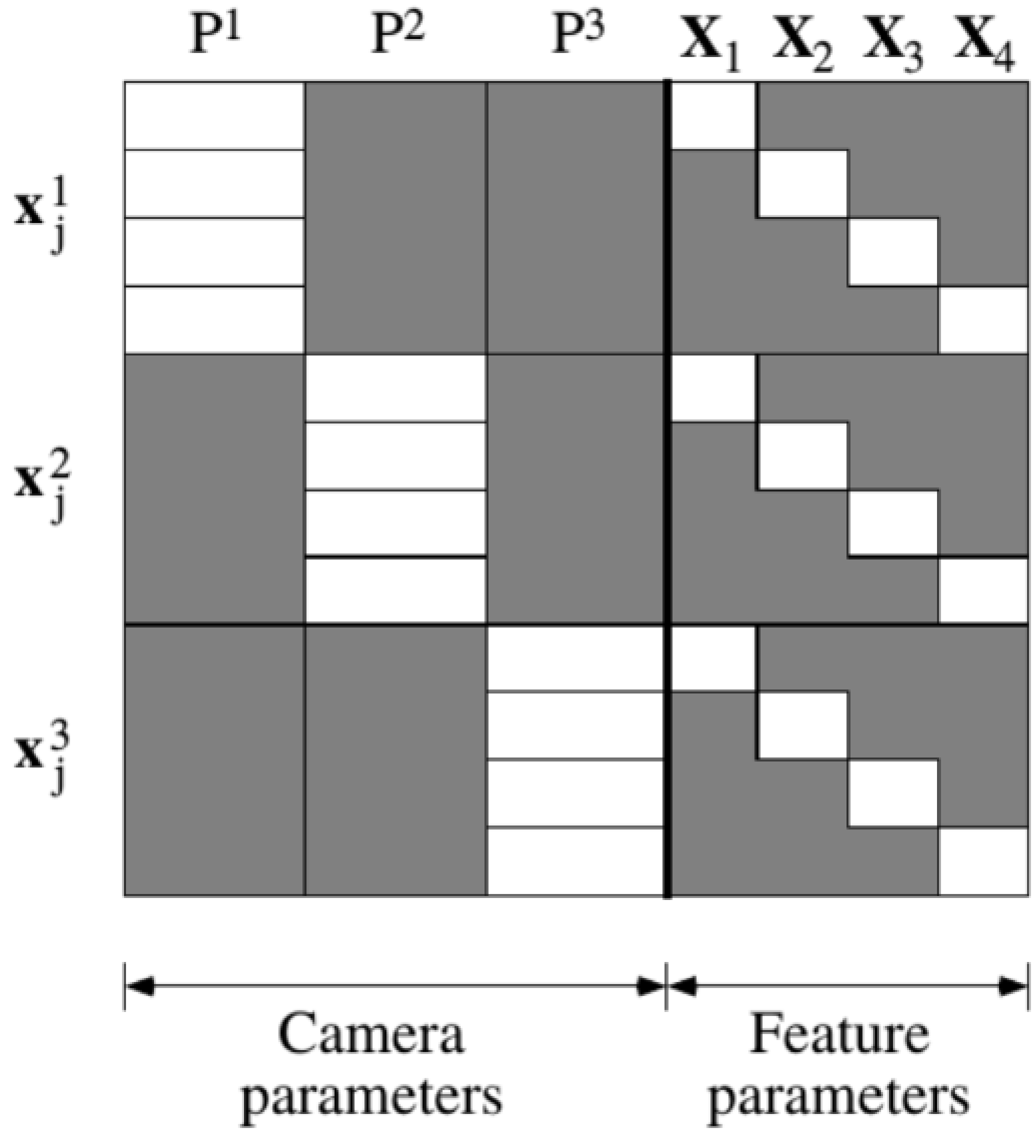
3D-3D

??

Bundle Adjustment

Given a set of images depicting a number of 3D points from different viewpoints, bundle adjustment can be defined as the problem of simultaneously refining the 3D coordinates describing the scene geometry, the parameters of the relative motion, and the optical characteristics of the camera(s) employed to acquire the images, according to an optimality criterion involving the corresponding image projections of all points.

$$arg \min_{X_j, P_i} \sum_{i=1}^{M} \sum_{j=1}^{N} \|x_{ij} = P_i X_j\|^2$$


P^1
P^2
P^3
$\mathbf{X}_1$
$\mathbf{X}_2$
$\mathbf{X}_3$
$\mathbf{X}_4$
$\mathbf{x}_j^1$
$\mathbf{x}_j^2$
$\mathbf{x}_j^3$
Camera parameters
Feature parameters

References

- https://www.uio.no/studier/emner/matnat/its/nedlagte-emner/UNIK4690/v16/forelesninger/lecture_7_2-triangulation.pdf
- Stachniss Lectures on Triangulation, PnP
- https://sites.google.com/site/jimdavidshome/research/finding-all-the-solutions-of-nonlinear-perspective-n-point-problem
- http://vr.cs.uiuc.edu/node292.html
- https://www.youtube.com/watch?v=N1aCvzFll6Q