CEE598 - Visual Sensing for Civil Infrastructure Eng. & Mgmt.

Session 14 – Affine Structure from Motion

Mani Golparvar-Fard
Department of Civil and Environmental Engineering
3129D, Newmark Civil Engineering Lab
e-mail: mgolpar@illinois.edu

Reminders

- Term Project (P2) is due Apr 4th!
  - Every group should ONLY prepare a 6-slide PPT:

Slide 1- Introduction to the Engineering problem your work is addressing;

Slide 2- Review of previous works, or the works you're following for your own implementation;

Slide 3- A summary of your technical solution;

Slide 4- Presenting experimental results and discuss validation approach (if you do not have experimental results by then, that's fine... prepare a detailed plan as to how you will implement and validate your algorithm; and

Slide 5- Plan and schedule of activities for the remainder of the project.

- Assignment "A3" will be out next Tuesday

State-of-the-Art

massachusetts institute of technology

MITnews

search

engineering science management architecture + planning humanities, arts, and social sciences campus video press

A camera that peers around corners

A new imaging system could use opaque walls, doors or floors as 'mirrors' to gather information about scenes outside its line of sight.

Larry Hardesty, MIT News Office

March 21, 2012

Share

In December, MIT Media Lab researchers caused a stir by releasing a slow-motion video of a burst of light traveling the length of a plastic bottle. But the experimental setup that enabled that video was designed for a much different application: a camera that can see around corners.

In a paper appearing this week in the journal *Nature Communications*, the

today's news

Water world

The 'weather in a tank' curriculum includes several rotating tank experiments, which are designed to help students understand how Earth's atmosphere and oceans work. An image from the dye stirring experiment is shown here.
Photo: John Marshall

'Weather in a tank' demonstration helps students grasp fluid dynamics.

multimedia

Video: How to see around corners
From *Nature*/YouTube

Video: Labcast #60: Cornar
From the Media Lab

Video: "Looking Around Corners: New Opportunities in Femto-Photography," invited talk

Video: Visualizing video at the speed of light — one trillion frames per second

related

CORNAR: Looking

http://web.mit.edu/newsoffice/2013/camera-sees-around-corners-0321.html#.T2nhvvKkxh4.twitter

State-of-the-Art

- http://video.mit.edu/watch/visualizing-video-at-the-speed-of-light-one-trillion-frames-per-second-9742/

- http://web.mit.edu/newsoffice/2013/camera-sees-around-corners-0321.html

Review – 3D Reconstruction

- 3D Reconstruction Using Structured Light

http://www.youtube.com/watch?v=t4TD9JgWTFs

- 3D reconstruction in Government Context

http://www.youtube.com/watch?v=vdt66tsPjCM&list=UU8vCnRtf8YXq05sXeAIZKSA&index=10&feature=plcp

http://www.youtube.com/watch?v=AQhyFnLg034&feature=autoplay&list=UU8vCnRtf8YXq05sXeAIZKSA&lf=plcp&playnext=3

- Early Structure from Motions

http://www.youtube.com/watch?v=oZ3YqZCYJT0

- Building Rome in a Day

http://grail.cs.washington.edu/projects/rome/

http://www.youtube.com/watch?v=i7ierVkXYa8
http://www.youtube.com/watch?v=SHa_LBIzDac

Outline

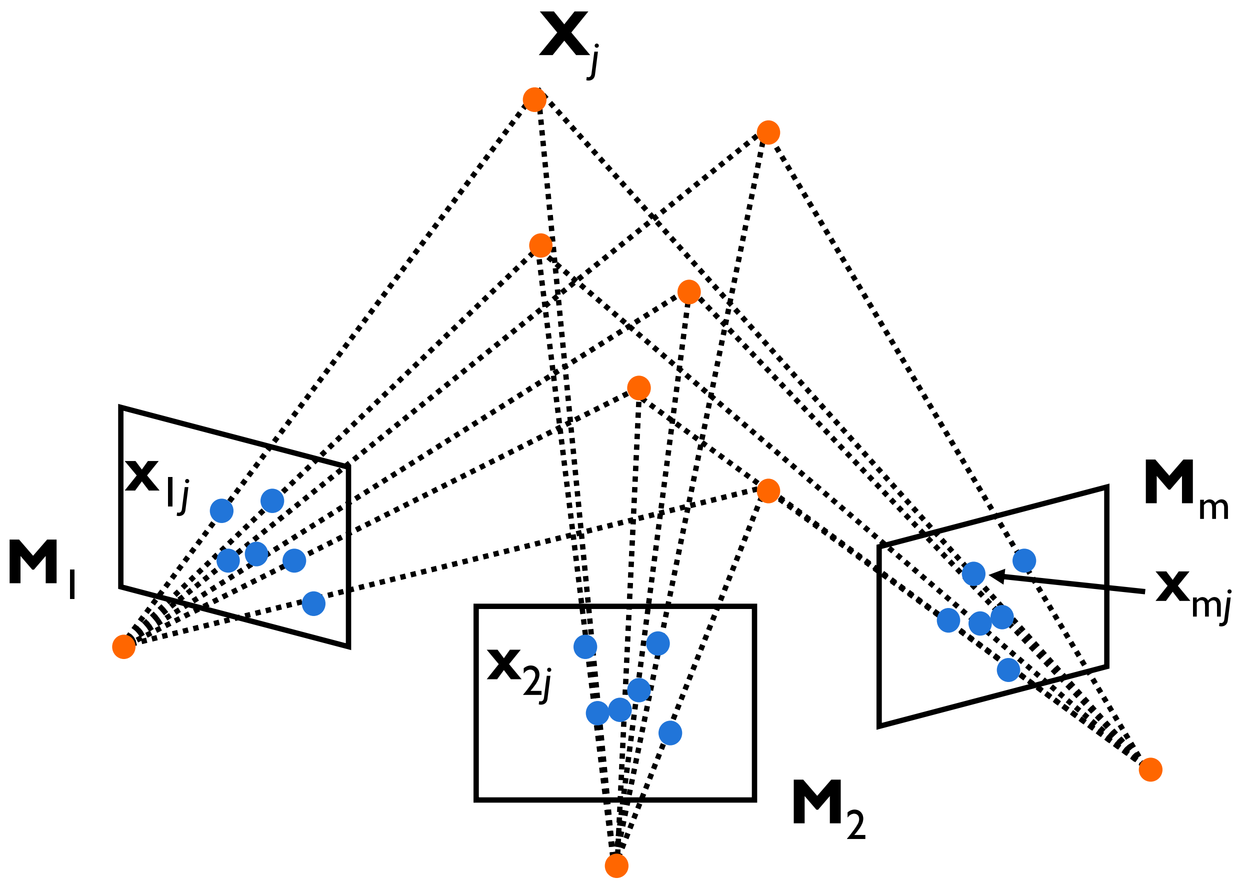
- Multiple view geometry
  Affine structure from Motion

- Affine structure from motion problem
- Algebraic methods
- Factorization methods

- Reading: [HZ] Chapters: 6,14,18
  [FP] Chapter: 12

Some slides of this lectures are courtesy of prof Savarese, prof. J. Ponce, prof FF Li, prof S. Lazebnik & prof. M. Hebert

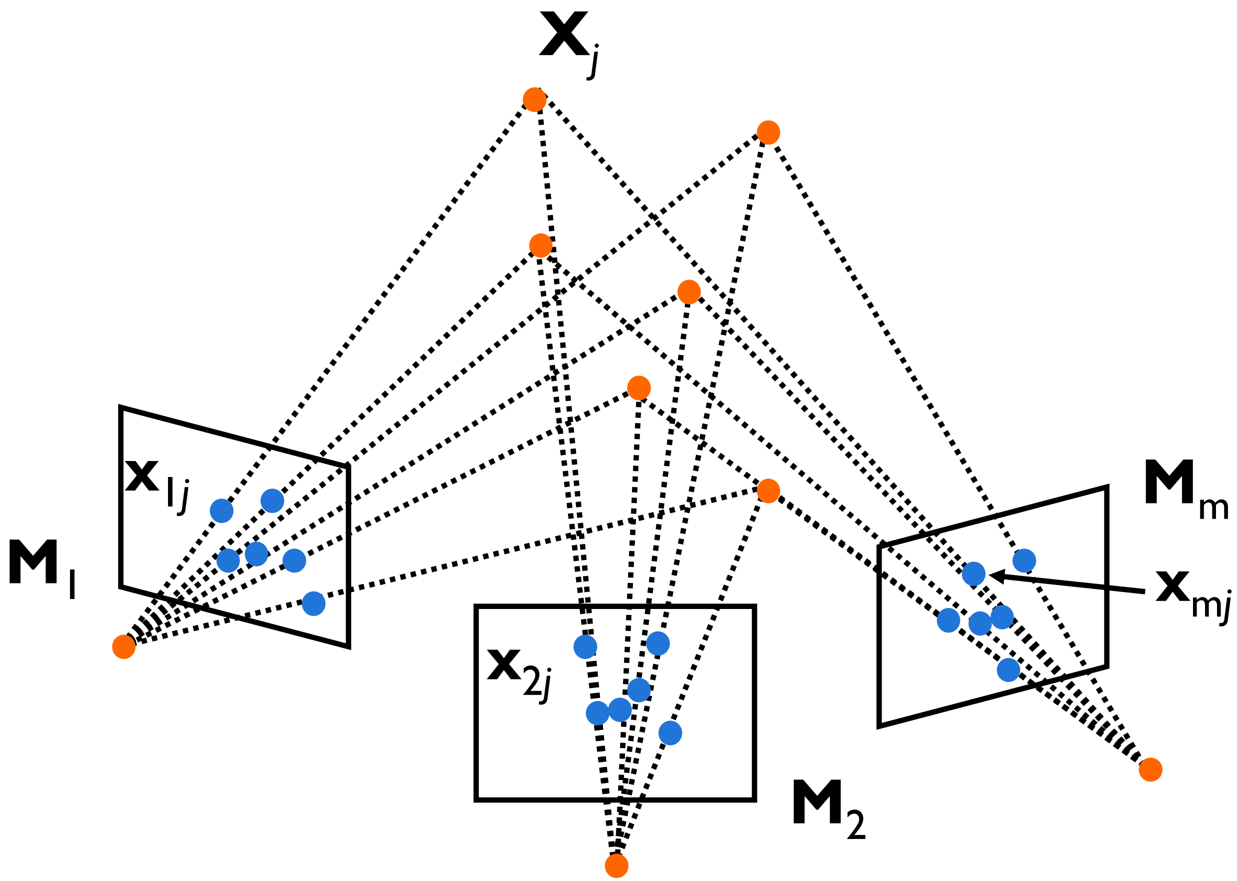
Structure from motion problem



Given *m* images of *n* fixed 3D points

- $\mathbf{x}_{ij} = \mathbf{M}_i \mathbf{X}_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n$

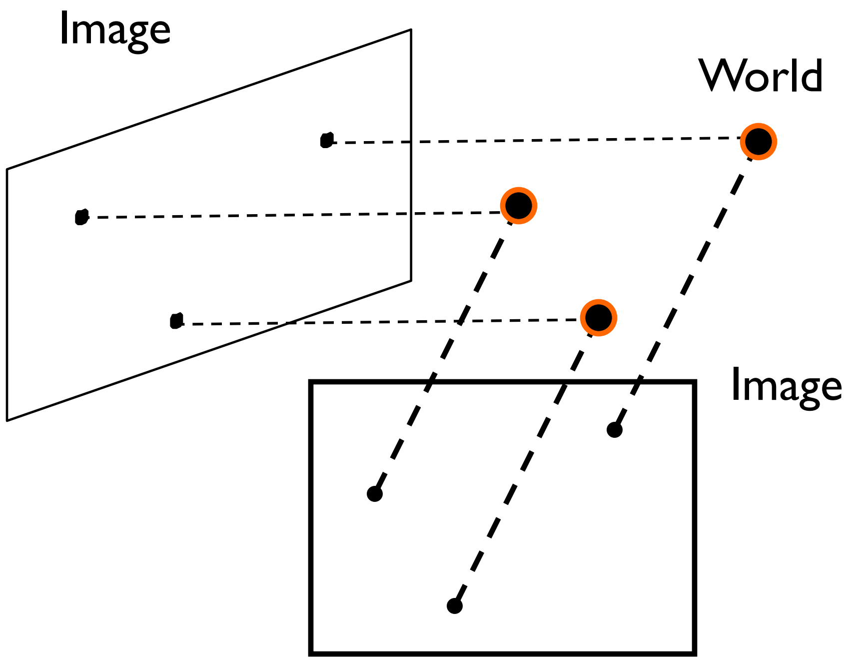
Structure from motion problem



From the m×n correspondences $\mathbf{x}_{ij}$, estimate:

• *m* projection matrices $\mathbf{M}_i$	motion
• *n* 3D points $\mathbf{X}_j$	structure

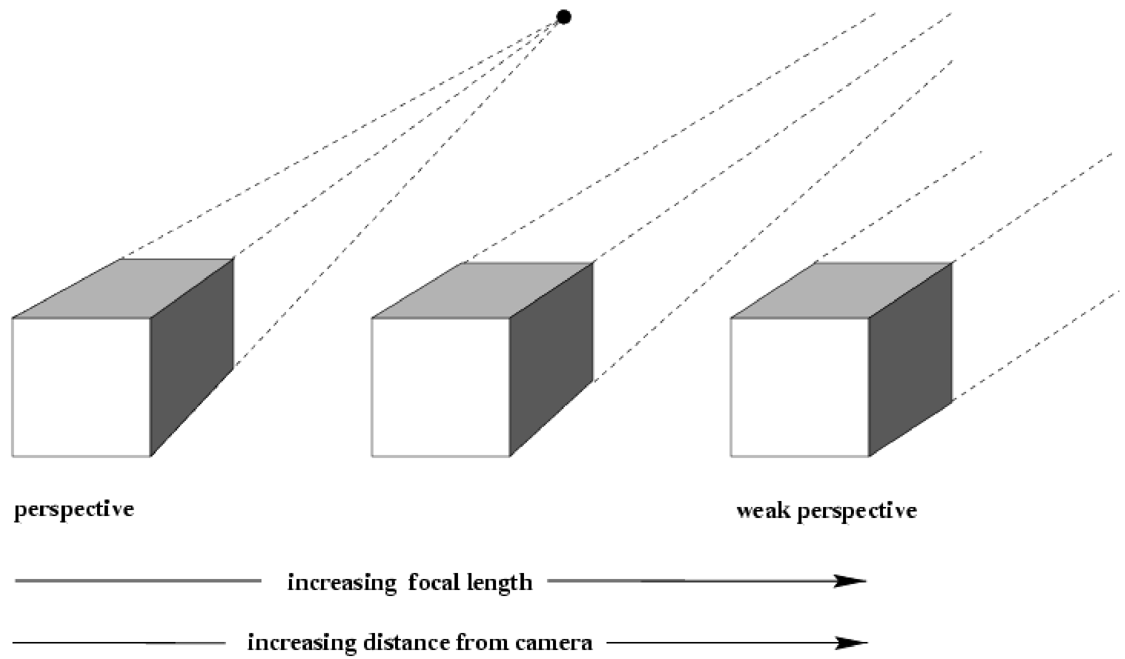
Affine structure from motion
(simpler problem)

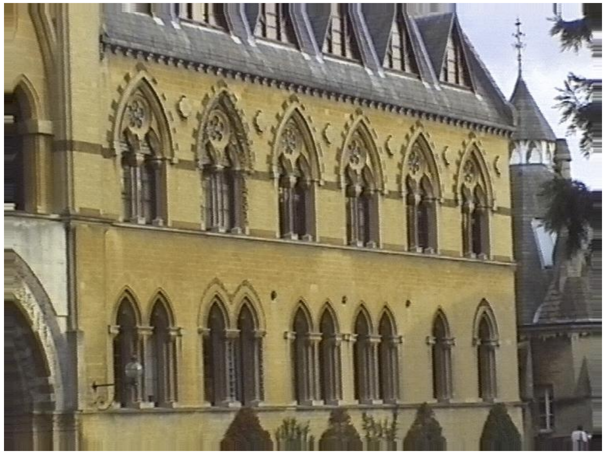


From the m×n correspondences $\mathbf{x}_{ij}$, estimate:

- *m* projection matrices $\mathbf{M}_i$ (affine cameras)
- *n* 3D points $\mathbf{X}_j$

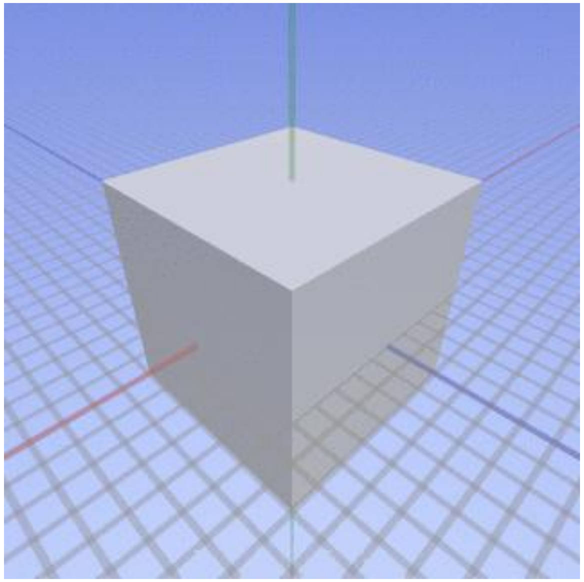
Affine Cameras





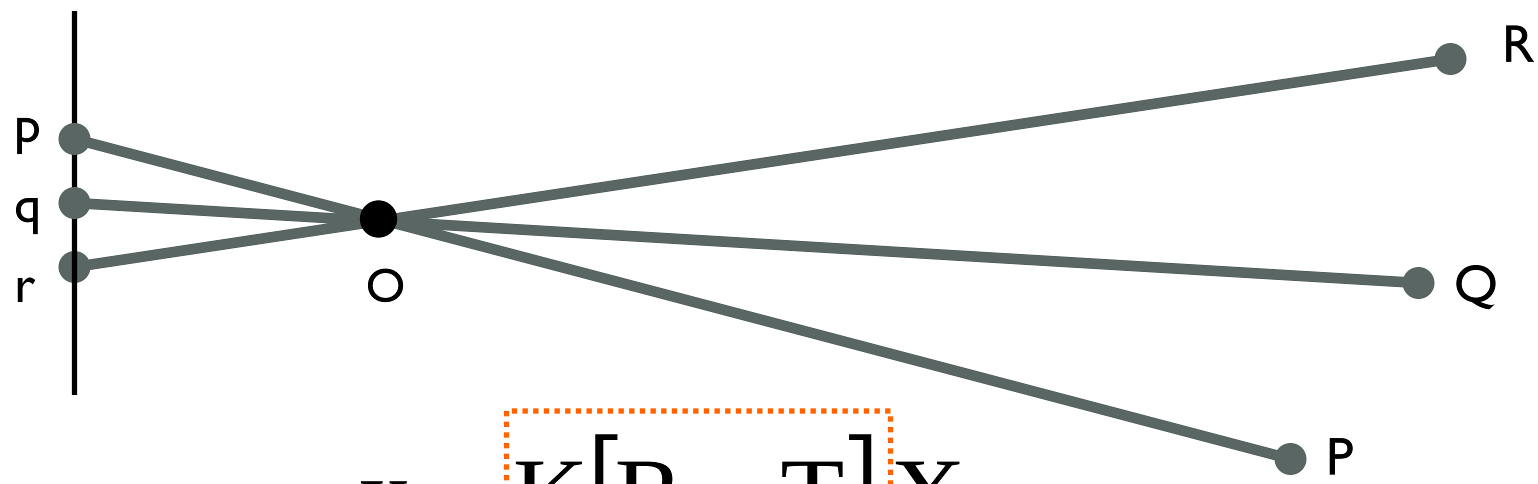
Changing Camera Focal Length

Increasing the focal length and distance of the camera in a perspective projection results in an approximation of orthographic projection



http://en.wikipedia.org/wiki/File:Orthographic_camera_distance_focal_length.gif

Finite cameras



$$x = \underbrace{K\begin{bmatrix} R & T \end{bmatrix}}_{M} X$$

Canonical perspective projection matrix

$$K = \begin{bmatrix} \alpha_x & s & x_o \\ 0 & \alpha_y & y_o \\ 0 & 0 & 1 \end{bmatrix}$$

$$M = K_{3x3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

Affine Homography (in 2D)

Affine homography (in 3D)

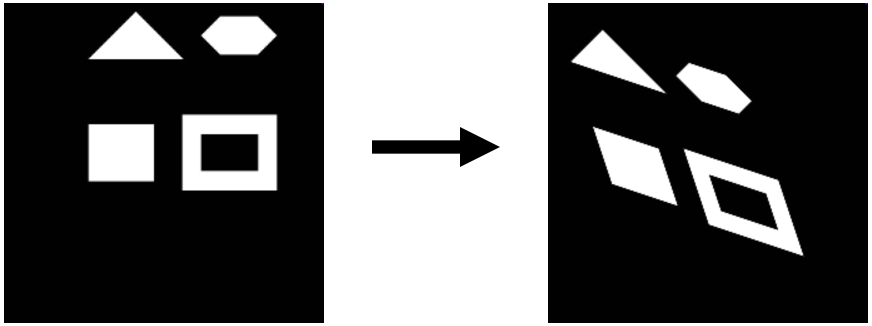
Transformation in 2D

Affinities:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_a \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

-Preserve:
- Parallel lines
- Ratio of areas
- Ratio of lengths on collinear lines
- others…

- How many DOF? 6



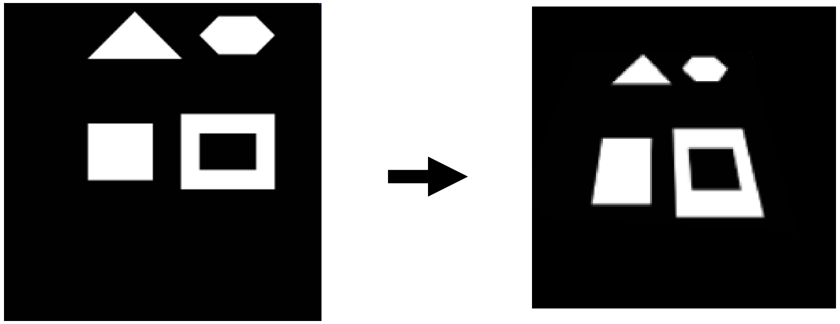
Transformation in 2D

Projective:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & t \\ v & b \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_p \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

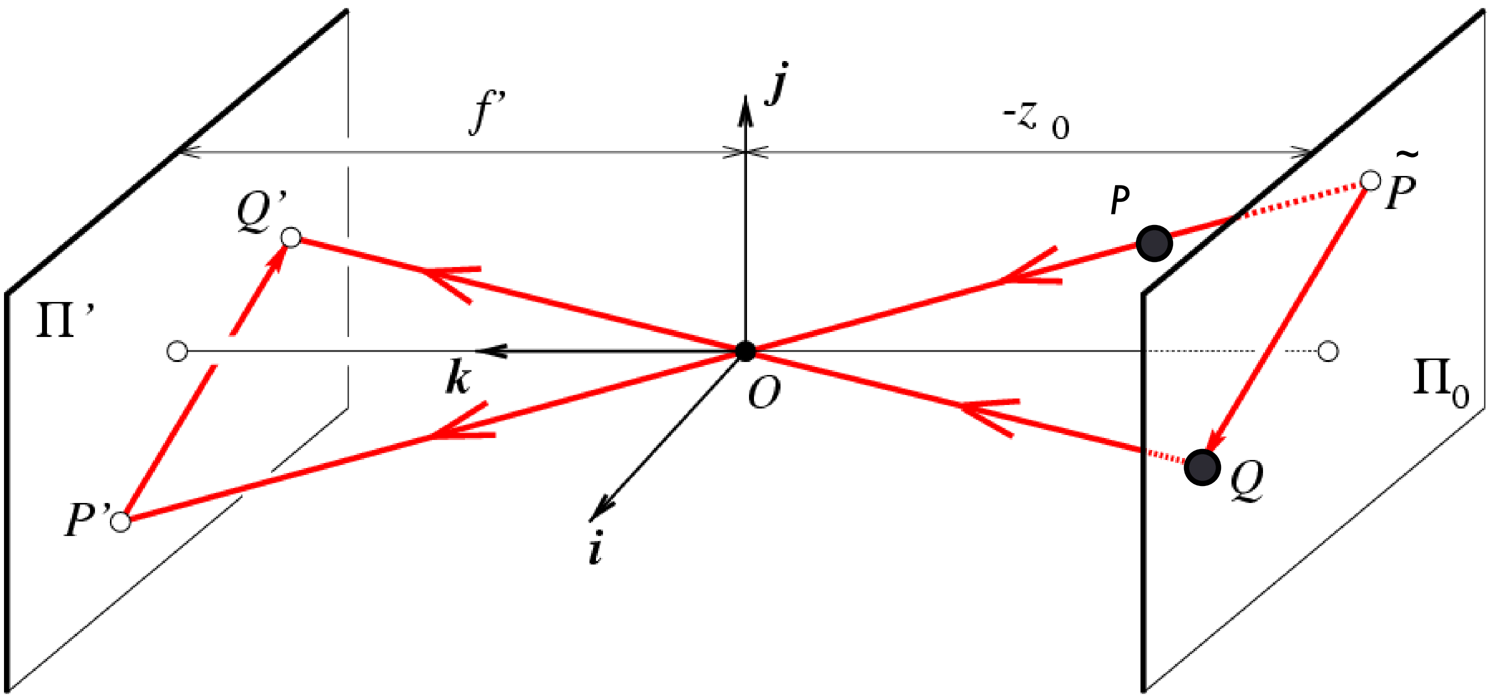
-Preserve:
- cross ratio of 4 collinear points
- collinearity
- and a few others…

- How many DOF? 8



Weak perspective projection

When the relative scene depth is small compared to its distance from the camera

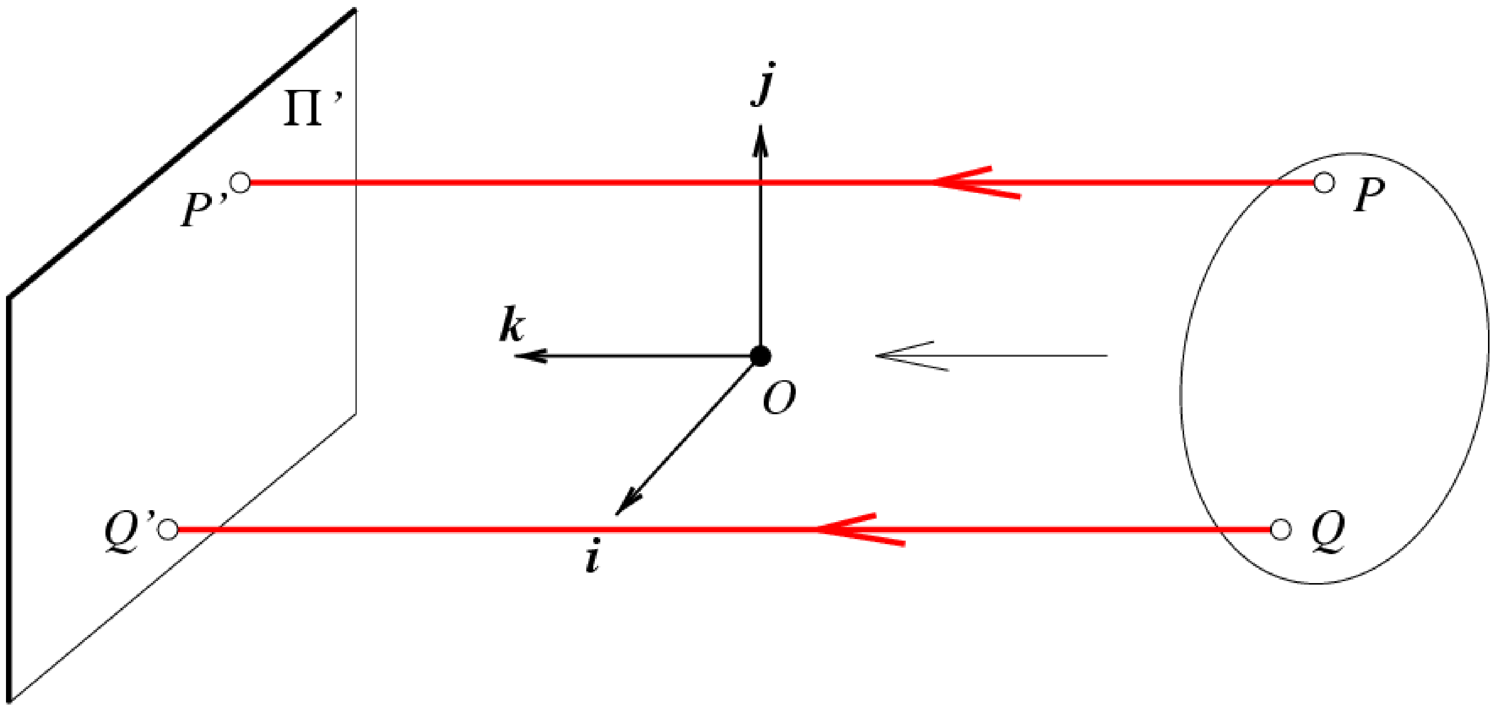


$$\begin{cases} x' = -mx \\ y' = -my \end{cases}$$

where $m = -\frac{f'}{z_0}$ is the magnification.

Orthographic (affine) projection

When the camera is at a (roughly constant) distance from the scene

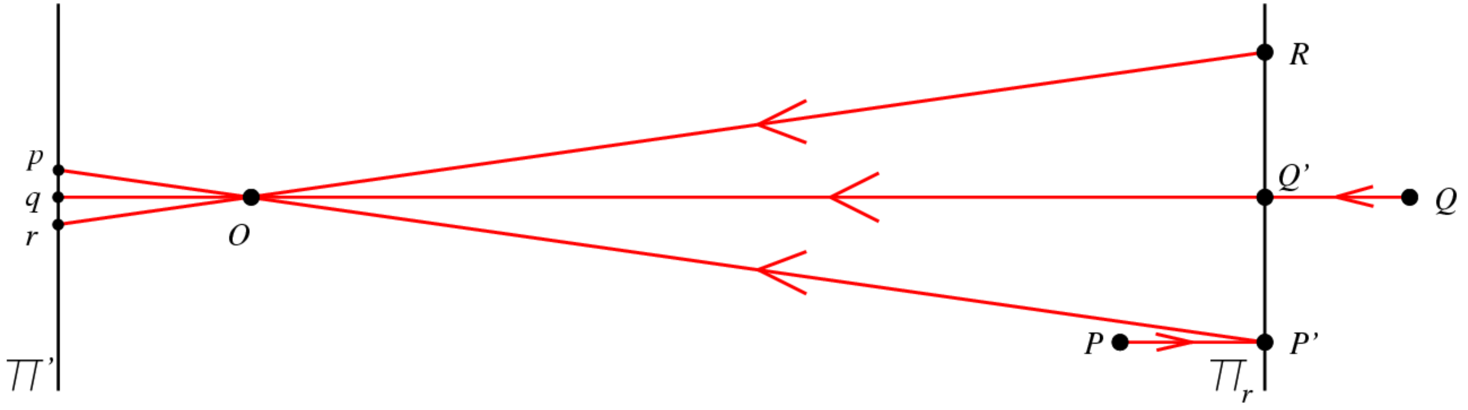


–Distance from center of projection to image plane is infinite

$$\begin{cases} x' = x \\ y' = y \end{cases}$$

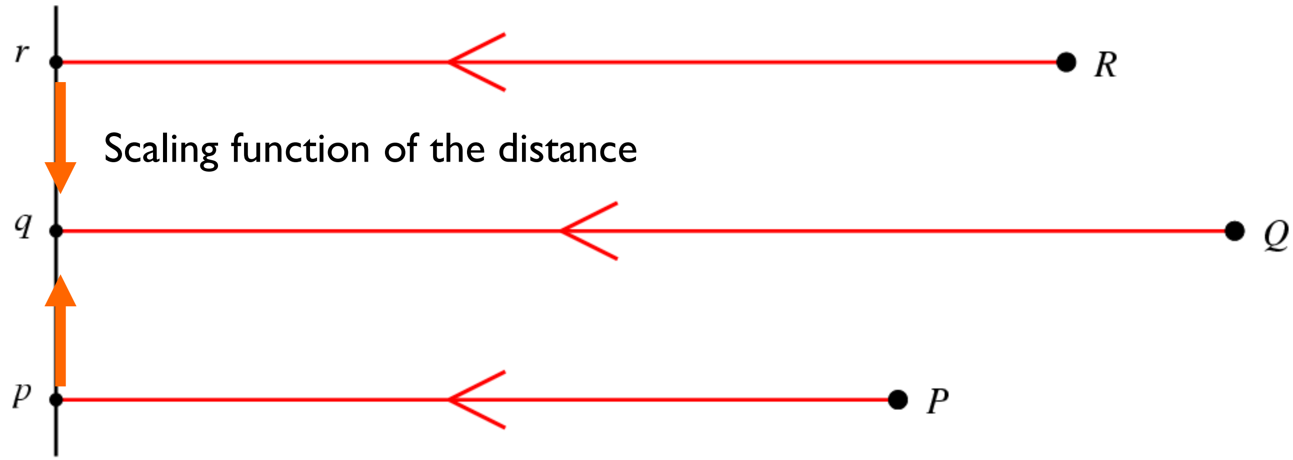
Weak-Perspective Projection

$$K = \begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



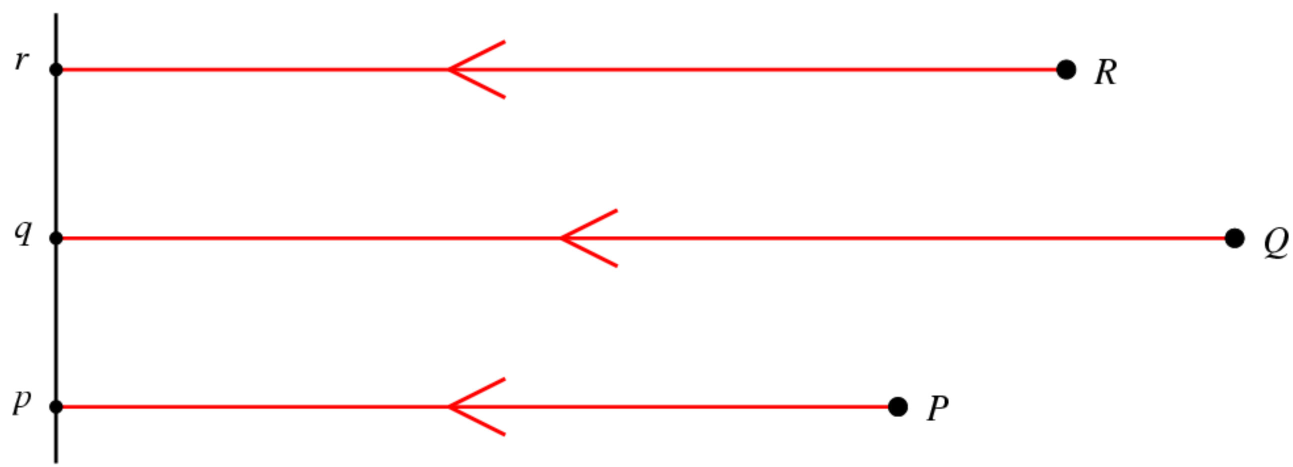
Weak-Perspective Projection

$$K = \begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Orthographic Projection

$$K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Affine cameras

$$x = K\begin{bmatrix} R & T \end{bmatrix} X$$

Projective case

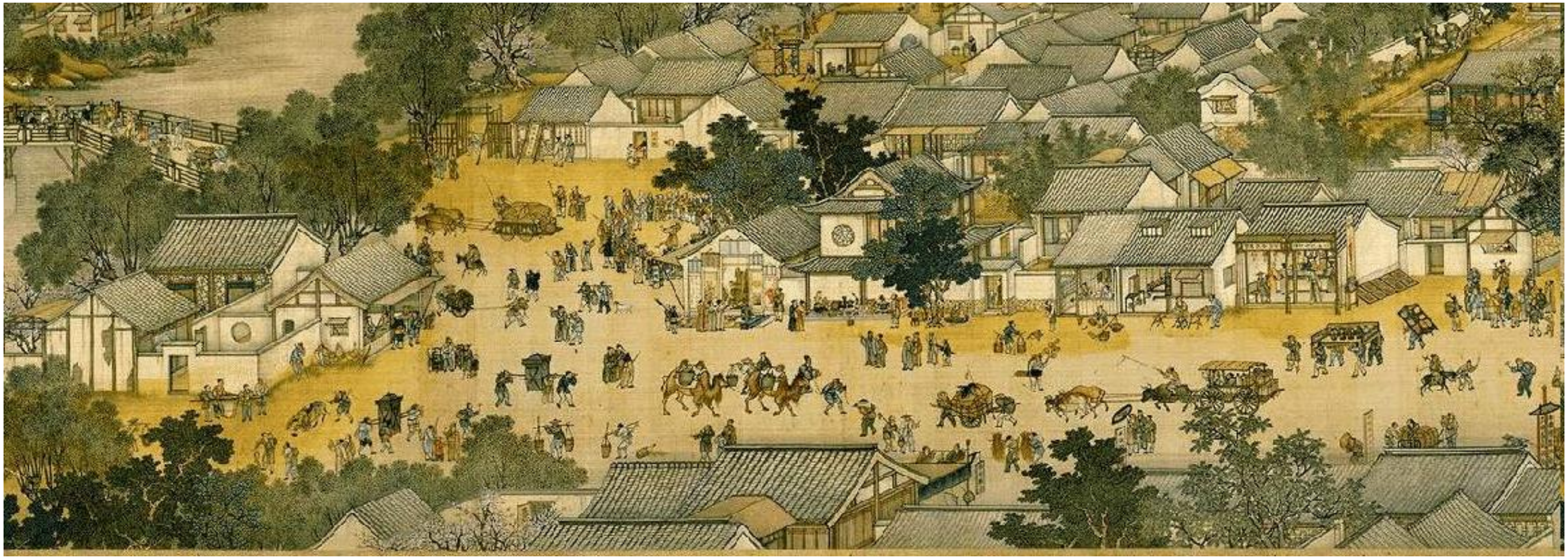
$$K = \begin{bmatrix} \alpha_x & s & x_o \\ 0 & \alpha_y & y_o \\ 0 & 0 & 1 \end{bmatrix} \qquad M = K \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

Affine case

$$K = \begin{bmatrix} \alpha_x & s & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad M = K \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

Parallel projection matrix
(points at infinity are mapped as points at infinity)

Weak perspective projection



Qingming Festival by the Riverside Zhang Zeduan ~900 AD

Affine cameras

$$\mathrm{x} = \mathrm{K}\begin{bmatrix}\mathrm{R} & \mathrm{T}\end{bmatrix}\mathrm{X}$$ [Homogeneous]

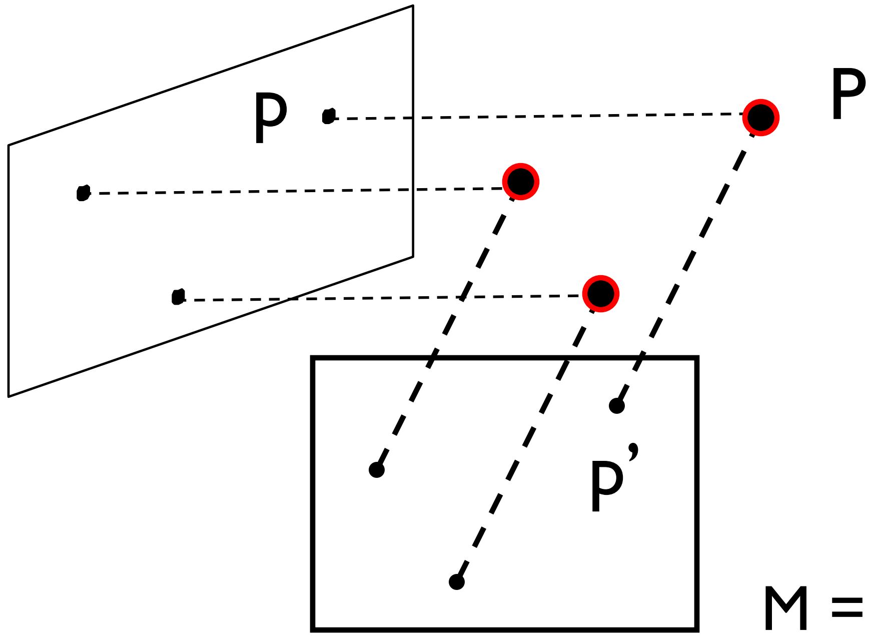
$$\mathrm{K} = \begin{bmatrix} \alpha_x & s & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathrm{M} = \mathrm{K}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \mathrm{R} & \mathrm{T} \\ 0 & 1 \end{bmatrix}$$

$$M = [3\times 3 \,\text{affine}]\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}[4\times 4\,\text{affine}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathbf{AX} + \mathbf{b} = M_{Euc}\begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix} = M_{Euc}\begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix};$$

$$\mathrm{M}_{\mathrm{Euc}} = \mathrm{M} = \begin{bmatrix}\mathrm{A} & \mathrm{b}\end{bmatrix}$$ [non-homogeneous image coordinates]

Affine cameras



M = camera matrix

To recap:

from now on we define M as the camera matrix for the affine case

$$\mathbf{p} = \begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{AP} + \mathbf{b} = M \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix}; \qquad \mathrm{M} = \begin{bmatrix} \mathrm{A} & \mathrm{b} \end{bmatrix}$$

The Affine Structure-from-Motion Problem

Given m images of n fixed points P_j (=X_i) we can write

$$\boldsymbol{p}_{ij} = \mathcal{M}_i \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix} = \mathcal{A}_i \boldsymbol{P}_j + \boldsymbol{b}_i \quad \text{for} \quad i = 1, \ldots, m \quad \text{and} \quad j = 1, \ldots, n.$$

Problem: estimate the m 2×4 matrices M_i and the n positions P_j from the m×n correspondences $\mathbf{p}_{ij}$.

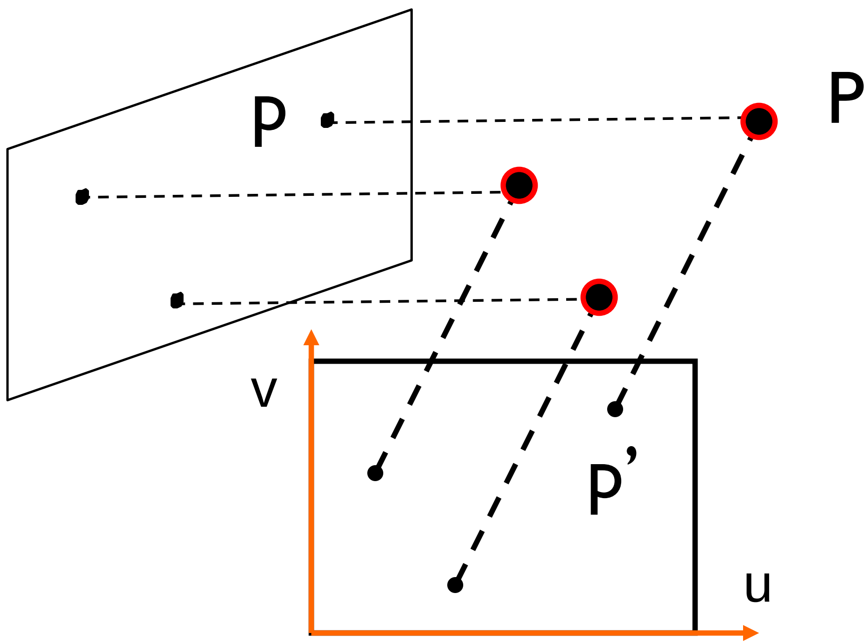
How many equations and how many unknown?

2m × n equations in 8m+3n unknowns

Two approaches:

- Algebraic approach (affine epipolar geometry; estimate F; cameras; points)
- Factorization method

Algebraic analysis (2-view case)



Homogeneous system

$$\begin{cases} \boldsymbol{p} = \mathcal{A}\boldsymbol{P} + \boldsymbol{b} \\ \boldsymbol{p}' = \mathcal{A}'\boldsymbol{P} + \boldsymbol{b}' \end{cases} \Rightarrow \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} \begin{pmatrix} \boldsymbol{P} \\ -1 \end{pmatrix} = \boldsymbol{0}$$

Full rank matrix; dim = ? 4x4

$$\Rightarrow \text{Det} \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} = 0 \Rightarrow \alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0$$

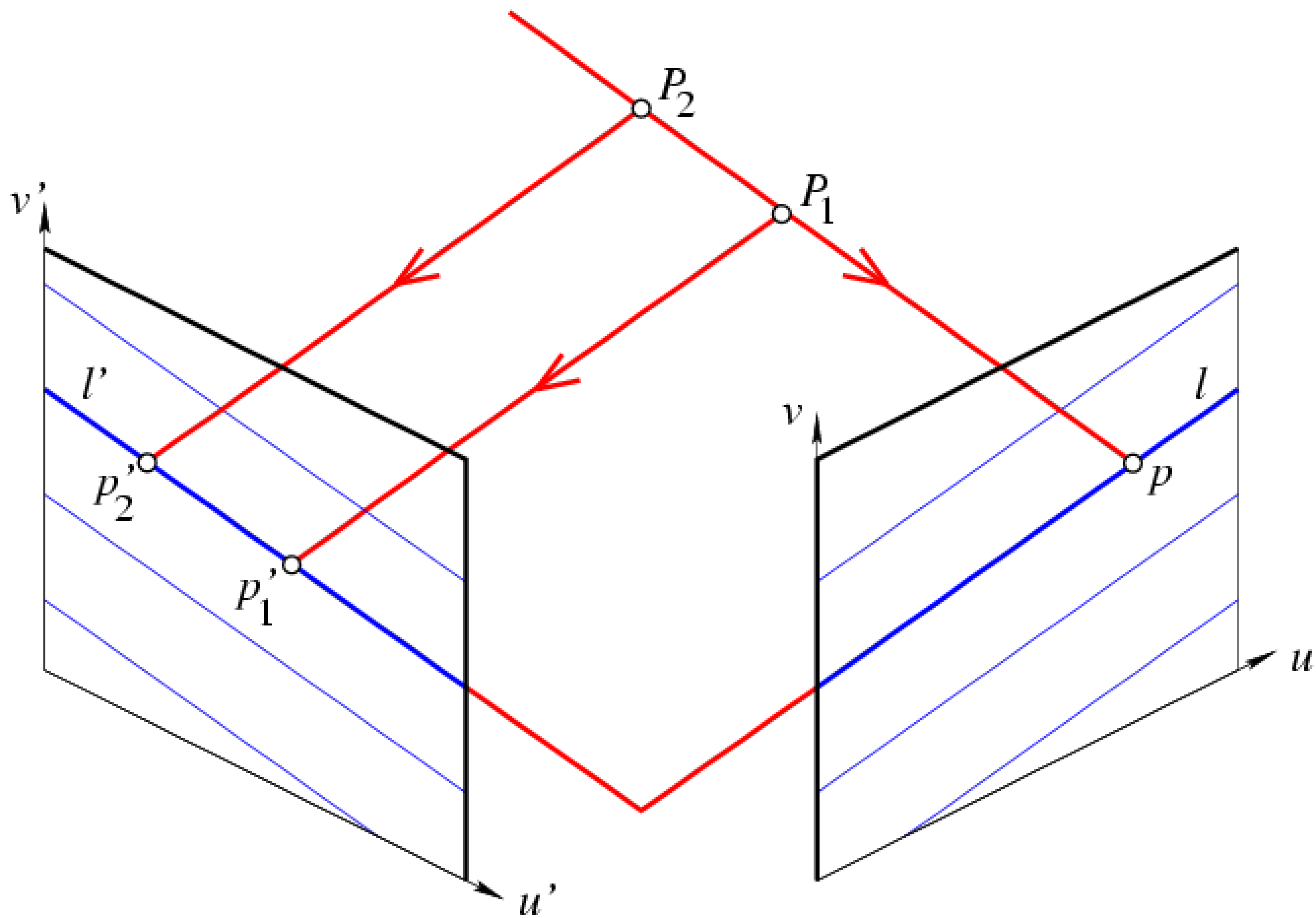
Algebraic analysis (2-view case)

$$\alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0$$

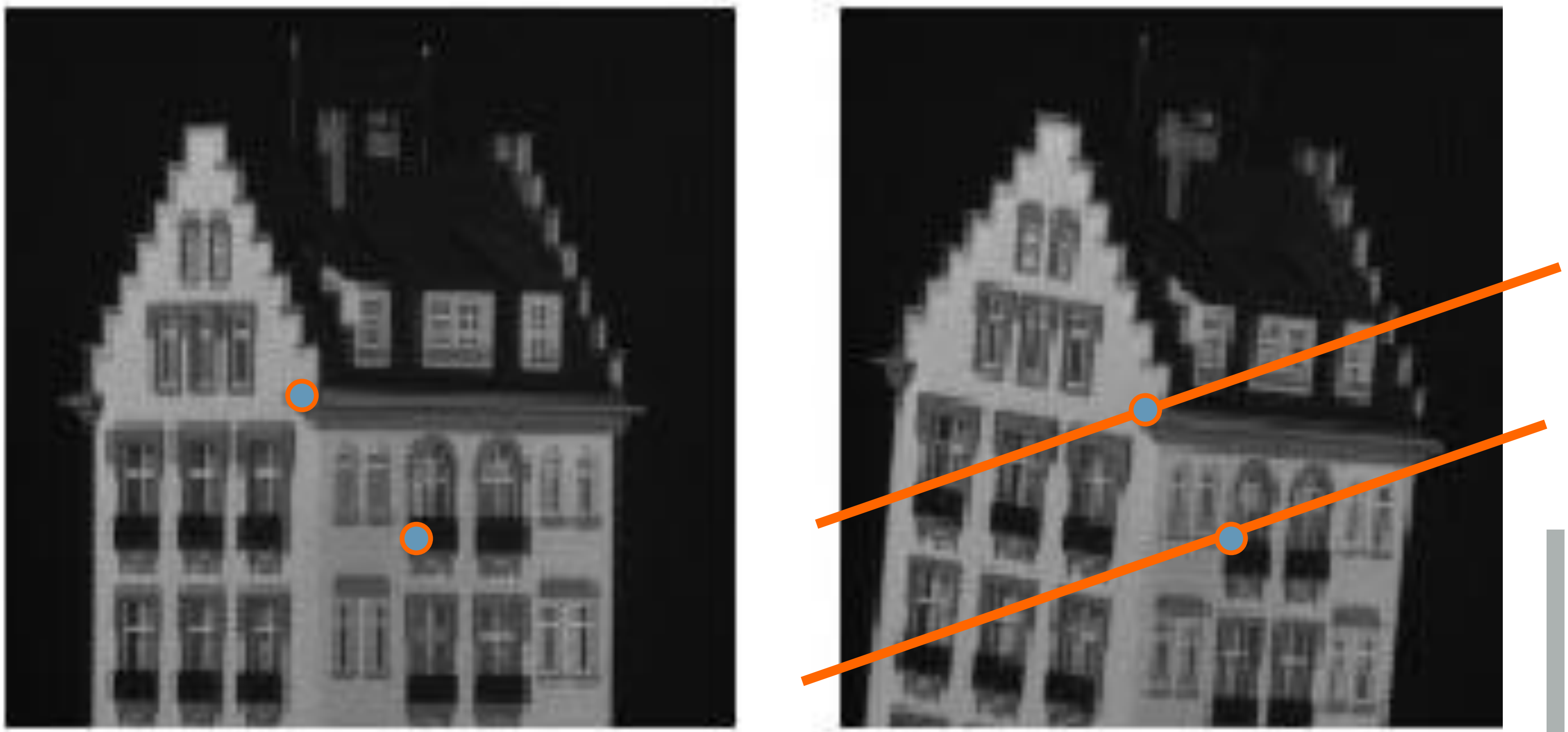
$$(u, v, 1)\mathcal{F}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \quad \text{where} \quad \mathcal{F} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & \alpha \\ 0 & 0 & \beta \\ \alpha' & \beta' & \delta \end{pmatrix}$$

The Affine Fundamental Matrix!

Affine Epipolar Geometry



Note: the epipolar lines are parallel.



Estimating F

$$\alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0$$

- Measurements: u, u', v, v'
- From at least 4 correspondences, we obtain a linear system on the unknown alpha, beta, etc…

$$\begin{bmatrix} u_1' & v_1' & u_1 & v_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_n' & v_n' & u_n & v_n & 1 \end{bmatrix} \mathbf{f} = 0$$

- Computed by least square and by enforcing |f|=1
- SVD

Estimating projection matrices from epipolar constraints

If M_i and P_i are solutions,
then M_i' and P_i' are also solutions,

where

$$\mathcal{M}'_i = \mathcal{M}_i \mathcal{Q} \quad \text{and} \quad \begin{pmatrix} \boldsymbol{P}'_j \\ 1 \end{pmatrix} = \mathcal{Q}^{-1} \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix}$$

and

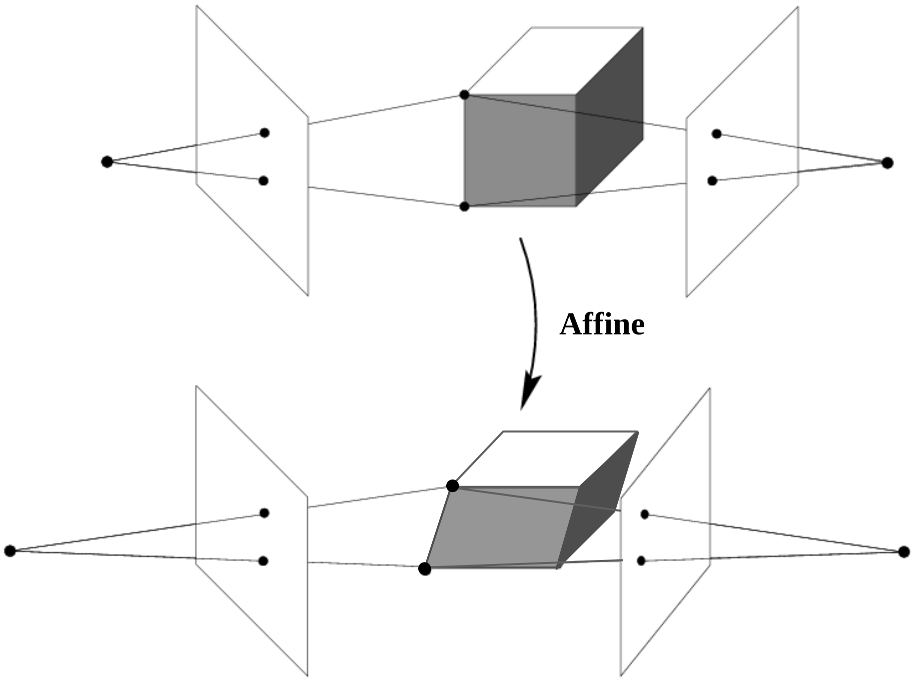
$$\mathcal{Q} = \begin{pmatrix} \mathcal{C} & \boldsymbol{d} \\ \boldsymbol{0}^T & 1 \end{pmatrix}$$

Q is an affine transformation.

Proof:

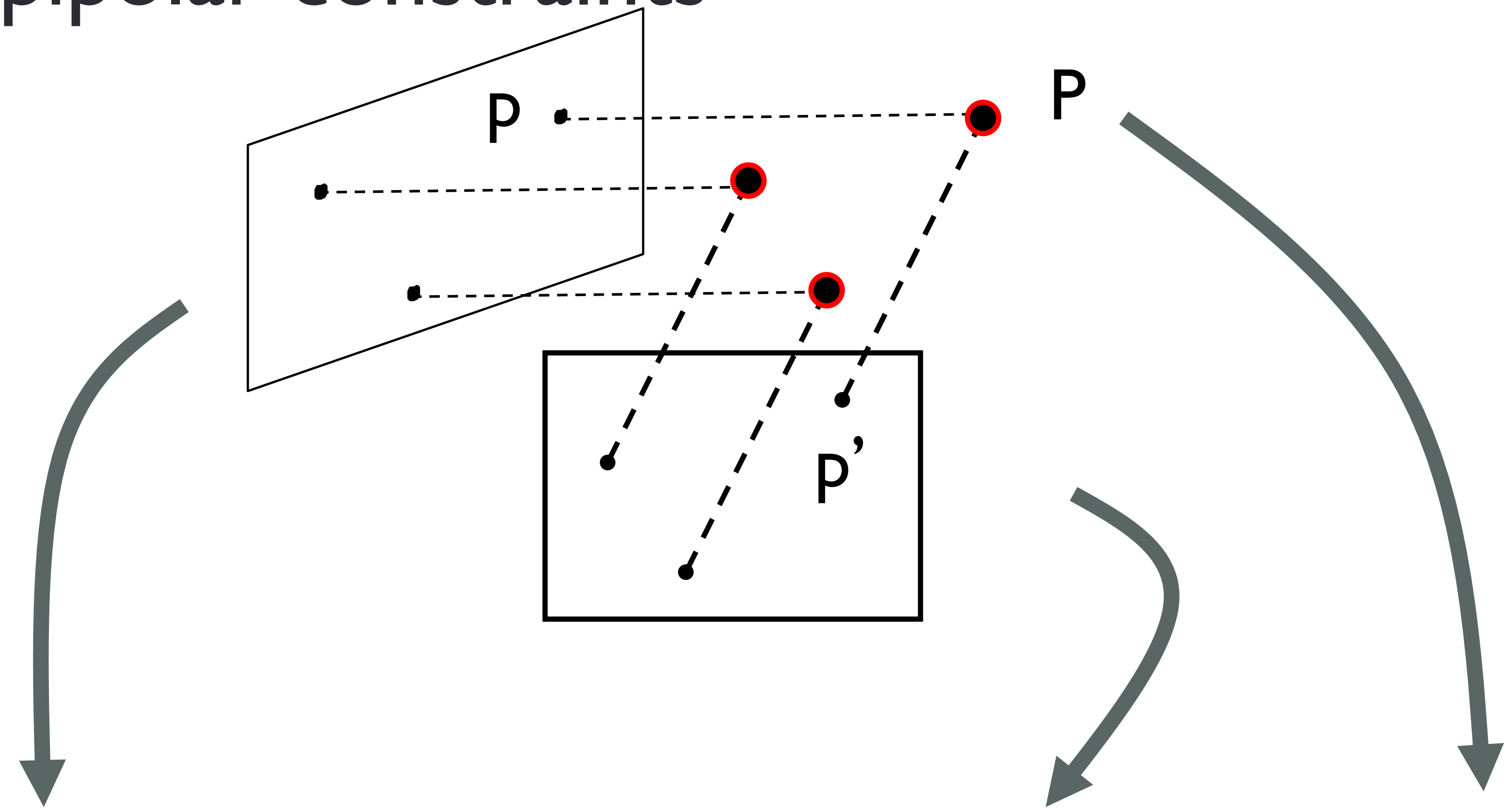
$$\boldsymbol{p}_{ij} = \mathcal{M}_i \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix} = (\mathcal{M}_i \mathcal{Q}) \left(\mathcal{Q}^{-1} \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix}\right) = \mathcal{M}'_i \begin{pmatrix} \boldsymbol{P}'_j \\ 1 \end{pmatrix} \quad \blacksquare$$

Affine ambiguity



$$\mathbf{x} = \mathbf{P}\mathbf{X} = \left(\mathbf{P}\mathbf{Q}_A^{-1}\right)\left(\mathbf{Q}_A\mathbf{X}\right)$$

Estimating projection matrices from epipolar constraints



$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

Estimating projection matrices from epipolar constraints

$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

Choose Q such that…

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \qquad \tilde{\boldsymbol{P}}$$

Canonical affine cameras

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad \tilde{A}' = \begin{bmatrix} 0 & 0 & 1 \\ a & b & c \end{bmatrix}$$

$$\tilde{b} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T \qquad \tilde{b}' = \begin{bmatrix} 0 & d \end{bmatrix}^T$$

Function of the parameters of F

Estimating projection matrices from epipolar constraints

$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

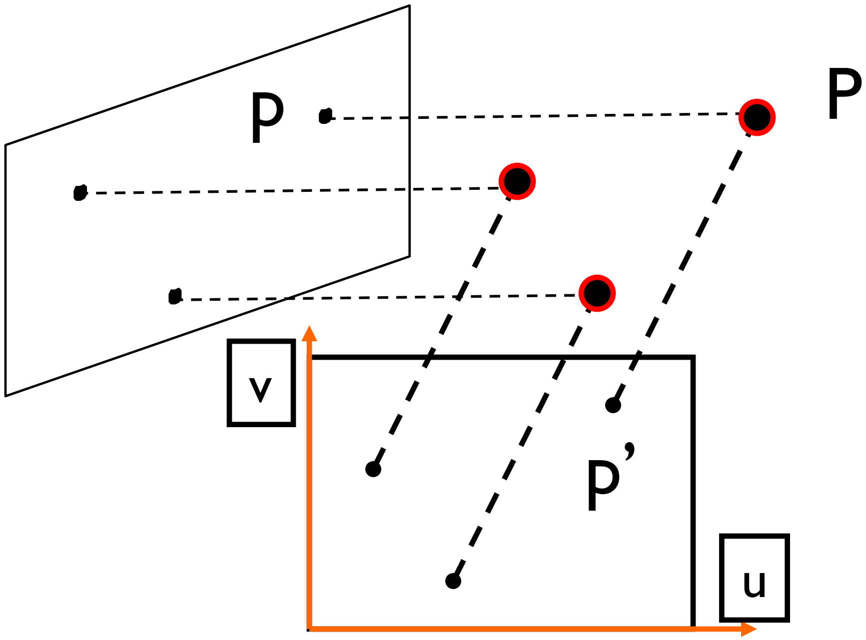
$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

Choose Q such that…

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \qquad \tilde{\boldsymbol{P}}$$

By re-enforcing the epipolar constraint, we can compute a, b, c, d directly from the measurements

Reminder: Epipolar constraint



Homogeneous system

$$\begin{cases} \boldsymbol{p} = \mathcal{A}\boldsymbol{P} + \boldsymbol{b} \\ \boldsymbol{p}' = \mathcal{A}'\boldsymbol{P} + \boldsymbol{b}' \end{cases} \Rightarrow \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} \begin{pmatrix} \boldsymbol{P} \\ -1 \end{pmatrix} = \mathbf{0}$$

$$\Rightarrow \boxed{\text{Det} \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} = 0} \Rightarrow \alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0$$

Estimating projection matrices from epipolar constraints

$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

Choose Q such that…

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \qquad \tilde{\boldsymbol{P}}$$

A b

Re-enforce the Epipolar constraint

$$\mathrm{Det}\begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} = 0 \rightarrow \mathrm{Det}\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v' - d \end{pmatrix} = 0$$

Estimating projection matrices from epipolar constraints

$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

Choose Q such that…

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \qquad \tilde{\boldsymbol{P}}$$

A b

$$\text{Det}\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v' - d \end{pmatrix} = au - bv + cu' + v' - d = 0$$

Estimating projection matrices from epipolar constraints

$$\text{Det}\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v'-d \end{pmatrix} = au - bv + cu' + v' - d = 0$$

- Linear relationship between measurements and unknown

  Unknown: a, b, c, d
  Measurements: u, u', v, v'

- From at least 4 correspondences, we can solve this linear system and compute a, b, c, d (via least square)

- The cameras can be computed

- How about the structure?

Estimating the structure from epipolar constraints

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \quad \tilde{\boldsymbol{P}}$$

A b

$$\begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} \begin{pmatrix} \boldsymbol{P} \\ -1 \end{pmatrix} = \boldsymbol{0} \quad \Rightarrow$$

$$\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v' - d \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{P}} \\ -1 \end{pmatrix} = 0 \quad \Rightarrow \quad \boxed{\tilde{\boldsymbol{P}} = \begin{pmatrix} u \\ v \\ u' \end{pmatrix}}$$

Can be solved by least square again

A factorization method – Tomasi & Kanade algorithm

C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *IJCV*, 9(2):137-154, November 1992.

- Centering the data
- Factorization

A factorization method - factorization

- Centering: subtract the centroid of the image points

$$\hat{\mathbf{x}}_{ij} = \mathbf{x}_{ij} - \frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_{ik}$$

A factorization method - factorization

- Centering: subtract the centroid of the image points

$$\hat{\mathbf{x}}_{ij} = \mathbf{x}_{ij} - \frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_{ik} = \mathbf{A}_i\mathbf{X}_j + \mathbf{b}_i - \frac{1}{n}\sum_{k=1}^{n}\left(\mathbf{A}_i\mathbf{X}_k + \mathbf{b}_i\right)$$

A factorization method - factorization

- Centering: subtract the centroid of the image points

$$\hat{\mathbf{x}}_{ij} = \mathbf{x}_{ij} - \frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_{ik} = \mathbf{A}_i\mathbf{X}_j + \mathbf{b}_i - \frac{1}{n}\sum_{k=1}^{n}\left(\mathbf{A}_i\mathbf{X}_k + \mathbf{b}_i\right)$$

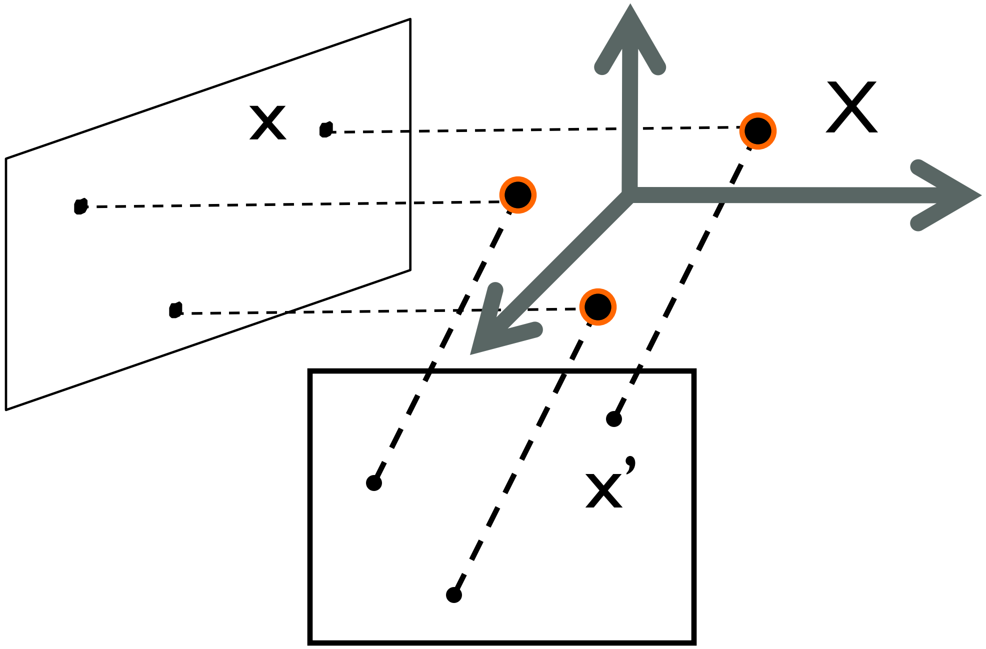
$$= \mathbf{A}_i\left(\mathbf{X}_j - \boxed{\frac{1}{n}\sum_{k=1}^{n}\mathbf{X}_k}\right)$$

Assume that the origin of the world coordinate system is at the centroid of the 3D points

After centering, each normalized point $\mathbf{x}_{ij}$ is related to the 3D point $\mathbf{X}_i$ by

$$\hat{\mathbf{x}}_{ij} = \mathbf{A}_i\mathbf{X}_j$$

A factorization method - Centering the data



$$\hat{\mathbf{x}}_{ij} = \mathbf{A}_i \mathbf{X}_j$$

A factorization method - factorization

- Let's create a $2m \times n$ data (measurement) matrix:

$$\mathbf{D} = \begin{bmatrix} \hat{\mathbf{x}}_{11} & \hat{\mathbf{x}}_{12} & \cdots & \hat{\mathbf{x}}_{1n} \\ \hat{\mathbf{x}}_{21} & \hat{\mathbf{x}}_{22} & \cdots & \hat{\mathbf{x}}_{2n} \\ & & \ddots & \\ \hat{\mathbf{x}}_{m1} & \hat{\mathbf{x}}_{m2} & \cdots & \hat{\mathbf{x}}_{mn} \end{bmatrix}$$

cameras ($2m$)

points (n)

A factorization method - factorization

- Let's create a 2m × n data (measurement) matrix:

$$\mathbf{D} = \begin{bmatrix} \hat{\mathbf{x}}_{11} & \hat{\mathbf{x}}_{12} & \cdots & \hat{\mathbf{x}}_{1n} \\ \hat{\mathbf{x}}_{21} & \hat{\mathbf{x}}_{22} & \cdots & \hat{\mathbf{x}}_{2n} \\ & & \ddots & \\ \hat{\mathbf{x}}_{m1} & \hat{\mathbf{x}}_{m2} & \cdots & \hat{\mathbf{x}}_{mn} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_n \end{bmatrix}}_{S}$$

(2m × n)

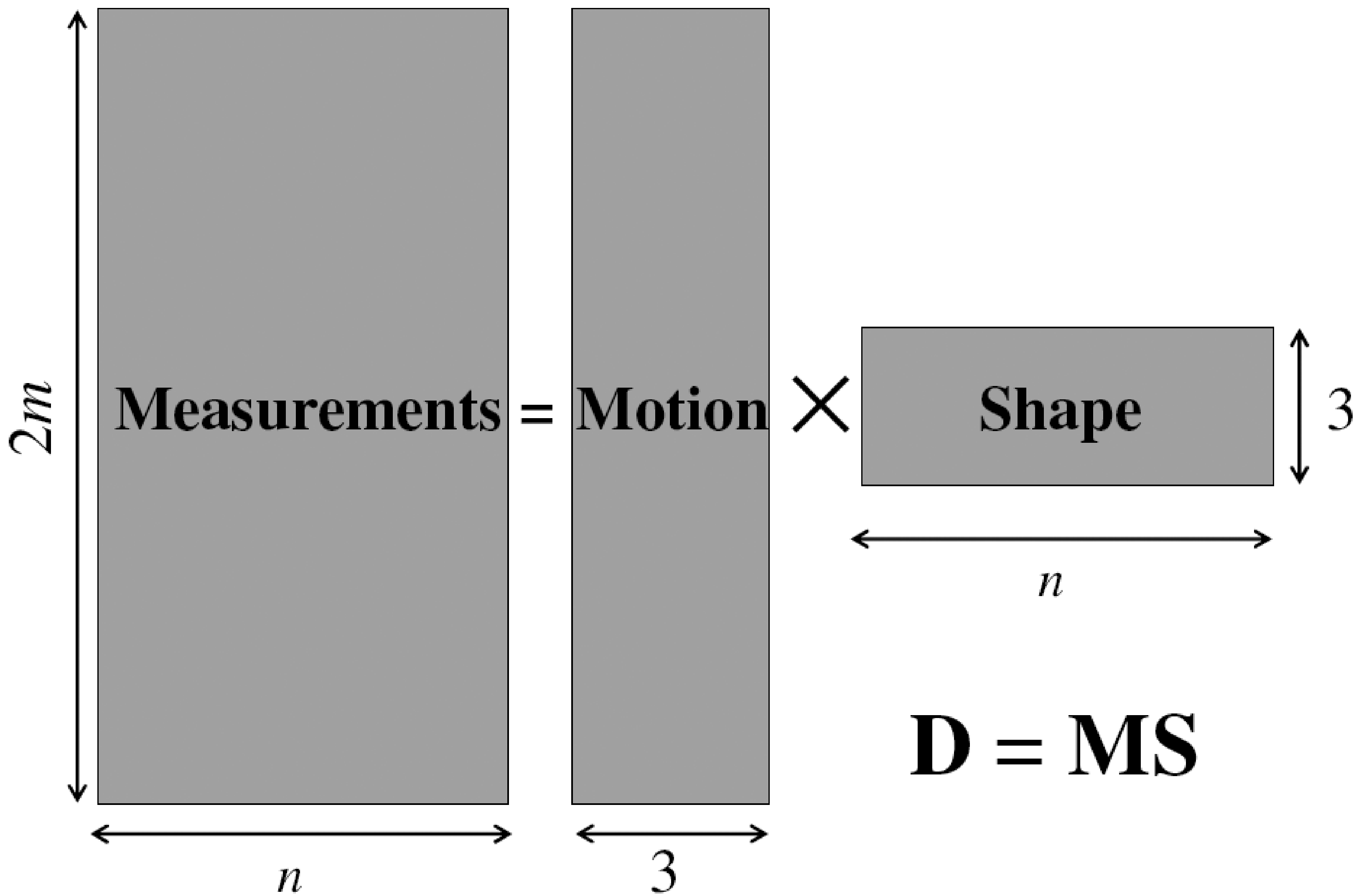
cameras (2m × 3)

points (3 × n)

The measurement matrix **D = M S** must have rank 3

(it's a product of a 2mx3 matrix and 3xn matrix)

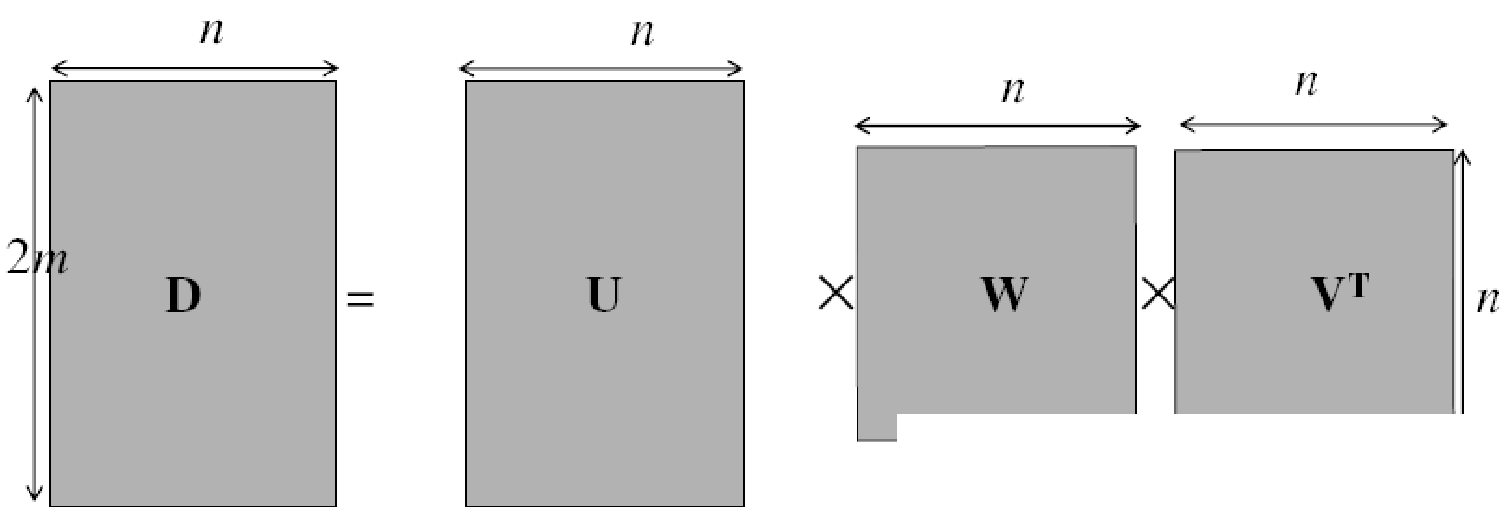
Factorizing the measurement matrix



Slide courtesy of M. Hebert

Factorizing the measurement matrix

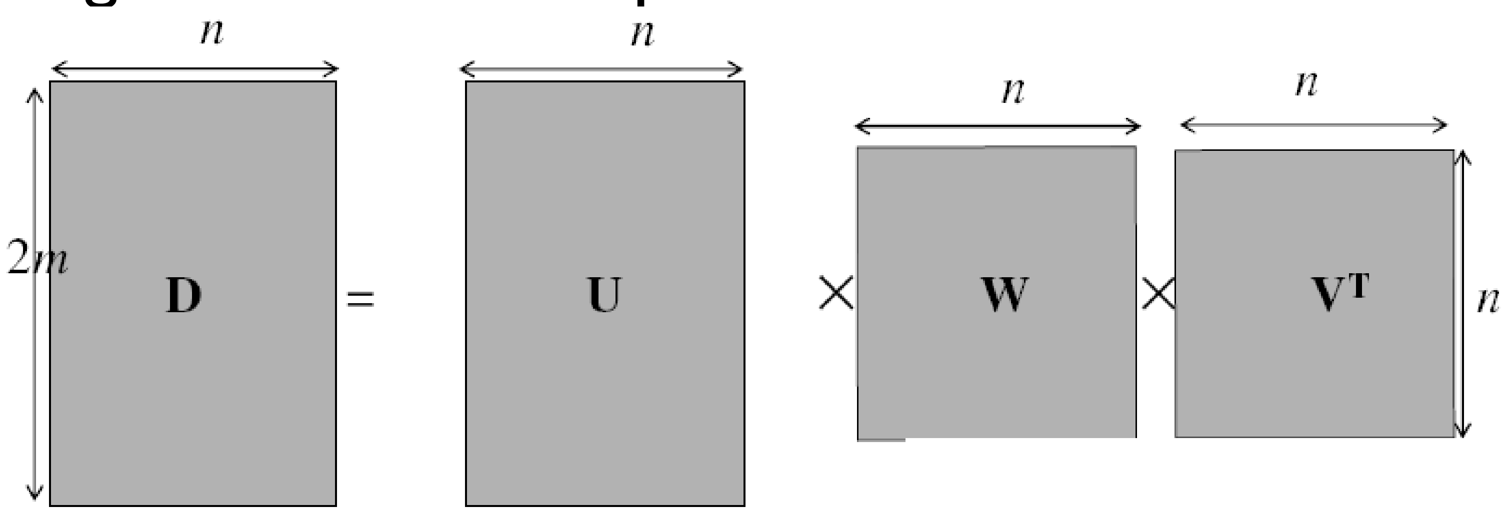
- Singular value decomposition of D:



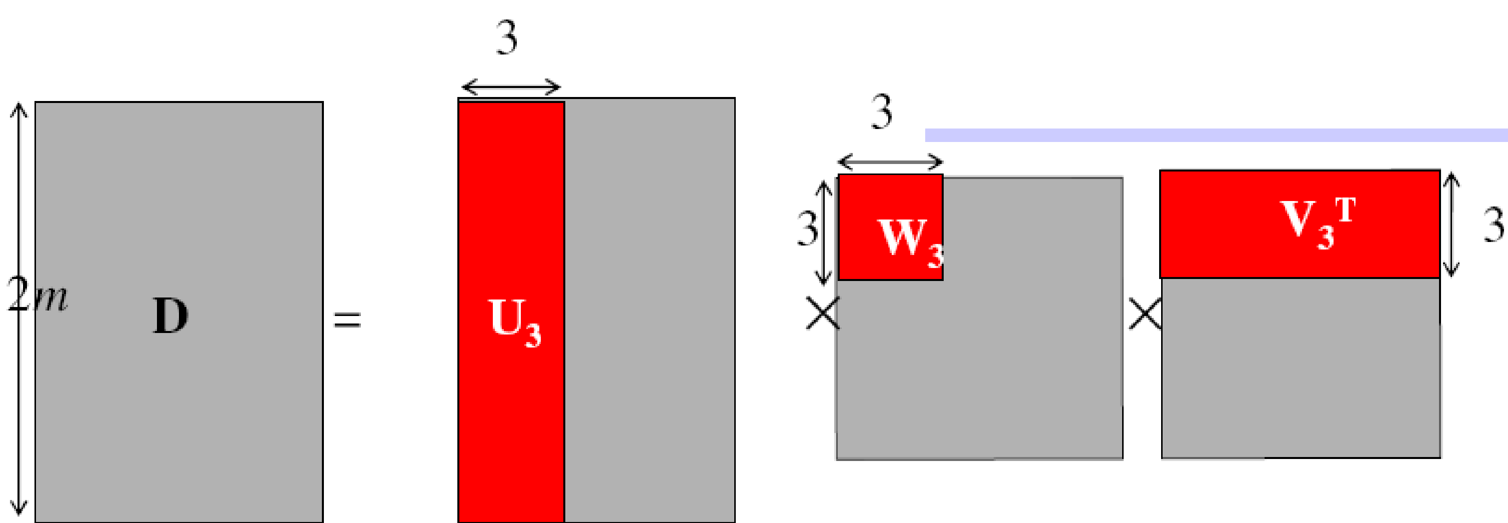
Slide courtesy of M. Hebert

Factorizing the measurement matrix

- Singular value decomposition of D:



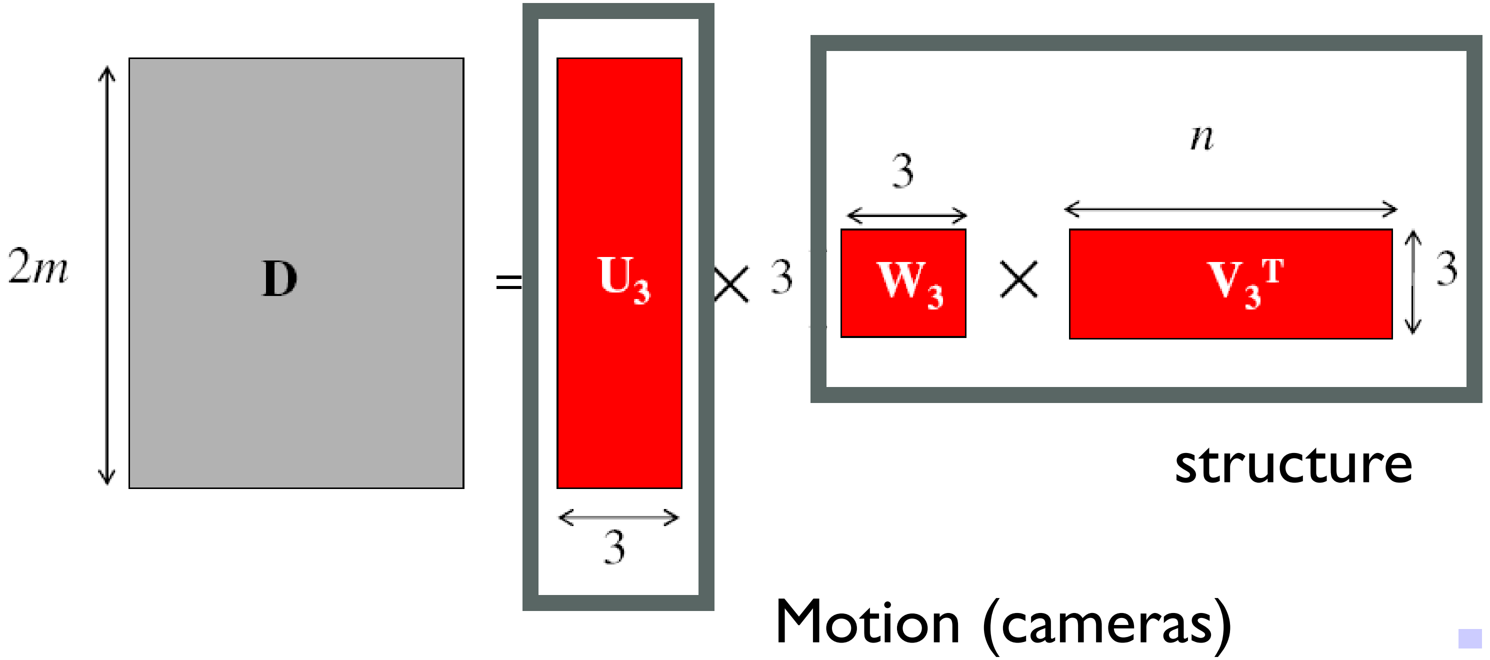
Since rank (D)=3, there are only 3 non-zero singular values



Slide courtesy of M. Hebert

Factorizing the measurement matrix

- Obtaining a factorization from SVD:

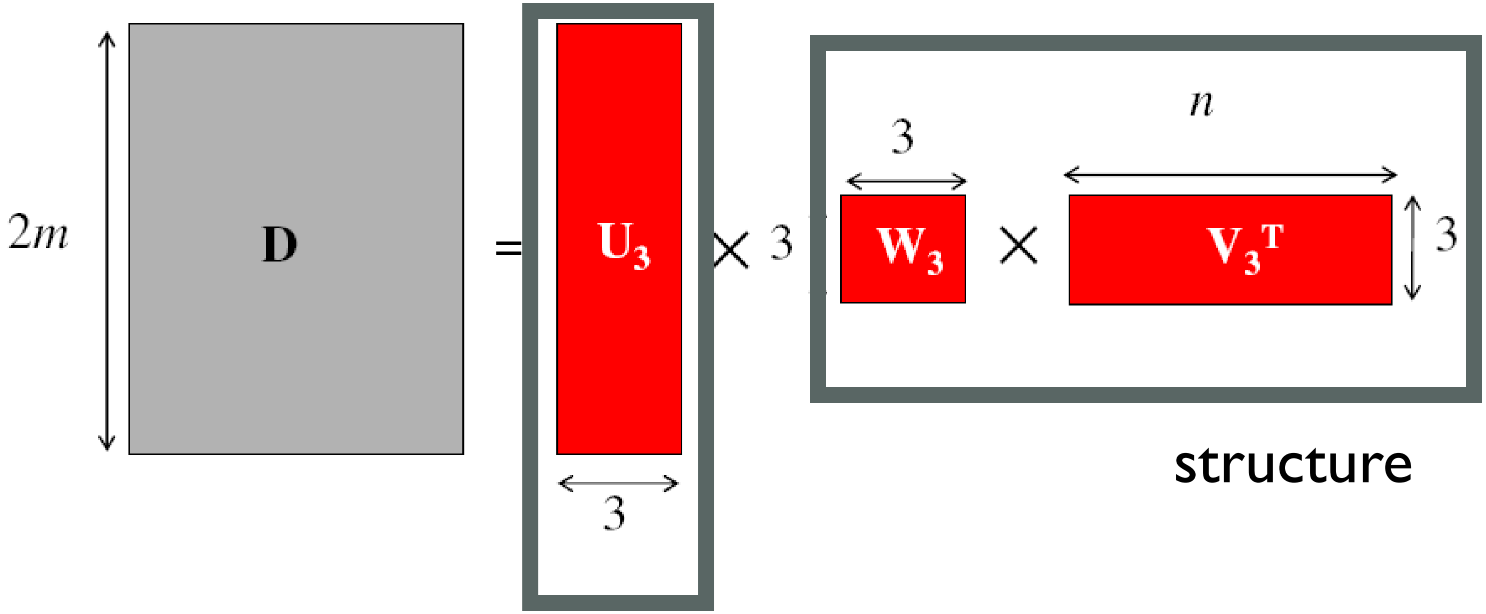


What is the issue here?

D has rank>3 because of - measurement noise
- affine approximation

Factorizing the measurement matrix

- Obtaining a factorization from SVD:

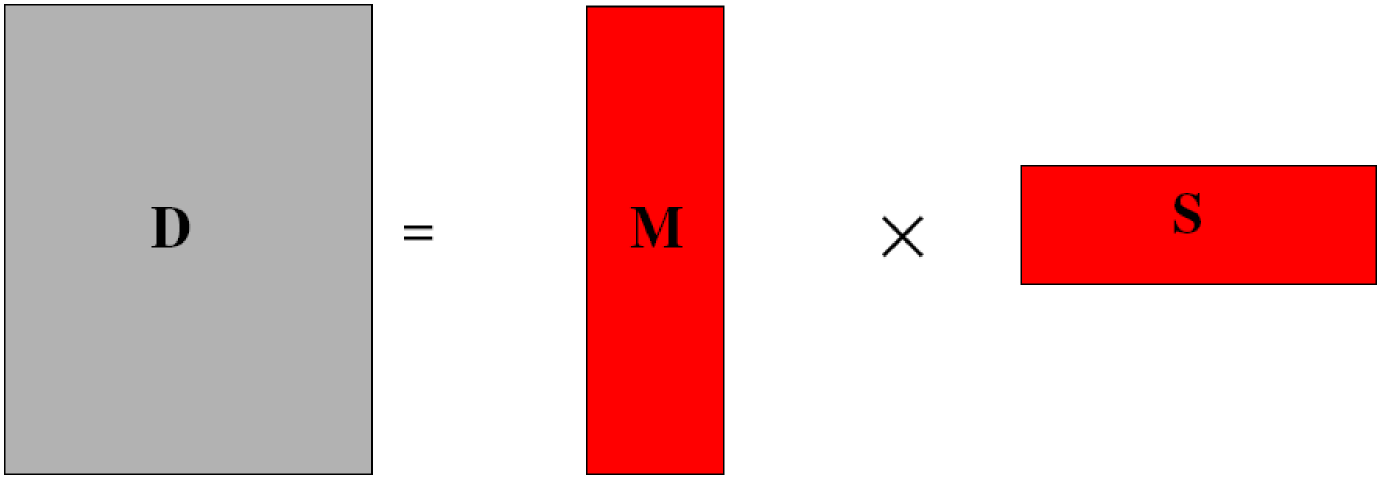


Theorem: When D has a rank greater than p, $\mathcal{U}_p \mathcal{W}_p \mathcal{V}_p^T$ is the best possible rank-p approximation of D in the sense of the Frobenius norm.

$$\mathcal{D} = \mathcal{U}_3 \mathcal{W}_3 \mathcal{V}_3^T$$

$$\begin{cases} \mathcal{A}_0 = \mathcal{U}_3 \\ \mathcal{P}_0 = \mathcal{W}_3 \mathcal{V}_3^T \end{cases}$$

Affine ambiguity



- The decomposition is not unique. We get the same **D** by using any 3×3 matrix **C** and applying the transformations **M** → **MC**, **S** → $\mathbf{C}^{-1}\mathbf{S}$

- We can enforce some Euclidean constraints to resolve the ambiguity (more on next lecture!)

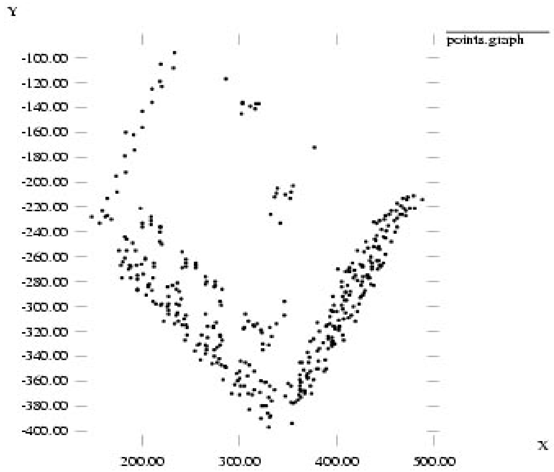
Algorithm summary

- Given: *m* images and *n* features $\mathbf{x}_{ij}$
- For each image *i*, center the feature coordinates
- Construct a 2*m* × *n* measurement matrix **D**:
  - Column *j* contains the projection of point *j* in all views
  - Row *i* contains one coordinate of the projections of all the *n* points in image *i*
- Factorize **D**:
  - Compute SVD: $\mathbf{D} = \mathbf{U}\,\mathbf{W}\,\mathbf{V}^T$
  - Create $\mathbf{U}_3$ by taking the first 3 columns of **U**
  - Create $\mathbf{V}_3$ by taking the first 3 columns of **V**
  - Create $\mathbf{W}_3$ by taking the upper left 3 × 3 block of **W**
- Create the motion and shape matrices:
  - $\mathbf{M} = \mathbf{U}_3\mathbf{W}_3^{1/2}$ and $\mathbf{S} = \mathbf{W}_3^{1/2}\,\mathbf{V}_3^T$ (**or** $\mathbf{M} = \mathbf{U}_3$ and $\mathbf{S} = \mathbf{W}_3\mathbf{V}_3^T$)
- Eliminate affine ambiguity

Reconstruction results

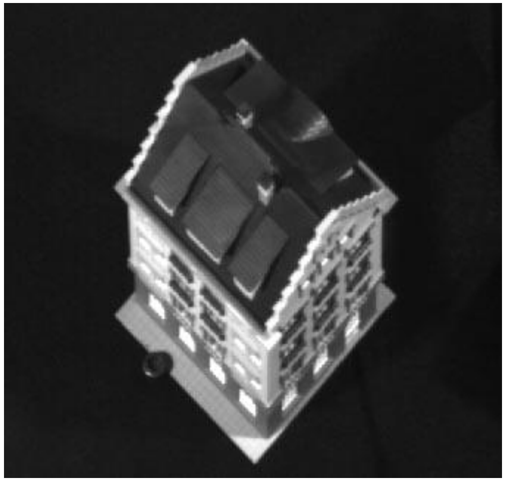
1

60



120

150



C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *IJCV*, 9(2):137-154, November 1992.

Next Lecture

- Multiple view geometry
  Perspective structure from Motion