The Fundamental Matrix $\mathbf{F}$

- Mapping of point in one image to epipolar line in other image $\mathbf{x}$! $\mathbf{l}'$ is expressed algebraically by the **fundamental matrix** $\mathbf{F}$

line　　point

- Write this as $\mathbf{l}' = \mathbf{F}\mathbf{x}$
- Since $\mathbf{x}'$ is on $\mathbf{l}'$, by the point-on-line definition we know that

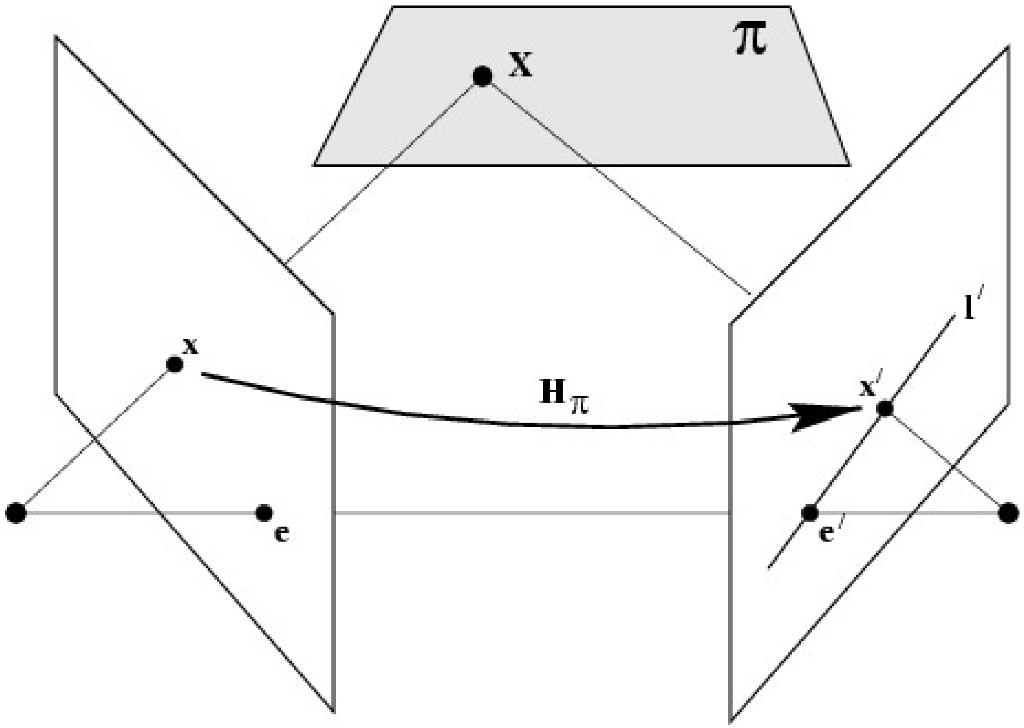
$$\mathbf{x}'^{T}\,\mathbf{l}' = 0$$

- Substitute $\mathbf{l}' = \mathbf{F}\mathbf{x}$, we can thus relate corresponding points in the camera pair $(\mathbf{P}, \mathbf{P}')$ to each other with the following:

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

The Fundamental Matrix **F**

geometric derivation



$$x' = H_\pi x$$

$$l' = e' \times x' = [e']_\times H_\pi x = Fx$$

Computing Fundamental Matrix

$$u^T F u' = 0$$

(u is same as x in the prev. slide,
u' is same as x')

Fundamental Matrix is 3x3 singular matrix with rank 2

In principle F has 7 DOF → can be estimated from 7 points correspondences

Direct simple method (8-point algorithm) requires 8 correspondences

Estimating Fundamental Matrix

The 8-point algorithm

$$u^T F u' = 0$$

Each point correspondence can be expressed as a linear equation

$$\begin{bmatrix} u & v & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} uu' & uv' & u & u'v & vv' & v & u' & v' & 1 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = 0$$

The 8-point Algorithm

8 corresponding points, 8 equations.

$$\begin{pmatrix} u_1u'_1 & u_1v'_1 & u_1 & v_1u'_1 & v_1v'_1 & v_1 & u'_1 & v'_1 \\ u_2u'_2 & u_2v'_2 & u_2 & v_2u'_2 & v_2v'_2 & v_2 & u'_2 & v'_2 \\ u_3u'_3 & u_3v'_3 & u_3 & v_3u'_3 & v_3v'_3 & v_3 & u'_3 & v'_3 \\ u_4u'_4 & u_4v'_4 & u_4 & v_4u'_4 & v_4v'_4 & v_4 & u'_4 & v'_4 \\ u_5u'_5 & u_5v'_5 & u_5 & v_5u'_5 & v_5v'_5 & v_5 & u'_5 & v'_5 \\ u_6u'_6 & u_6v'_6 & u_6 & v_6u'_6 & v_6v'_6 & v_6 & u'_6 & v'_6 \\ u_7u'_7 & u_7v'_7 & u_7 & v_7u'_7 & v_7v'_7 & v_7 & u'_7 & v'_7 \\ u_8u'_8 & u_8v'_8 & u_8 & v_8u'_8 & v_8v'_8 & v_8 & u'_8 & v'_8 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Invert and solve for $\mathcal{F}$. the system Af = b

…or build Af = 0 and solve it by minimizing ||Af||
Subject to ||f||=1
This works even for N points > 8!

Computing F: The Eight-point Algorithm

- Input: n point correspondences (n >= 8)
  - Construct homogeneous system Af= 0 from $\mathbf{u}_i'^T \mathbf{F} \mathbf{u}_i = 0$
    - f = (f_{11},f_{12}, ,f_{13}, f_{21},f_{22},f_{23} f_{31},f_{32}, f_{33}) : entries in F
    - Each correspondence gives one equation
    - A is a nx9 matrix (in homogenous format)
  - Obtain estimate F^ by SVD of A $\mathbf{A} = \mathbf{UDV}^T$
    - f (up to a scale) is column of V corresponding to the least singular value
  - Enforce singularity constraint: since Rank (F) = 2
    - Compute SVD of F^ $\hat{\mathbf{F}} = \mathbf{UDV}^T$
    - Set the smallest singular value to 0: D -> D'
    - Correct estimate of F : $\mathbf{F}' = \mathbf{UD}'\mathbf{V}^T$
- Output: the estimate of the fundamental matrix, F'
- Similarly we can compute E given intrinsic parameters

Computing F: The Eight-point Algorithm

Additional considerations:

- As points coordinates in the image space can range from 0 to >1000, the A matrix could be ill-conditioned. →normalize!
- A quick and easy normalization is the scaling of points coordinates into the range [-1;+1]
- RANSAC algorithm could be used to find a set of correspondence points without outliers

Locating the Epipoles from F

- Input: Fundamental Matrix F
  - Find the SVD of F
  - The epipole e_l is the column of V corresponding to the null singular value (as shown above)
  - The epipole e_r is the column of U corresponding to the null singular value
- Output: Epipole e_l and e_r

$$\mathbf{F} = \mathbf{UDV}^T$$

Reconstruction Ambiguity

- With uncalibrated cameras, reconstruction is defined up to a projective transformation x←→x' equal to Hx←→Hx'
- The same ambiguity holds for the 3D points X
- With calibrated cameras, up to an affine transformation

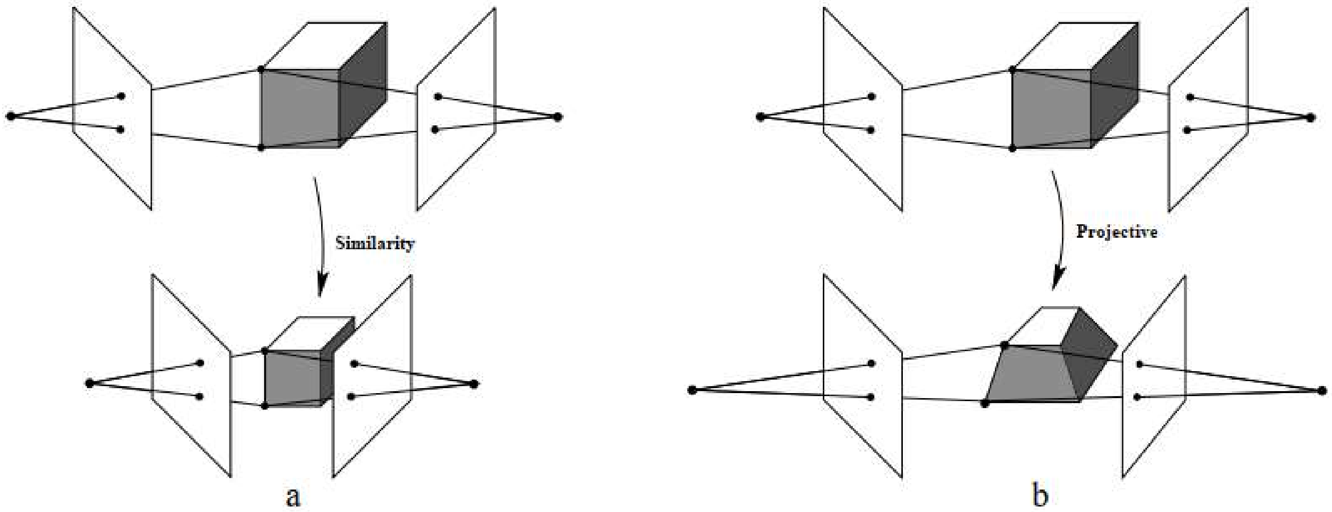


Fig. 10.2. **Reconstruction ambiguity.** *(a) If the cameras are calibrated then any reconstruction must respect the angle between rays measured in the image. A similarity transformation of the structure and camera positions does not change the measured angle. The angle between rays and the baseline (epipoles) is also unchanged. (b) If the cameras are uncalibrated then reconstructions must only respect the image points (the intersection of the rays with the image plane). A projective transformation of the structure and camera positions does not change the measured points, although the angle between rays is altered. The epipoles are also unchanged (intersection with baseline).*

3D Reconstruction:
problem statement

- Given F or the corresponding camera matrices P/P’, triangulate the 3D point X which mapped to x=PX, x’=P’X
- Since there are estimating errors in F or P/P’, the back-projected rays do not intersect! X and x’ does not satisfy the epipolar constraint $x'^T F x=0$

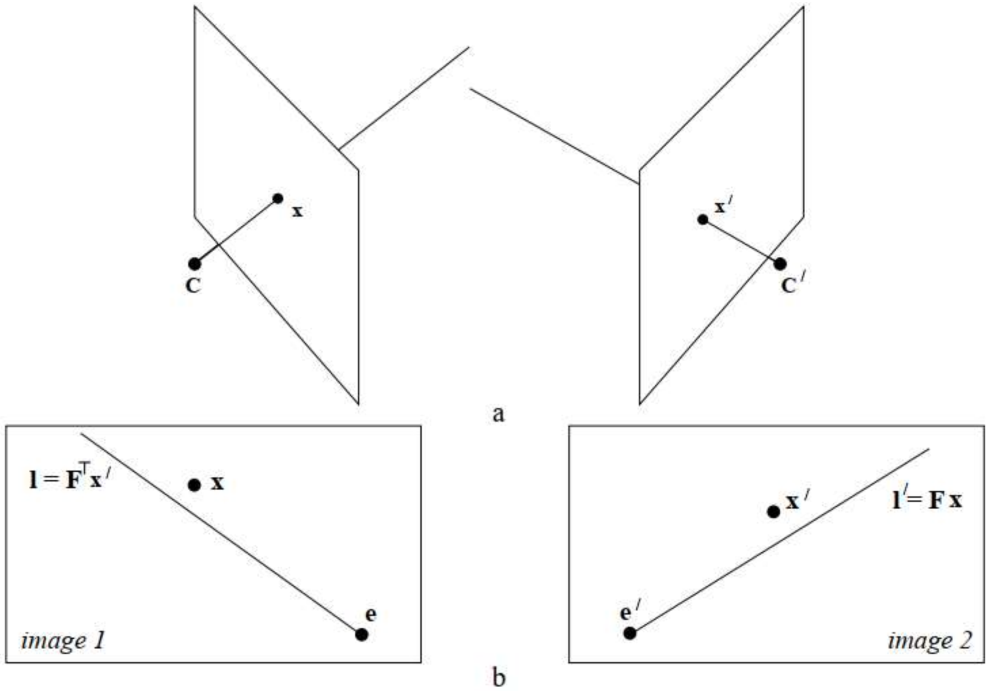


Fig. 12.1. *(a) Rays back-projected from imperfectly measured points* $\mathbf{x}, \mathbf{x}'$ *are skew in 3-space in general. (b) The epipolar geometry for* $\mathbf{x}, \mathbf{x}'$. *The measured points do not satisfy the epipolar constraint. The epipolar line* $\mathbf{l}' = \mathtt{F}\mathbf{x}$ *is the image of the ray through* $\mathbf{x}$, *and* $\mathbf{l} = \mathtt{F}^{\mathsf{T}}\mathbf{x}'$ *is the image of the ray through* $\mathbf{x}'$. *Since the rays do not intersect,* $\mathbf{x}'$ *does not lie on* $\mathbf{l}'$, *and* $\mathbf{x}$ *does not lie on* $\mathbf{l}$.

3D Reconstruction:
problem statement

- Moreover, in the uncalibrated case it makes no sense minimizing a 3D reconstruction distance – distance is not defined in 3D projective space!
- We need a triangulation method which is projective-invariant in the uncalibrated case, and affine-invariant in the calibrated case
- Key idea: instead of minimizing distance in 3D projective space, lets minimize it in the 2D image space!

Linear triangulation: homogeneous method

x=PX can be written as the cross product of x and PX = 0

This gives as result:

$$\begin{aligned} x(\mathbf{p}^{3\mathsf{T}}\mathbf{X}) - (\mathbf{p}^{1\mathsf{T}}\mathbf{X}) &= 0 \\ y(\mathbf{p}^{3\mathsf{T}}\mathbf{X}) - (\mathbf{p}^{2\mathsf{T}}\mathbf{X}) &= 0 \\ x(\mathbf{p}^{2\mathsf{T}}\mathbf{X}) - y(\mathbf{p}^{1\mathsf{T}}\mathbf{X}) &= 0 \end{aligned}$$

where $\mathbf{p}^{i\mathsf{T}}$ are the rows of $\mathtt{P}$. These equations are *linear* in the components of $\mathbf{X}$.

An equation of the form $\mathtt{A}\mathbf{X} = \mathbf{0}$ can then be composed, with

$$\mathtt{A} = \begin{bmatrix} x\mathbf{p}^{3\mathsf{T}} - \mathbf{p}^{1\mathsf{T}} \\ y\mathbf{p}^{3\mathsf{T}} - \mathbf{p}^{2\mathsf{T}} \\ x'\mathbf{p}'^{3\mathsf{T}} - \mathbf{p}'^{1\mathsf{T}} \\ y'\mathbf{p}'^{3\mathsf{T}} - \mathbf{p}'^{2\mathsf{T}} \end{bmatrix}$$

This can be solved with usual method min ||AX||, SVD blabla ;-) (p88,p90,p592 of the book.

It is not projective/affine invariant! :-(

Linear triangulation: inhomogeneous method

Same AX=0 as before, but working with $X=(X,Y,Z,1)^T$

This means that the equation system is reduced to a set of 4 inhomogeneous equation in 3 unknowns. (p90,p588 of the book).

This supposes that last coordinate is equal to 1, but a projective transformation can put it to 0 --> not projective invariant!

...but it works for affine transformations :-)

Take home message: inhomogeneous method is "good" for calibrated cameras. Homogeneous method is only "acceptable", for any camera.

3D Reconstruction:

geometric error

- Key idea: instead of minimizing distance in 3D projective space, lets minimize <u>a distance function</u> the 2D image space: $\hat{\mathbf{x}} = \mathrm{P}\widehat{\mathbf{X}} \quad \hat{\mathbf{x}}' = \mathrm{P}'\widehat{\mathbf{X}}$

$$\mathcal{C}(\mathbf{x}, \mathbf{x}') = d(\mathbf{x}, \hat{\mathbf{x}})^2 + d(\mathbf{x}', \hat{\mathbf{x}}')^2 \quad \text{subject to} \quad \hat{\mathbf{x}}'^{\mathsf{T}}\mathrm{F}\hat{\mathbf{x}} = 0$$

3D Reconstruction:

optimal triangulation method

Objective

Given a measured point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$, and a fundamental matrix $\mathtt{F}$, compute the corrected correspondences $\hat{\mathbf{x}} \leftrightarrow \hat{\mathbf{x}}'$ that minimize the geometric error (12.1) subject to the epipolar constraint $\hat{\mathbf{x}}'^{\mathsf{T}}\mathtt{F}\hat{\mathbf{x}} = 0$.

Algorithm

(i) Define transformation matrices

$$\mathtt{T} = \begin{bmatrix} 1 & & -x \\ & 1 & -y \\ & & 1 \end{bmatrix} \text{ and } \mathtt{T}' = \begin{bmatrix} 1 & & -x' \\ & 1 & -y' \\ & & 1 \end{bmatrix}.$$

These are the translations that take $\mathbf{x} = (x, y, 1)^{\mathsf{T}}$ and $\mathbf{x}' = (x', y', 1)^{\mathsf{T}}$ to the origin.

(ii) Replace $\mathtt{F}$ by $\mathtt{T}'^{-\mathsf{T}}\mathtt{F}\mathtt{T}^{-1}$. The new $\mathtt{F}$ corresponds to translated coordinates.

(iii) Compute the right and left epipoles $\mathbf{e} = (e_1, e_2, e_3)^{\mathsf{T}}$ and $\mathbf{e}' = (e'_1, e'_2, e'_3)^{\mathsf{T}}$ such that $\mathbf{e}'^{\mathsf{T}}\mathtt{F} = \mathbf{0}$ and $\mathtt{F}\mathbf{e} = \mathbf{0}$. Normalize (multiply by a scale) $\mathbf{e}$ such that $e_1^2 + e_2^2 = 1$ and do the same to $\mathbf{e}'$.

(iv) Form matrices

$$\mathtt{R} = \begin{bmatrix} e_1 & e_2 & \\ -e_2 & e_1 & \\ & & 1 \end{bmatrix} \text{ and } \mathtt{R}' = \begin{bmatrix} e'_1 & e'_2 & \\ -e'_2 & e'_1 & \\ & & 1 \end{bmatrix}$$

and observe that $\mathtt{R}$ and $\mathtt{R}'$ are rotation matrices, and $\mathtt{R}\mathbf{e} = (1, 0, e_3)^{\mathsf{T}}$ and $\mathtt{R}'\mathbf{e}' = (1, 0, e'_3)^{\mathsf{T}}$.

(v) Replace $\mathtt{F}$ by $\mathtt{R}'\mathtt{F}\mathtt{R}^{\mathsf{T}}$. The resulting $\mathtt{F}$ must have the form (12.3).

(vi) Set $f = e_3$, $f' = e'_3$, $a = F_{22}$, $b = F_{23}$, $c = F_{32}$ and $d = F_{33}$.

(vii) Form the polynomial $g(t)$ as a polynomial in t according to (12.7). Solve for t to get 6 roots.

(viii) Evaluate the cost function (12.5) at the real part of each of the roots of $g(t)$ (alternatively evaluate at only the real roots of $g(t)$). Also, find the asymptotic value of (12.1) for $t = \infty$, namely $1/f^2 + c^2/(a^2 + f'^2c^2)$. Select the value $t_{\min}$ of t that gives the smallest value of the cost function.

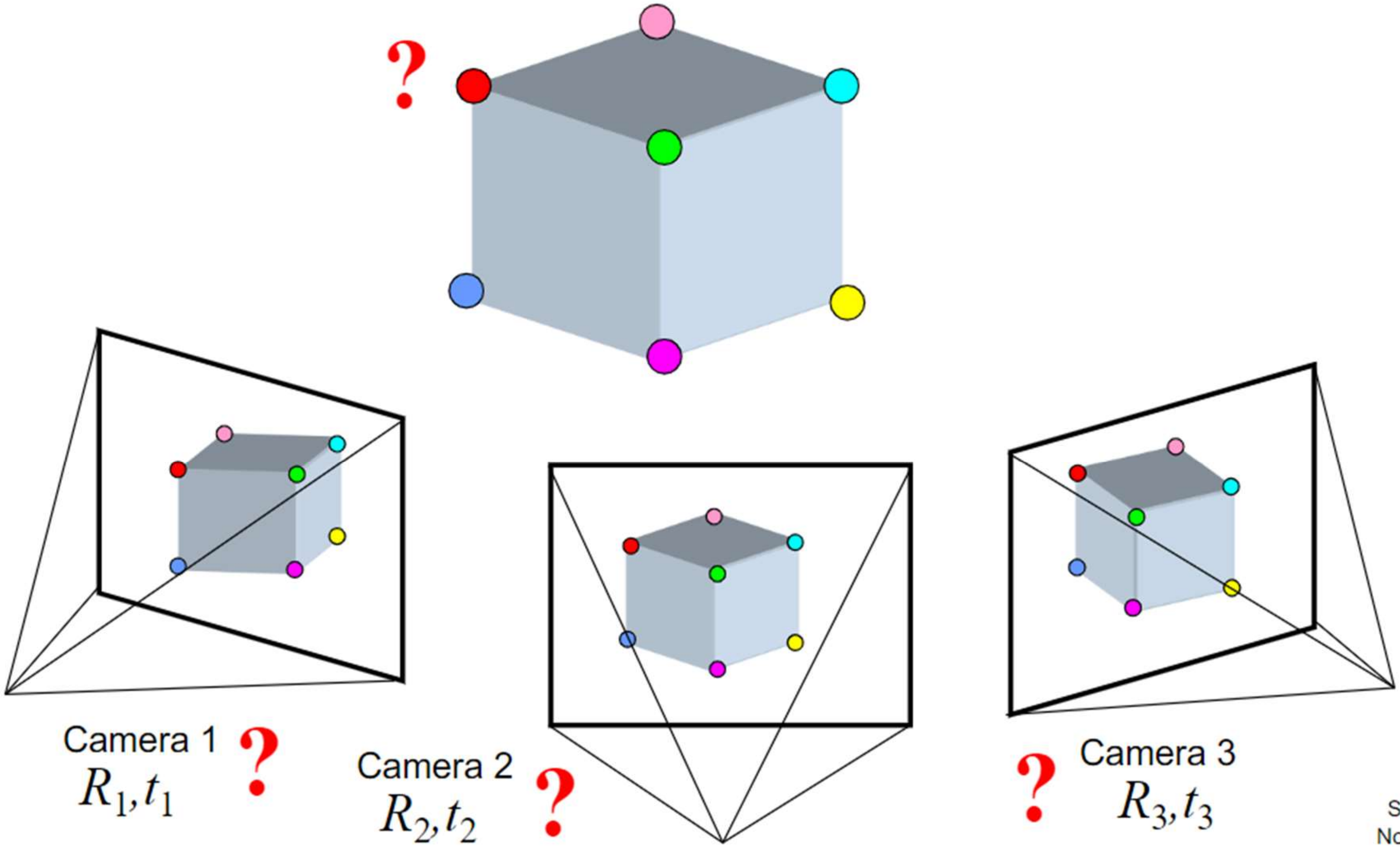
(ix) Evaluate the two lines $\mathbf{l} = (tf, 1, -t)$ and $\mathbf{l}'$ given by (12.4) at $t_{\min}$ and find $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ as the closest points on these lines to the origin. For a general line (λ, μ, ν), the formula for the closest point on the line to the origin is $(-\lambda\nu, -\mu\nu, \lambda^2 + \mu^2)$.

(x) Transfer back to the original coordinates by replacing $\hat{\mathbf{x}}$ by $\mathtt{T}^{-1}\mathtt{R}^{\mathsf{T}}\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ by $\mathtt{T}'^{-1}\mathtt{R}'^{\mathsf{T}}\hat{\mathbf{x}}'$.

(xi) The 3-space point $\widehat{\mathbf{X}}$ may then be obtained by the homogeneous method of section 12.2.

Structure from motion

- Given a set of corresponding points in two or more images, compute the camera parameters and the 3D point coordinates

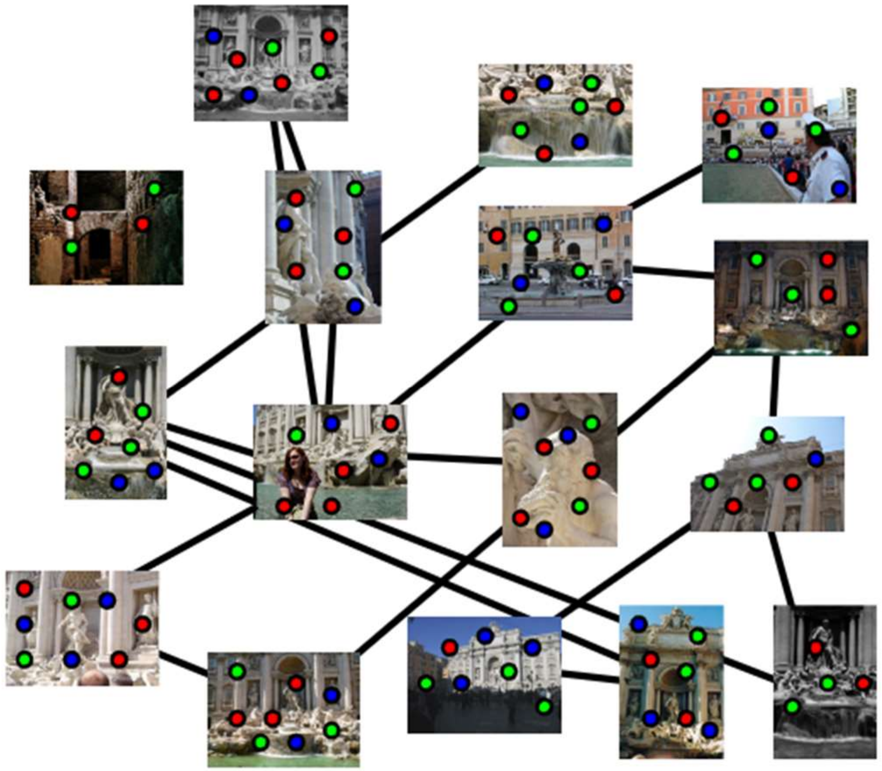


Slide credit:
Noah Snavely

Representative SFM pipeline

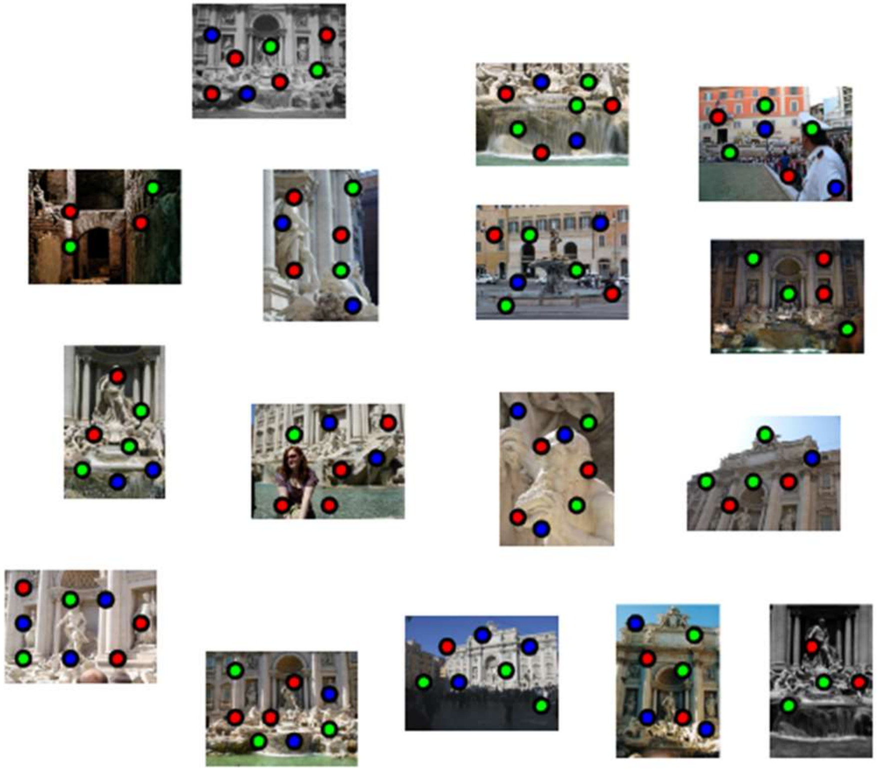
Feature matching

Match features between each pair of images



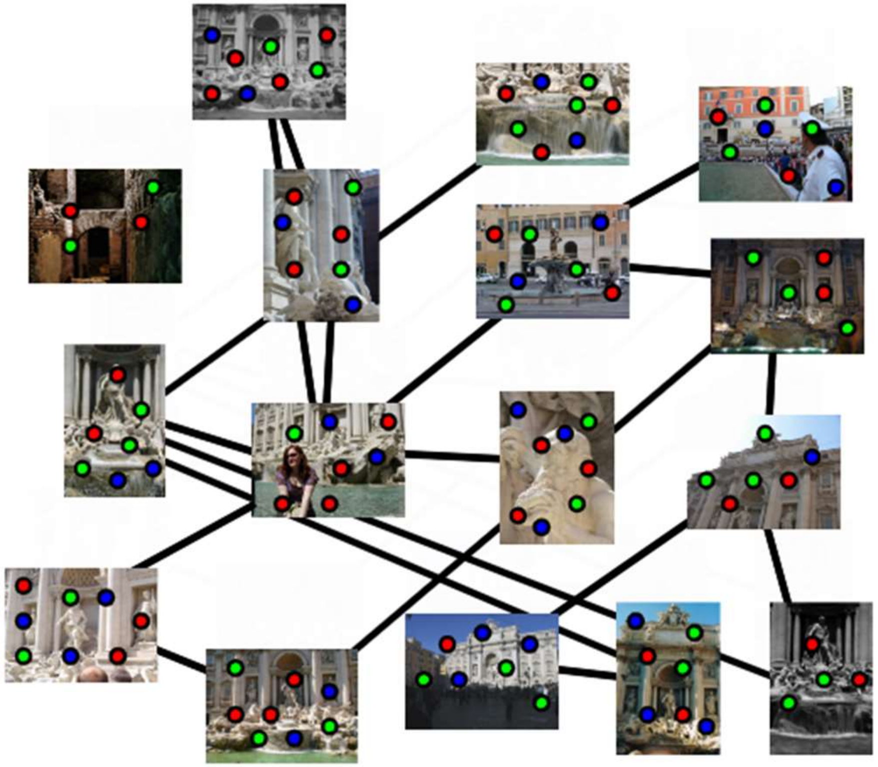
Feature detection

Detect SIFT features



Feature matching

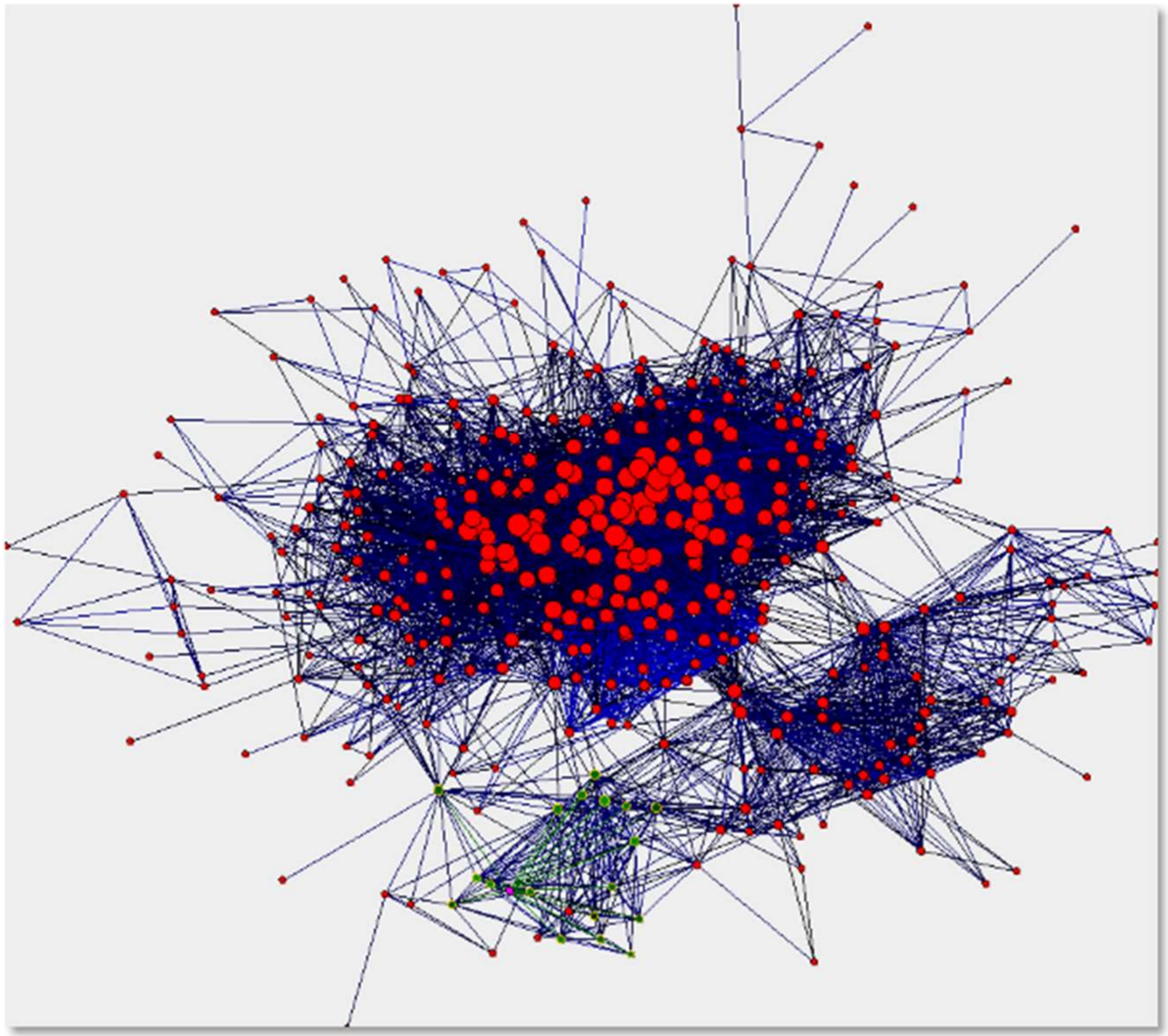
Use RANSAC to estimate fundamental matrix between each pair



Feature matching

Use RANSAC to estimate fundamental matrix between each pair


Image connectivity graph



(graph layout produced using the Graphviz toolkit: http://www.graphviz.org/)

Structure from motion

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

$$\lambda_{ij}\mathbf{x}_{ij} = \mathbf{P}_i\mathbf{X}_j, \quad i = 1, \dots, m, \quad j = 1, \dots, n$$

- Problem: estimate *m* projection matrices $\mathbf{P}_i$ and *n* 3D points $\mathbf{X}_j$ from the *mn* correspondences $\mathbf{x}_{ij}$

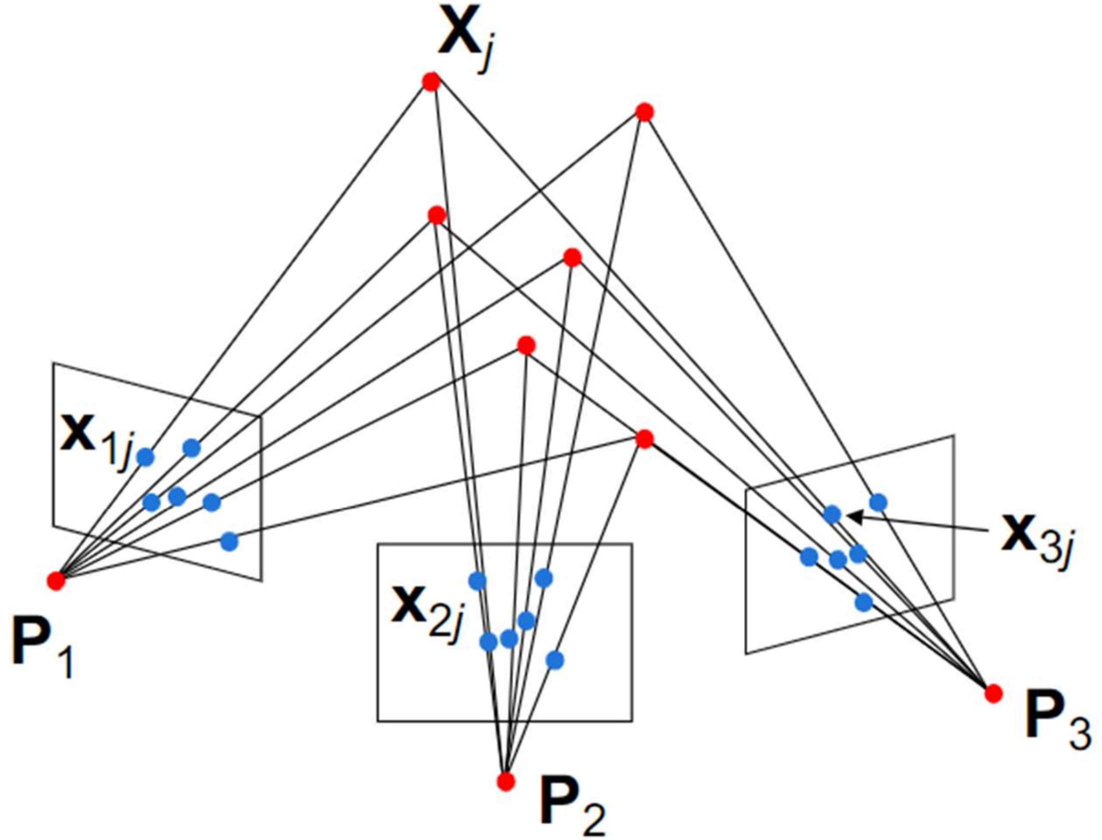
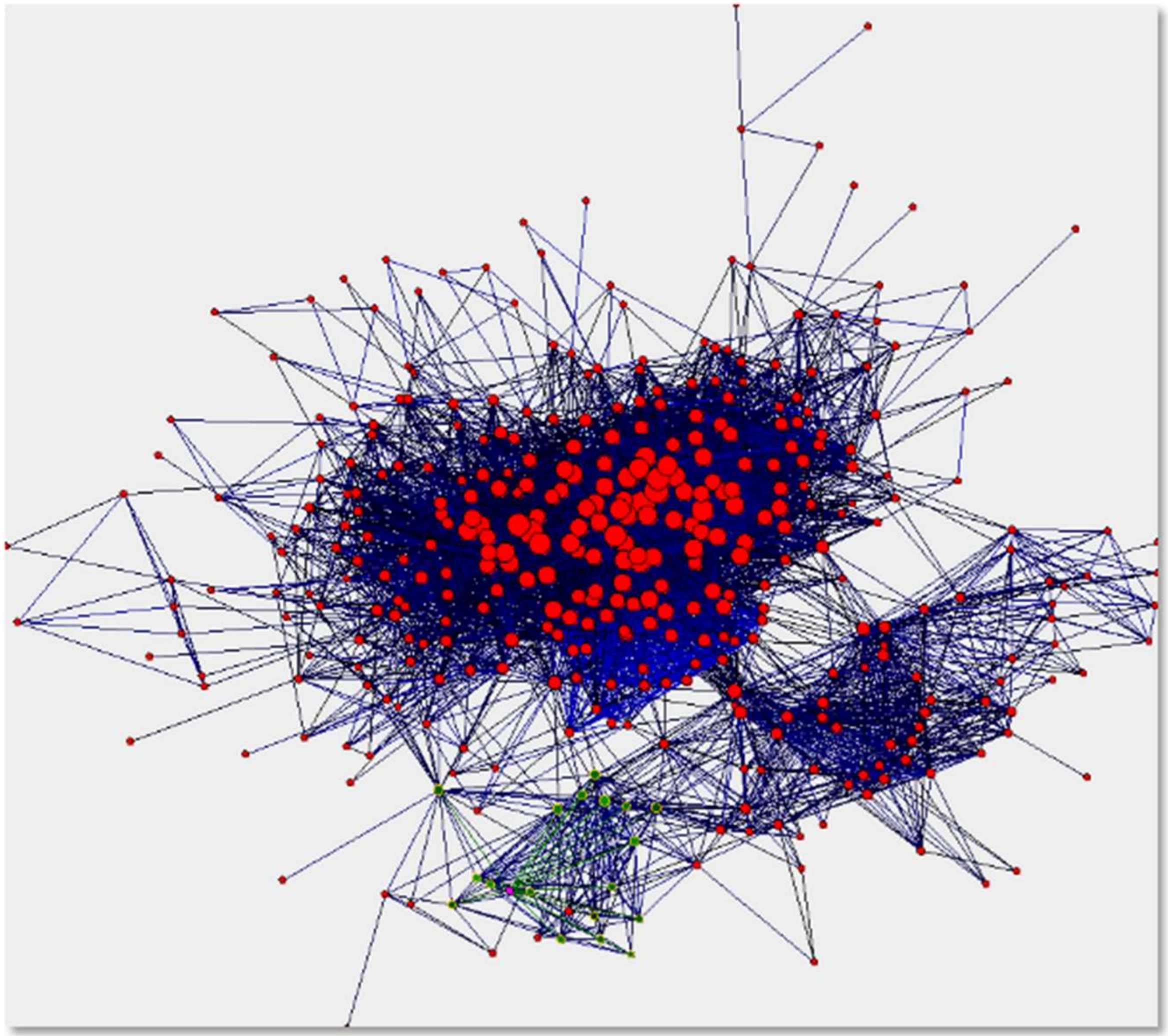


Image connectivity graph



(graph layout produced using the Graphviz toolkit: http://www.graphviz.org/)

Projective structure from motion

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

$$\lambda_{ij}\mathbf{x}_{ij} = \mathbf{P}_i\mathbf{X}_j, \quad i = 1, \dots, m, \quad j = 1, \dots, n$$

- Problem: estimate *m* projection matrices $\mathbf{P}_i$ and *n* 3D points $\mathbf{X}_j$ from the *mn* correspondences $\mathbf{x}_{ij}$
- With no calibration info, cameras and points can only be recovered up to a 4x4 projective transformation **Q**:

$$\mathbf{X} \rightarrow \mathbf{QX}, \ \mathbf{P} \rightarrow \mathbf{PQ}^{-1}$$

- We can solve for structure and motion when

$$2mn >= 11m + 3n - 15$$

- For two cameras, at least 7 points are needed

Projective SFM: Two-camera case

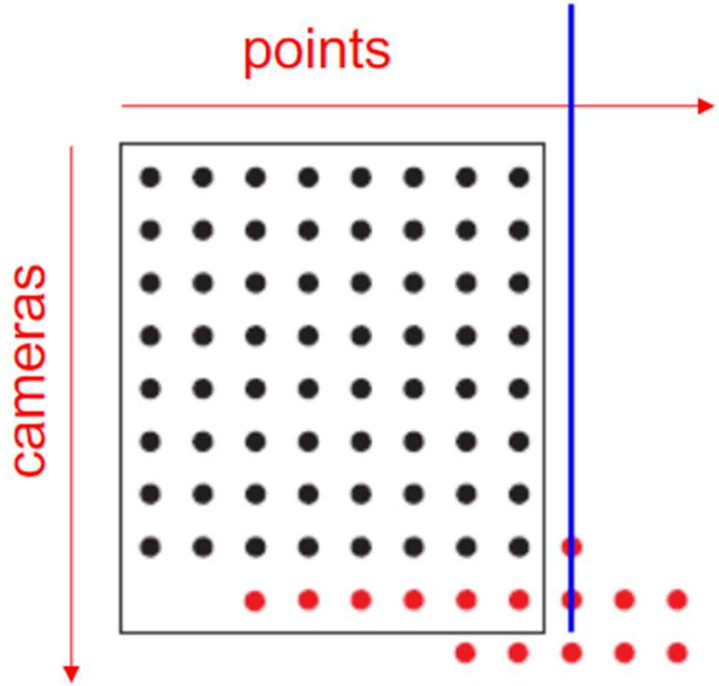
- Compute fundamental matrix $\mathbf{F}$ between the two views
- First camera matrix: $[\mathbf{I} \mid \mathbf{0}]$
- Second camera matrix: $[\mathbf{A} \mid \mathbf{b}]$
- Then $\mathbf{b}$ is the epipole ($\mathbf{F}^T\mathbf{b} = 0$), $\mathbf{A} = -[\mathbf{b}_\times]\mathbf{F}$

Incremental structure from motion

- Initialize motion from two images using fundamental matrix

- Initialize structure by triangulation

- For each additional view:
  - Determine projection matrix of new camera using all the known 3D points that are visible in its image – *calibration*
  - Refine and extend structure: compute new 3D points, re-optimize existing points that are also seen by this camera – *triangulation*



Bundle adjustment

- Non-linear method for refining structure and motion
- Minimize reprojection error

$$\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} \left\| \mathbf{x}_{ij} - \frac{1}{\lambda_{ij}} \mathbf{P}_i \mathbf{X}_j \right\|^2$$

visibility flag: is point j visible in view i?