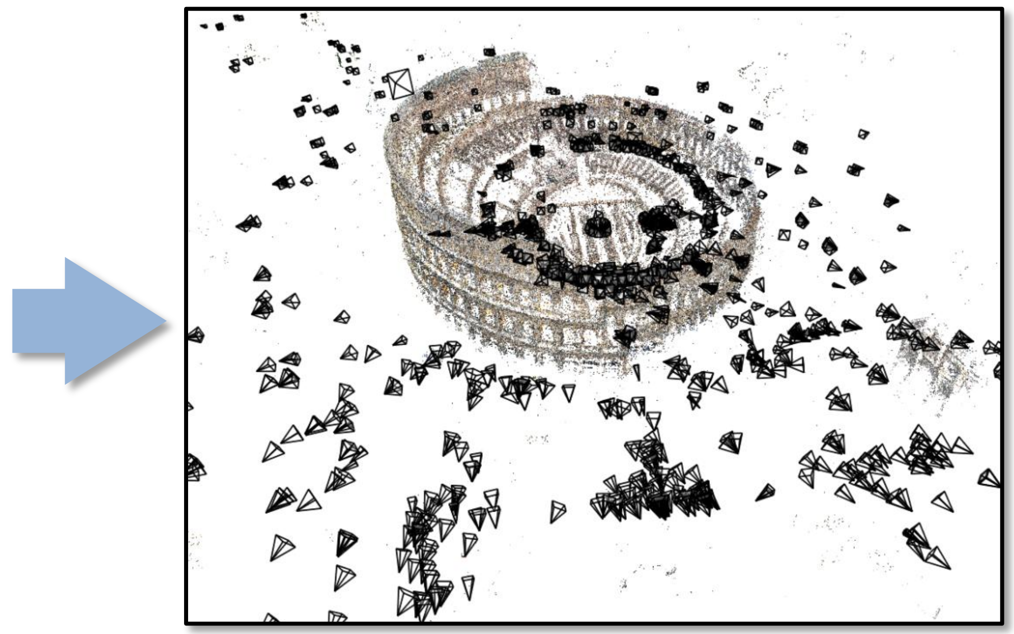
CS6670: Computer Vision

Noah Snavely

Lecture 11: Structure from motion, part 2



Announcements

- Project 2 due tonight at 11:59pm
  - Artifact due tomorrow (Friday) at 11:59pm
- Questions?

Final project

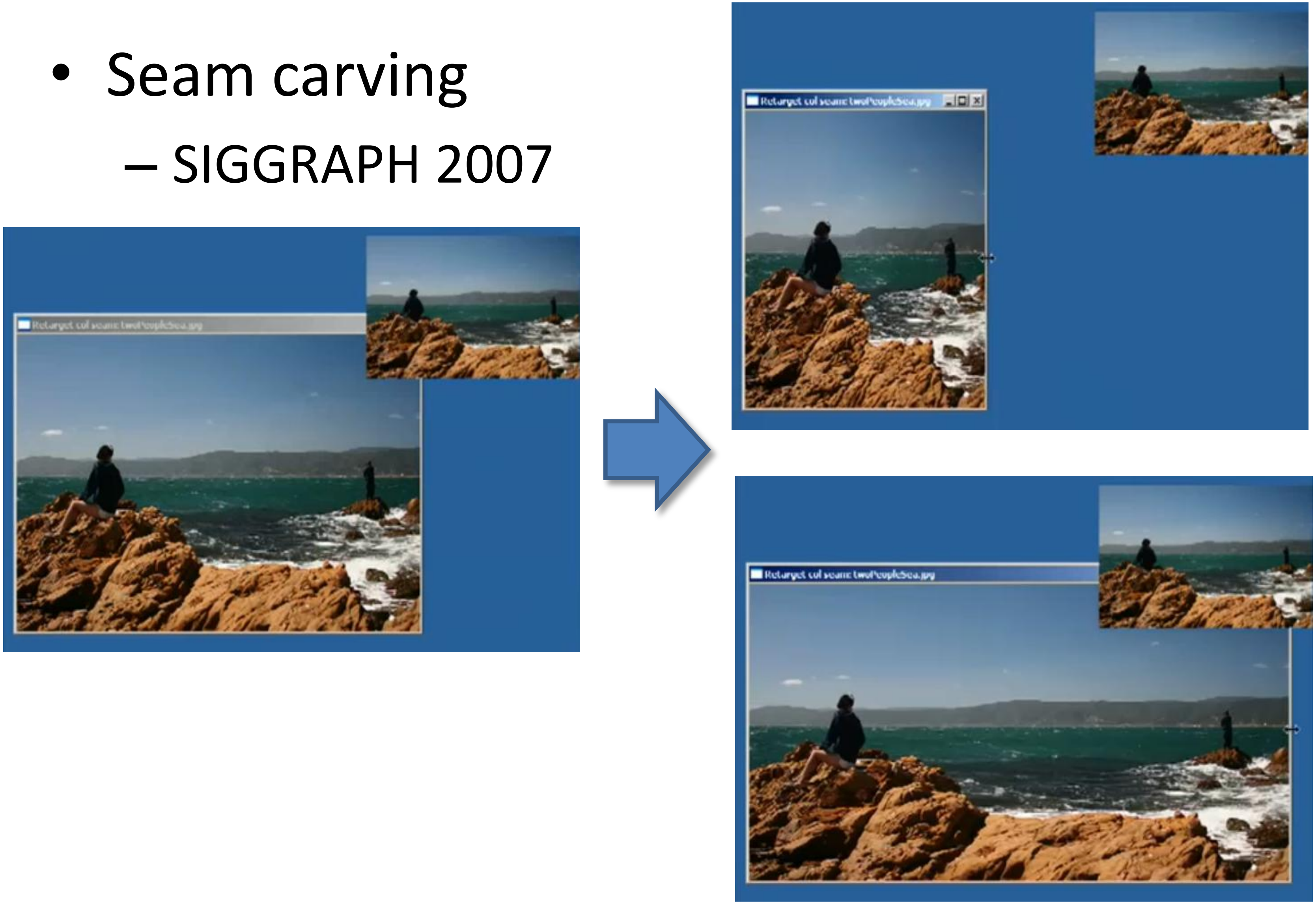
- Form your groups and submit a proposal
- Work on final project
- Final project presentations the last few classes
- Final report due the last day of classes

Final project

- One-person project: implement a recent research paper

Example one person projects

- Seam carving
  - SIGGRAPH 2007



Example one person projects

- What does the sky tell us about the camera?
  - Lalonde, et al, ECCV 2008

Other ideas

- Will post more ideas online
- Look at recent ICCV/CVPR/ECCV/SIGGRAPH papers

Example two/three person projects

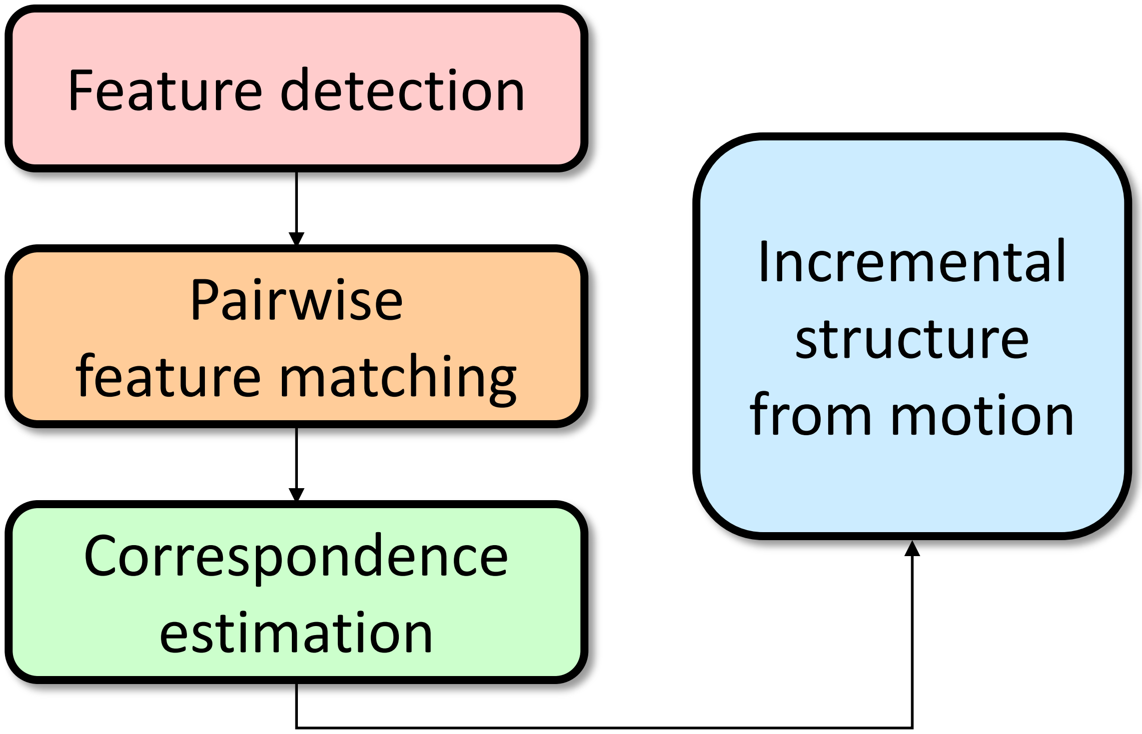
- Automatic calibration of your camera's clock
- In roughly what year was a given photo taken?
- Matching historical images to modern-day images
- Matching images on news sites / blogs for article matching
- Will post others online soon…

Final projects

- Think about your groups and project ideas

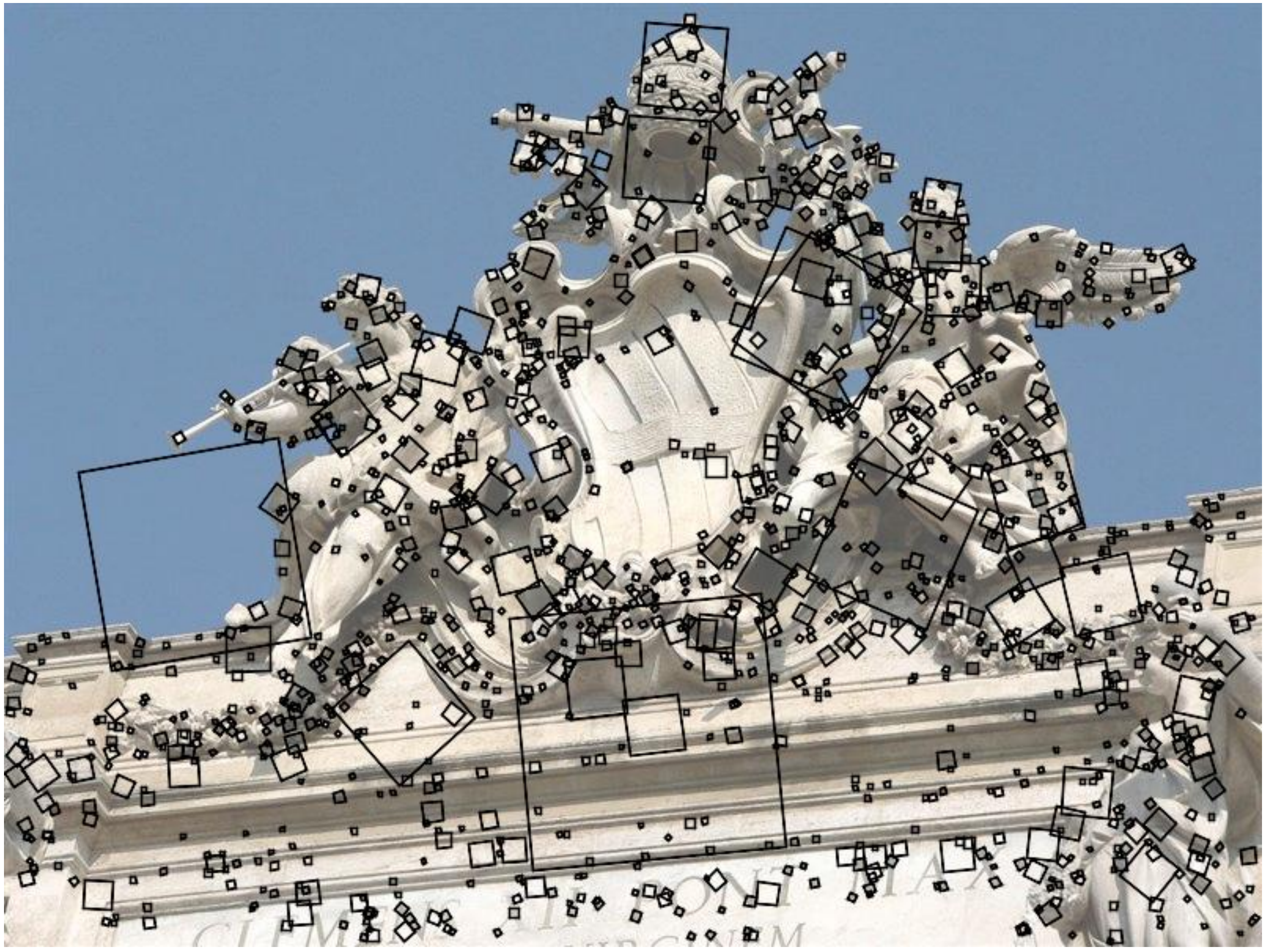
- Proposals will be due on Tuesday, Oct. 27

Scene reconstruction



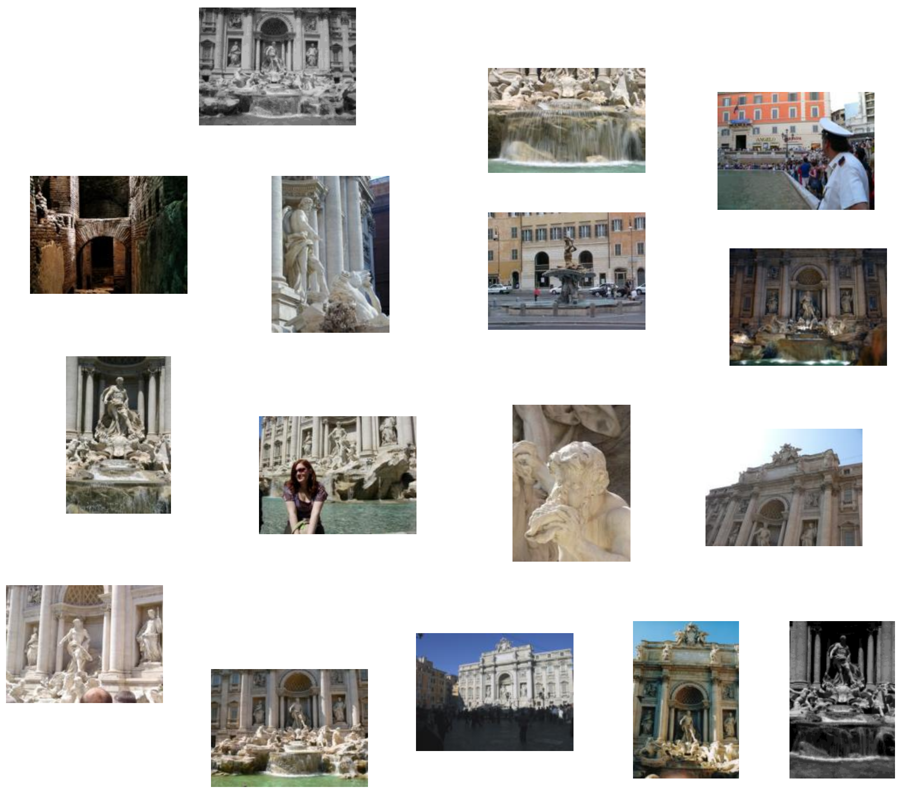
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



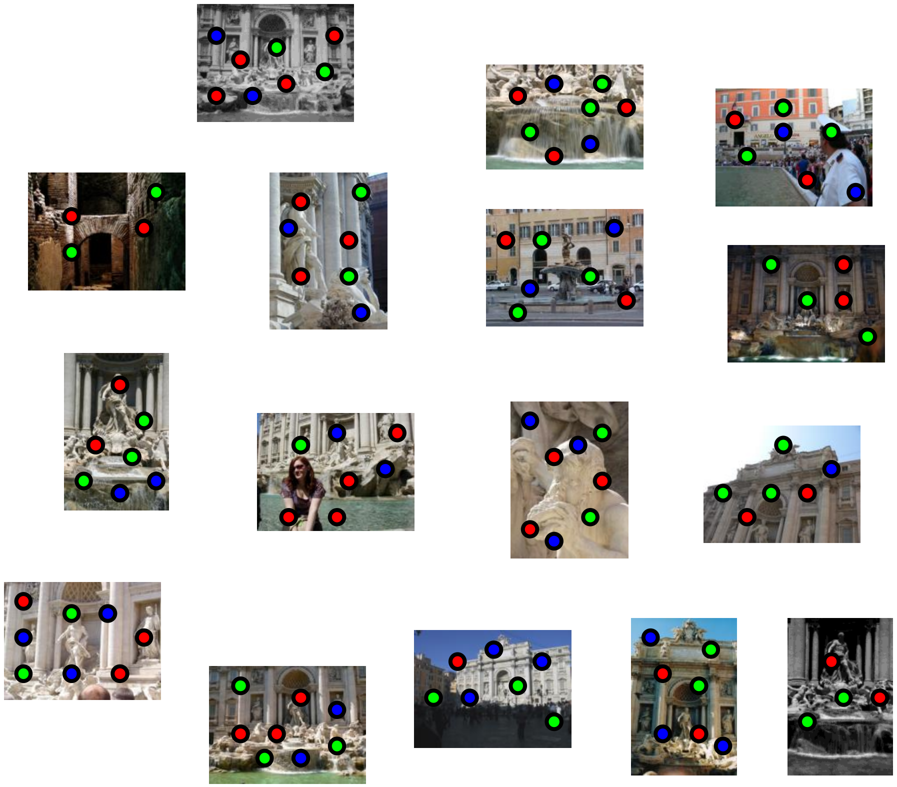
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



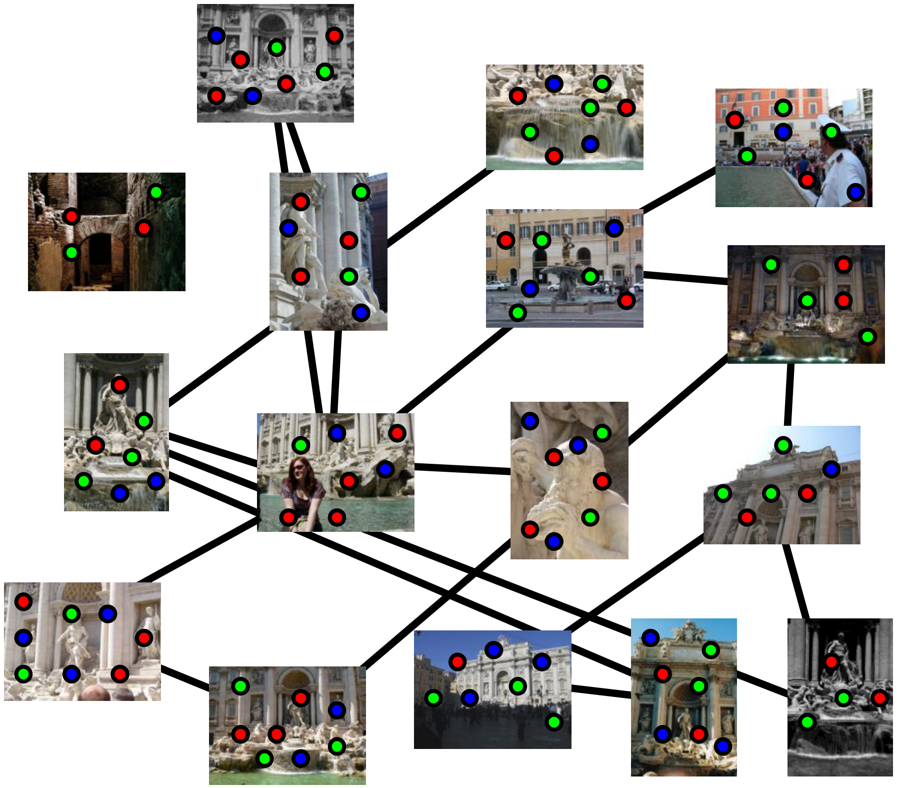
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



Feature matching

Match features between each pair of images



Feature matching

Refine matching using RANSAC [Fischler & Bolles 1987] to estimate fundamental matrices between pairs

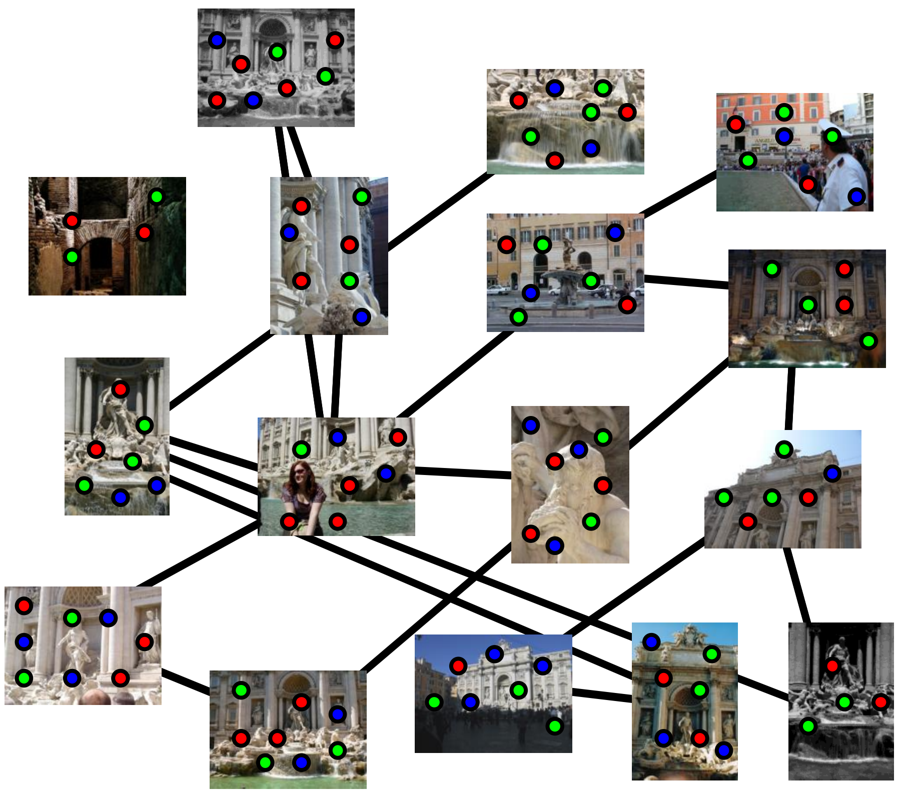
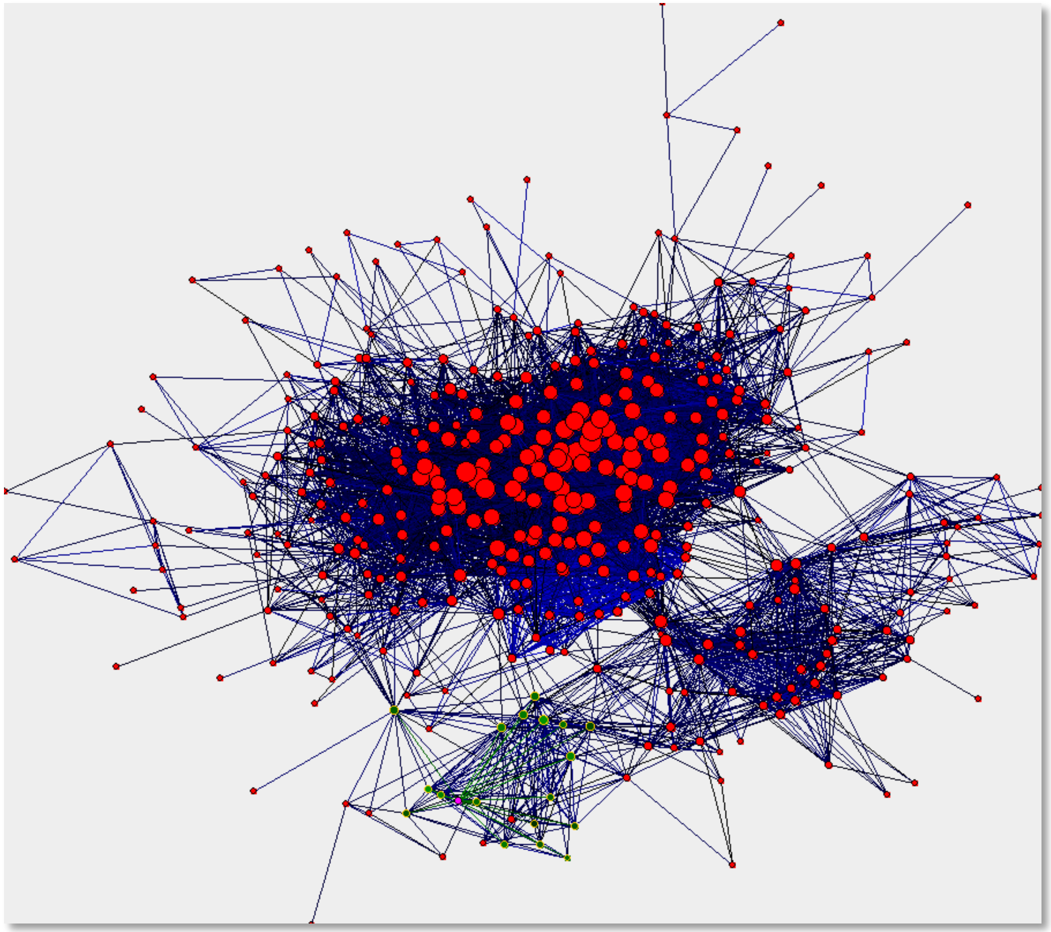
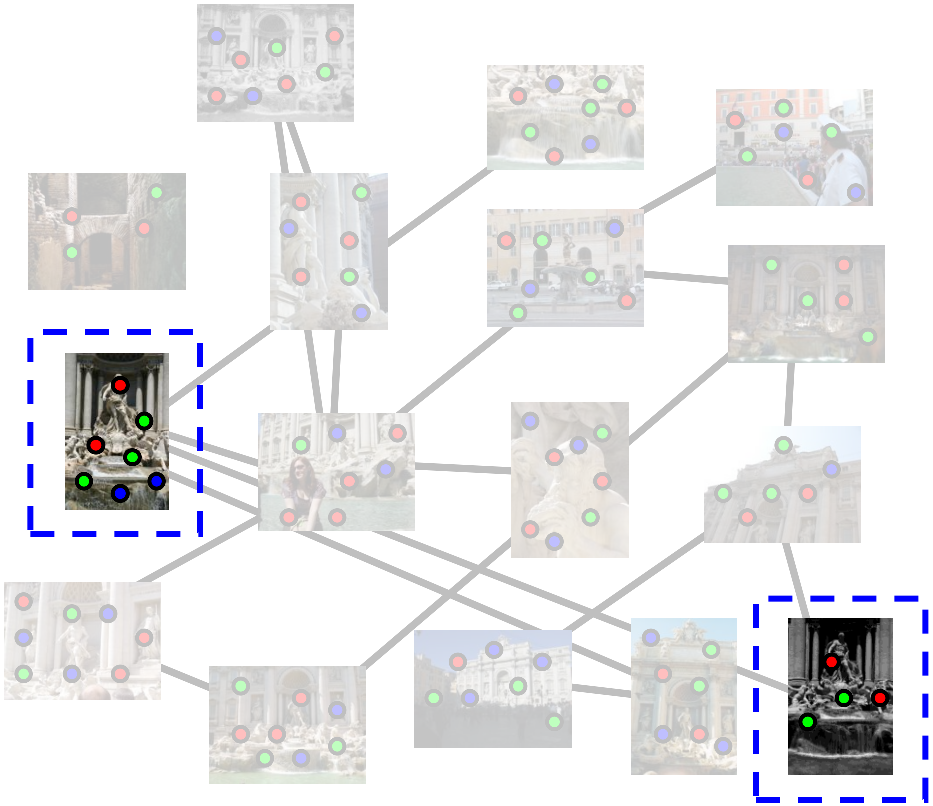


Image connectivity graph



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

Incremental structure from motion



Incremental structure from motion

Incremental structure from motion

Problem size

- Trevi Fountain collection

  466 input photos

  + > 100,000 3D points

  = very large optimization problem

Photo Tourism overview

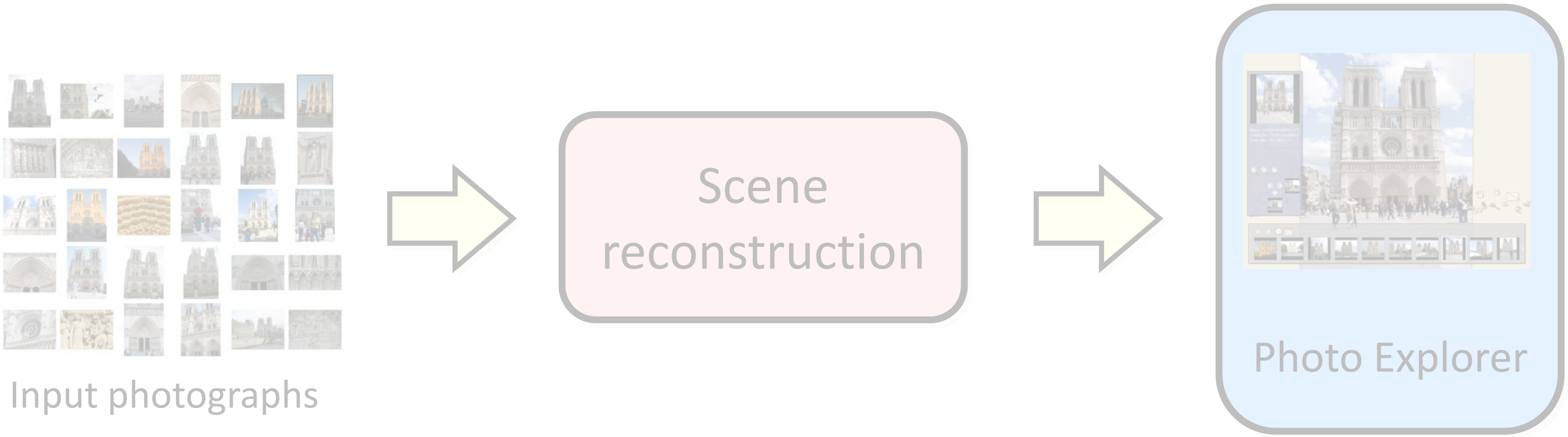
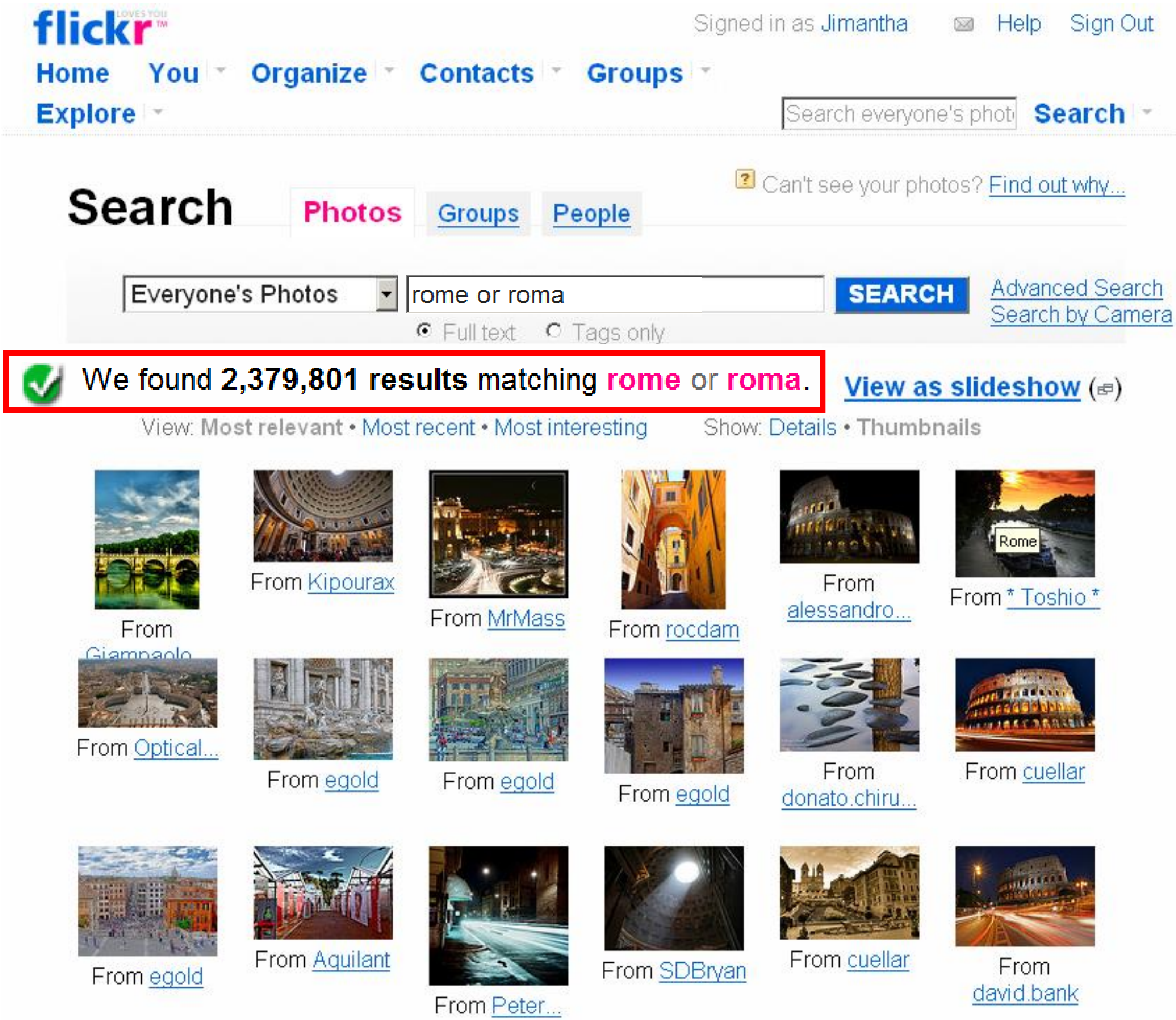
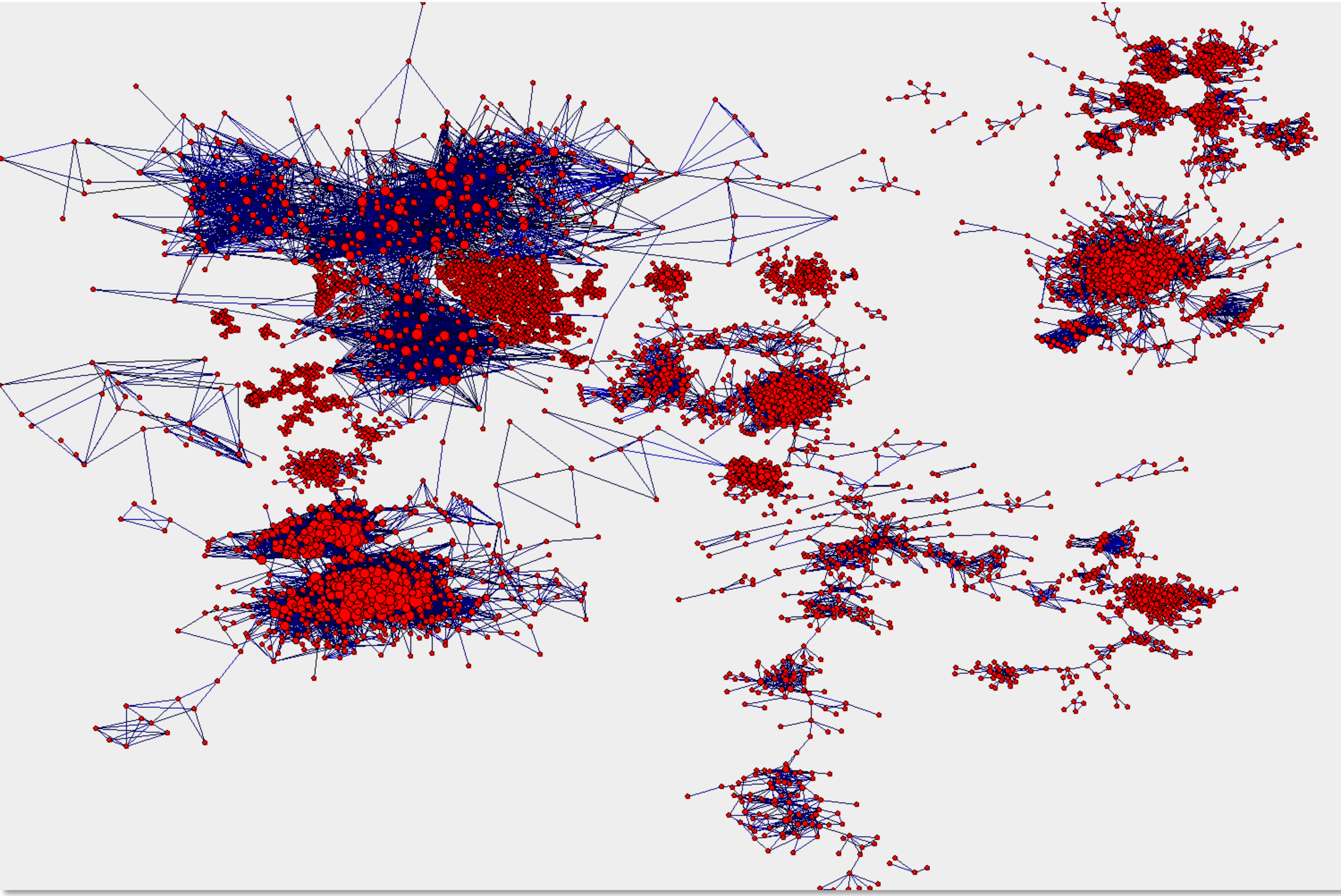


Photo Explorer

Can we reconstruct entire cities?


flickr
Signed in as Jimantha
Help
Sign Out
Home
You
Organize
Contacts
Groups
Explore
Search everyone's phot
Search
Search
Photos
Groups
People
Can't see your photos? Find out why...
Everyone's Photos
rome or roma
SEARCH
Advanced Search
Search by Camera
Full text
Tags only
We found 2,379,801 results matching rome or roma.
View as slideshow
View: Most relevant • Most recent • Most interesting
Show: Details • Thumbnails
From Kipourax
From MrMass
From rocdam
From alessandro...
Rome
From * Toshio *
From
From Optical...
From egold
From egold
From egold
From donato.chiru...
From cuellar
From egold
From Aquilant
From Peter...
From SDBryan
From cuellar
From david.bank



Gigantic matching problem

- 1,000,000 images → 500,000,000,000 pairs
  - Matching all of these on a 1,000-node cluster would take more than a year, even if we match 10,000 every second
  - And involves TBs of data

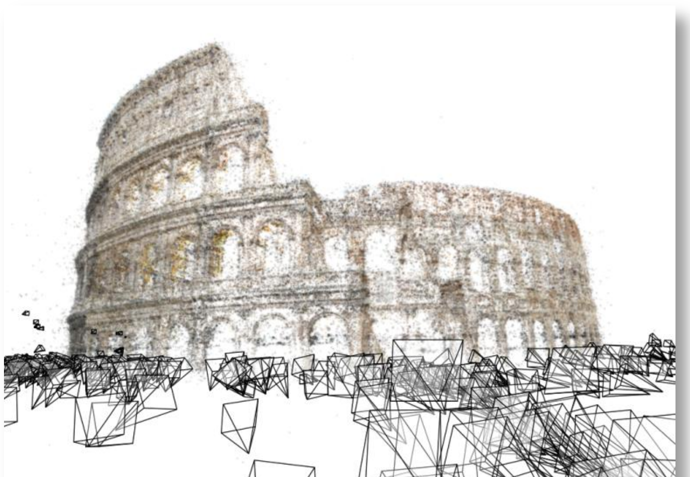
- The vast majority (>99%) of image pairs do *not* match
- There are better ways of finding matching images (more on this later)

Gigantic SfM problem

- Largest problem size we've seen:
  - 15,000 cameras
  - 4 million 3D points
  - more than 12 million parameters
  - more than 25 million equations
- Huge optimization problem
- Requires sparse least squares techniques

Building Rome in a Day

St. Peter's Basilica



Colosseum

Trevi Fountain

Rome, Italy. Reconstructed 150,000 in 21 hours on 496 machines

Dubrovnik, Croatia. 4,619 images (out of an initial 57,845).
Total reconstruction time: 23 hours
Number of cores: 352

Dubrovnik

Dubrovnik, Croatia. 4,619 images (out of an initial 57,845).
Total reconstruction time: 23 hours
Number of cores: 352

San Marco Square

San Marco Square and environs, Venice. 14,079 photos, out of an initial 250,000.
Total reconstruction time: 3 days. Number of cores: 496.

Libration

From Wikipedia, the free encyclopedia

Not to be confused with Liberation or Libation.

In astronomy **libration** (from the Latin verb *librare* "to balance, to sway", cf. *libra* "scales") refers to the various orbital conditions which make it possible to see more than 50% of the moon's surface over time, even though the front of the Moon is tidally locked to always face towards Earth. By extension, libration can also be used to describe the same phenomenon for other orbital bodies that are nominally locked to present the same face. As the orbital processes are repetitive, libration is manifested as a slow rocking back and forth (or up and down) of the face of the orbital body as viewed from the parent body, much like the rocking of a pair of scales about the point of balance.

Simulated views of the Moon over one month, demonstrating librations in latitude and longitude.

In the specific case of the Moon's librations, this motion permits a terrestrial observer to see slightly differing halves of the Moon's surface at different times. This means that a total of 59% of the Moon's surface can be observed from Earth.

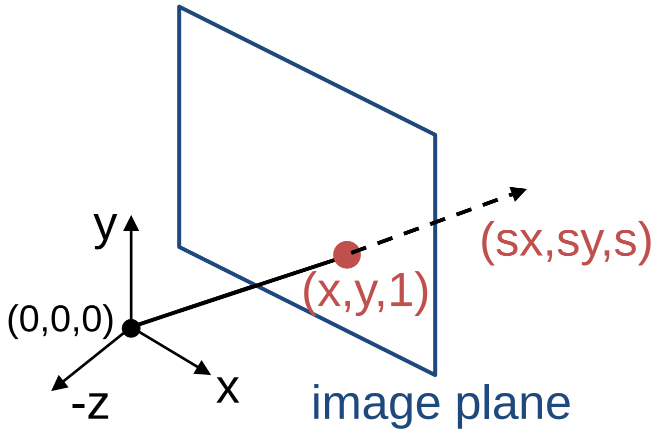
There are three types of libration:

- *Libration in longitude* is a consequence of the Moon's orbit around Earth being somewhat eccentric, so that the Moon's rotation sometimes leads and sometimes lags its orbital position.
- *Libration in latitude* is a consequence of the Moon's axis of rotation being slightly inclined to the normal to the plane of its orbit around Earth. Its origin is analogous to the way in which the seasons arise from Earth's revolution about the Sun.
- *Diurnal libration* is a small daily oscillation due to the Earth's rotation, which carries an observer first to one side and then to the other side of the straight line joining Earth's center to the Moon's center, allowing the observer to look first around one side of the Moon and then around the other. This is because the observer is on the surface of the Earth, not at its centre.

Questions?

A few more preliminaries

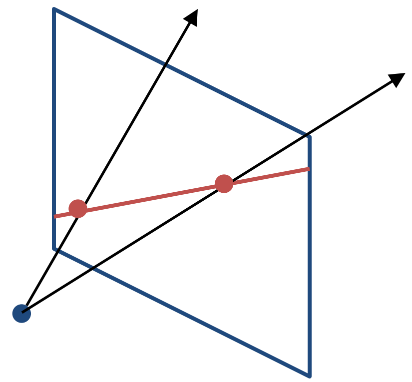
- Let’s revisit homogeneous coordinates
- What is the geometric intuition?
  – a point in the image is a *ray* in projective space



- Each *point* (x,y) on the plane is represented by a *ray* (sx,sy,s)
  – all points on the ray are equivalent: (x, y, 1) ≅ (sx, sy, s)

Projective lines

- What does a line in the image correspond to in projective space?



- A line is a *plane* of rays through origin
  - all rays (x,y,z) satisfying: ax + by + cz = 0

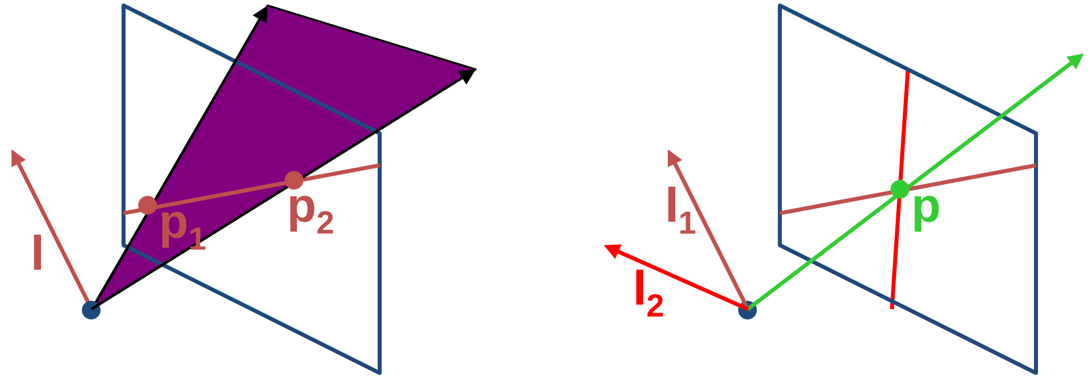
$$\text{in vector notation}: \quad 0 = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$\mathbf{l}$ $\quad$ $\mathbf{p}$

- A line is also represented as a homogeneous 3-vector **l**

Point and line duality

- A line **l** is a homogeneous 3-vector
- It is ⊥ to every point (ray) **p** on the line: **l·p**=0



What is the line **l** spanned by rays $\mathbf{p_1}$ and $\mathbf{p_2}$?

- **l** is ⊥ to $\mathbf{p_1}$ and $\mathbf{p_2}$ $\Rightarrow$ $\mathbf{l} = \mathbf{p_1} \times \mathbf{p_2}$
- **l** is the plane normal

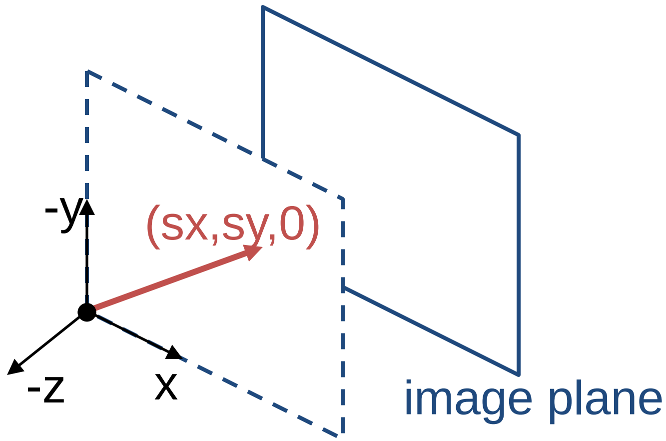
What is the intersection of two lines $\mathbf{l_1}$ and $\mathbf{l_2}$?

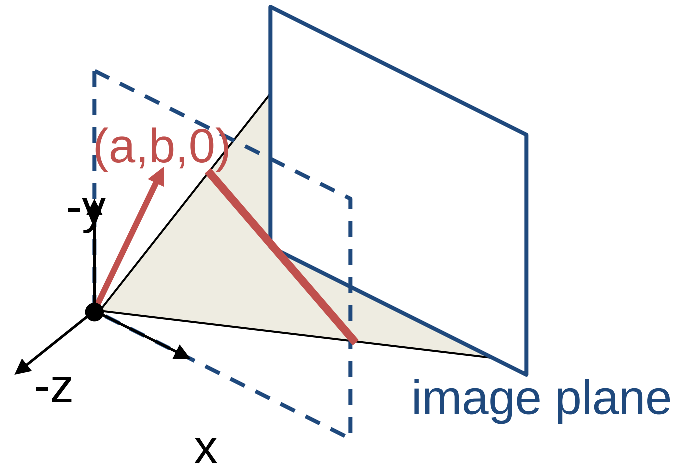
- **p** is ⊥ to $\mathbf{l_1}$ and $\mathbf{l_2}$ $\Rightarrow$ $\mathbf{p} = \mathbf{l_1} \times \mathbf{l_2}$

Points and lines are *dual* in projective space

- given any formula, can switch the meanings of points and lines to get another formula

Ideal points and lines

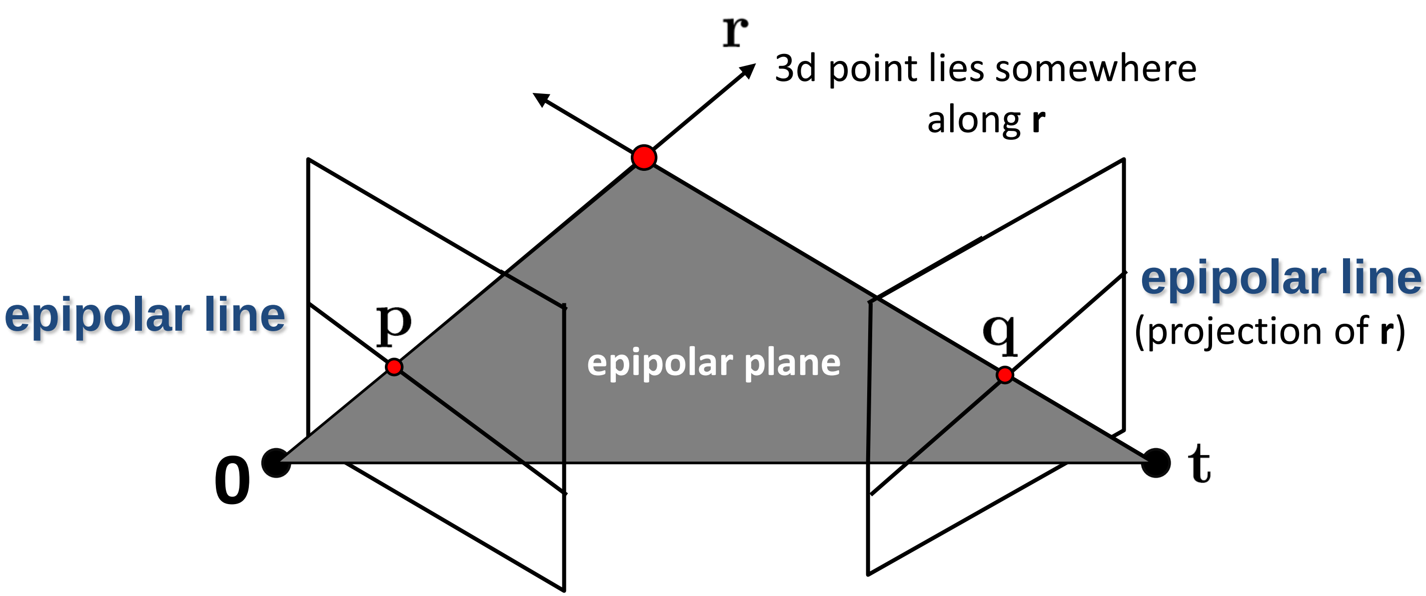




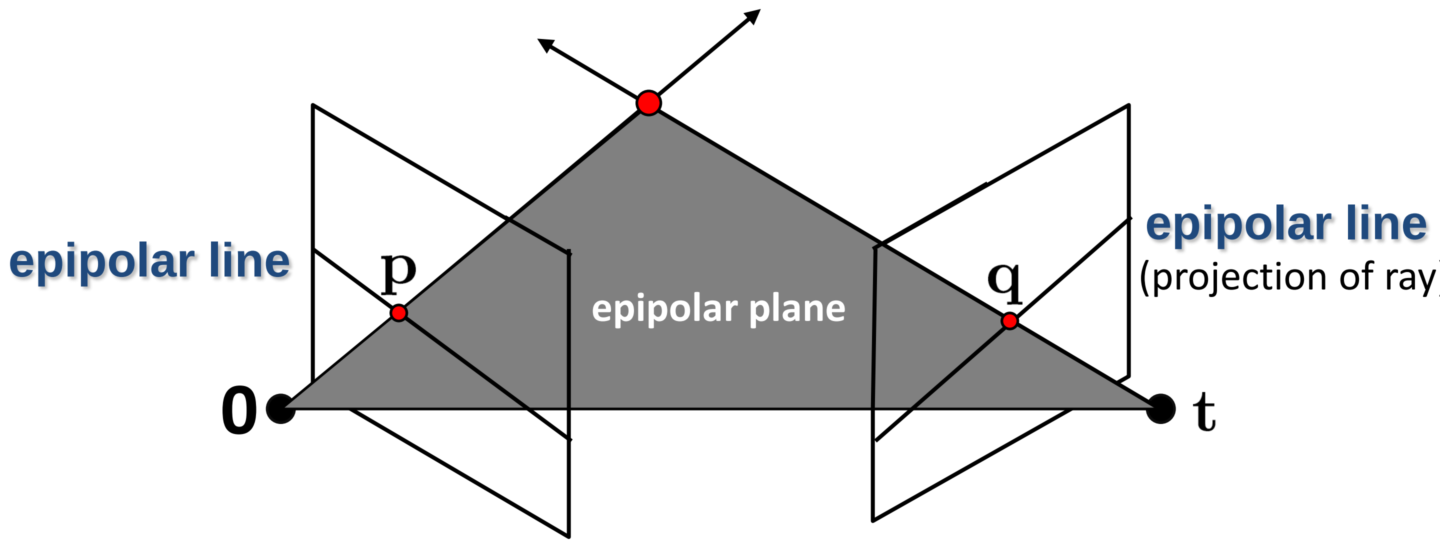
- Ideal point ("point at infinity")
  - p ≅ (x, y, 0) – parallel to image plane
  - It has infinite image coordinates
- Ideal line
  - l ≅ (a, b, 0) – parallel to image plane
  - Corresponds to a line in the image (finite coordinates)
    - goes through image origin (*principle point*)

Two-view geometry

- Where do epipolar lines come from?

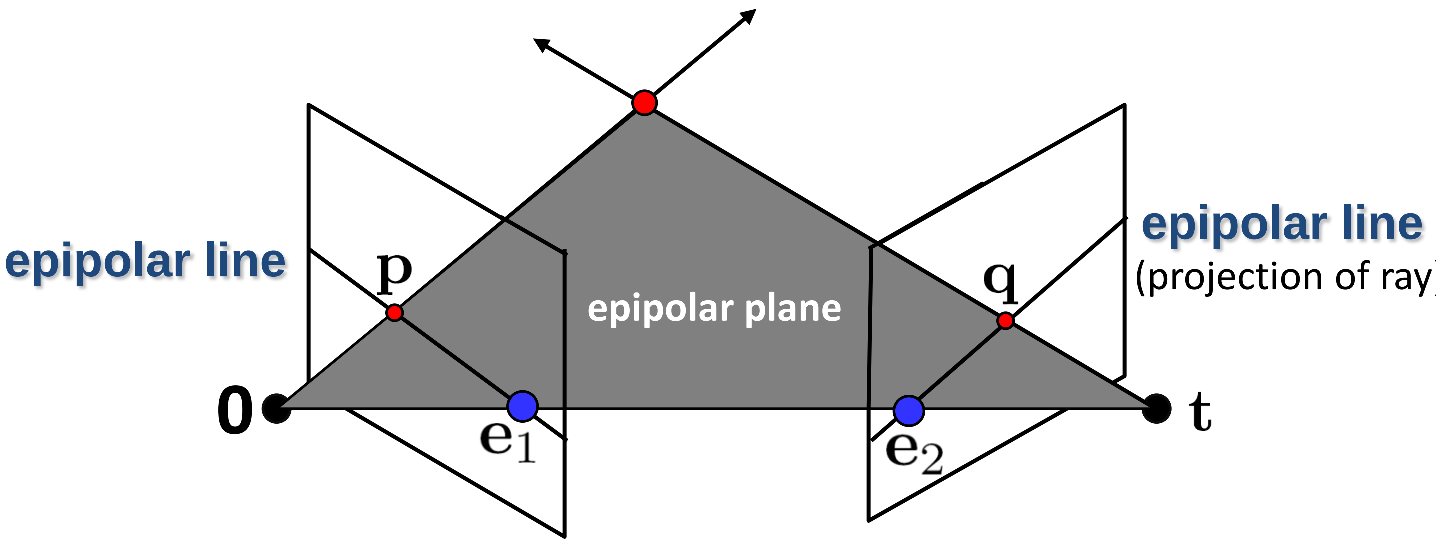


Fundamental matrix



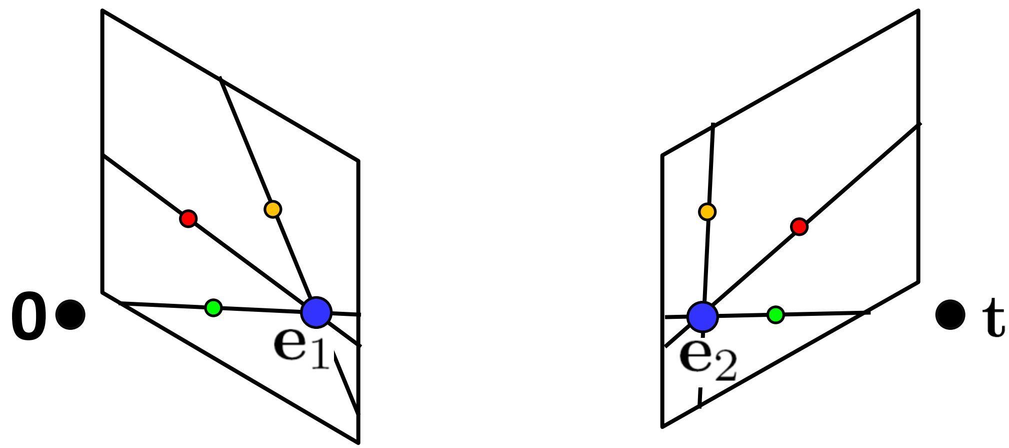
- This *epipolar geometry* is described by a special 3x3 matrix $\mathbf{F}$, called the *fundamental matrix*
- Epipolar constraint on corresponding points: $\mathbf{q}^T\mathbf{F}\mathbf{p} = 0$
- The epipolar line of point **p** in image **J** is: $\mathbf{F}\mathbf{p}$

Fundamental matrix



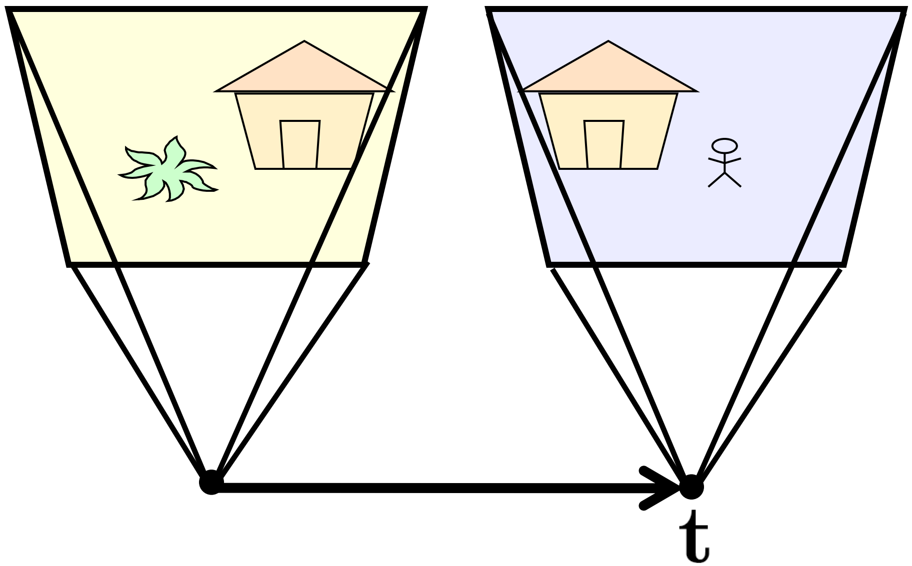
- Two special points: $\mathbf{e}_1$ and $\mathbf{e}_2$ (the *epipoles*): projection of one camera into the other

Fundamental matrix



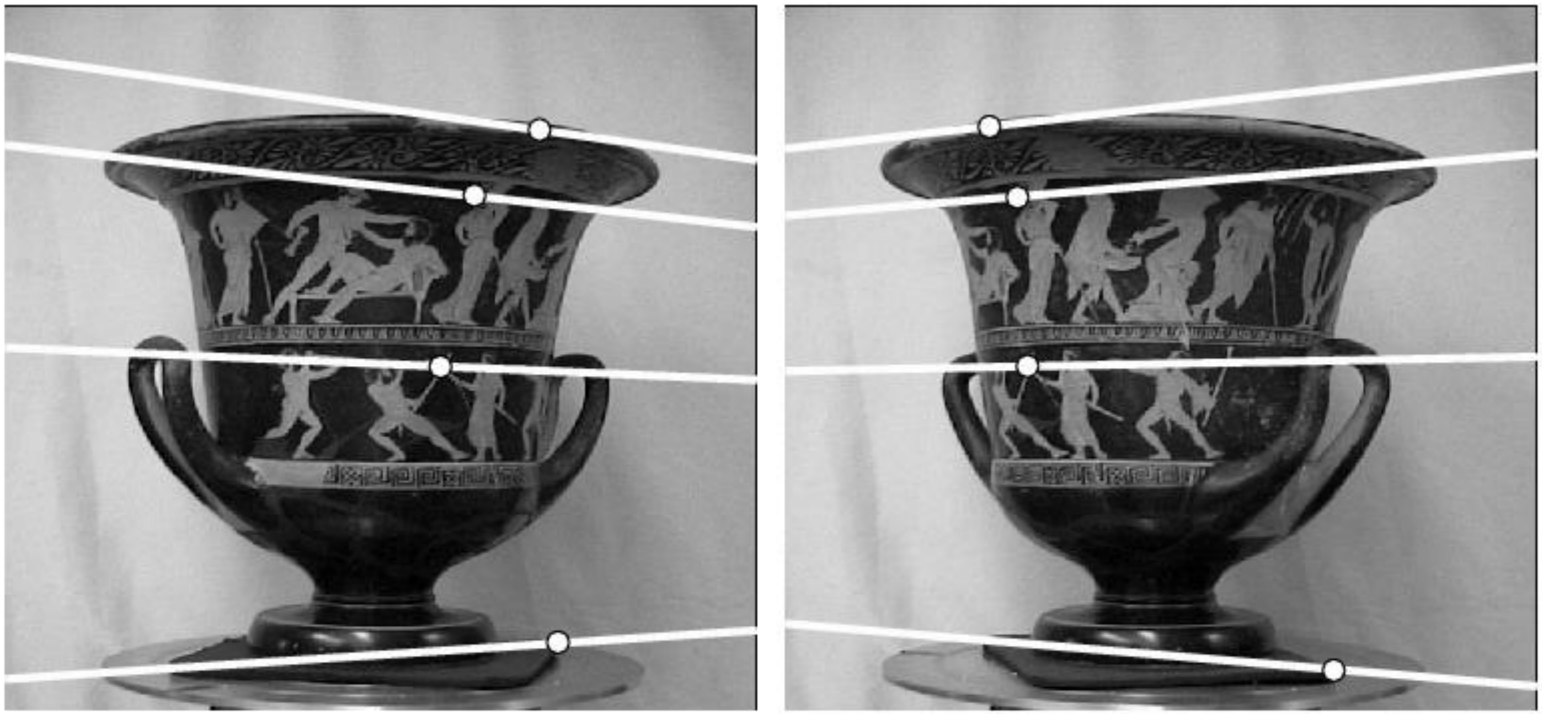
- Two special points: $\mathbf{e}_1$ and $\mathbf{e}_2$ (the *epipoles*): projection of one camera into the other

Rectified case

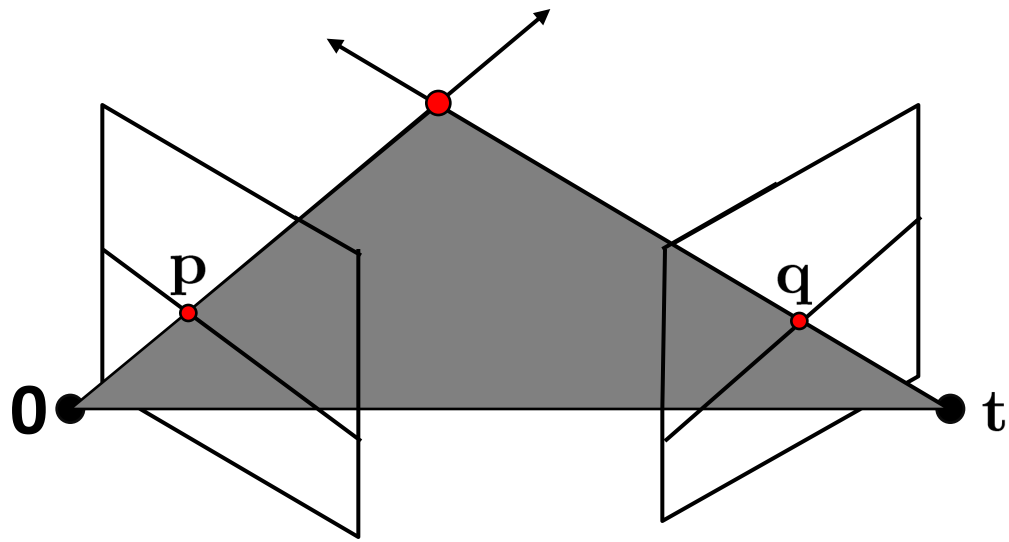


- Images have the same orientation, **t** parallel to image planes
- Where are the epipoles?

Epipolar geometry demo

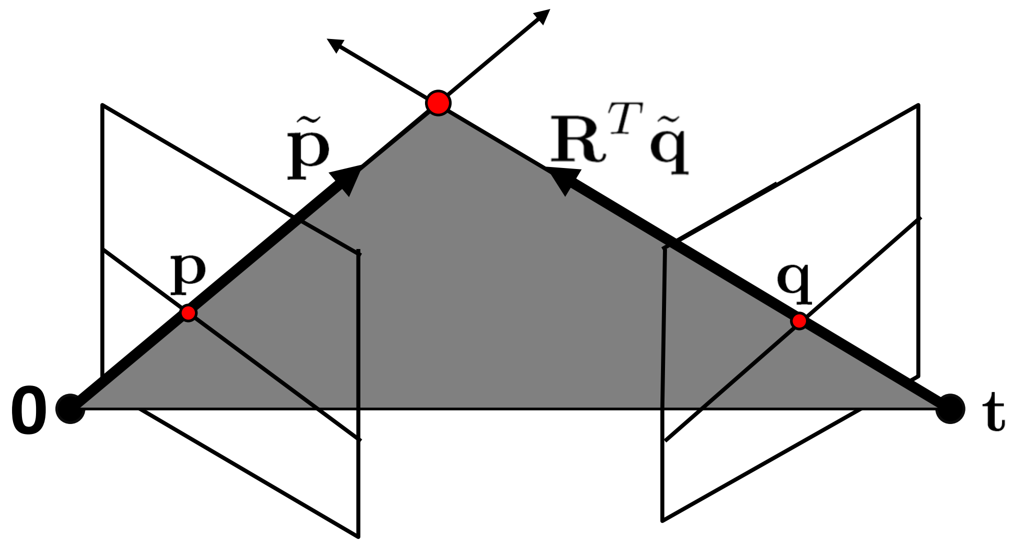


Fundamental matrix



- Why does **F** exist?
- Let's derive it…

Fundamental matrix – calibrated case



$\mathbf{K}_1$: intrinsics of camera 1

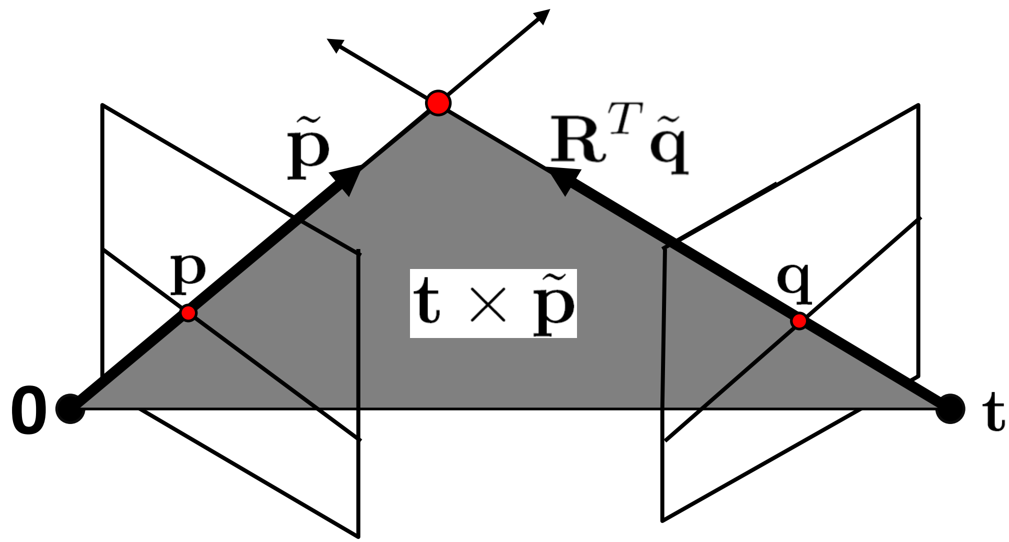
$\mathbf{K}_2$: intrinsics of camera 2

$\mathbf{R}$: rotation of image 2 w.r.t. camera 1

$\tilde{\mathbf{p}} = \mathbf{K}_1^{-1}\mathbf{p}$: ray through **p** in camera 1's (and world) coordinate system

$\tilde{\mathbf{q}} = \mathbf{K}_2^{-1}\mathbf{q}$: ray through **q** in camera 2's coordinate system

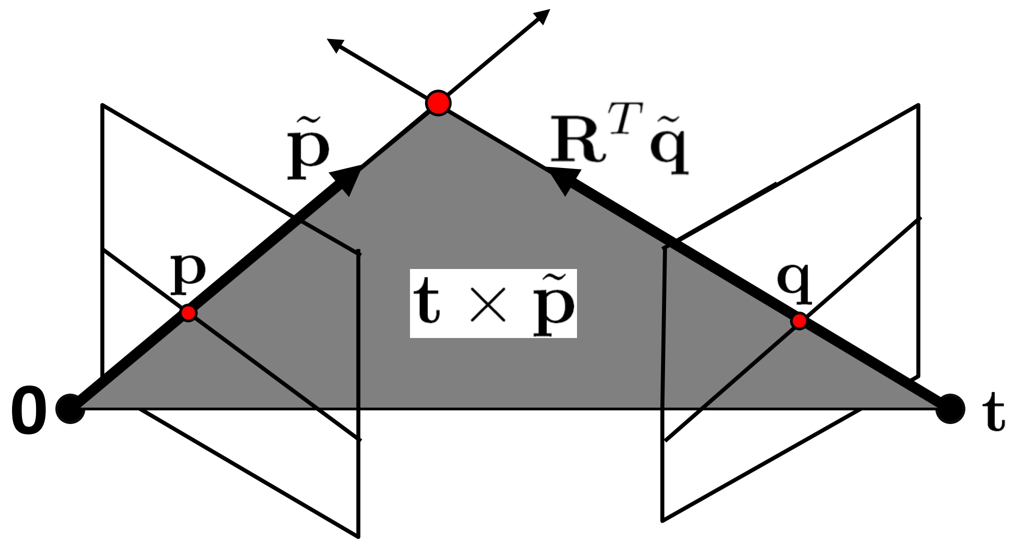
Fundamental matrix – calibrated case



- $\tilde{\mathbf{p}}$, $\mathbf{R}^T\tilde{\mathbf{q}}$, and $\mathbf{t}$ are coplanar
- epipolar plane can be represented as $\mathbf{t} \times \tilde{\mathbf{p}}$

$$(\mathbf{R}^T\tilde{\mathbf{q}})^T(\mathbf{t} \times \tilde{\mathbf{p}}) = 0$$

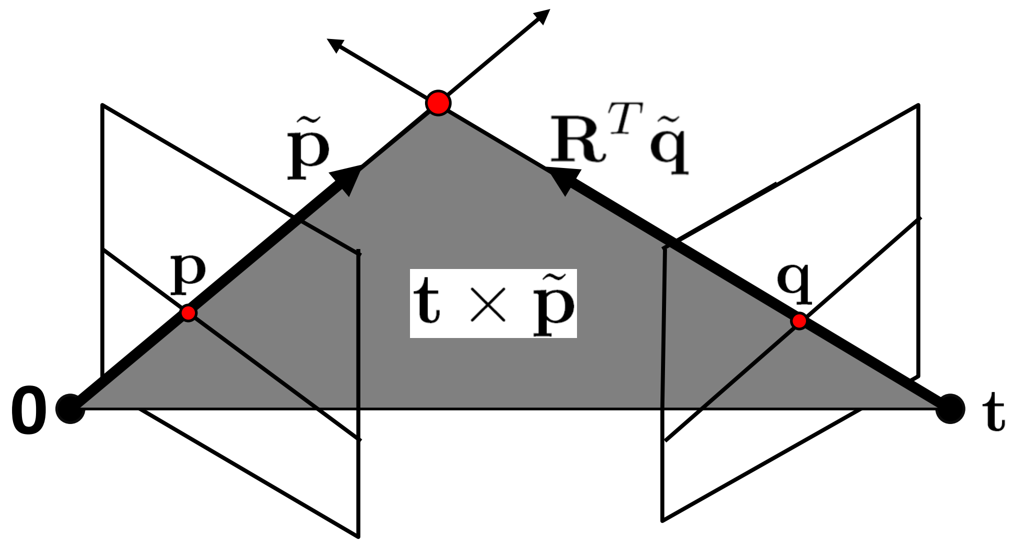
Fundamental matrix – calibrated case



$$(\mathbf{R}^T\tilde{\mathbf{q}})^T(\mathbf{t} \times \tilde{\mathbf{p}}) = 0$$

$$\tilde{\mathbf{q}}^T\mathbf{R}(\mathbf{t} \times \tilde{\mathbf{p}}) = 0$$

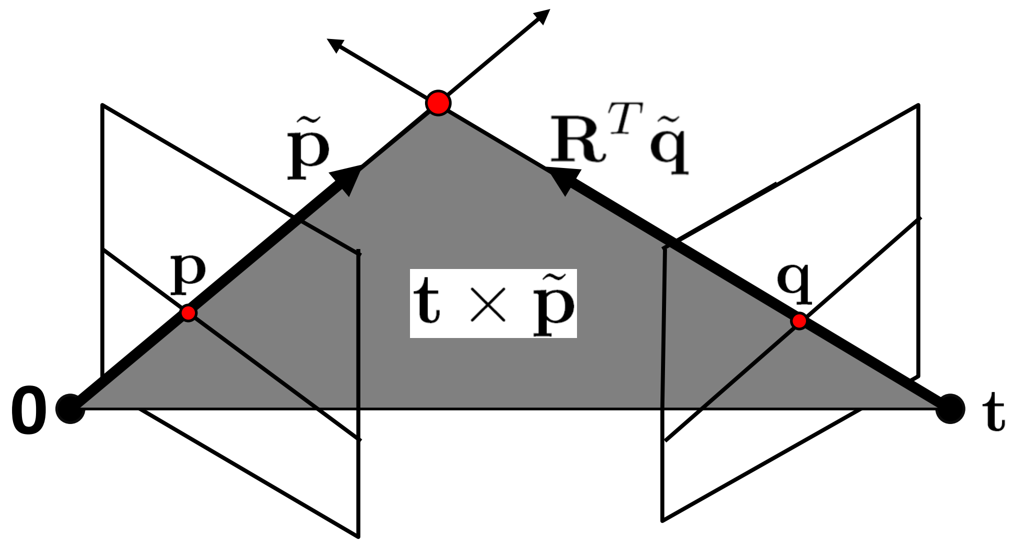
Fundamental matrix – calibrated case



- One more substitution:
  - Cross product with **t** can be represented as a 3x3 matrix

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \qquad \mathbf{t} \times \tilde{\mathbf{p}} = [\mathbf{t}]_\times \tilde{\mathbf{p}}$$

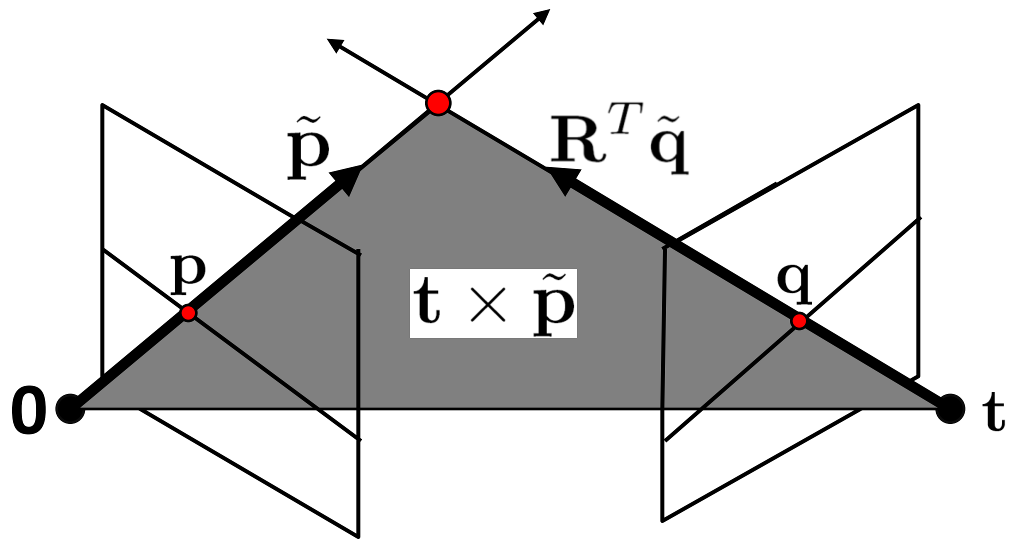
Fundamental matrix – calibrated case



$$\tilde{\mathbf{q}}^T \mathbf{R}(\mathbf{t} \times \tilde{\mathbf{p}}) = 0$$

$$\tilde{\mathbf{q}}^T \mathbf{R} \left[\mathbf{t}\right]_\times \tilde{\mathbf{p}} = 0$$

Fundamental matrix – calibrated case

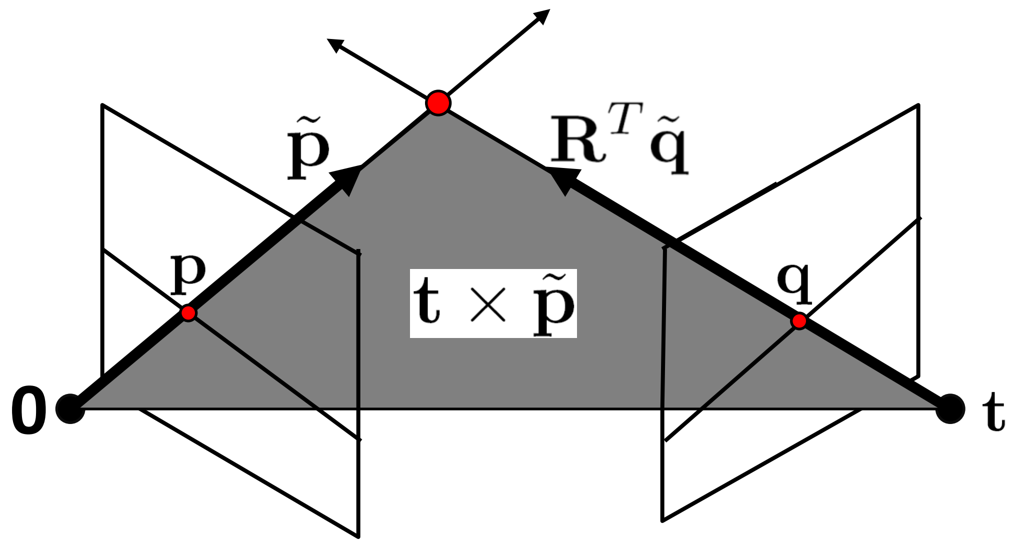


$$\tilde{\mathbf{q}}^T \underbrace{\mathbf{R}\,[\mathbf{t}]_\times}_{\mathbf{E}} \tilde{\mathbf{p}} = 0$$

$$\tilde{\mathbf{q}}^T \mathbf{E} \tilde{\mathbf{p}} = 0$$

the Essential matrix

Fundamental matrix – uncalibrated case



$$\tilde{\mathbf{q}}^T \mathbf{R} [\mathbf{t}]_\times \tilde{\mathbf{p}} = 0$$

$$\mathbf{q}^T \underbrace{\mathbf{K}_2^{-T} \mathbf{R} [\mathbf{t}]_\times \mathbf{K}_1^{-1}}_{\mathbf{F}} \mathbf{p} = 0$$

$\mathbf{F}$ ← the Fundamental matrix

Properties of the Fundamental Matrix

- $\mathbf{F}\mathbf{p}$ is the epipolar line associated with $\mathbf{p}$
- $\mathbf{F}^T\mathbf{q}$ is the epipolar line associated with $\mathbf{q}$
- $\mathbf{F}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{F}^T\mathbf{e}_2 = \mathbf{0}$
- $\mathbf{F}$ is rank 2
- How many parameters does **F** have?