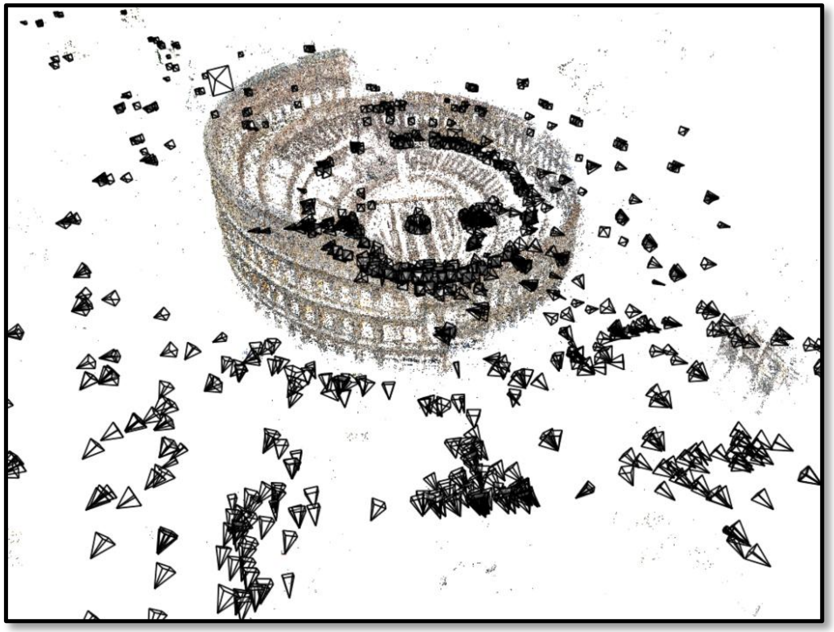
CS6670: Computer Vision

Noah Snavely

Lecture 22: Structure from motion



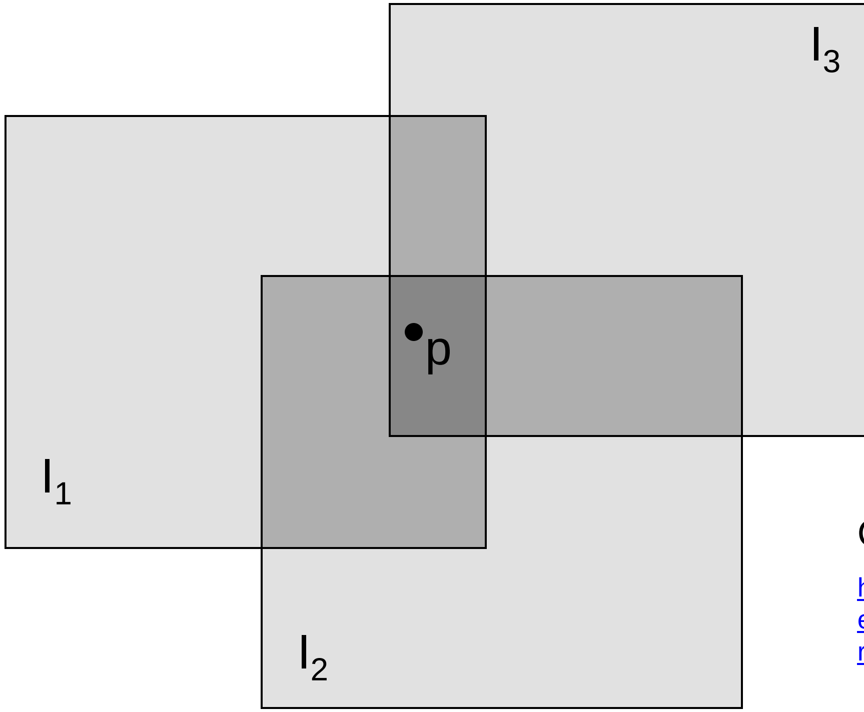
Readings

- Szeliski, Chapter 7.1 – 7.4

Announcements

- Project 2b due on Tuesday by 10:59pm
- Final project proposals due today at 11:59pm
- Course survey

Alpha Blending



Optional: see Blinn (CGA, 1994) for details:

http://ieeexplore.ieee.org/iel1/38/7531/00310740.pdf?isNumber=7531&prod=JNL&arnumber=310740&arSt=83&ared=87&arAuthor=Blinn%2C+J.F.

Encoding blend weights: $I(x,y) = (\alpha R, \alpha G, \alpha B, \alpha)$

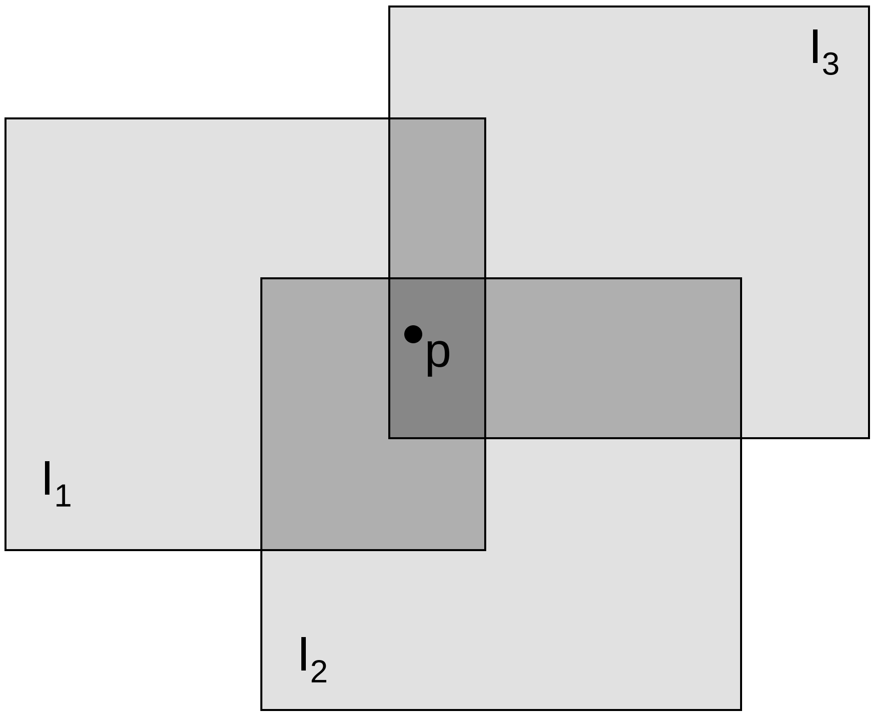
color at p = $$\frac{(\alpha_1 R_1,\ \alpha_1 G_1,\ \alpha_1 B_1) + (\alpha_2 R_2,\ \alpha_2 G_2,\ \alpha_2 B_2) + (\alpha_3 R_3,\ \alpha_3 G_3,\ \alpha_3 B_3)}{\alpha_1 + \alpha_2 + \alpha_3}$$

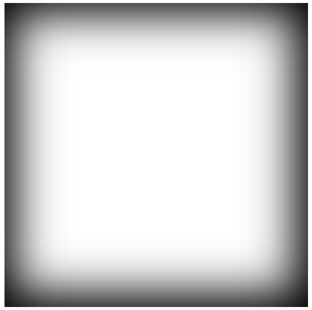
Implement this in two steps:

1. accumulate: add up the (α premultiplied) RGBα values at each pixel
2. normalize: divide each pixel's accumulated RGB by its α value

Q: what if $\alpha = 0$?

Blend weights





Feathered alpha channel

How should we set the alpha values of I_1, I_2, I_3 ?

Simplest choice: set all alpha values to one (gives discontinuities)

Better choice: use *feathering* to ramp alpha values to zero near the edges

What about more than two views?

- The geometry of three views is described by a 3 x 3 x 3 tensor called the *trifocal tensor*
- The geometry of four views is described by a 3 x 3 x 3 x 3 tensor called the *quadrifocal tensor*
- After this it starts to get complicated…

New approach

- These matrices, tensors, etc, model geometry as a matrix that depends (in complicated ways) on the camera parameters alone
- Instead, we will *explicitly* model both cameras and points

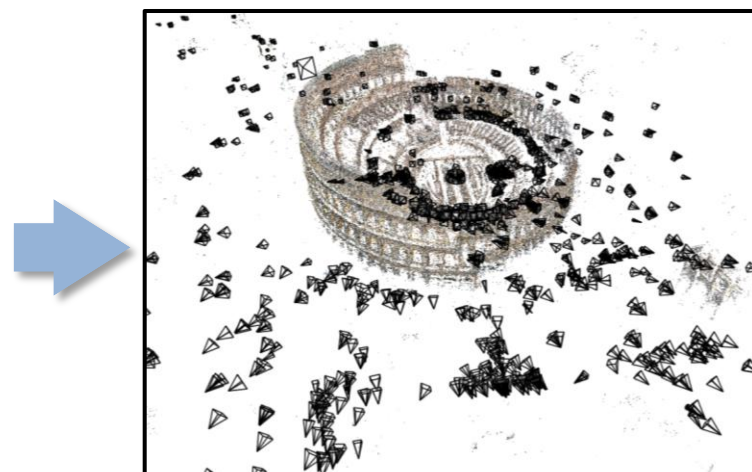
Large-scale structure from motion

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

Questions?

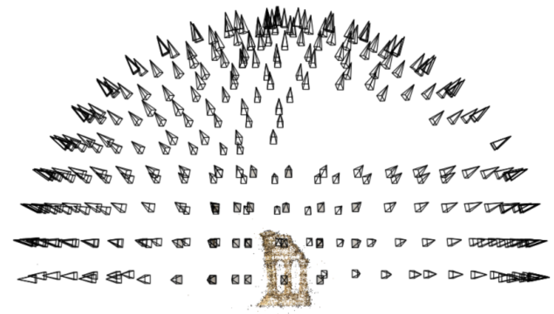
Structure from motion

- Given many images, how can we
  a) figure out where they were all taken from?
  b) build a 3D model of the scene?

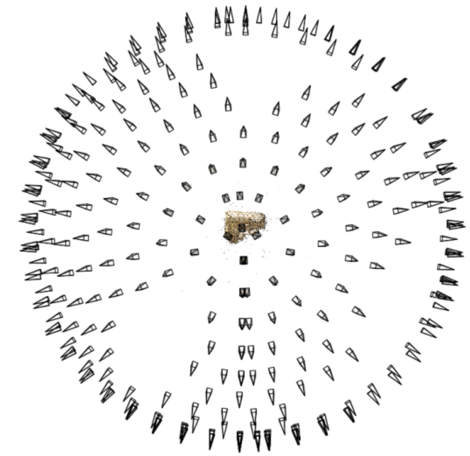


This is (roughly) the **structure from motion** problem

Structure from motion



Reconstruction (side)



(top)

- Input: images with points in correspondence
  $p_{i,j} = (u_{i,j}, v_{i,j})$

- Output
  - structure: 3D location $\mathbf{x}_i$ for each point p_i
  - motion: camera parameters $\mathbf{R}_j$, $\mathbf{t}_j$ possibly $\mathbf{K}_j$

- Objective function: minimize *reprojection error*

Camera calibration and triangulation

- Suppose we know 3D points
  - And have matches between these points and an image
  - How can we compute the camera parameters?

- Suppose we have know camera parameters, each of which observes a point
  - How can we compute the 3D location of that point?

Structure from motion

- SfM solves both of these problems *at once*
- A kind of chicken-and-egg problem
  - (but solvable)

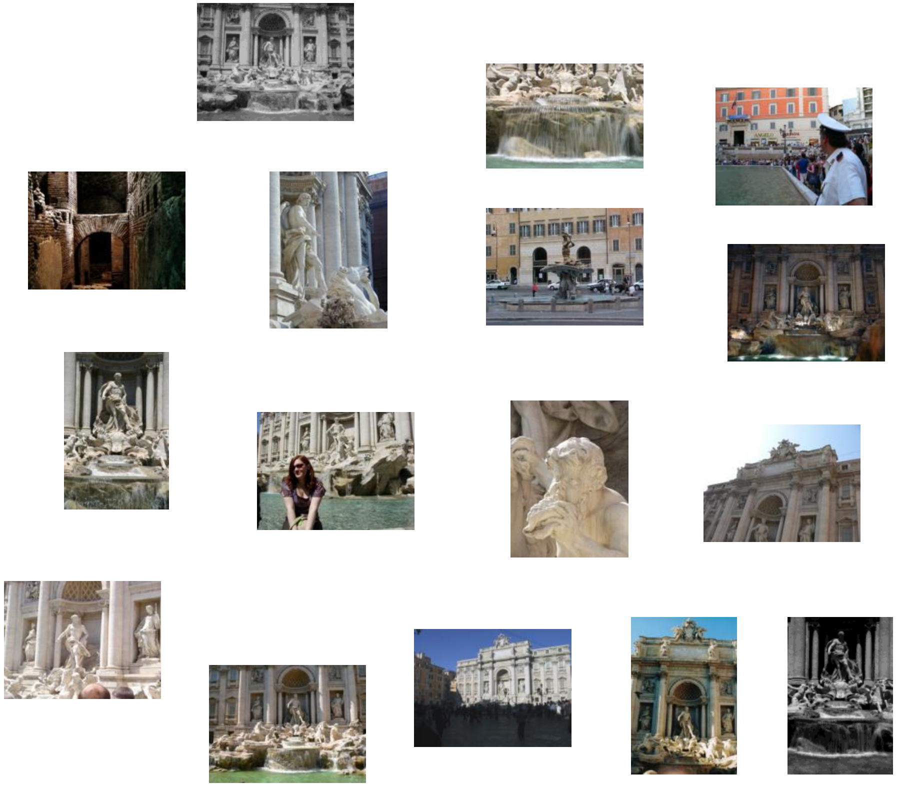
Photo Tourism

First step: how to get correspondence?

- Feature detection and matching

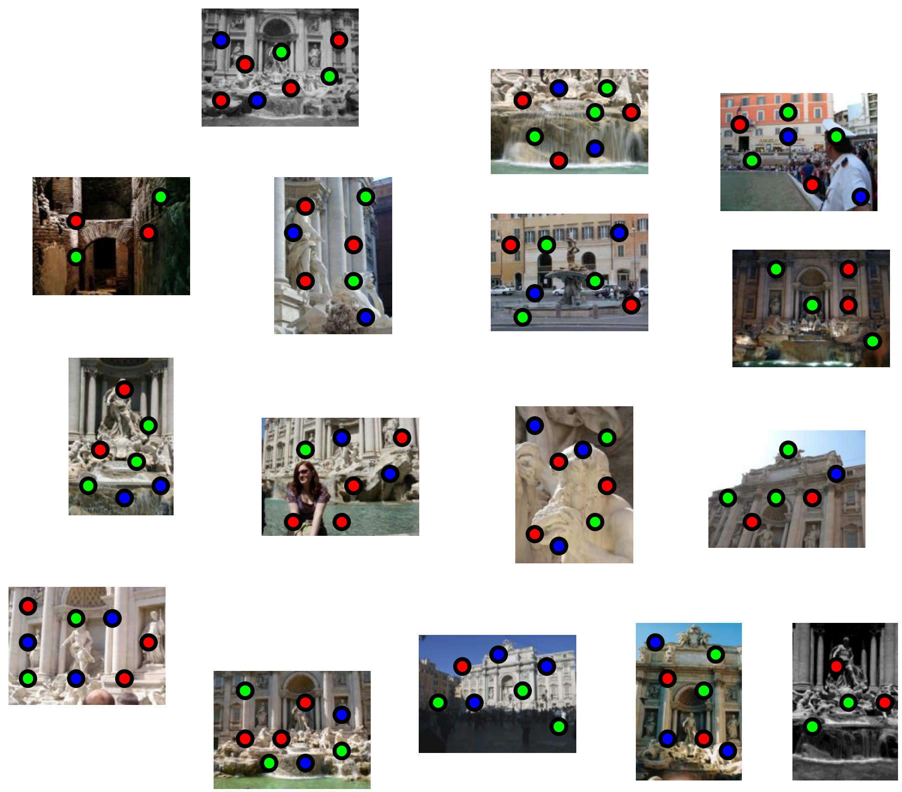
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



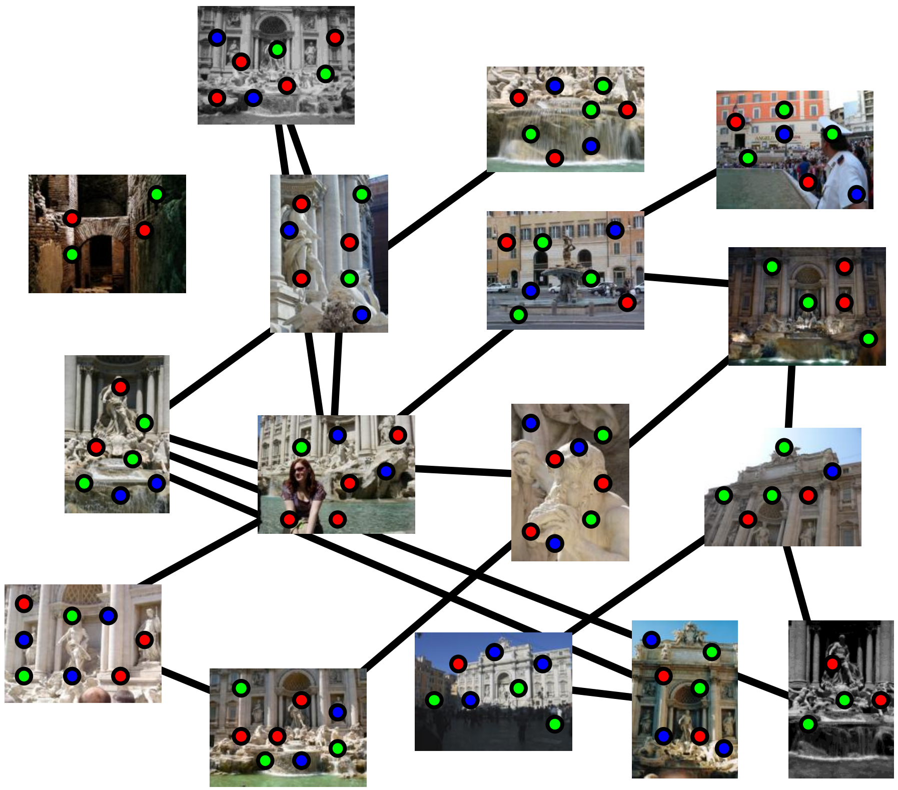
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



Feature matching

Match features between each pair of images



Feature matching

Refine matching using RANSAC to estimate fundamental matrix between each pair

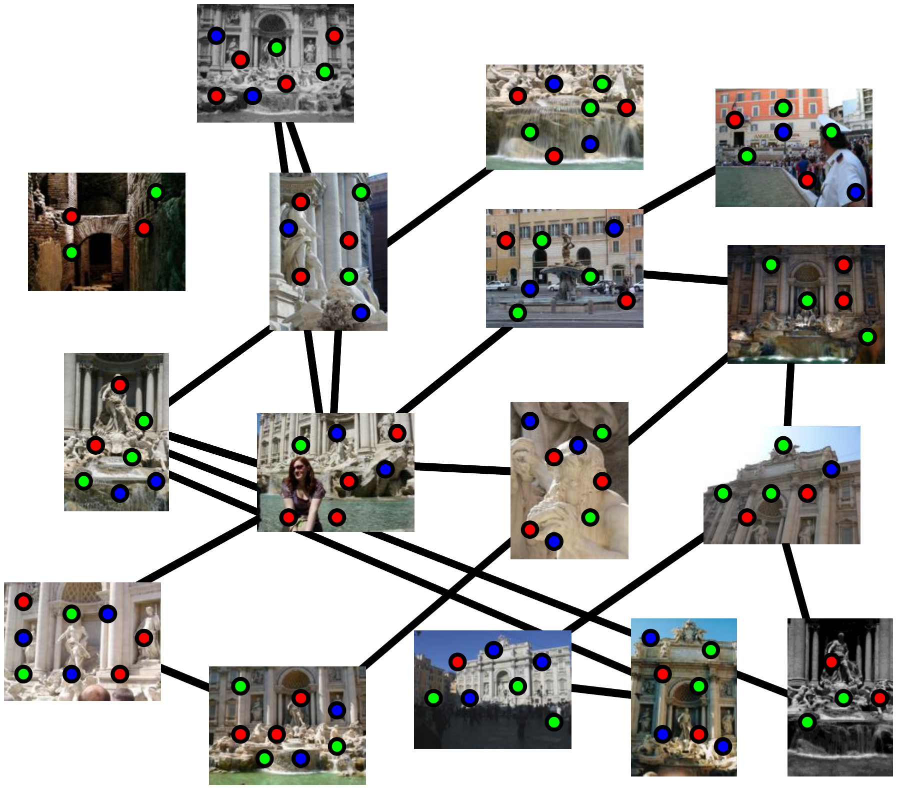
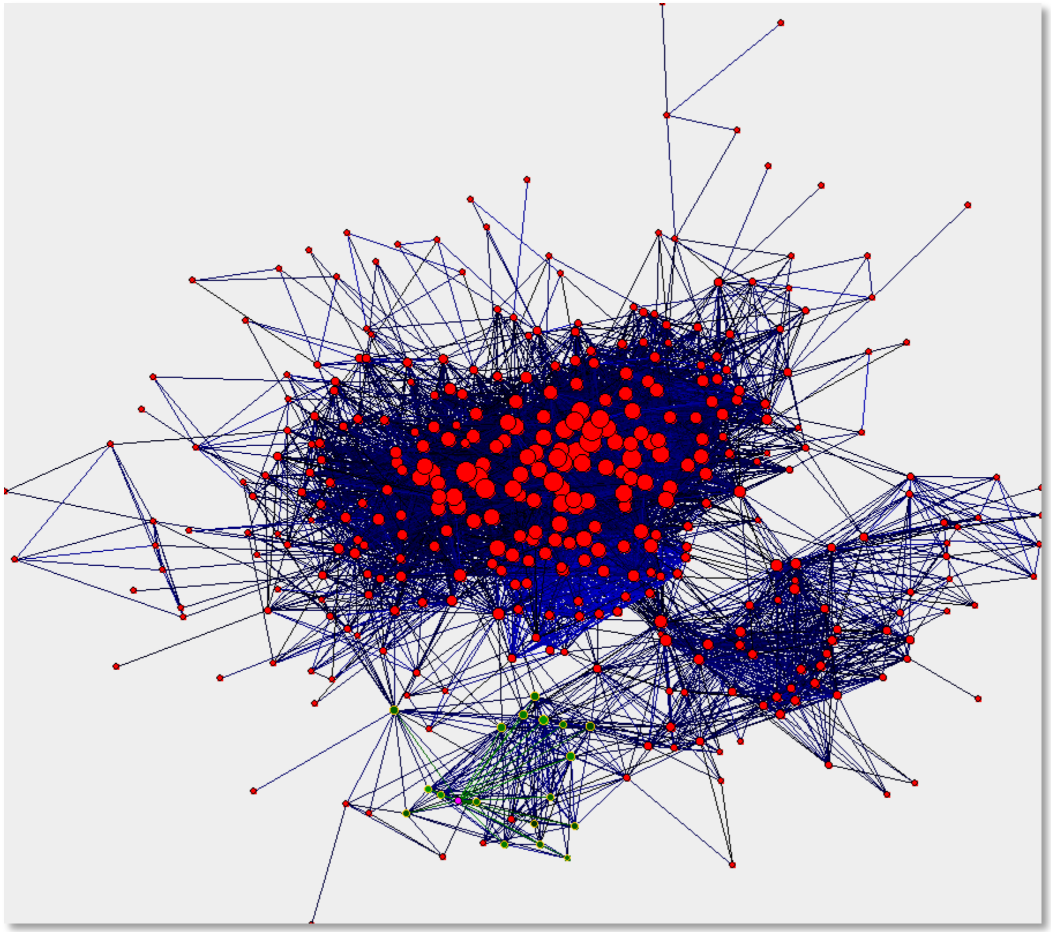
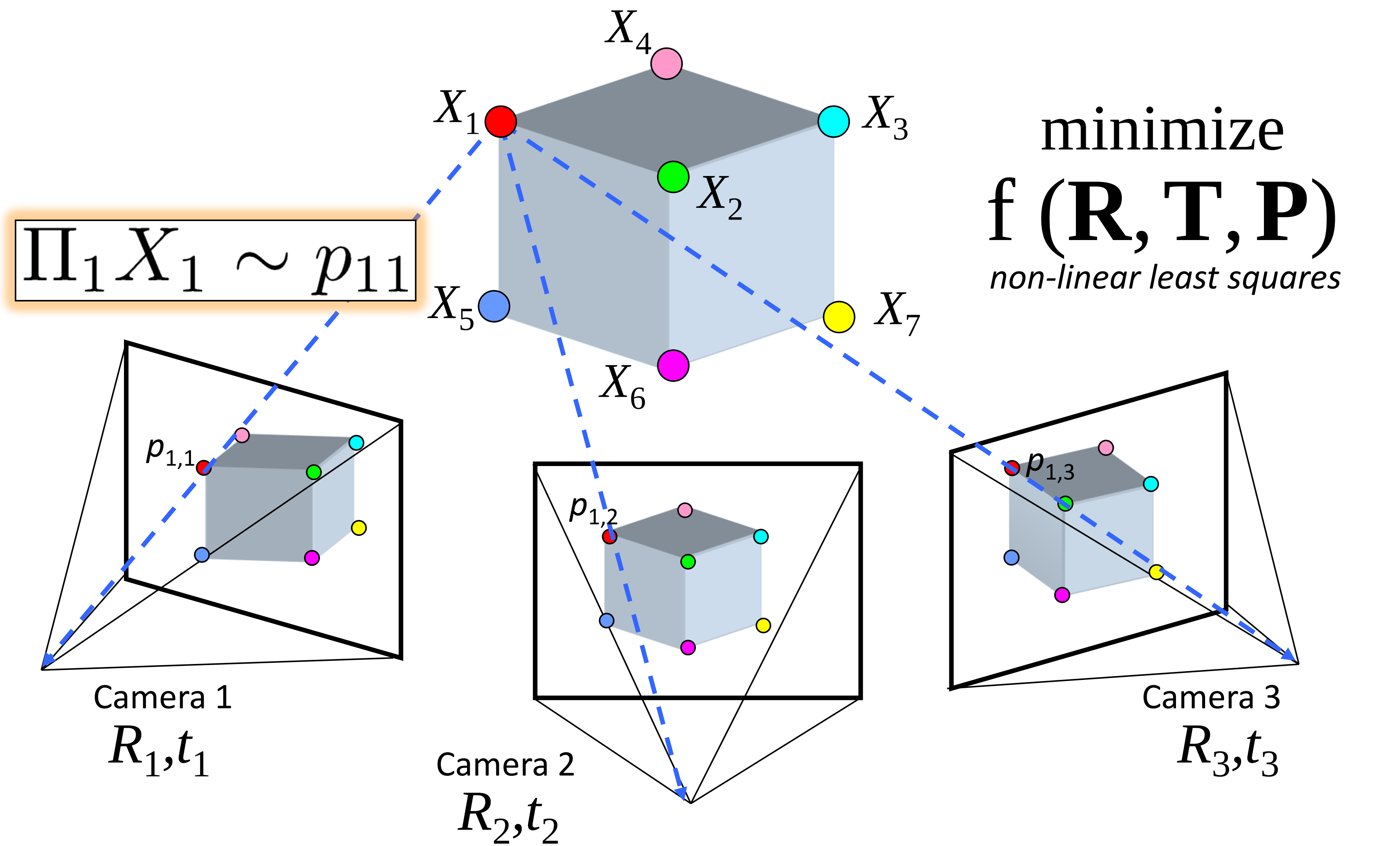


Image connectivity graph



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

Structure from motion



Problem size

- What are the variables?
- How many variables per camera?
- How many variables per point?

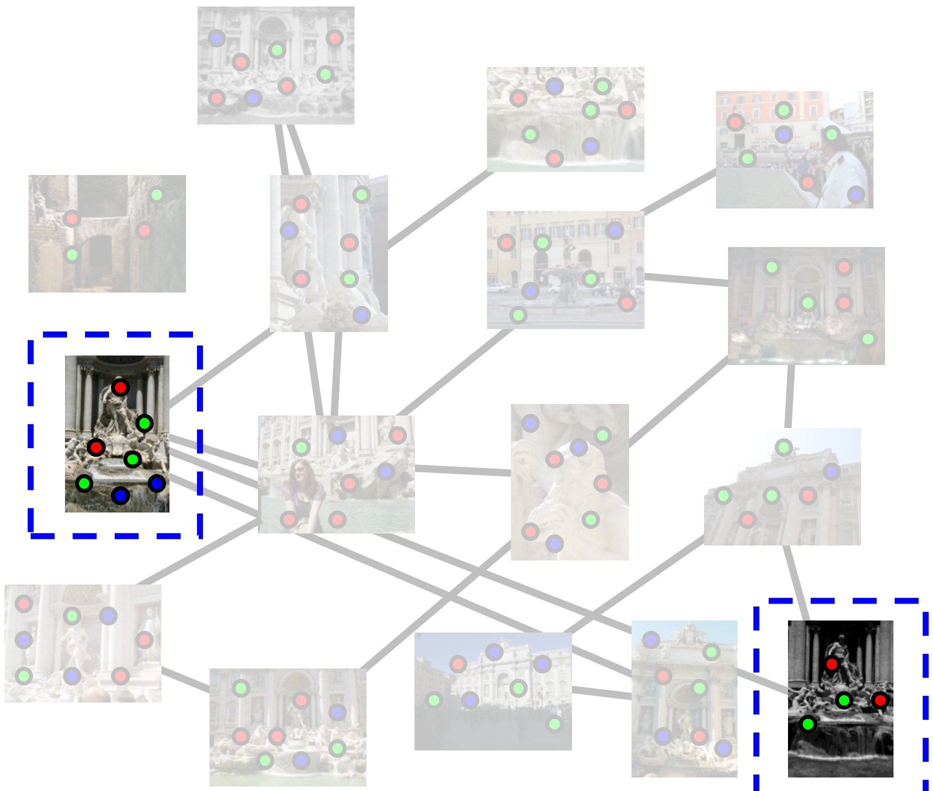
- Trevi Fountain collection

466 input photos

\+ > 100,000 3D points

= very large optimization problem

Incremental structure from motion



Incremental structure from motion

Incremental structure from motion

Photo Explorer

Demo

WIKIPEDIA
The Free Encyclopedia

navigation
- Main page
- Contents
- Featured content
- Current events
- Random article

search
Go Search

interaction
- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

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?