

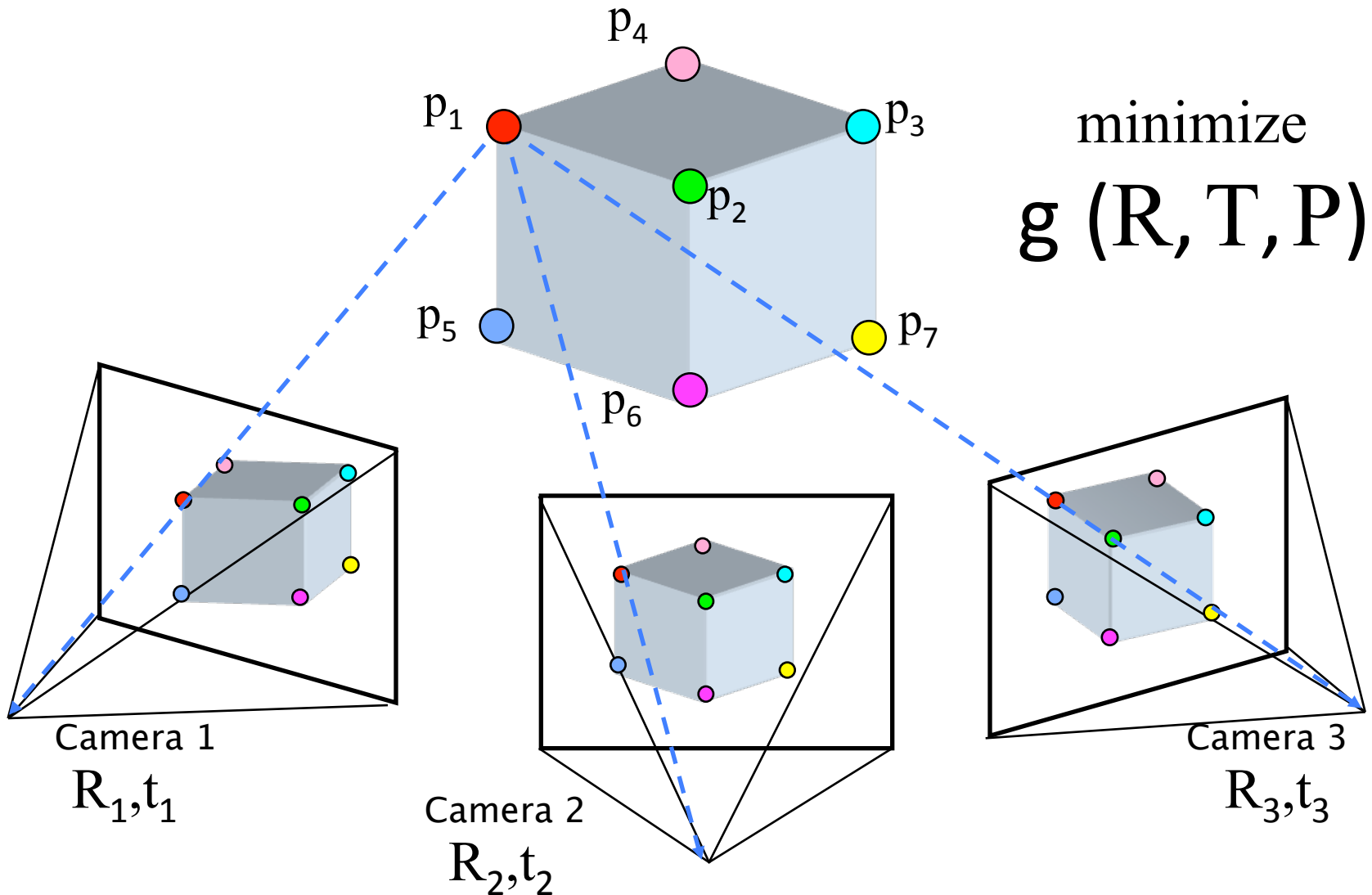
# Structure From Motion

Ali Farhadi

CSE 576

# Structure from motion

- aka “bundle adjustment” (texts: [Zisserman](#); [Faugeras](#))



# SfM objective function

---

- Given point  $\mathbf{x}$  and rotation and translation  $\mathbf{R}, \mathbf{t}$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathbf{R}\mathbf{x} + \mathbf{t} \quad \begin{matrix} u' = \frac{fx'}{z'} \\ v' = \frac{fy'}{z'} \end{matrix} \quad \begin{bmatrix} u' \\ v' \end{bmatrix} = \mathbf{P}(\mathbf{x}, \mathbf{R}, \mathbf{t})$$

- Minimize sum of squared reprojection errors:

$$g(\mathbf{X}, \mathbf{R}, \mathbf{T}) = \sum_{i=1}^m \sum_{j=1}^n w_{ij} \cdot \left\| \underbrace{\mathbf{P}(\mathbf{x}_i, \mathbf{R}_j, \mathbf{t}_j)}_{\text{predicted image location}} - \underbrace{\begin{bmatrix} u_{i,j} \\ v_{i,j} \end{bmatrix}}_{\text{observed image location}} \right\|^2$$

# Solving structure from motion

---

- Minimizing  $g$  is difficult:
  - $g$  is non-linear due to rotations, perspective division
  - lots of parameters: 3 for each 3D point, 6 for each camera
  - difficult to initialize
  - gauge ambiguity: error is invariant to a similarity transform (translation, rotation, uniform scale)
- Many techniques use non-linear least-squares optimization (bundle adjustment)
  - Levenberg-Marquardt is a popular algorithm
  - [http://en.wikipedia.org/wiki/Levenberg-Marquardt\\_algorithm](http://en.wikipedia.org/wiki/Levenberg-Marquardt_algorithm)
- Good code online
  - Bundler: <http://phototour.cs.washington.edu/bundler/>
  - Multicore: <http://grail.cs.washington.edu/projects/mcba/>

Suppose we know 3D points and affine camera parameters ...

then, we can compute the observed 2d positions of each point

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

Camera Parameters (2mx3)      3D Points (3xn)      2D Image Points (2mxn)

# What if we instead observe corresponding 2d image points?

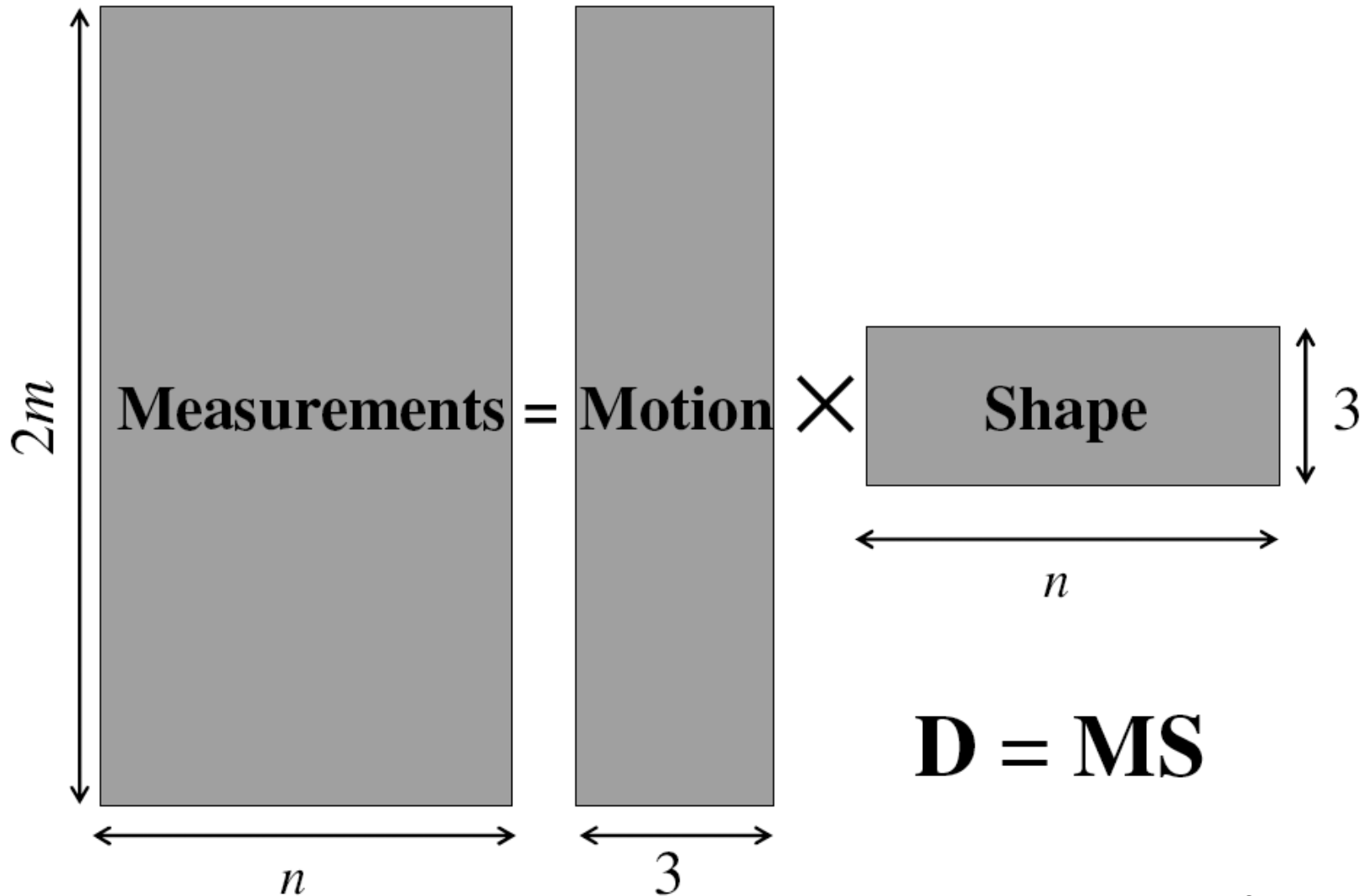
Can we recover the camera parameters and 3d points?

$$\mathbf{D} = \begin{array}{c} \text{cameras } (2m) \\ \left[ \begin{array}{cccc} \hat{\mathbf{X}}_{11} & \hat{\mathbf{X}}_{12} & \cdots & \hat{\mathbf{X}}_{1n} \\ \hat{\mathbf{X}}_{21} & \hat{\mathbf{X}}_{22} & \cdots & \hat{\mathbf{X}}_{2n} \\ & & \ddots & \\ \hat{\mathbf{X}}_{m1} & \hat{\mathbf{X}}_{m2} & \cdots & \hat{\mathbf{X}}_{mn} \end{array} \right] \end{array} \xRightarrow{?} \begin{array}{c} \left[ \begin{array}{c} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{array} \right] \end{array} \left[ \begin{array}{cccc} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_n \end{array} \right]$$

points ( $n$ )

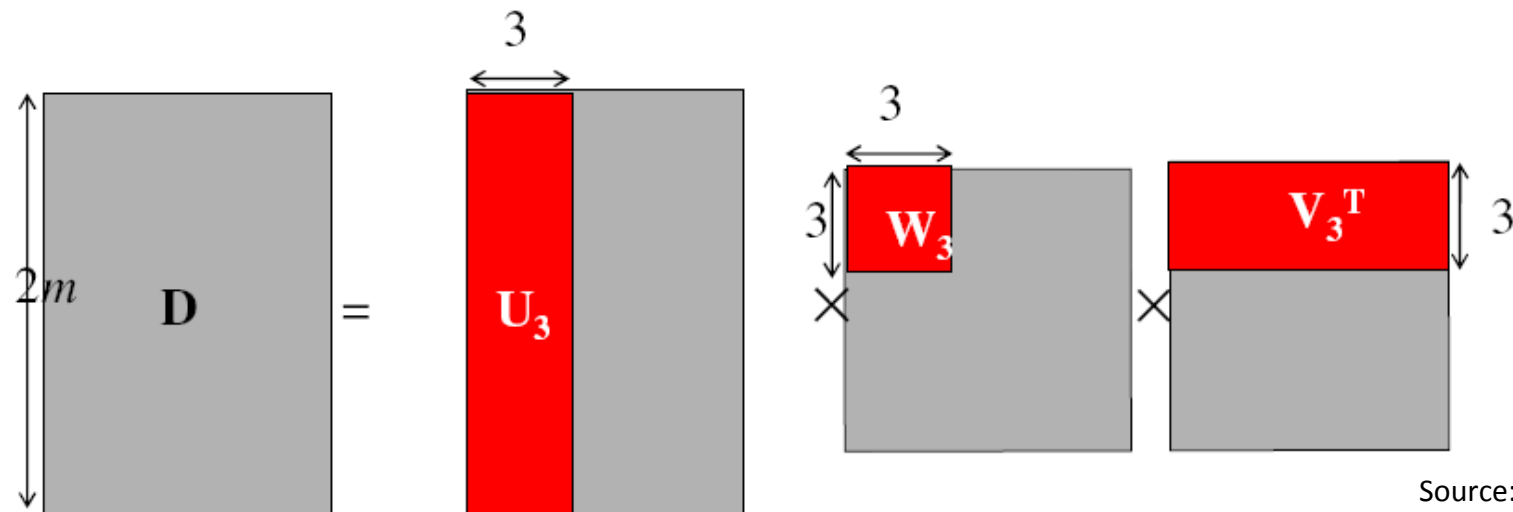
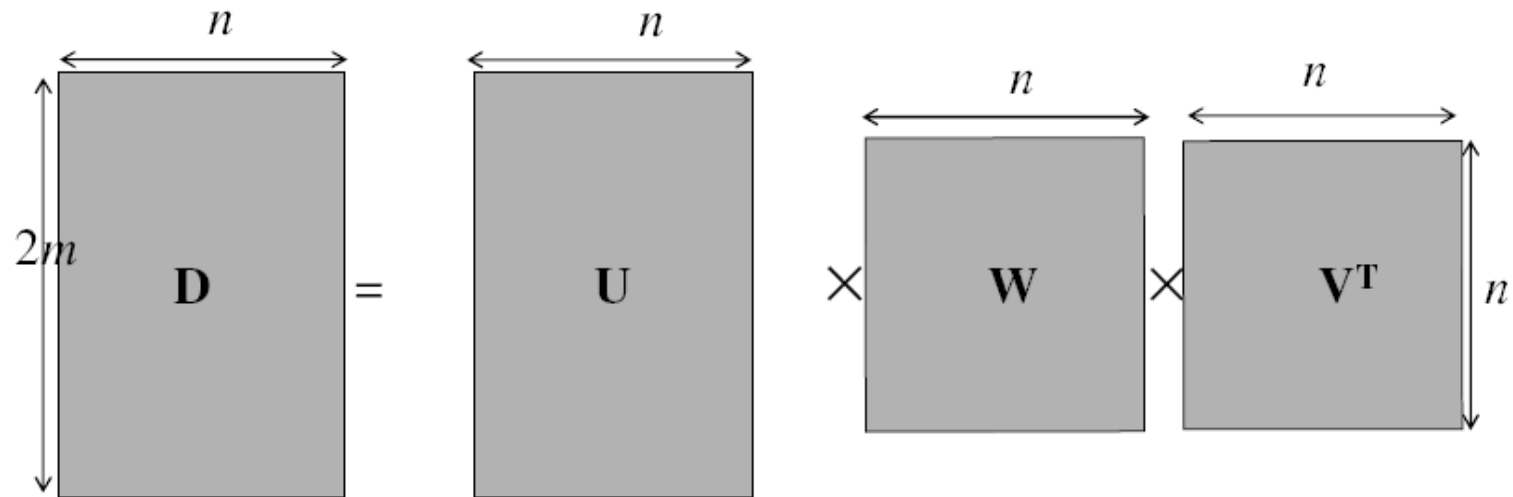
What rank is the matrix of 2D points?

# Factorizing the measurement matrix



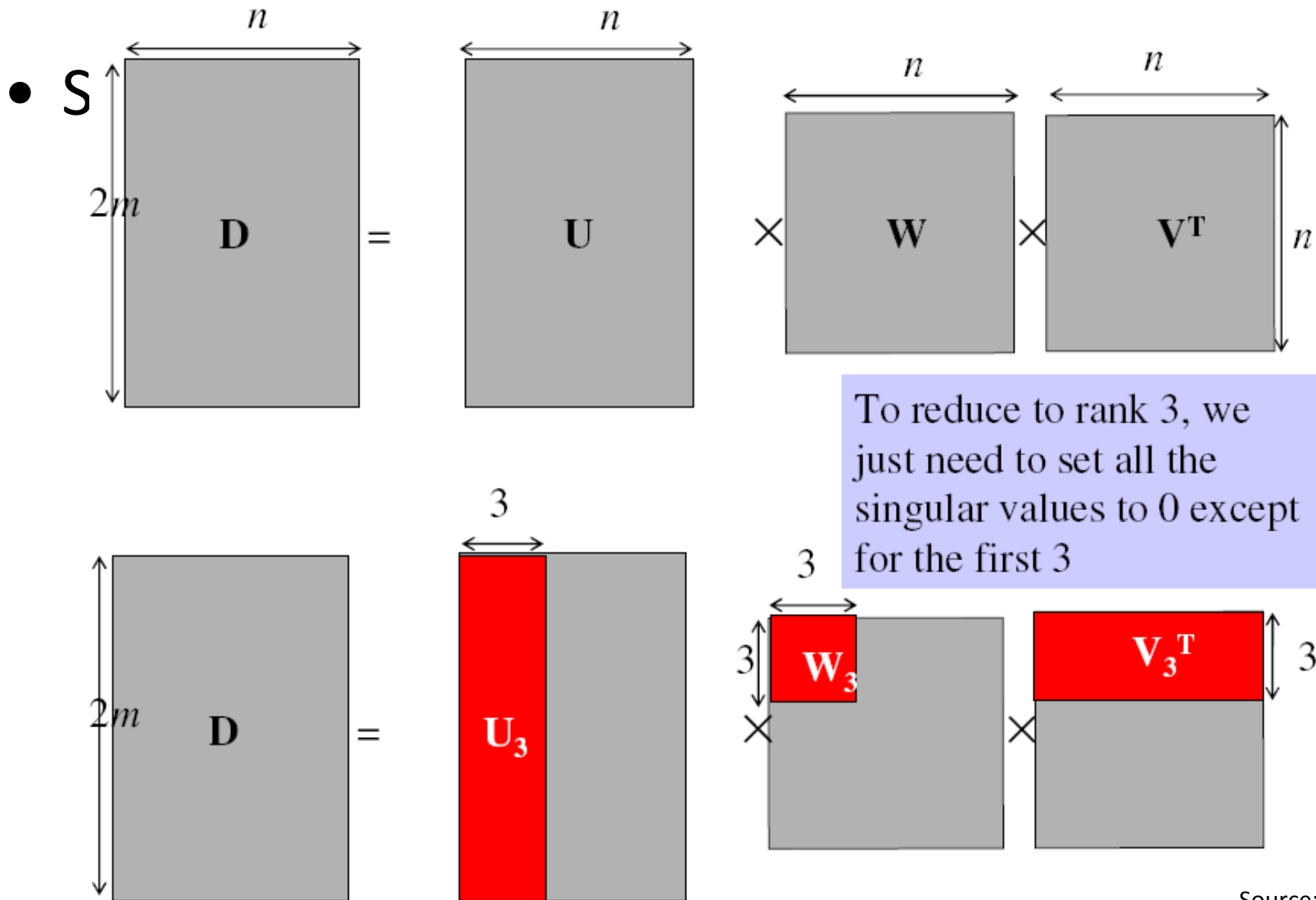
# Factorizing the measurement matrix

- Singular value decomposition of  $D$ :

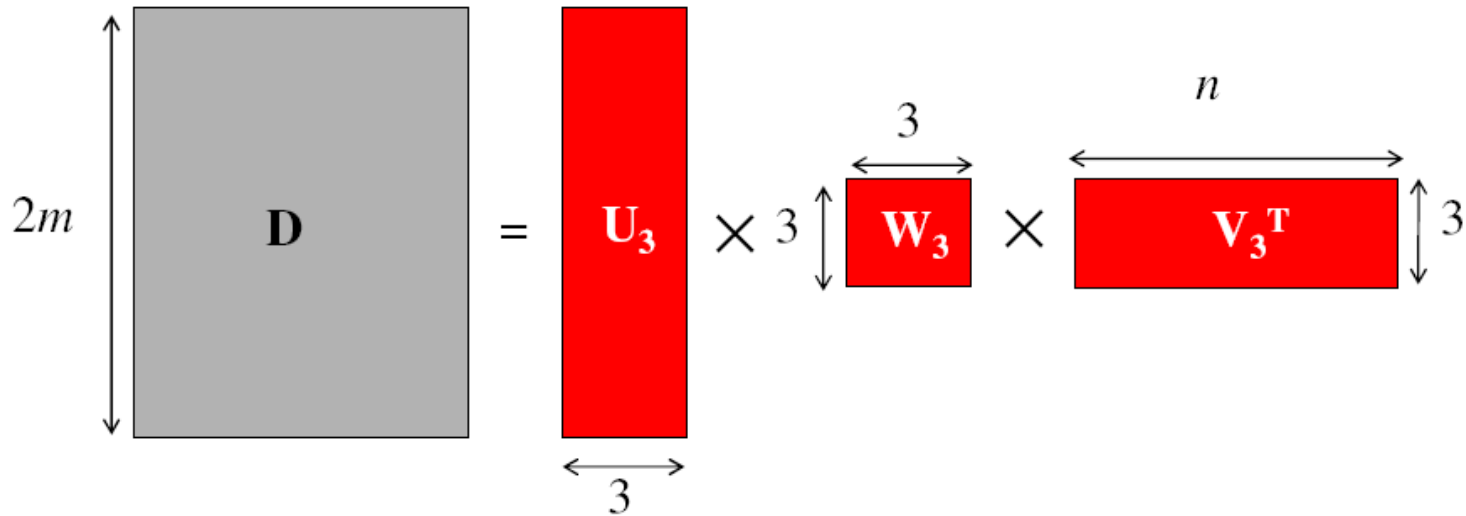




# Factorizing the measurement matrix

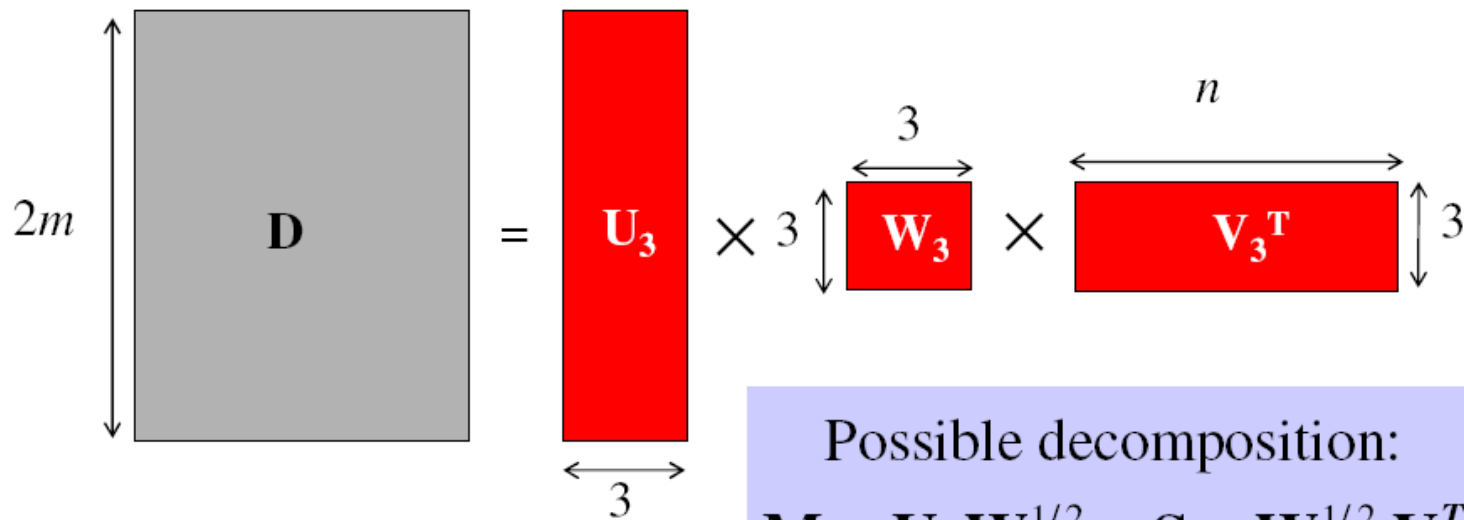


# Factorizing the measurement matrix



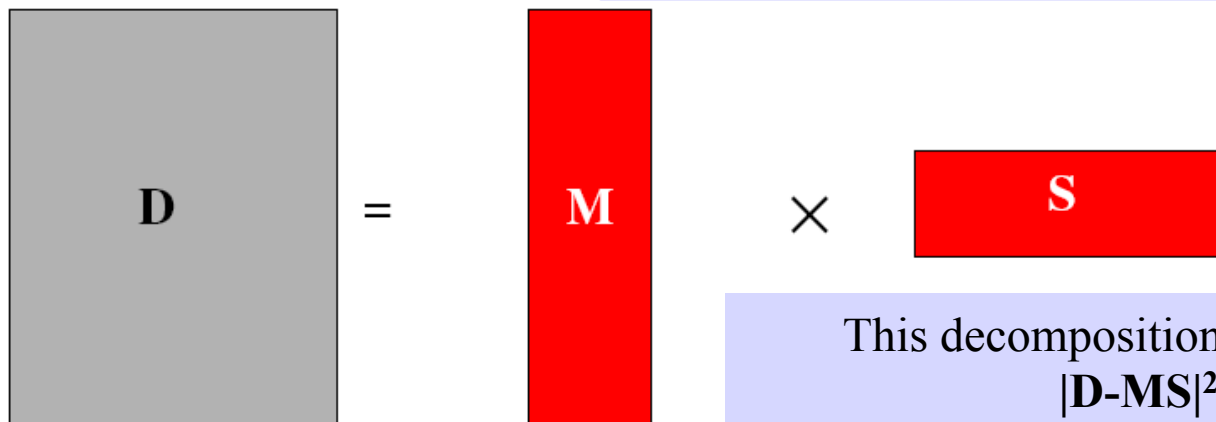
# Factorizing the measurement matrix

- Obtaining a factorization from SVD:



Possible decomposition:

$$\mathbf{M} = \mathbf{U}_3 \mathbf{W}_3^{1/2} \quad \mathbf{S} = \mathbf{W}_3^{1/2} \mathbf{V}_3^T$$



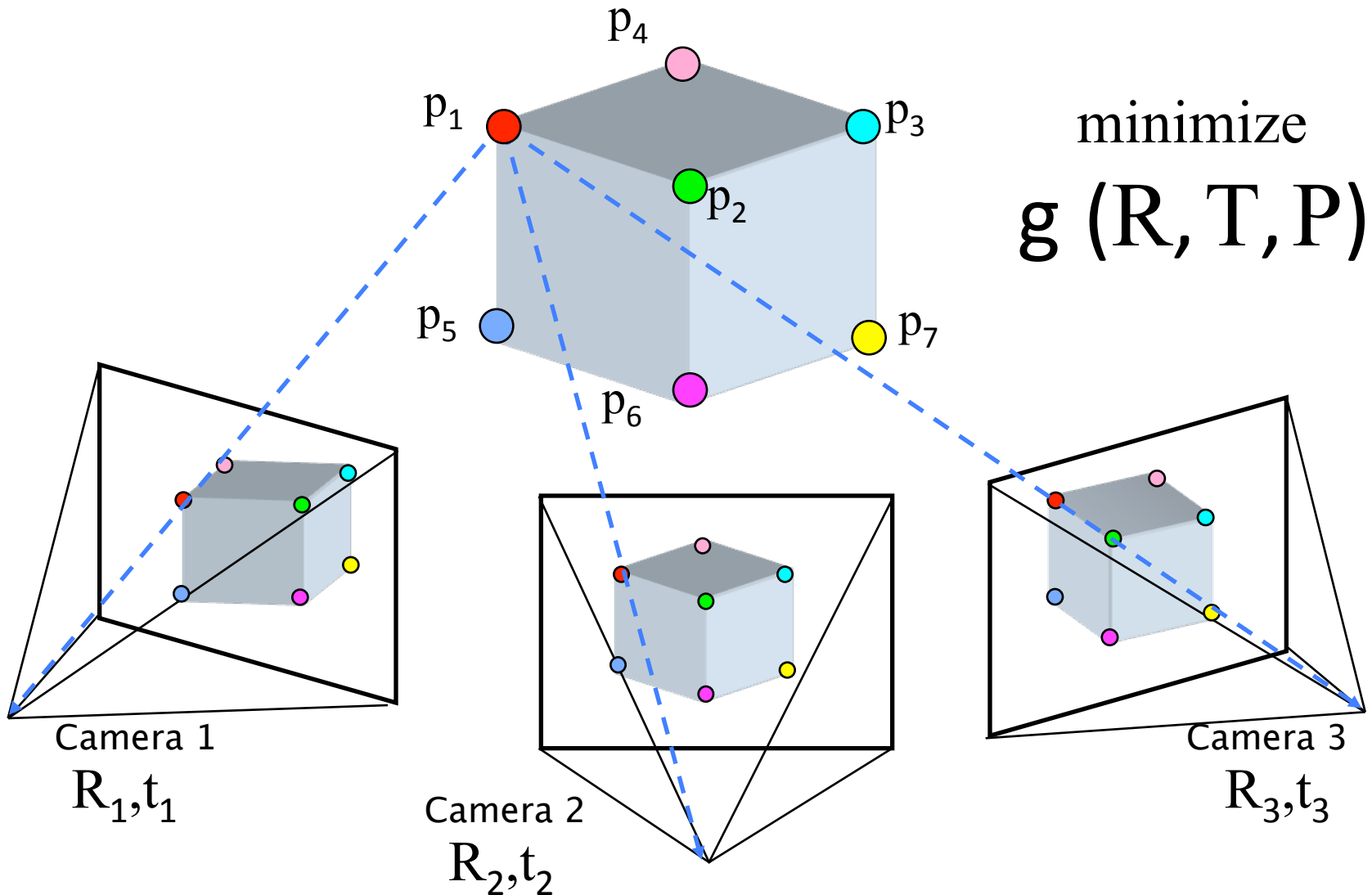
This decomposition minimizes  $|\mathbf{D} - \mathbf{MS}|^2$

# Algorithm summary

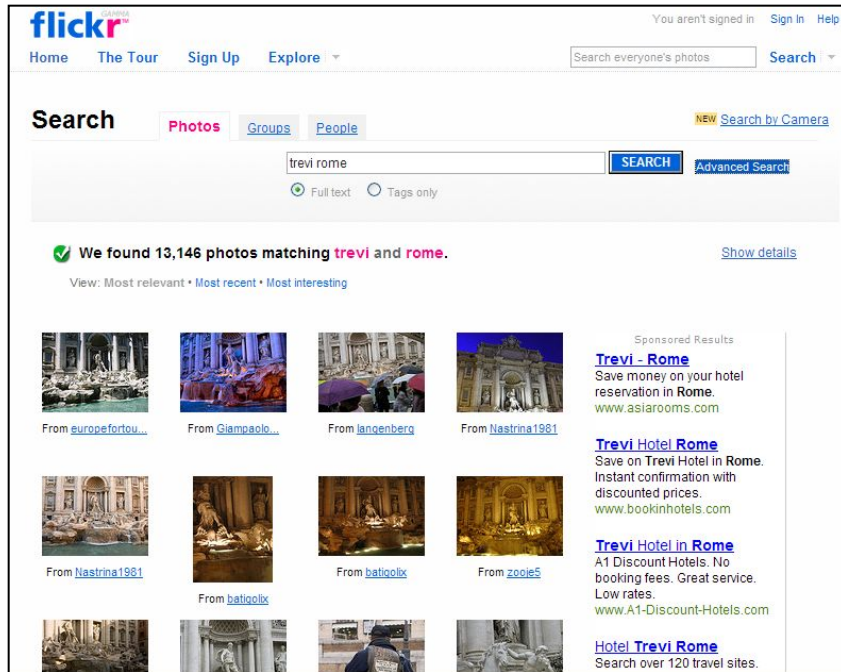
- Given:  $m$  images and  $n$  features  $\mathbf{x}_{ij}$
- For each image  $i$ , center the feature coordinates
- Construct a  $2m \times n$  measurement matrix  $\mathbf{D}$ :
  - Column  $j$  contains the projection of point  $j$  in all views
  - Row  $i$  contains one coordinate of the projections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

# Structure from motion

- aka “bundle adjustment” (texts: [Zisserman](#); [Faugeras](#))



# Structure from motion



Images on the Internet



Computed 3D structure



# Photo Tourism

Exploring photo collections in 3D

**Microsoft®**



(a)



(b)



(c)

- Photo tourism video: <http://www.youtube.com/watch?v=5Ji84zb2r8s>
- Photosynth: <http://photosynth.net/>

# Search

**Photos** [Groups](#) [People](#)

rome

**SEARCH**

[Advanced Search](#)  
[Search by Camera](#)

Full text  Tags only

✓ We found **1,281,363 results** matching **rome**.

[Slideshow](#)

View: [Most relevant](#) • [Most recent](#) • [Most interesting](#) Show: [Details](#) • [Thumbnails](#)



From [Dmitry...](#)



From [Lisa...](#)



From [Lisa...](#)



From [jerkhead29](#)



From [Lisa...](#)



From [Lisa...](#)



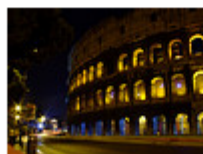
From [Lisa...](#)



From [Dmitry...](#)



From [Lisa...](#)



From [formicheit](#)



From [Lisa...](#)



From [Rick's Picks](#)

### Sponsored Results

#### [\\$387 Italy Flights](#)

Fly one-way to Italy from \$387.  
Book your flight before April 24.  
[www.AirFrance.us](http://www.AirFrance.us)

#### [What to Do Rome](#)

55+ Things to do in **Rome** on Viator. See the Vatican, Colosseum & more.  
[www.viator.com/Rome](http://www.viator.com/Rome)

#### [Rome with Kids](#)

Get Extra Savings on a Family Vacation to **Rome**. Limited Time Only.  
[www.AdventuresByDisney.com](http://www.AdventuresByDisney.com)

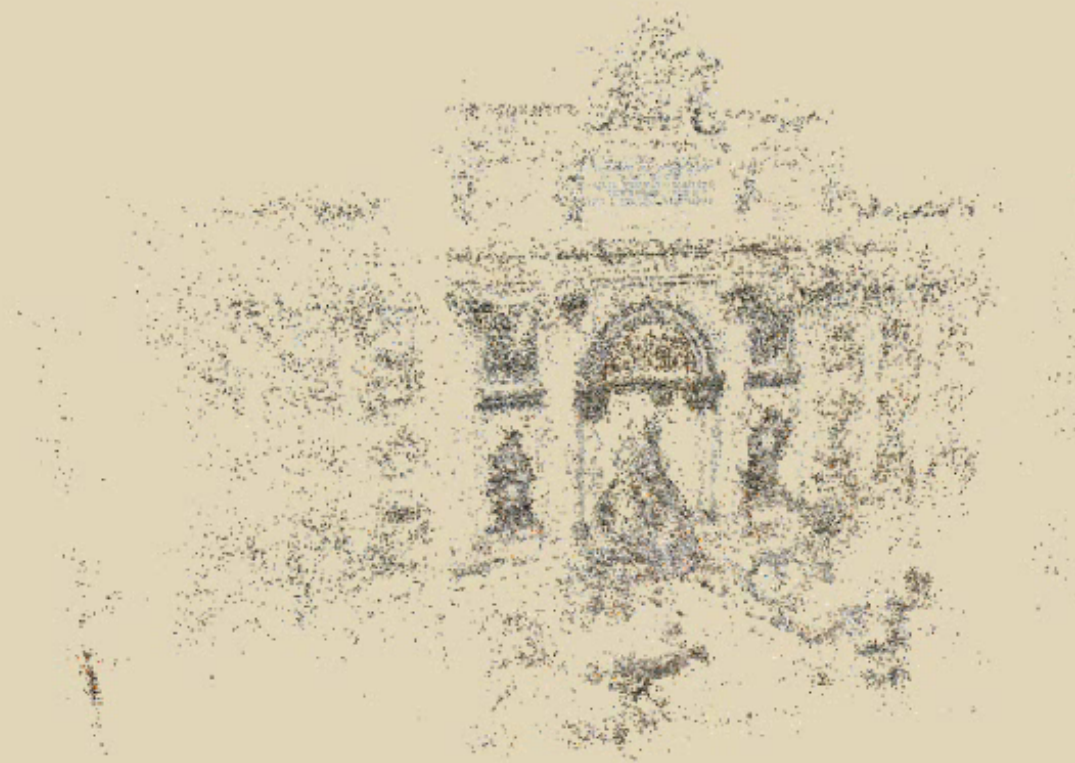
#### [Fly to Italy from \\$499 Round Trip](#)

Travel to Italy between May 1 - 31 from \$499 per Person.  
[www.EuroFlyusa.com](http://www.EuroFlyusa.com)

#### [Rome Italy Day Tours](#)

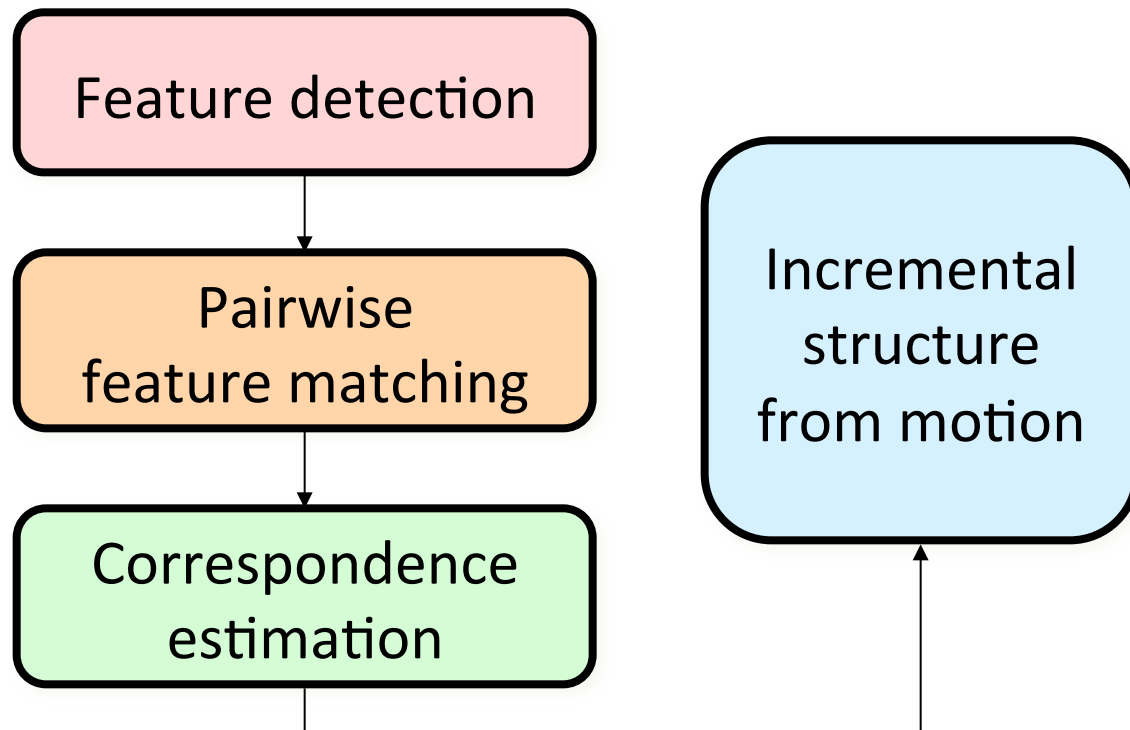
Private tours of **Rome** A variety of tours available....





# Scene reconstruction

- Automatically estimate
  - position, orientation, and focal length of cameras
  - 3D positions of feature points



# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



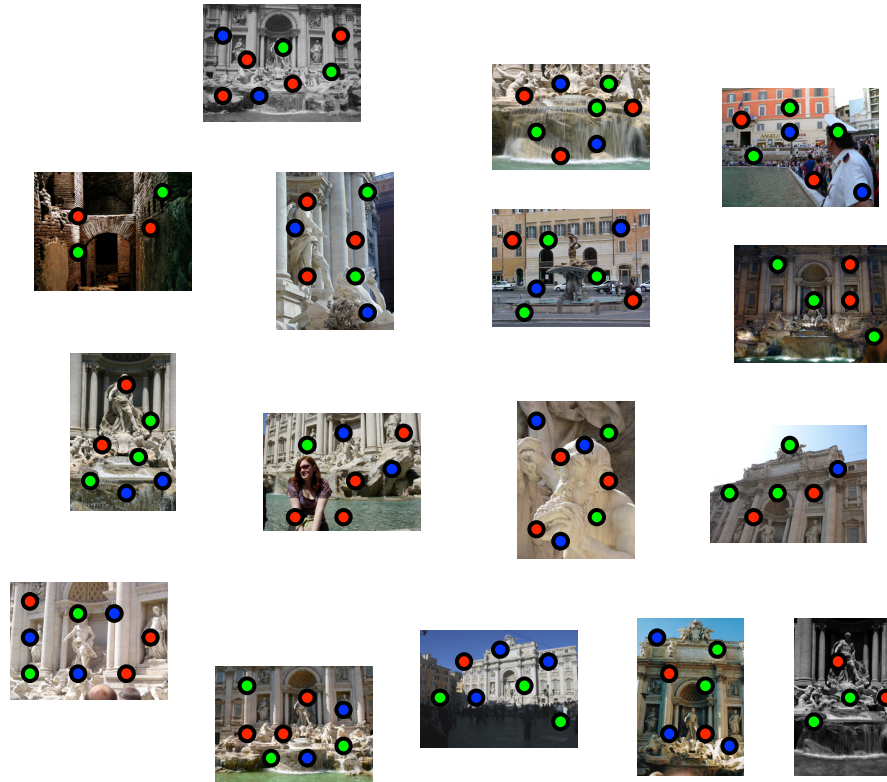
# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



# Correspondence estimation

- Link up pairwise matches to form connected components of matches across several images



Image 1



Image 2

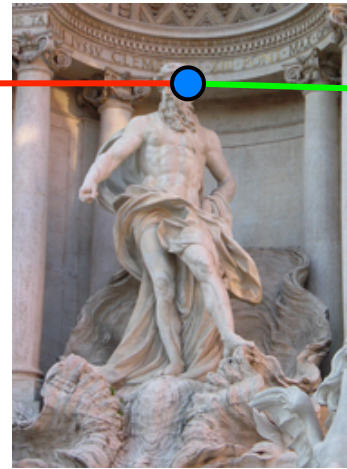


Image 3

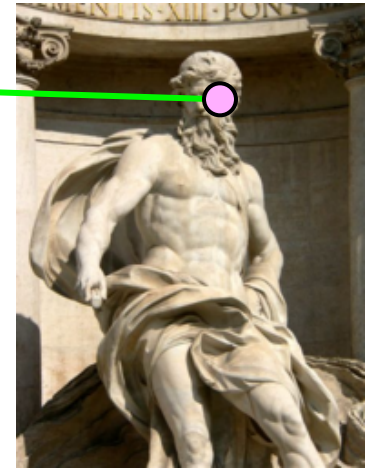
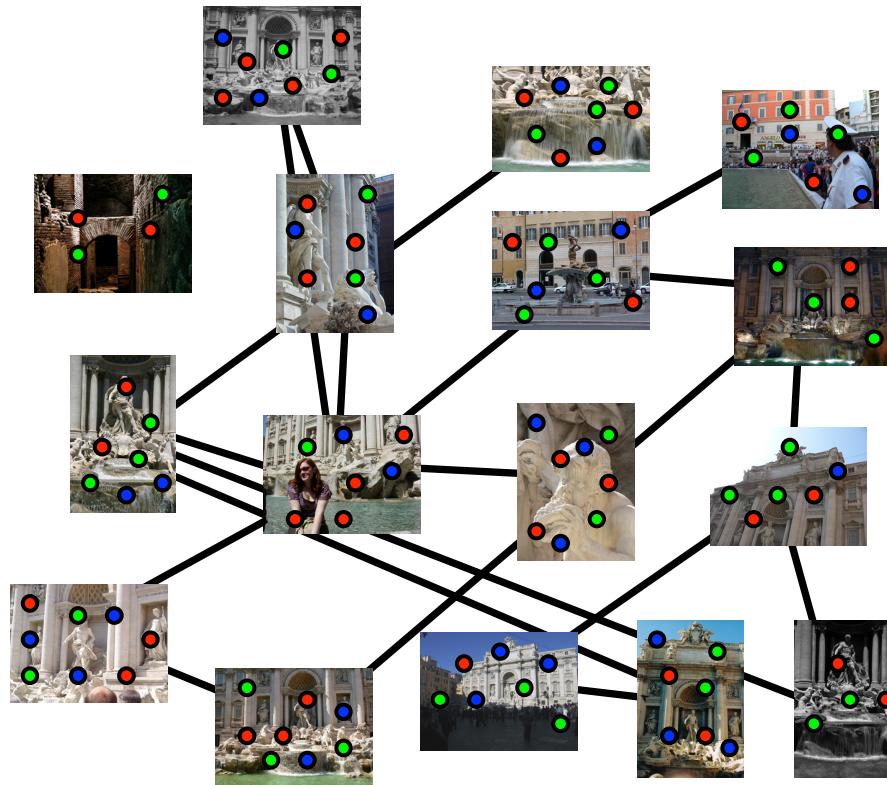


Image 4



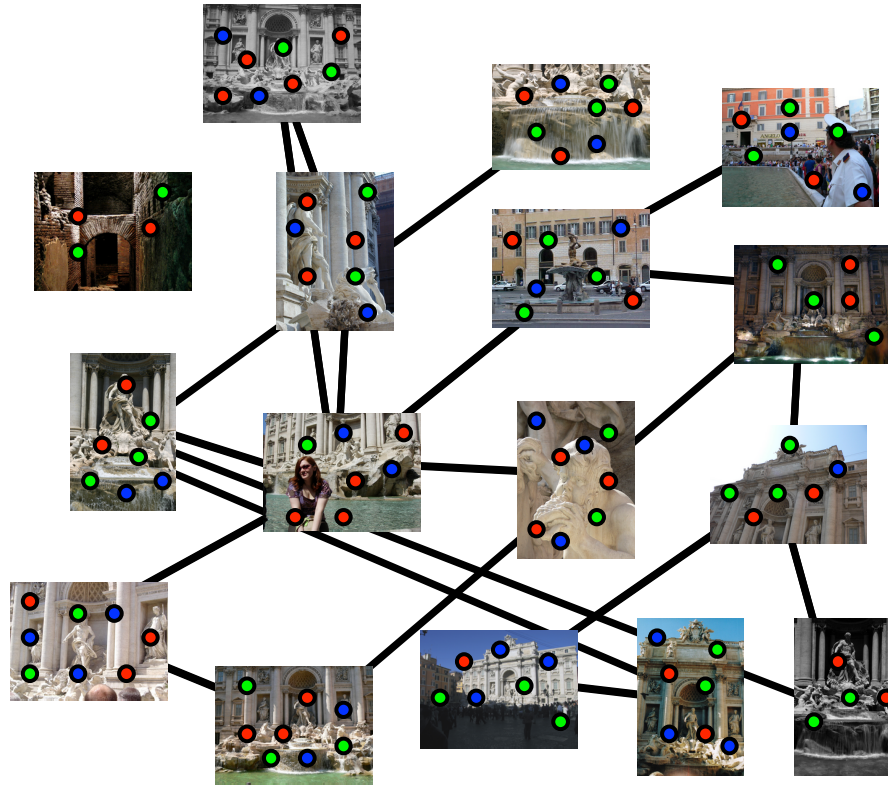
# Feature matching

Match features between each pair of images



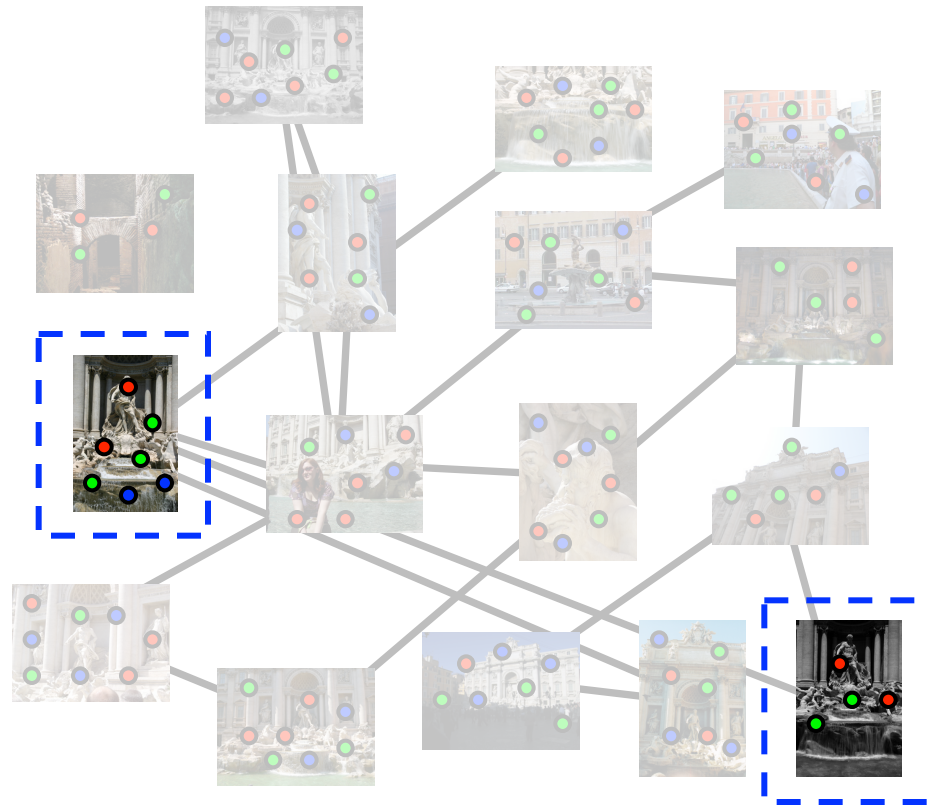
# Feature matching

Refine matching using RANSAC [Fischler & Bolles 1987]

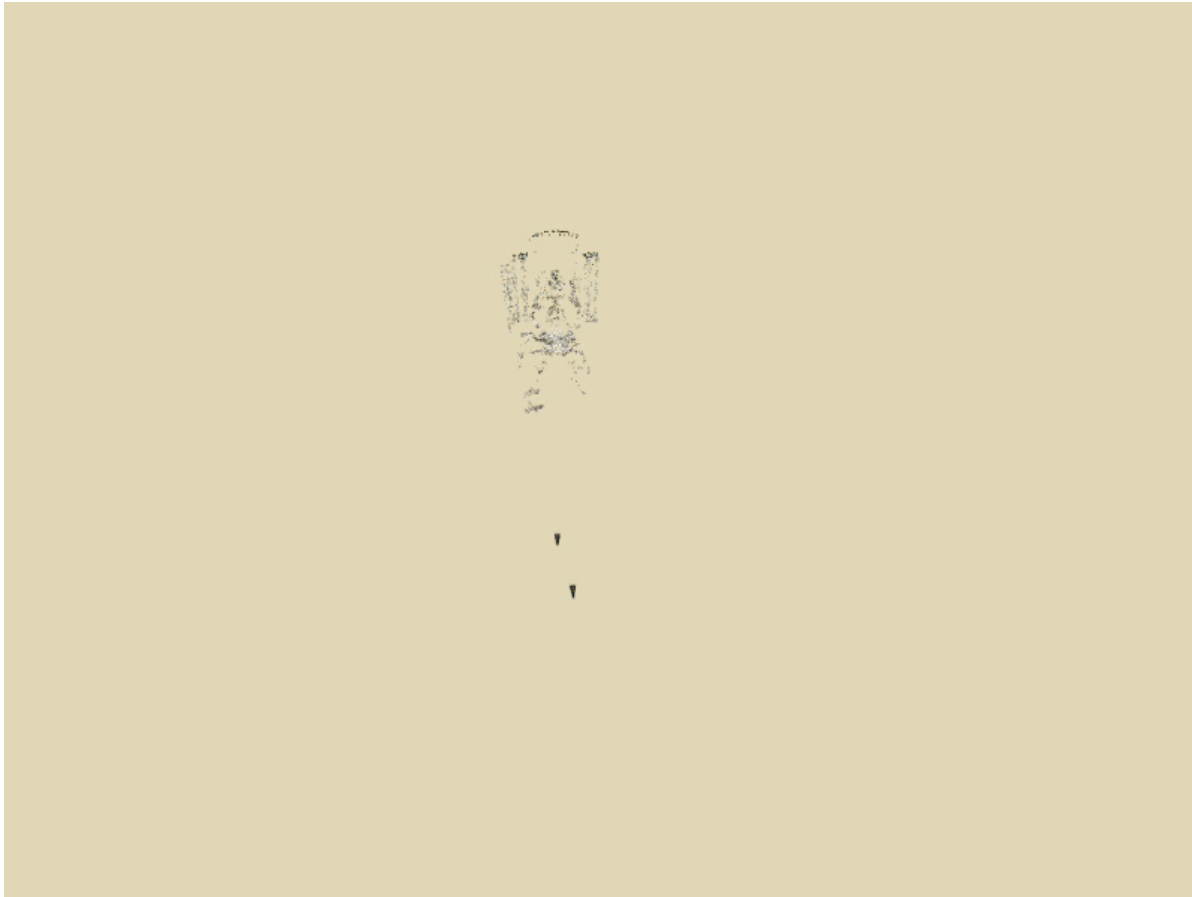




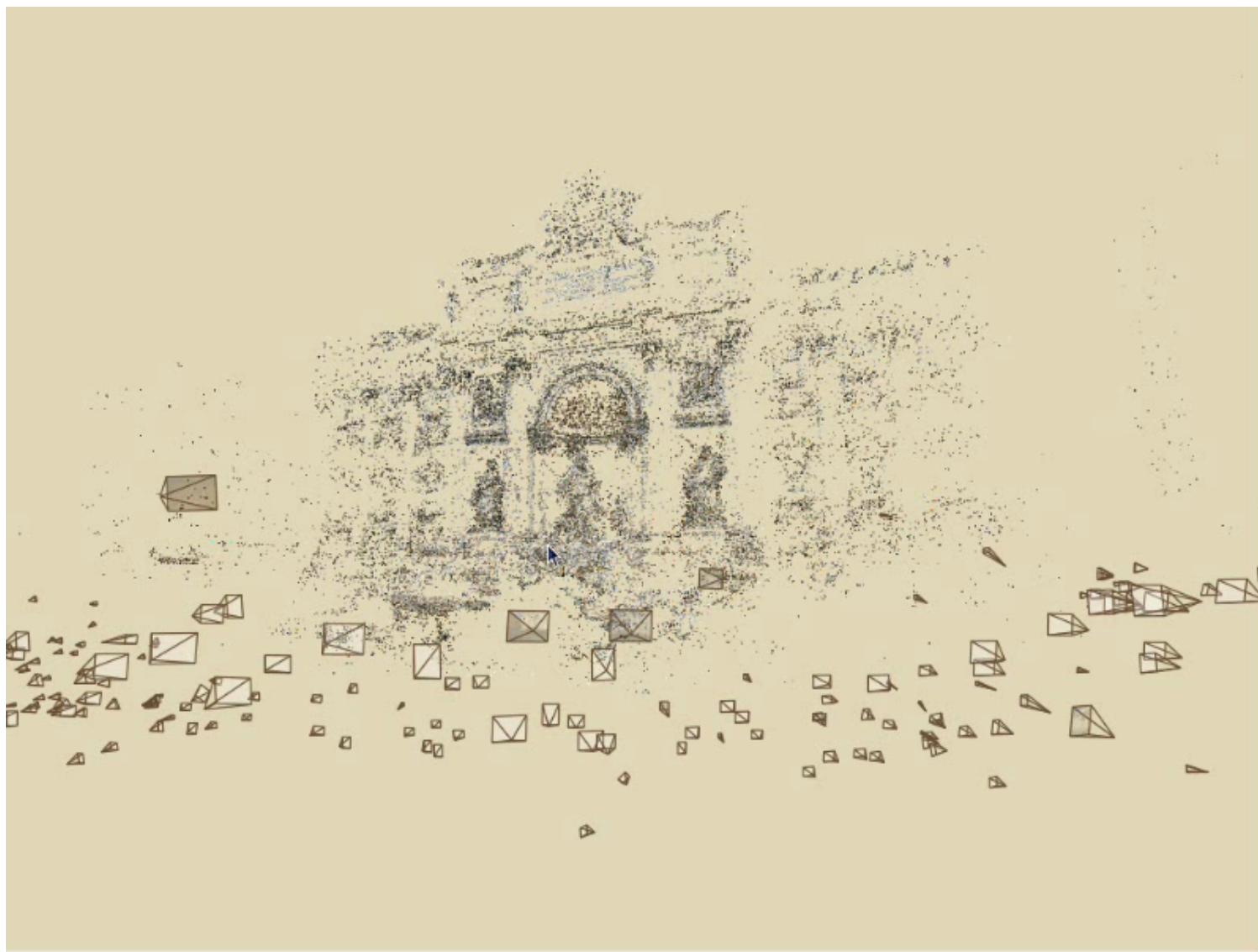
# Incremental structure from motion

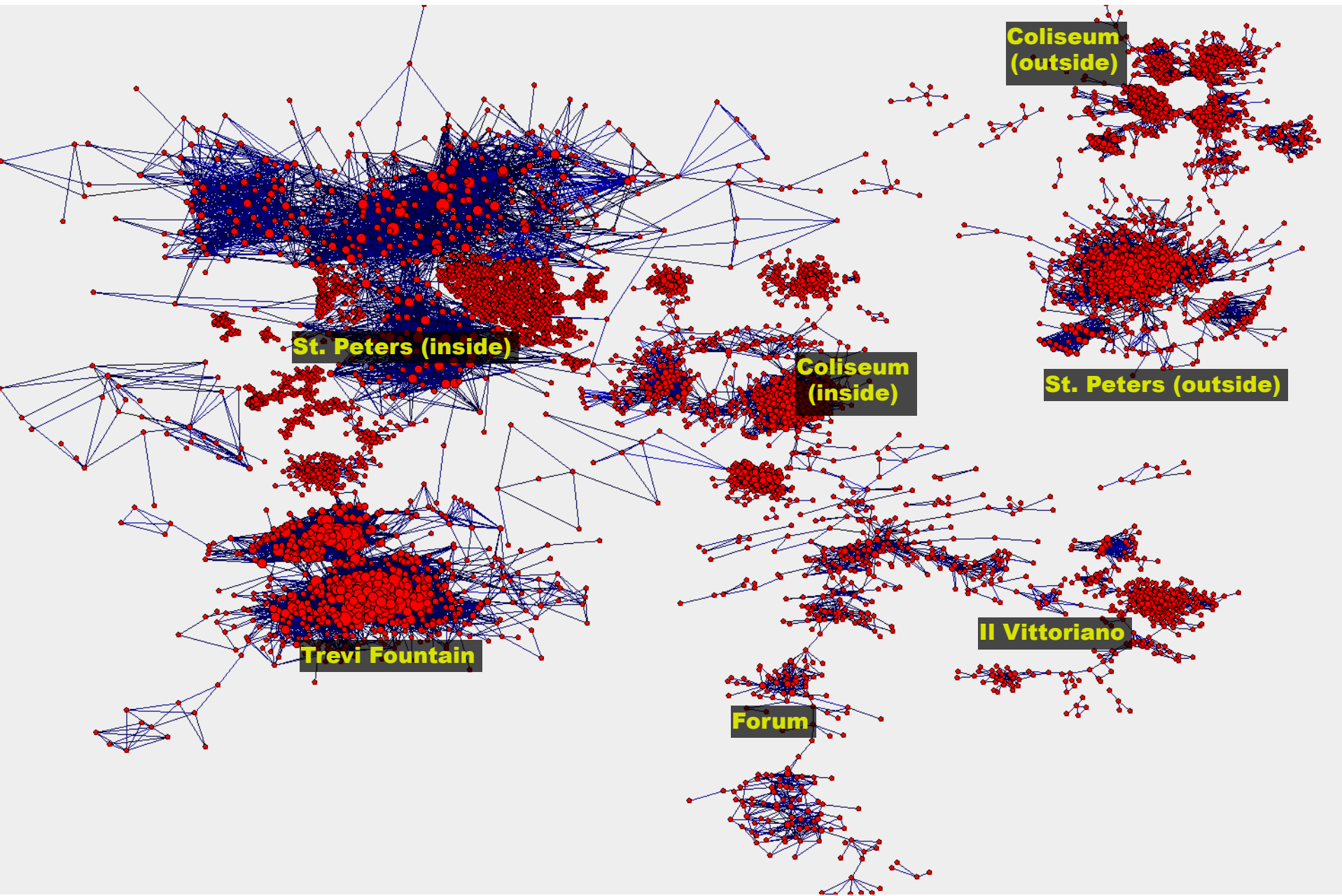


# Incremental structure from motion



# Photo Explorer





**Coliseum (outside)**

**St. Peters (inside)**

**Coliseum (inside)**

**St. Peters (outside)**

**Trevi Fountain**

**Il Vittoriano**

**Forum**

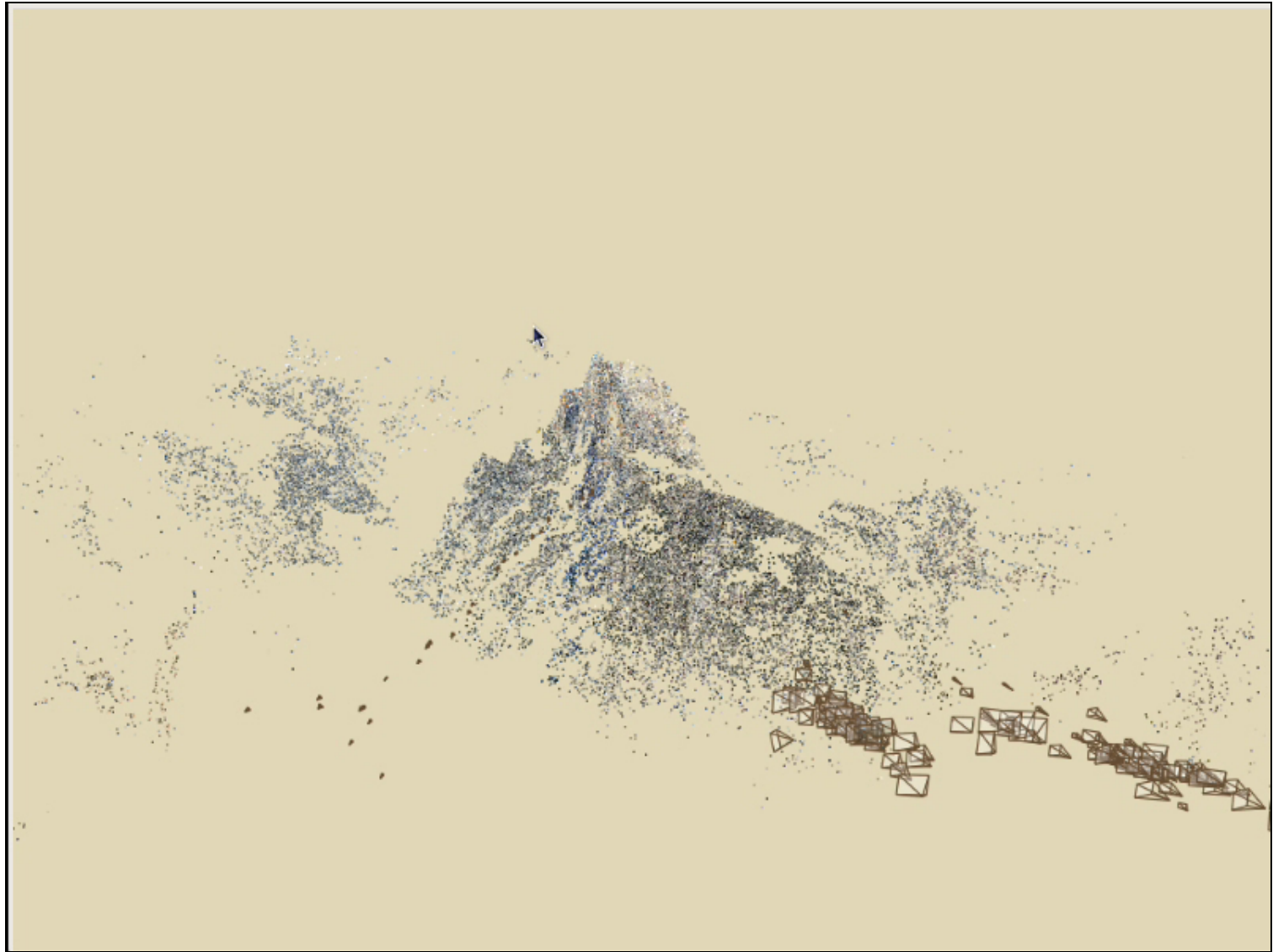
# Navigation: Prague Old Town Square



# Hierarchical annotations



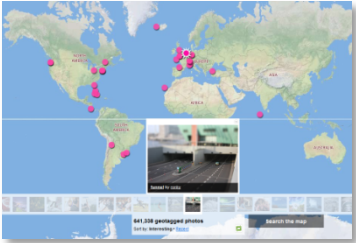
# Locking the camera (stabilization)



# Applications



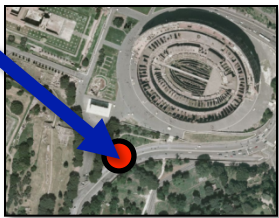
# Community photo collections



- “Wikipedia for photos” – visual record of world through community of photographers
  - *Geograph British Isles*  
<http://www.geograph.org.uk/>



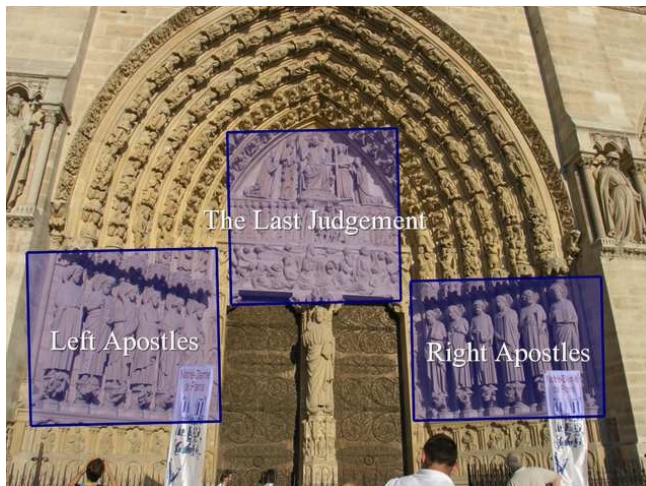
- Users can tag and comment on photos, link to other content
  - *World-wide telescope*



- “Where should I take a photo?”  
<http://photocitygame.com/>

# Community photo collections

- Leveraging large databases of photos, large number of users
  - Annotations / augmented reality



Space Needle - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://en.wikipedia.org/wiki/Space\_Needle

Getting Started Latest Headlines Minamoto no Yoshir...

Sign in / create account

37,751

You can help Wikipedia change the world! Donate now!

From the fundraising blog - Wikibooks and the Future of Free Education

"Kunskap är maktt!" - Anon...

**Space Needle**

From Wikipedia, the free encyclopedia

The **Space Needle** is a tower in Seattle, Washington. It is a major landmark of the Pacific Northwest region of the United States and the symbol of Seattle. Located in Seattle Center, it was built for the 1962 World's Fair, during which time nearly 20,000 people a day used the elevators — 2.3 million visitors in all for the World Fair. The Space Needle is 605 feet (184 m) high and 138 feet (42 m) wide at its widest point and weighs 9,550 tons. When it was completed it was the tallest building west of the Mississippi

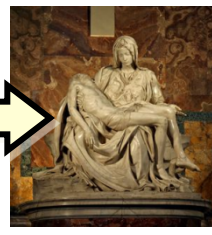
Space Needle from Volunteer

# Virtual tour guide scenario

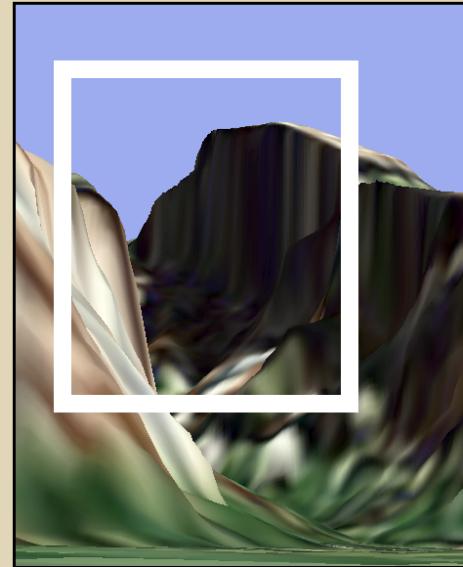
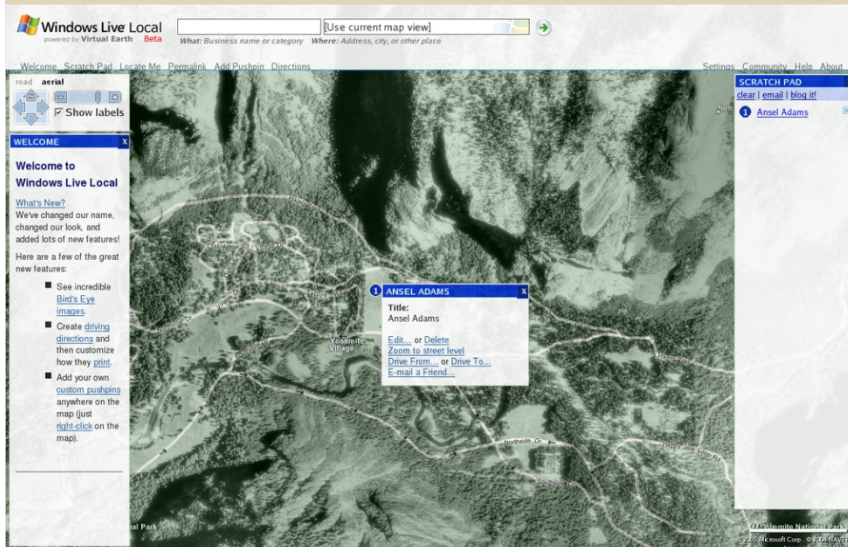
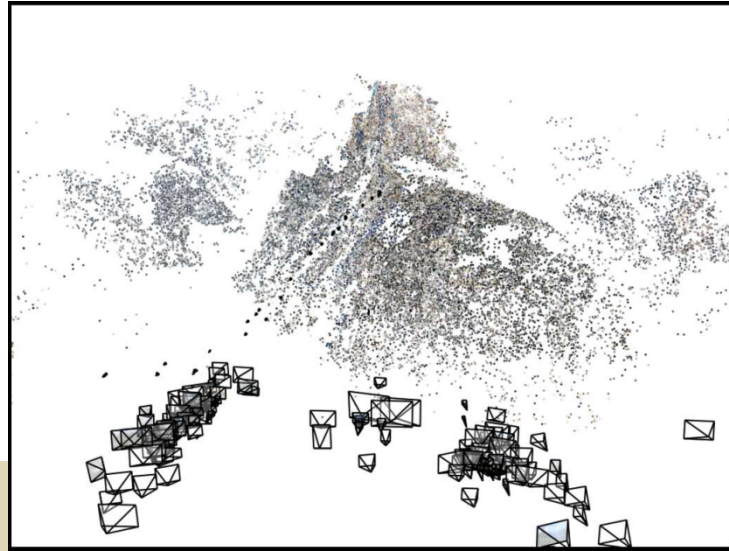
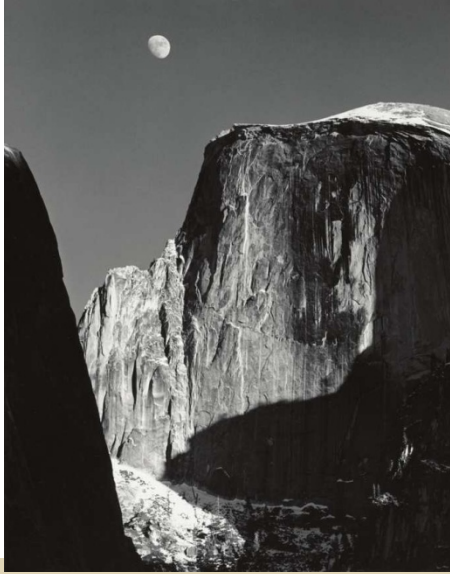


## St. Peter's Basilica

- Built: 1506-1626
- [http://en.wikipedia.org/wiki/St.\\_Peter's\\_Basilica](http://en.wikipedia.org/wiki/St._Peter's_Basilica)



# Rephotography



Topographic data courtesy USGS