

Structure From Motion

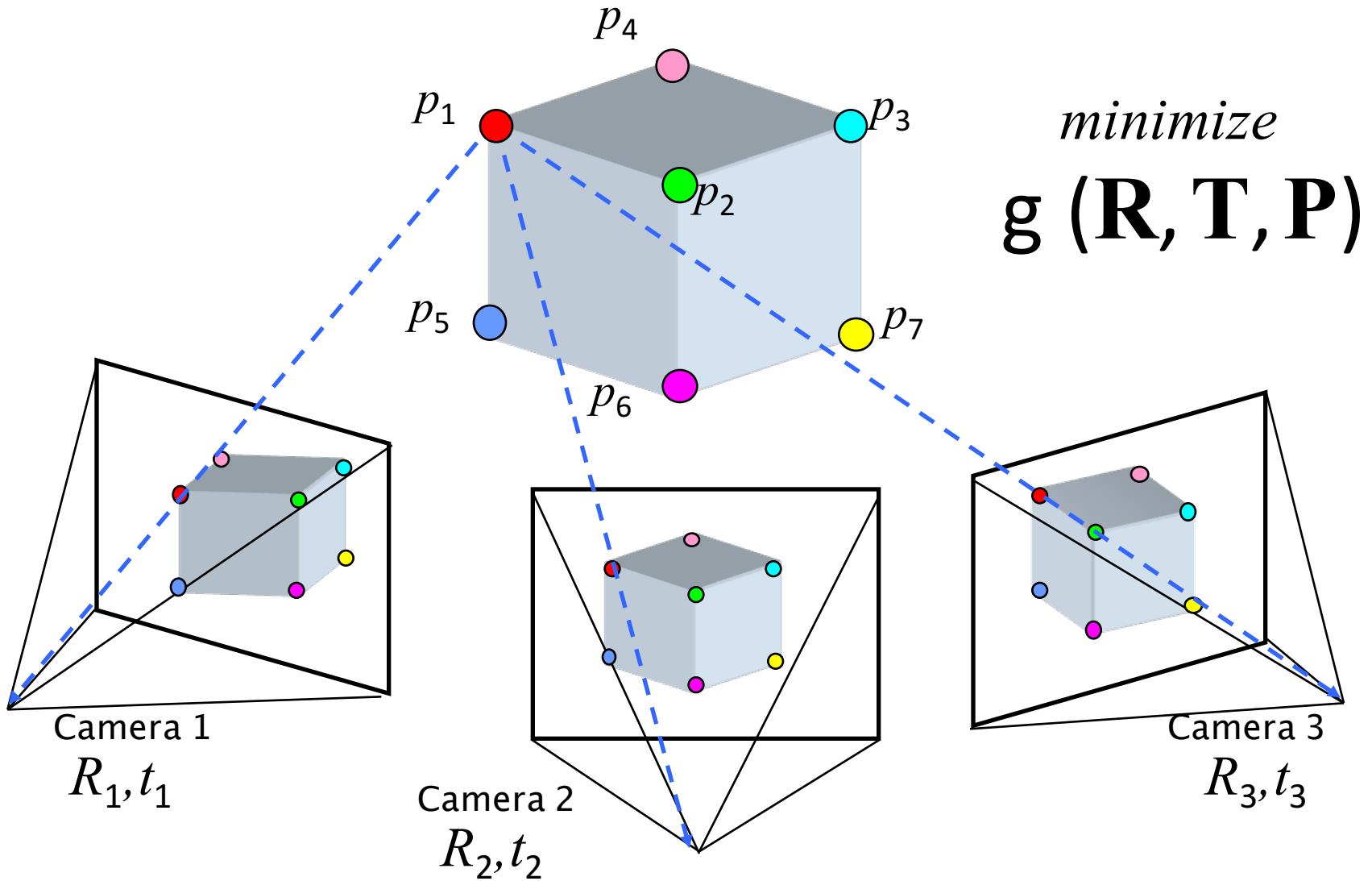
Ali Farhadi

CSE 576

Several slides from Steve Seitz, Rick Szeliski, Martial Hebert, and Noha Snavely

Structure from motion

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



SfM objective function

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

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

- Minimize sum of squared reprojection errors:

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

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
- 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

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

3D Points ($3 \times n$)

Camera Parameters ($2m \times 3$) 2D Image Points ($2m \times n$)

The diagram illustrates the matrix multiplication process. On the left, a vertical stack of matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$ is multiplied by a horizontal vector of 3D points $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$. This results in a vertical stack of matrices $\hat{\mathbf{x}}_{11}, \hat{\mathbf{x}}_{12}, \dots, \hat{\mathbf{x}}_{1n}, \hat{\mathbf{x}}_{21}, \hat{\mathbf{x}}_{22}, \dots, \hat{\mathbf{x}}_{2n}, \dots, \hat{\mathbf{x}}_{m1}, \hat{\mathbf{x}}_{m2}, \dots, \hat{\mathbf{x}}_{mn}$, which represent the observed 2D image points. Blue arrows indicate the correspondence between the labels and the respective parts of the equation.

What if we instead observe corresponding 2d image points?

Can we recover the camera parameters and 3d points?

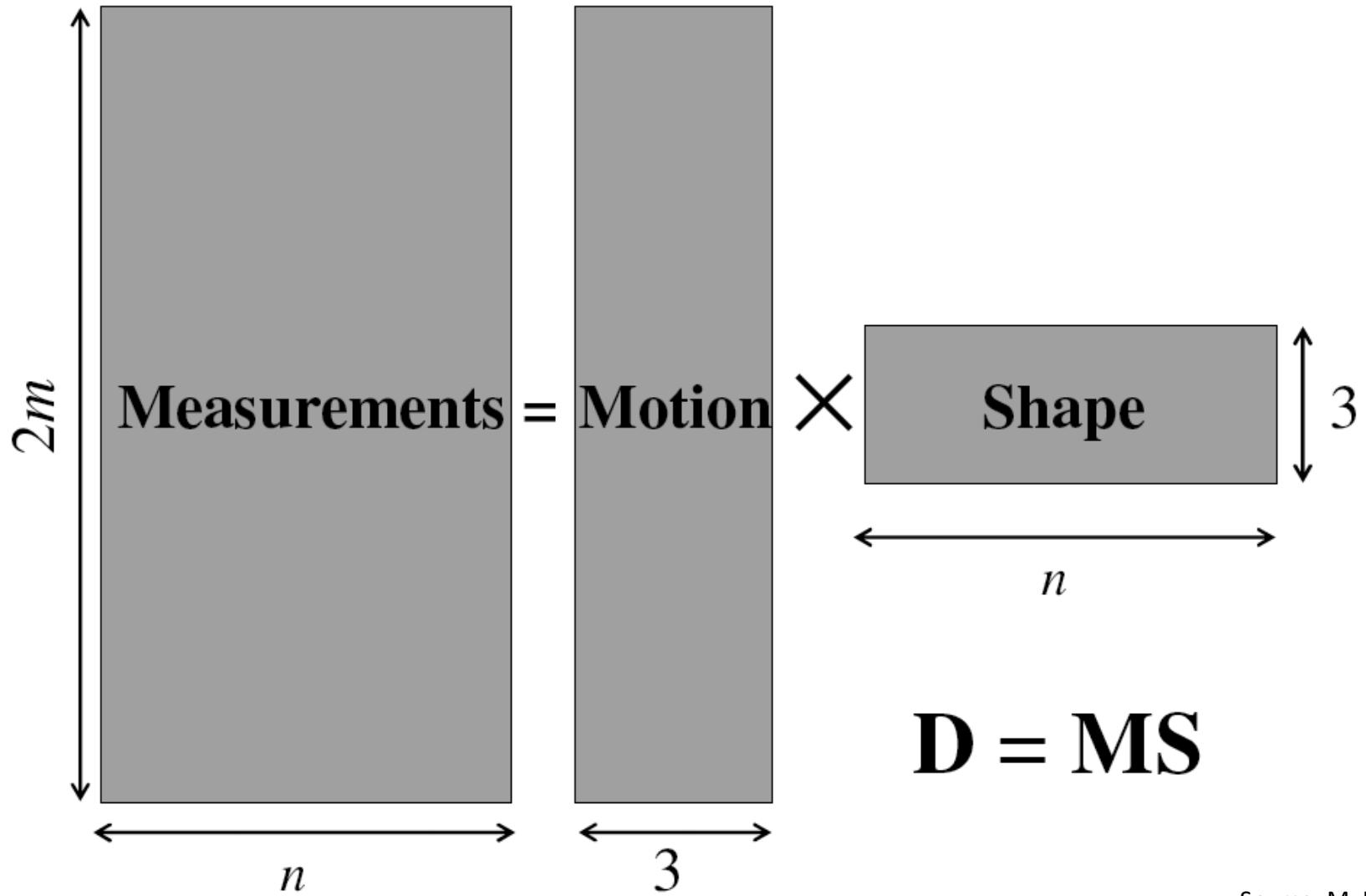
$$\mathbf{D} = \begin{bmatrix} \hat{\mathbf{x}}_{11} & \hat{\mathbf{x}}_{12} & \cdots & \hat{\mathbf{x}}_{1n} \\ \hat{\mathbf{x}}_{21} & \hat{\mathbf{x}}_{22} & \cdots & \hat{\mathbf{x}}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{x}}_{m1} & \hat{\mathbf{x}}_{m2} & \cdots & \hat{\mathbf{x}}_{mn}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

cameras ($2m$)

points (n)

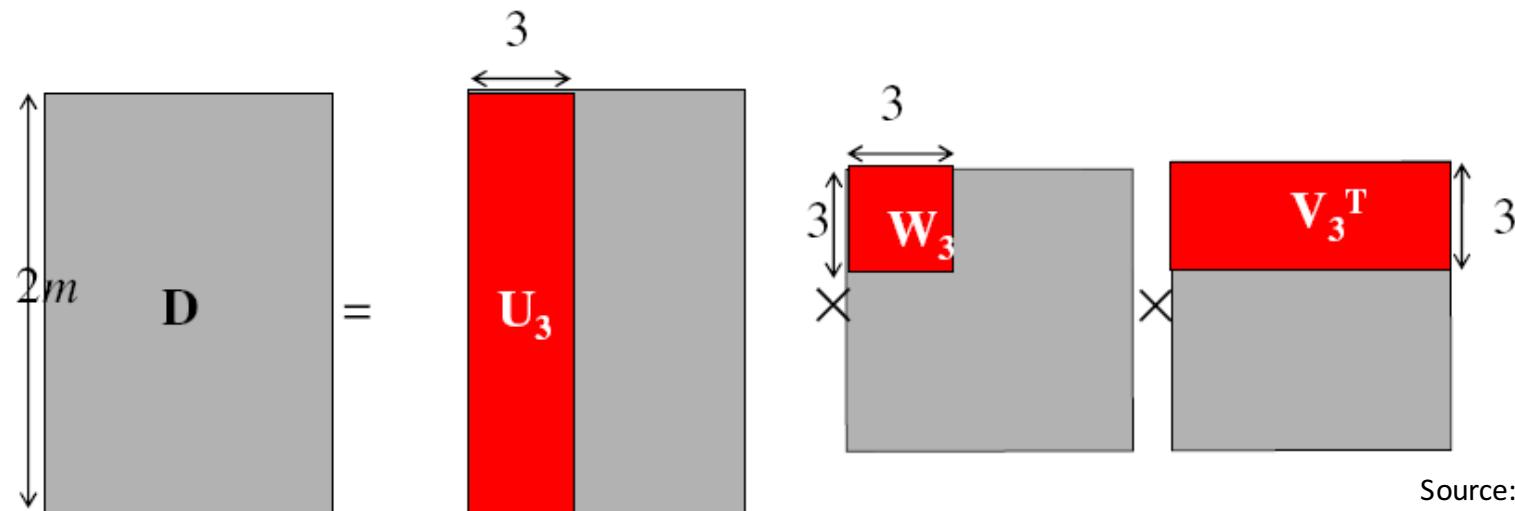
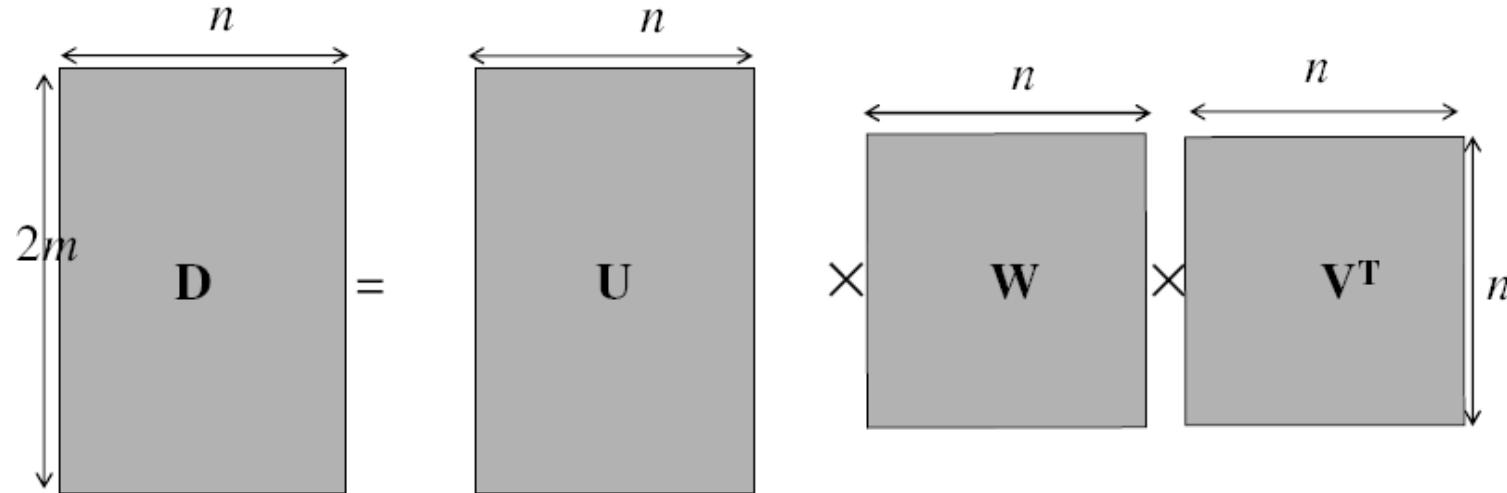
What rank is the matrix of 2D points?

Factorizing the measurement matrix



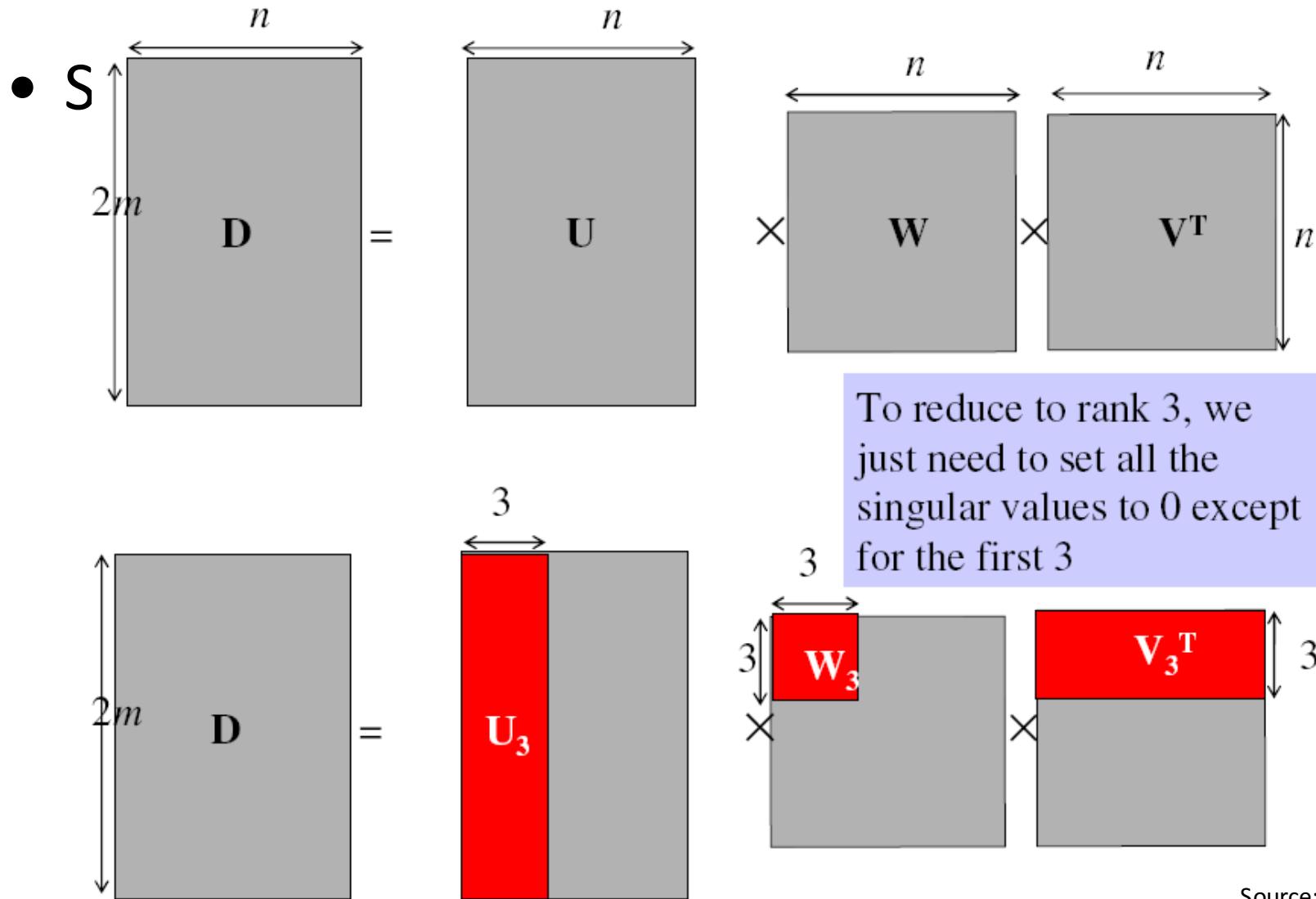
Factorizing the measurement matrix

- Singular value decomposition of D:

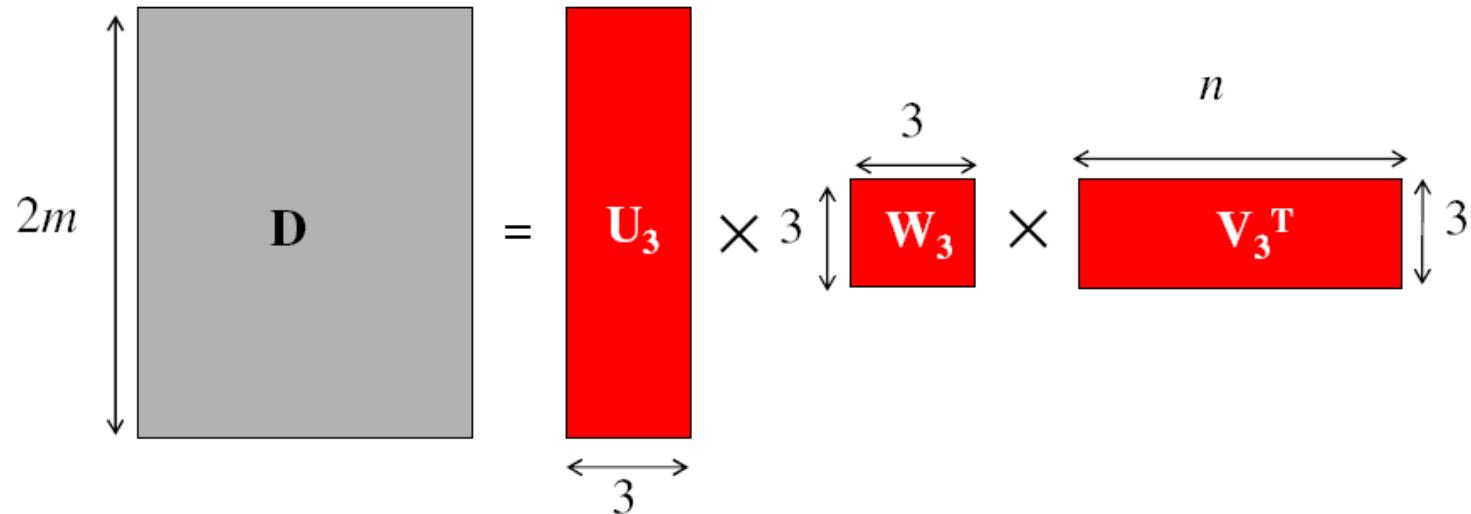


Source: M. Hebert

Factorizing the measurement matrix



Factorizing the measurement matrix



Factorizing the measurement matrix

- Obtaining a factorization from SVD:

$$\begin{matrix} \text{D} \\ \leftarrow 2m \end{matrix} = \begin{matrix} \text{U}_3 \\ \leftarrow 3 \end{matrix} \times \begin{matrix} \text{W}_3 \\ 3 \end{matrix} \times \begin{matrix} \text{V}_3^T \\ \leftarrow n \end{matrix}$$

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

$$\begin{matrix} \text{D} \\ \leftarrow 2m \end{matrix} = \begin{matrix} \text{M} \\ \leftarrow 3 \end{matrix} \times \begin{matrix} \text{S} \\ \leftarrow 3 \end{matrix}$$

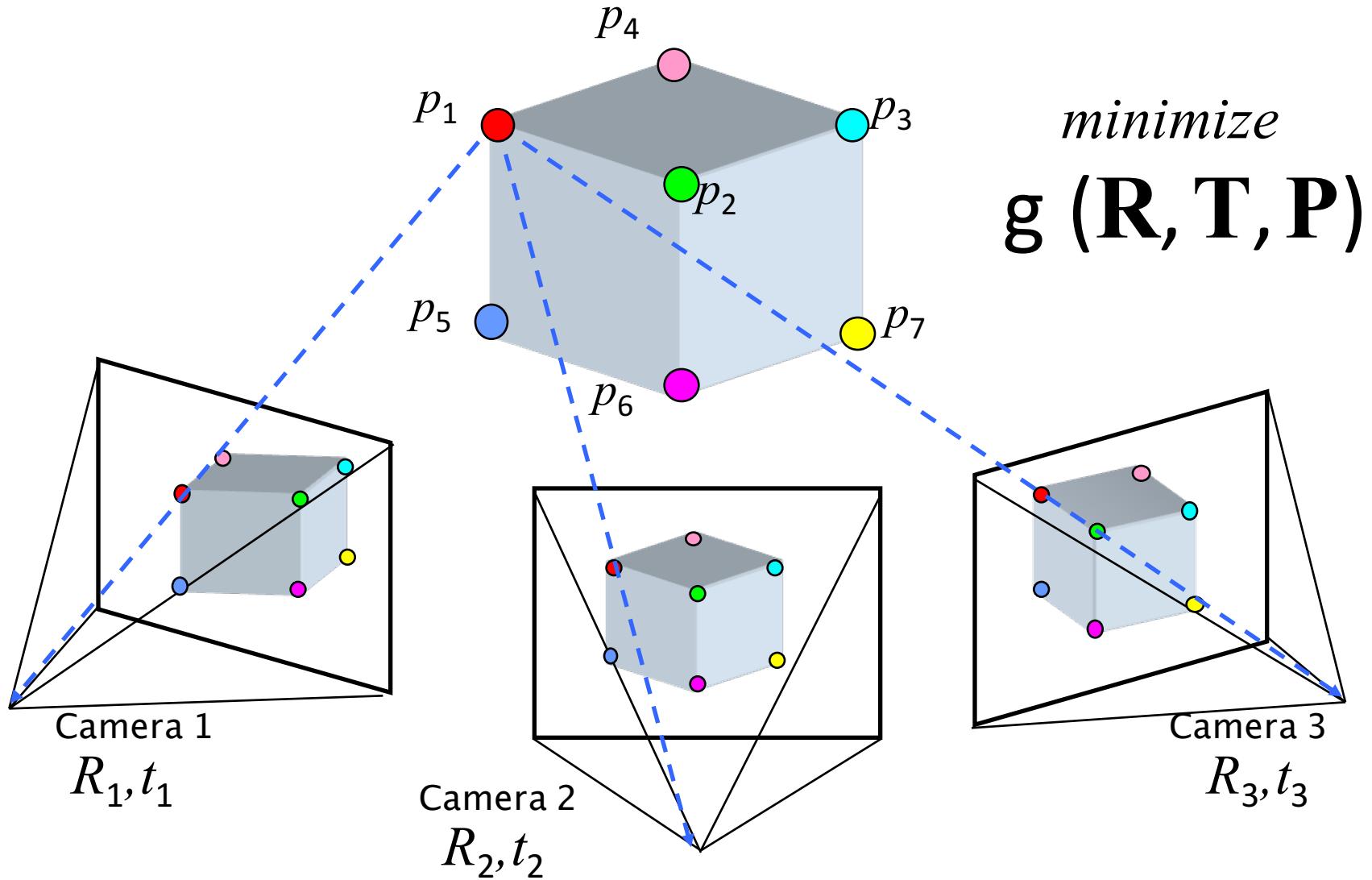
This decomposition minimizes
 $|\mathbf{D}-\mathbf{MS}|^2$

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 3×3 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

flickr

You aren't signed in [Sign In](#) [Help](#)

Home [The Tour](#) [Sign Up](#) [Explore](#) [▼](#)

[Search](#) [▼](#)

Search [Photos](#) [Groups](#) [People](#) [NEW Search by Camera](#)

[SEARCH](#) [Advanced Search](#)

Full text Tags only

We found 13,146 photos matching trevi and rome. [Show details](#)

View: Most relevant • Most recent • Most interesting


From [europefortou_](#)


From [Giampaolo_](#)


From [lanzenberg](#)


From [Nastrina1981](#)

Sponsored Results

Trevi - Rome
Save money on your hotel reservation in Rome.
[www.asiarooms.com](#)

Trevi Hotel Rome
Save on Trevi Hotel in Rome. Instant confirmation with discounted prices.
[www.bookinhotels.com](#)

Trevi Hotel in Rome
A1 Discount Hotels. No booking fees. Great service. Low rates.
[www.A1-Discount-Hotels.com](#)

Hotel Trevi Rome
Search over 120 travel sites.

Images on the Internet



Computed 3D structure



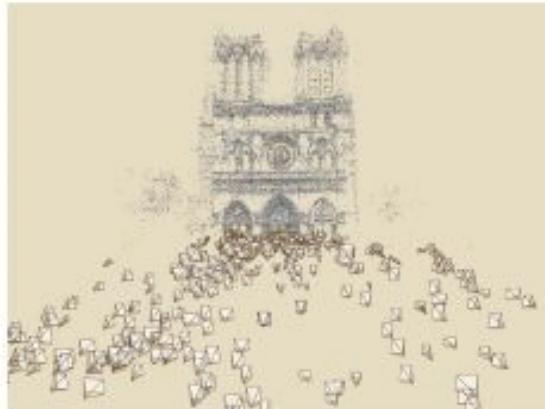
Photo Tourism

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](#)[SEARCH](#)[Advanced Search](#)[Search by Camera](#) Full text Tags only

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

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

Show: [Details](#) • [Thumbnails](#)

[Slideshow](#)



From [Dmitry...](#)



From [Lisa...](#)



From [Lisa...](#)



From [jerkhead29](#)



From [Lisa...](#)



From [Lisa...](#)



From [Lisa...](#)



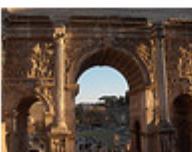
From [Dmitry...](#)



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](#)

[What to Do Rome](#)

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

[Rome with Kids](#)

Get Extra Savings on a Family Vacation to Rome. Limited Time Only.

[www.AdventuresByDisney.com](#)

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

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

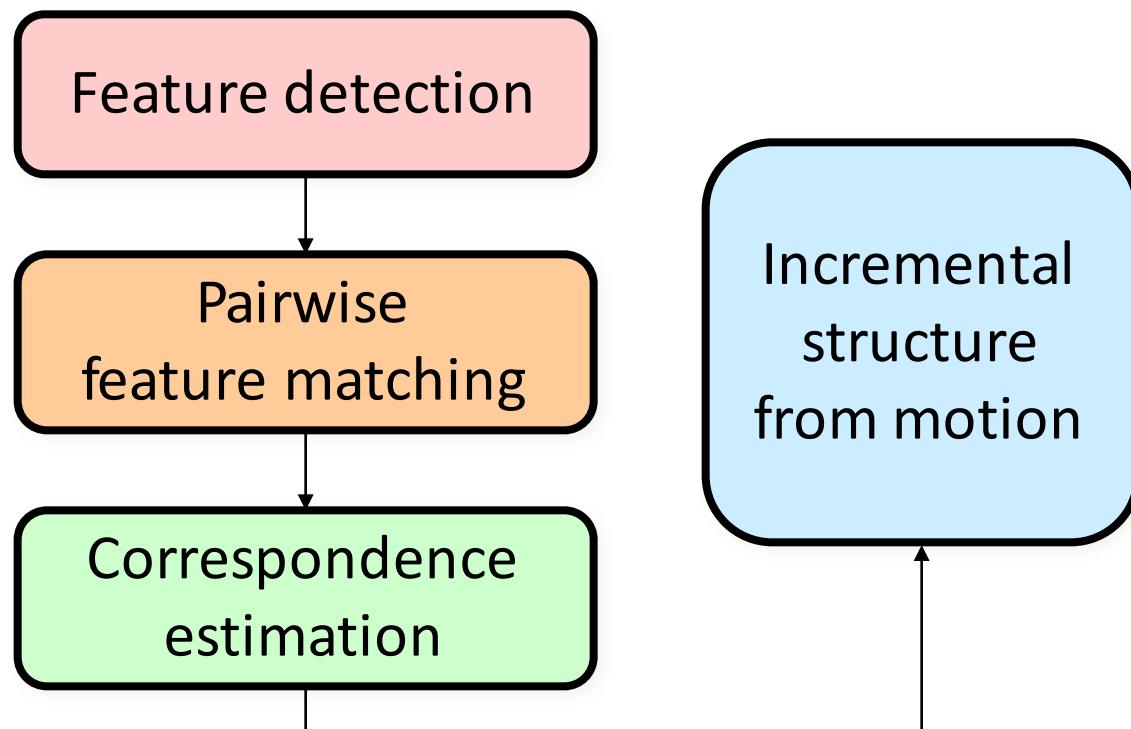
[Rome Italy Day Tours](#)

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



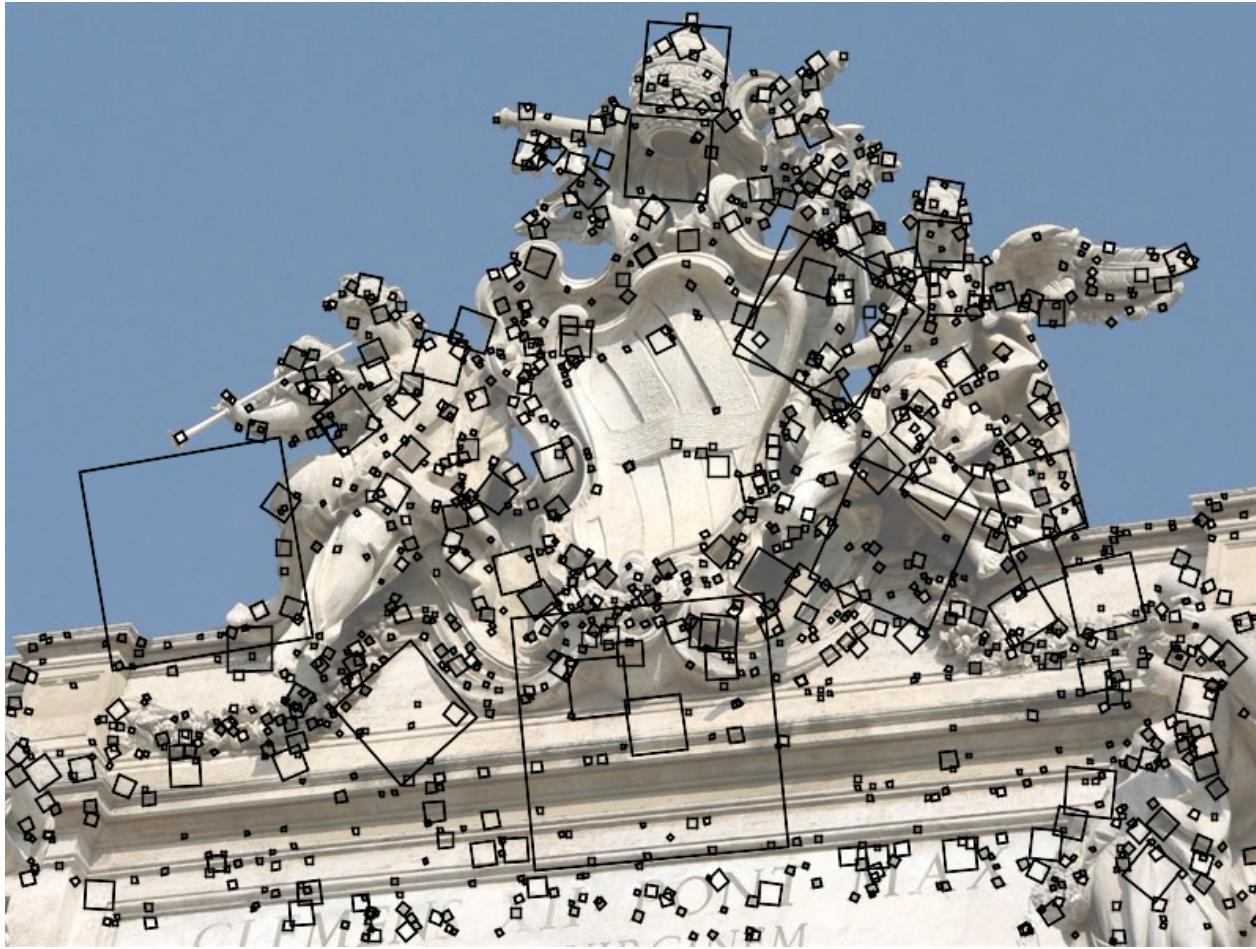
Scene reconstruction

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



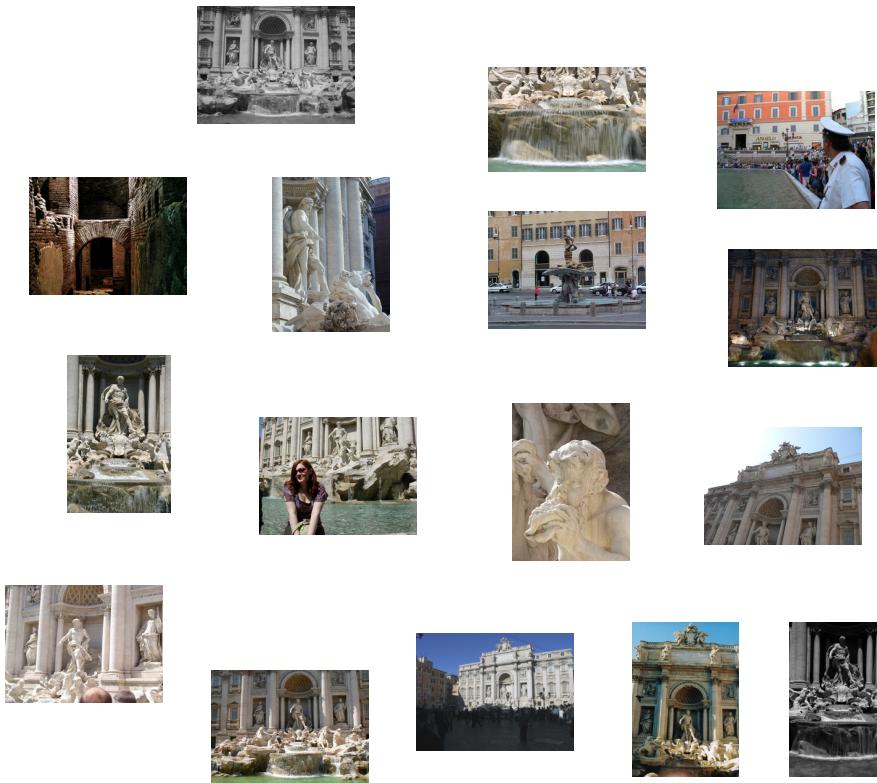
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



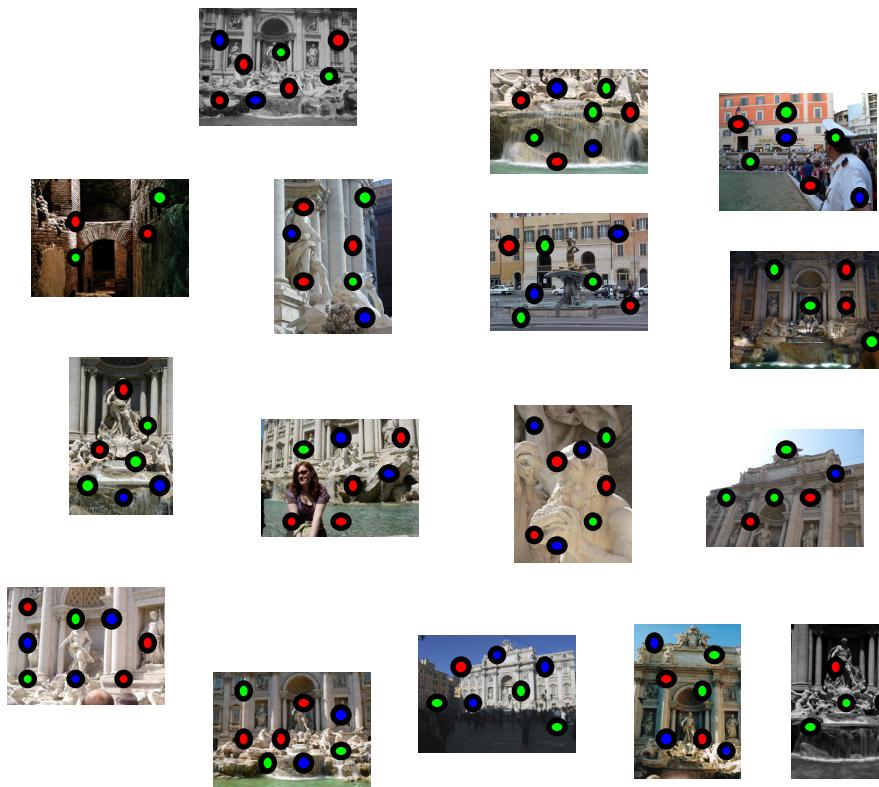
Feature detection

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

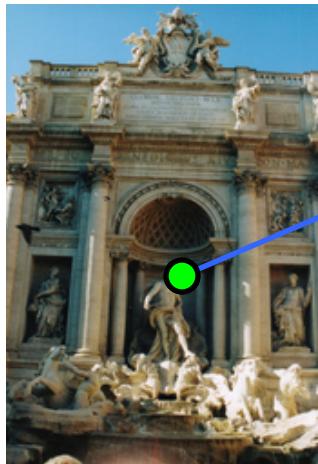


Image 1

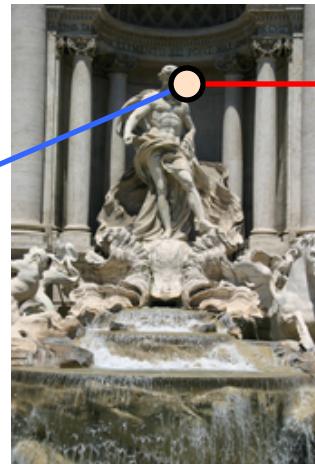


Image 2

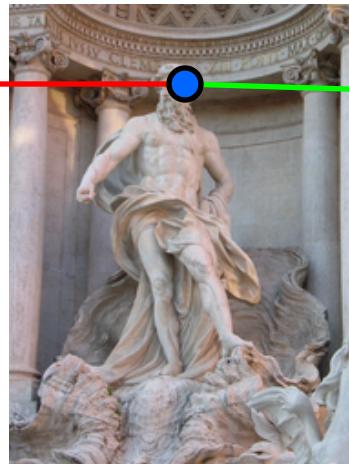


Image 3

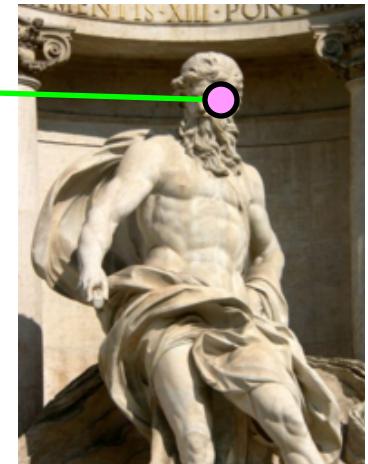
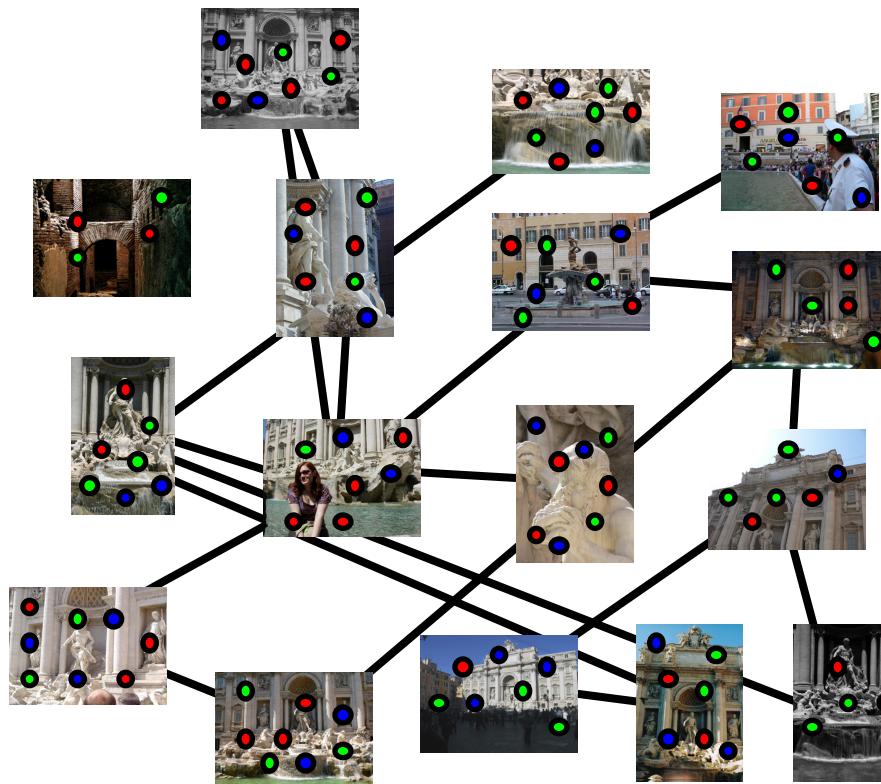


Image 4



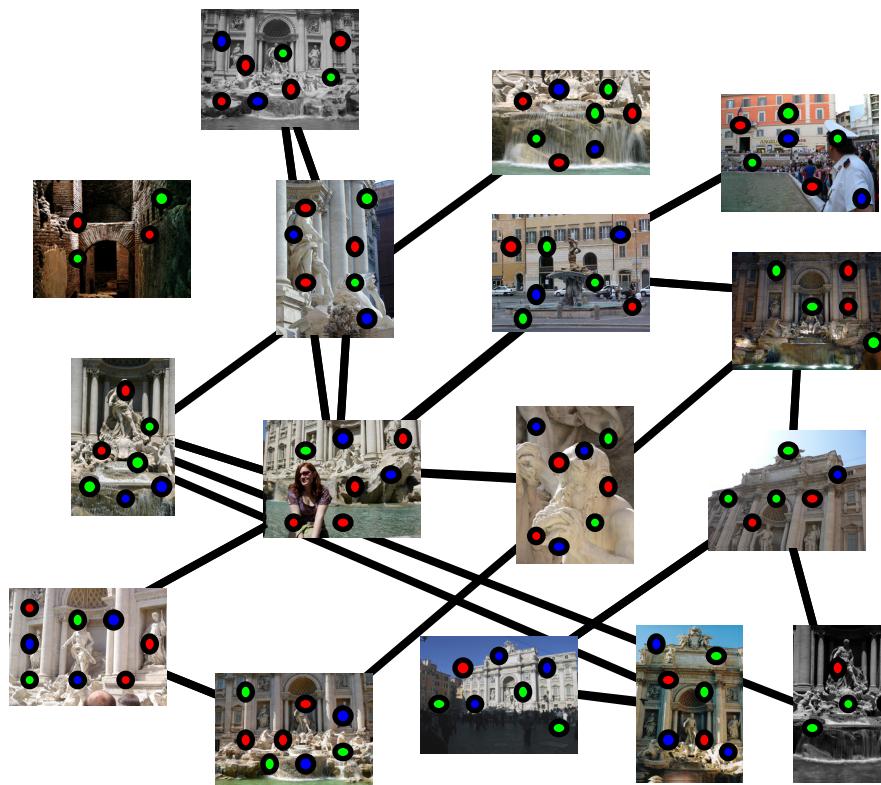
Feature matching

Match features between each pair of images

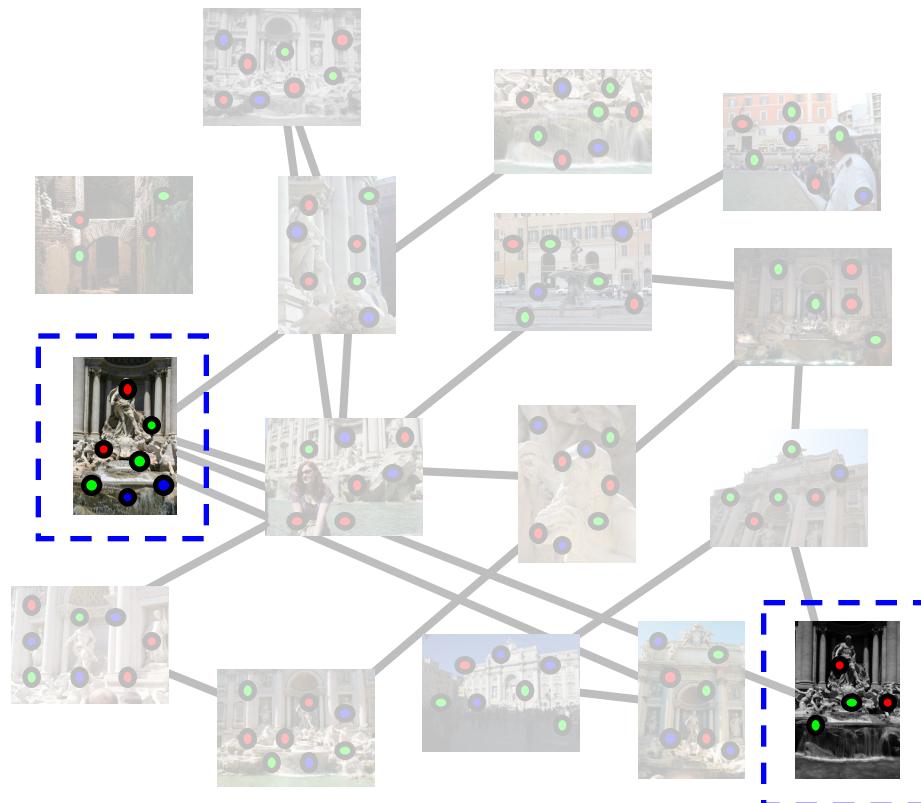


Feature matching

Refine matching using RANSAC [Fischler & Bolles 1987]



Incremental structure from motion



Incremental structure from motion

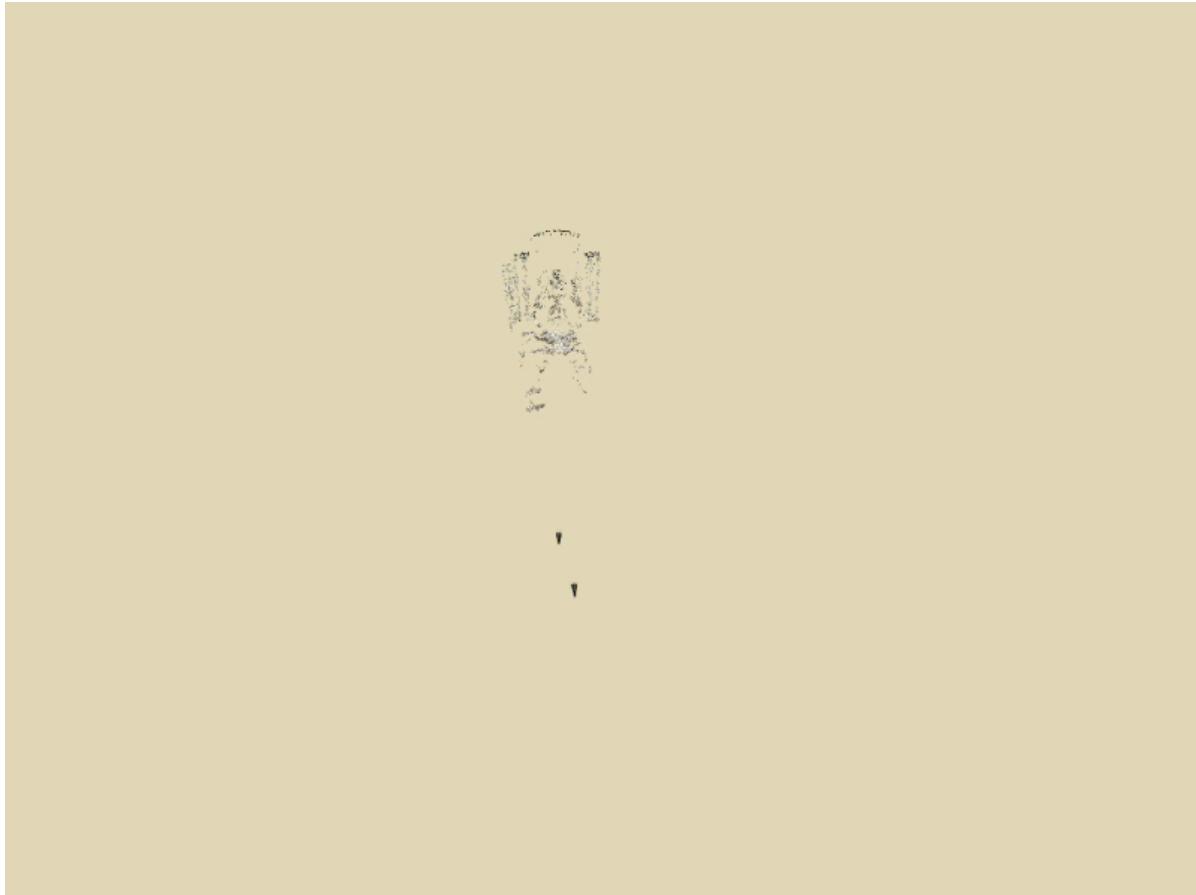
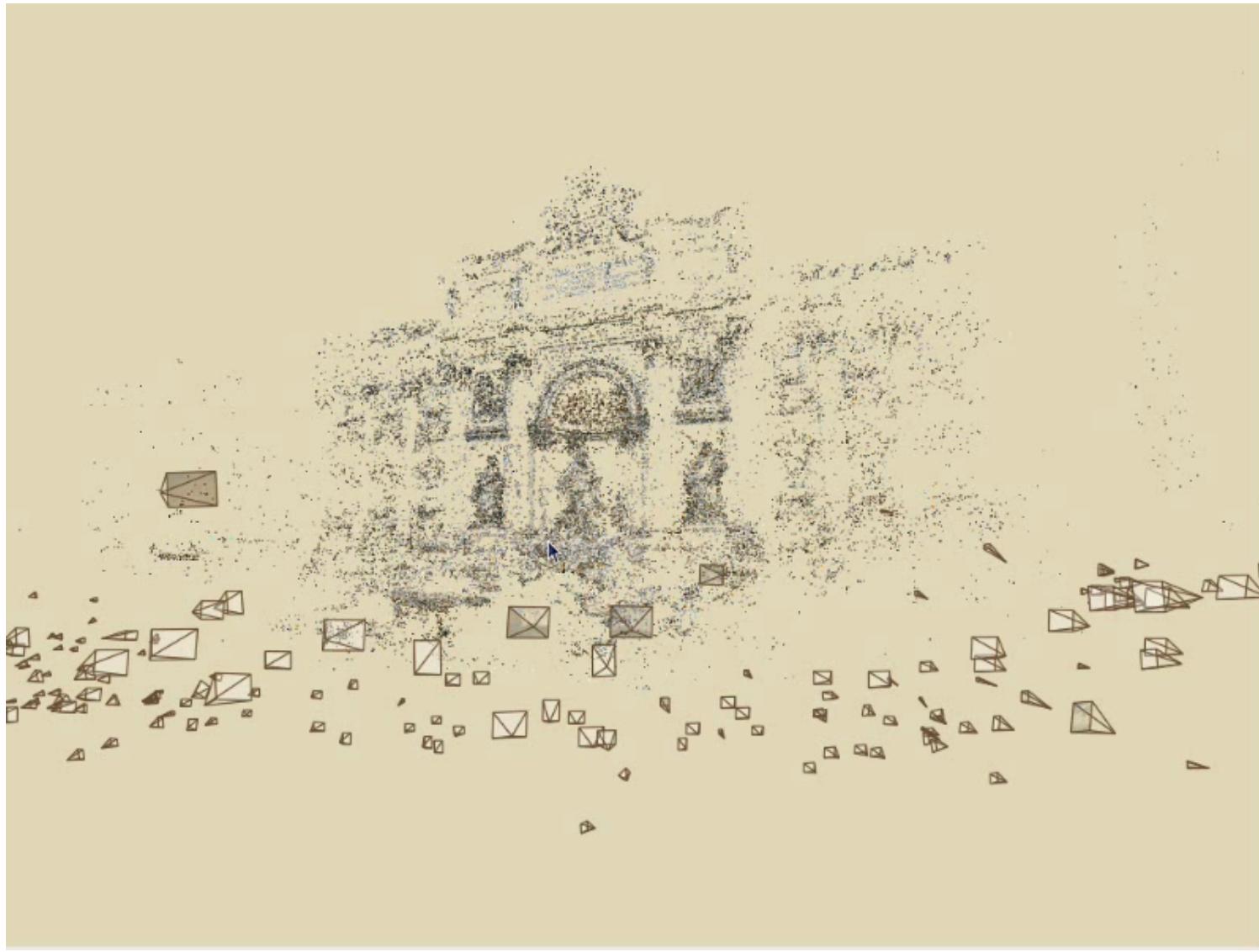
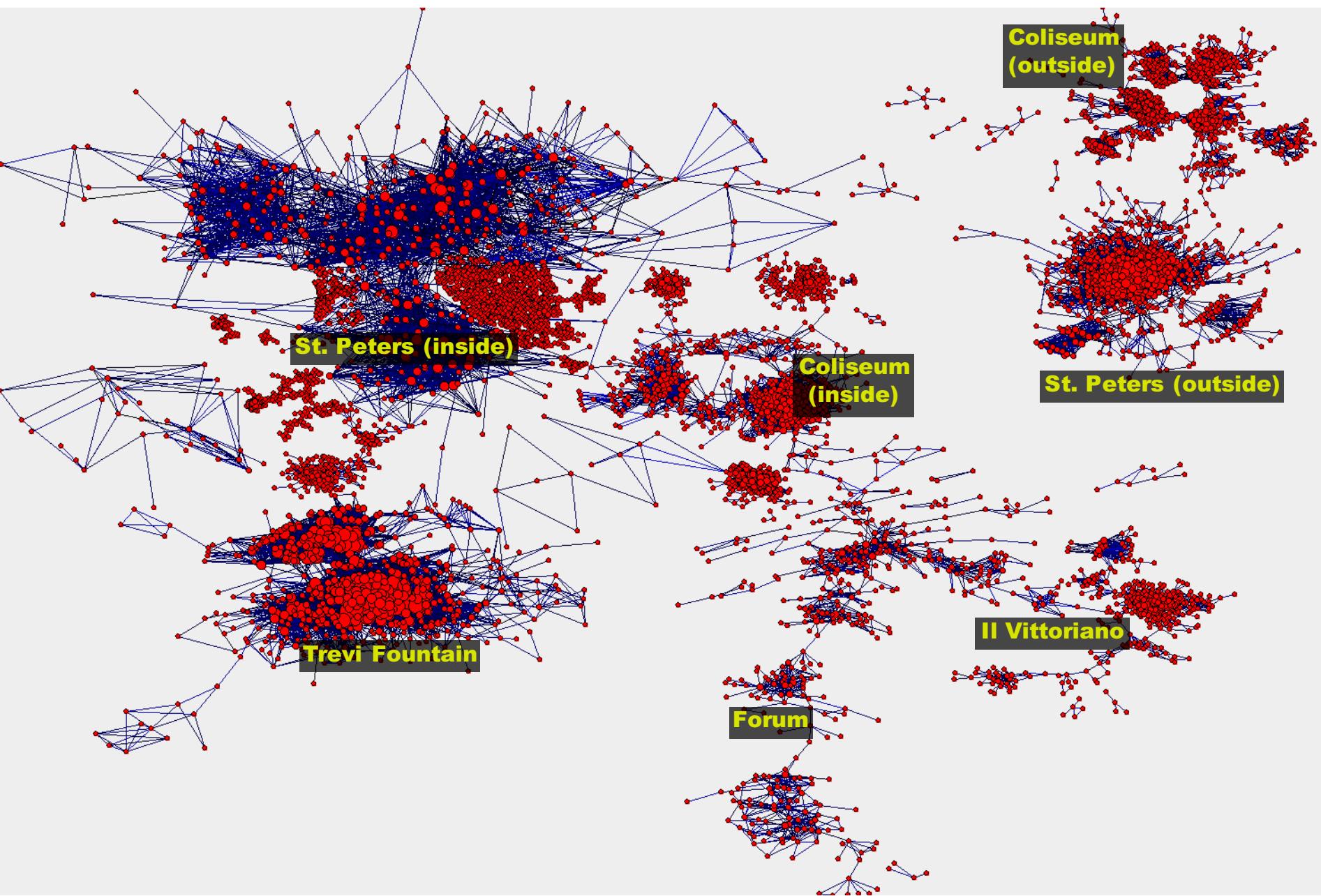


Photo Explorer





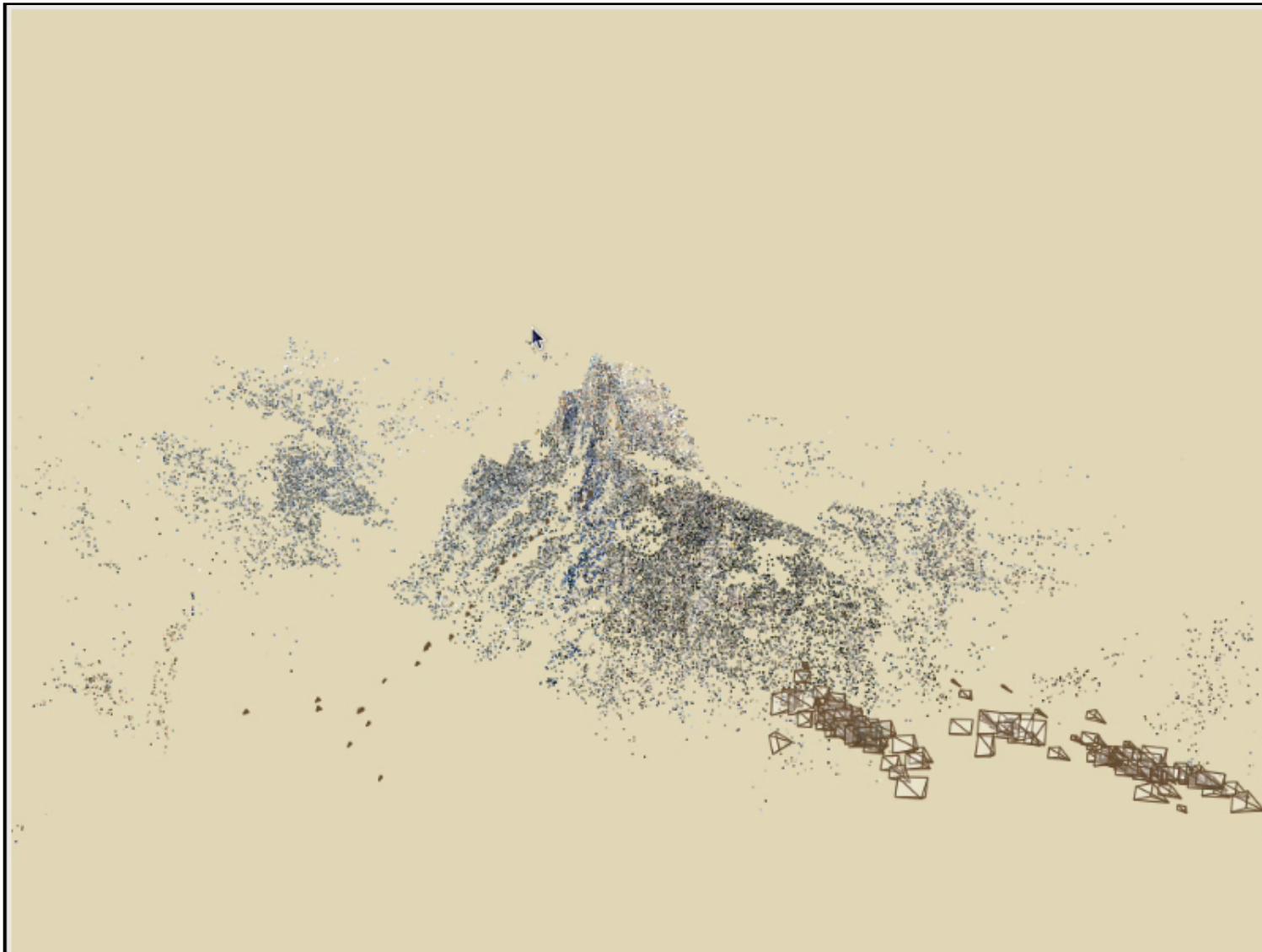
Navigation: Prague Old Town Square



Hierarchical annotations



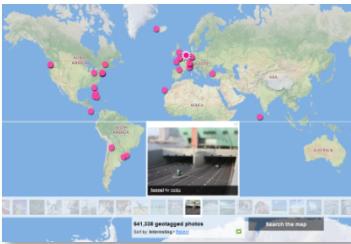
Locking the camera (stabilization)



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



Screenshot of a Mozilla Firefox browser window displaying the Wikipedia article for the Space Needle.

The page shows the following details:

- Page title: Space Needle - Wikipedia, the free encyclopedia
- Page URL: http://en.wikipedia.org/wiki/Space_Needle
- Page views: 37,751
- Coordinates: 47°48'05.1''N 122°34'03.3''W
- Summary: You can help Wikipedia change the world!
- Content: From the fundraising blog – [Wikibooks and the Future of Free Education](#)
- Image: A small thumbnail image of the Space Needle.
- Text: 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 136 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.
- Navigation: Main Page, Contents, Featured content, Recent changes, Contact Wikipedia, Donate to Wikipedia, Help.
- Search: Go, Search.
- Toolbox: What links here, Related changes, Upload file.

Virtual tour guide scenario

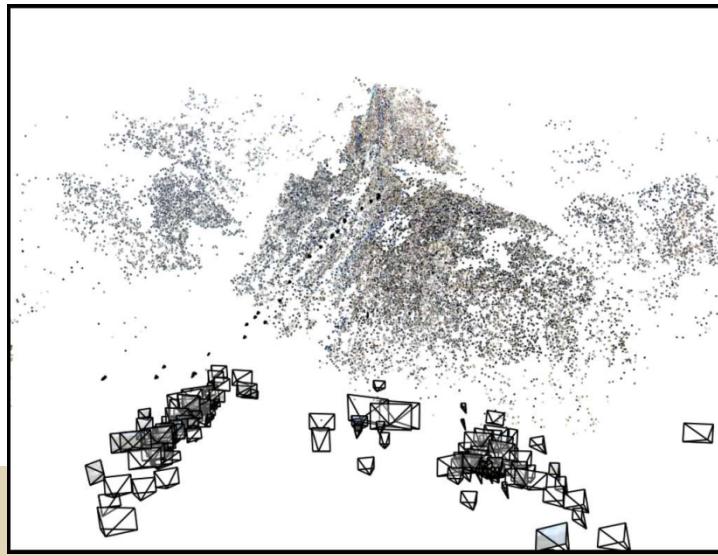
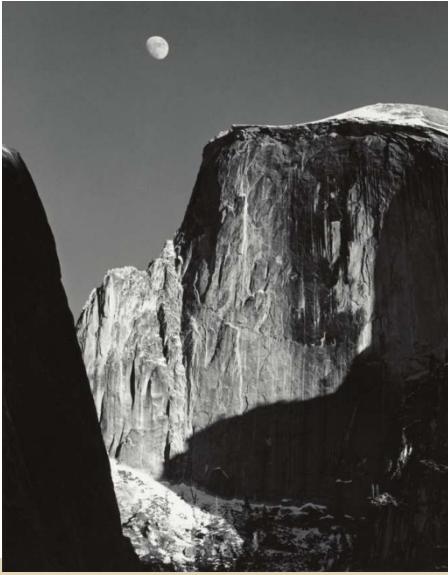


St. Peter's Basilica

- Built: 1506-1626
- http://en.wikipedia.org/wiki/St._Peter's_Basilica



Rephotography



Windows Live Local
powered by Virtual Earth Beta

Welcome to Windows Live Local

What's New?
We've changed our name, changed our look, and added lots of new features!

Here are a few of the great new features:

- See incredible [Bird's Eye images](#)
- Create [driving directions](#), then customize how they [print](#).
- Add your own [custom pushpins](#) anywhere on the map (just right-click on the map).

road aerial

What: Business name or category Where: Address, city, or other place

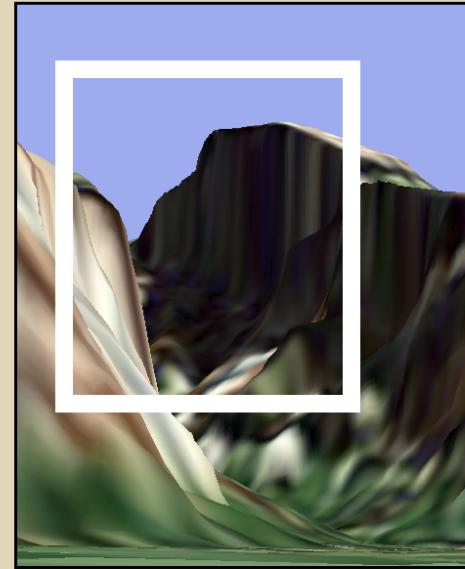
Scratch Pad

Title: Ansel Adams

Edit... or Delete Zoom to street level Drive From... or Drive To... E-mail a Friend...

GRANITE DOME, MOUNTAIN LANE, CALIFORNIA

© 2007 Microsoft Corp. All rights reserved.



Topographic data courtesy USGS