Studying the world and human activity by mining photo-sharing websites

David Crandall
School of Informatics and Computing
Indiana University
Bloomington, Indiana

Social photo sharing websites

facebook

flickr

Picasa™ Web Albums

photobucket

Panoramio from Google

Kodak Gallery

FOTOLOG™

Cyclists climbing past Vaduz Castle

Share This

ADD TO FAVES | BLOG THIS | ALL SIZES | ADD TO GALLERY

Uploaded on May 24, 2007
by claustral

claustral's photostream

966
uploads

browse

This photo also belongs to:

+ Top 200 (Set)

+ Germany & Liechtenstein (Set)

+ cycling (Pool)

- 3 people call this photo a favorite

Tags

- Canon AE-1
- film
- cyclists
- racing
- Vaduz
- castle
- cobbles
- alps

Comments

veronikis says:

i dont know why northern europeans (austrians, germans) insist in punishing themselves with such hard work during the weekeneds! Its horrifying! lol this picture represents the place, with its natural beauty, construction / history and social... (issues ;-)

Posted 33 months ago. (permalink)

claustral pro says:

While I recognize the comment was tongue in cheek, the real masochists are the Italians and Spaniards. They are the road racing nutters of Europe (with some help from the French, Dutch and Belgians). It really isn't so much northern Europe that gets on their bike to hurt themselves.

--

Seen on your photo stream. (?)

Posted 33 months ago. (permalink)

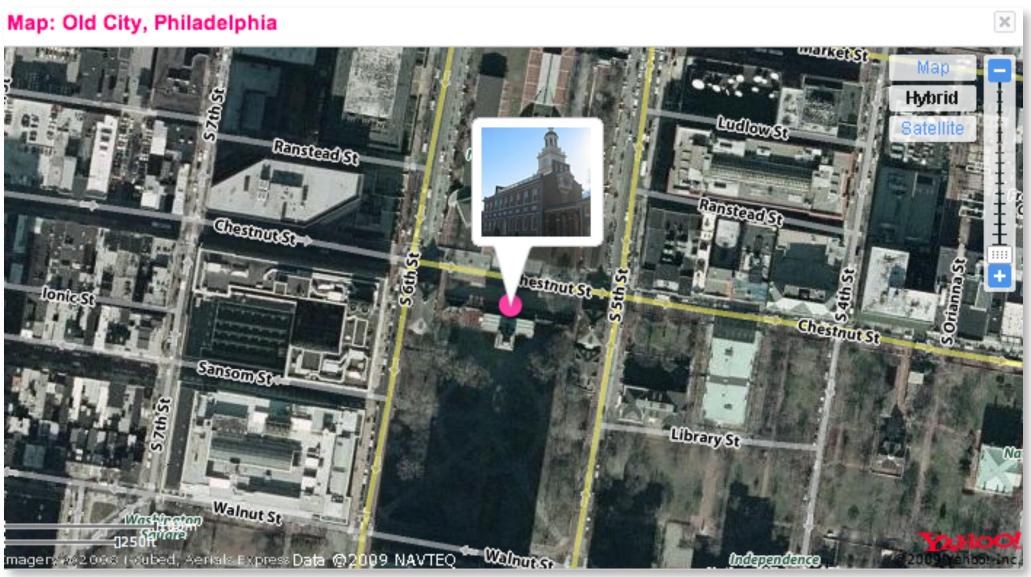
Additional Information

Anyone can see this photo

- Taken in **Vaduz**, Vaduz (map)
- Taken in June 1983
- Viewed **908** times

Geo-tagging

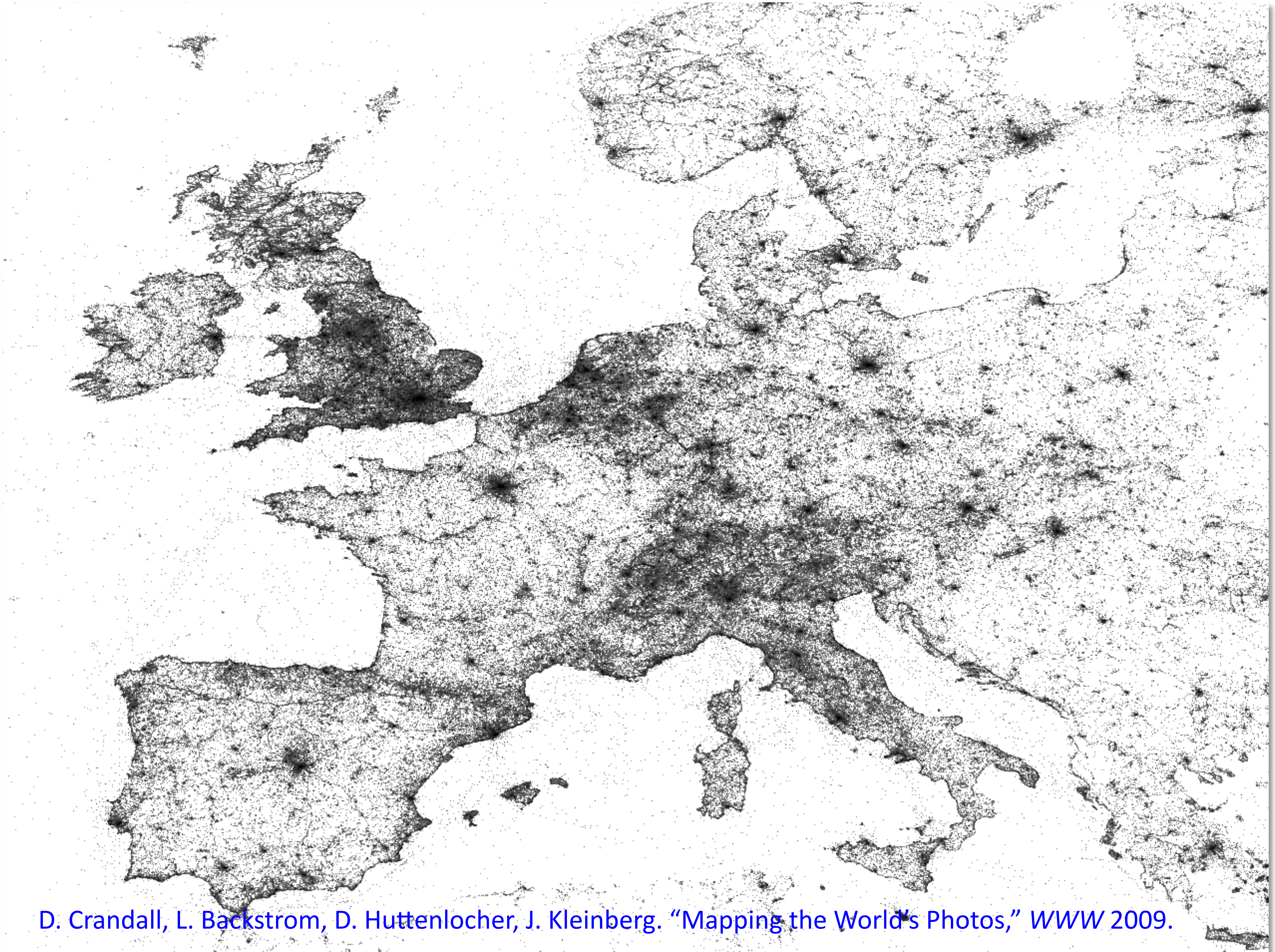
- Many photo-sharing sites allow photos to be annotated with geographic locations (geo-tags)
  - i.e. latitude-longitude coordinates
  - input via a map user interface, or from GPS



Broad research directions

How do we combine visual and non-visual analysis to:

1. organize vast collections of digital photos?
2. mine photos to study the world and its people?



D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. "Mapping the World's Photos," *WWW* 2009.

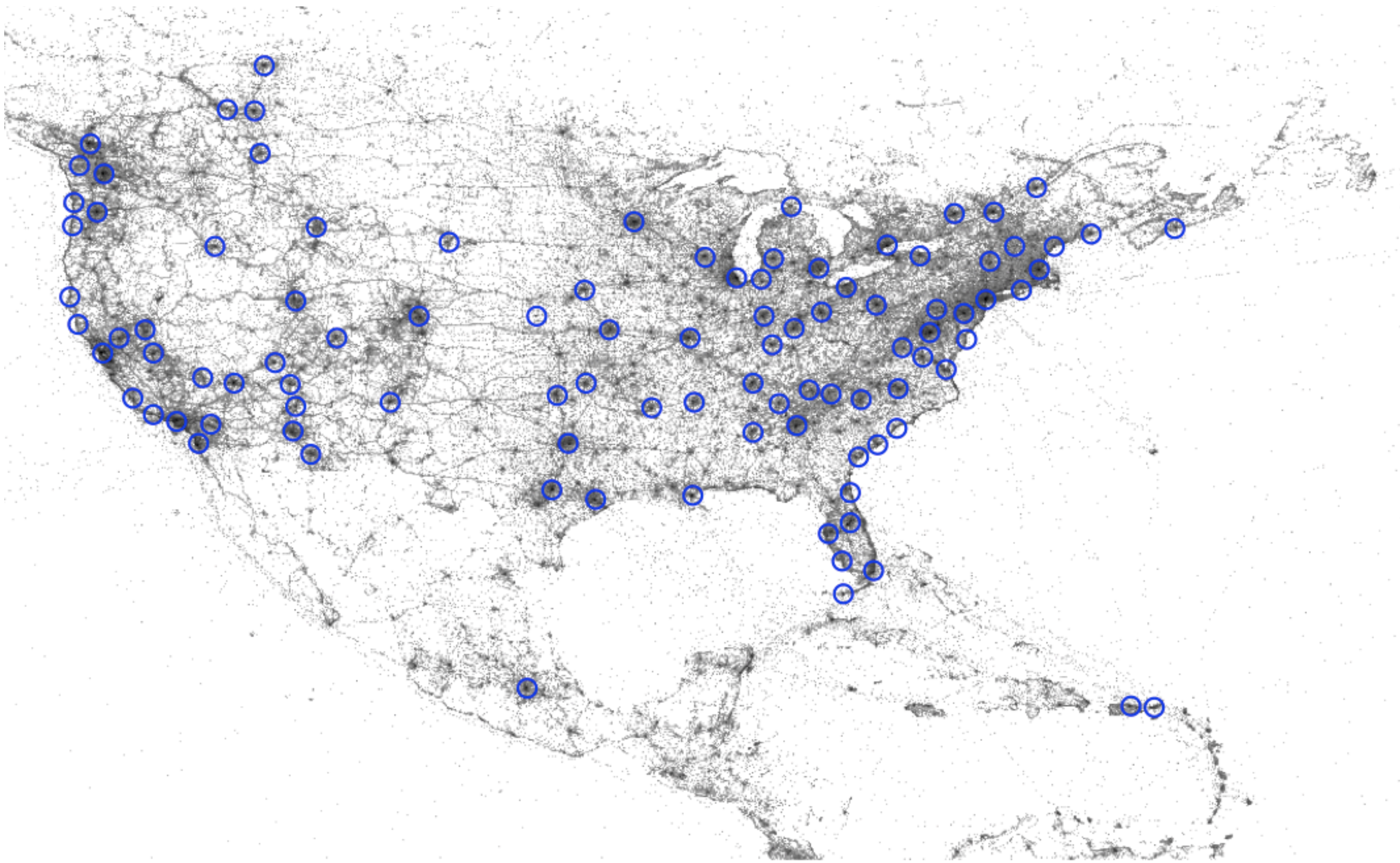
Remainder of talk

- Reconstructing annotated maps of the world
- Reconstructing 3D scene structure
- Studying human mobility
- Future directions

Reconstructing annotated maps of the world

D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. "Mapping the World's Photos," *WWW* 2009.

Finding peaks in the distribution



Finding representative tags

- Automatically find tags that are distinct to a spatial region
  - Consider tags that occur in >5% of the photos in the region
  - Score these tags according to a likelihood ratio,

$$\frac{P(\text{photo } p \text{ has tag } t \mid p \text{ inside region})}{P(\text{photo } p \text{ has tag } t)}$$

  - Limit any single user's contribution to these scores
  - Similar to approaches in [Ahern07], [Kennedy08]

Top city-scale (100 km) peaks

Rank	Users	Photos	Most distinctive tags
1	35860	1204137	newyorkcity nyc newyork
2	29152	1122476	london england
3	25694	1115870	sanfrancisco california
4	18940	586203	paris france
5	17729	775061	losangeles california
6	12025	515884	chicago illinois
7	11834	571698	washingtondc dc washington
8	11346	535671	seattle washington
9	9839	243726	rome roma italy italia
10	9607	280549	amsterdam holland netherlands
11	9318	402658	boston massachusetts
12	9229	258926	barcelona spain
13	9132	304720	sandiego california
14	8369	236818	berlin germany
15	7652	206670	lasvegas vegas nevada
16	7438	112204	firenze florence italy italia tuscany toscana
20	6586	164454	madrid spain españa
47	3620	156693	montreal canada quebec
61	2731	131367	hongkong china
73	2312	122972	pittsburgh pennsylvania
121	1591	20319	yellowstonenationalpark yellowstone wyoming
151	1308	61971	mexicocity df mexico
202	951	27754	ithaca newyork ny
301	579	19551	iowacity iowa
374	383	9580	nassau atlantis bahamas cruise

Top landmark-scale (100m) peaks

1. eiffeltower, paris
2. trafalgarsquare, london
3. bigben, london
4. londoneye, london
5. notredame, paris
6. tatemodern, london
7. empirestatebuilding, newyorkcity
8. venice, venezia
9. colosseum, rome
10. louvre, paris
11. timessquare, newyorkcity
12. rockefeller, newyorkcity
13. cloudgate, chicago
14. vaticano, rome
15. sacrecoeur, paris
16. piccadillycircus, london
17. buckingham, london
18. timessquare, newyorkcity
19. arcdetriomphe, paris
20. grandcentralstation, newyorkcity
21. pantheon, rome
22. sagradafamilia, barcelona
23. towerbridge, london
24. lincolnmemorial, washingtondc
25. britishmuseum, london
26. brandenburggate, berlin
27. toweroflondon, london
28. rialto, venezia
29. applestore, newyorkcity
30. spaceneedle, seattle

Finding representative images

- Can we find *visual* descriptions of hotspots?
  - i.e. an image representative of the place

  - random Flickr images geo-tagged <50m from Independence Hall:

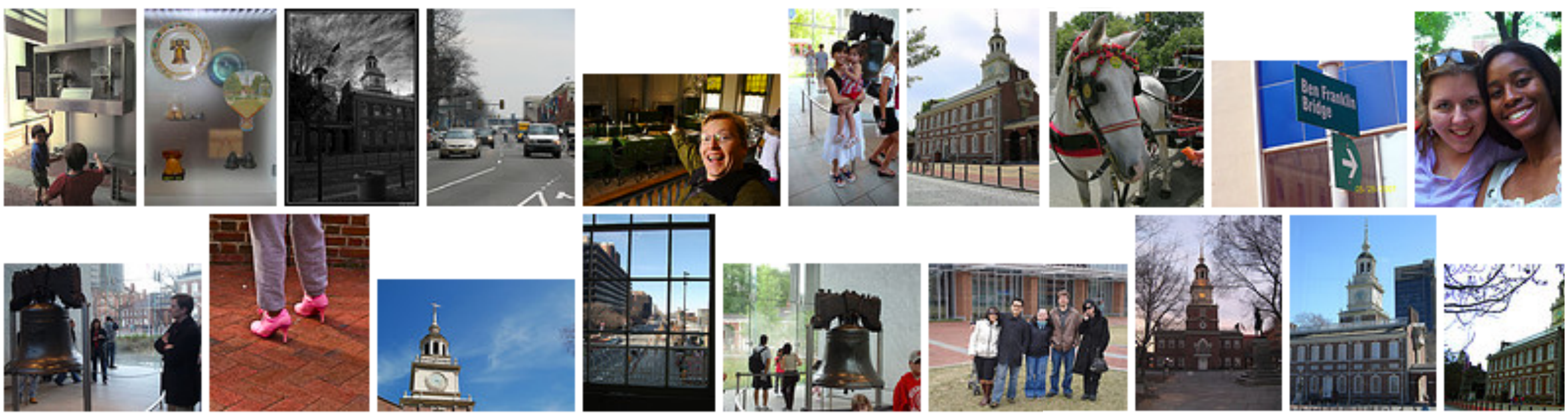
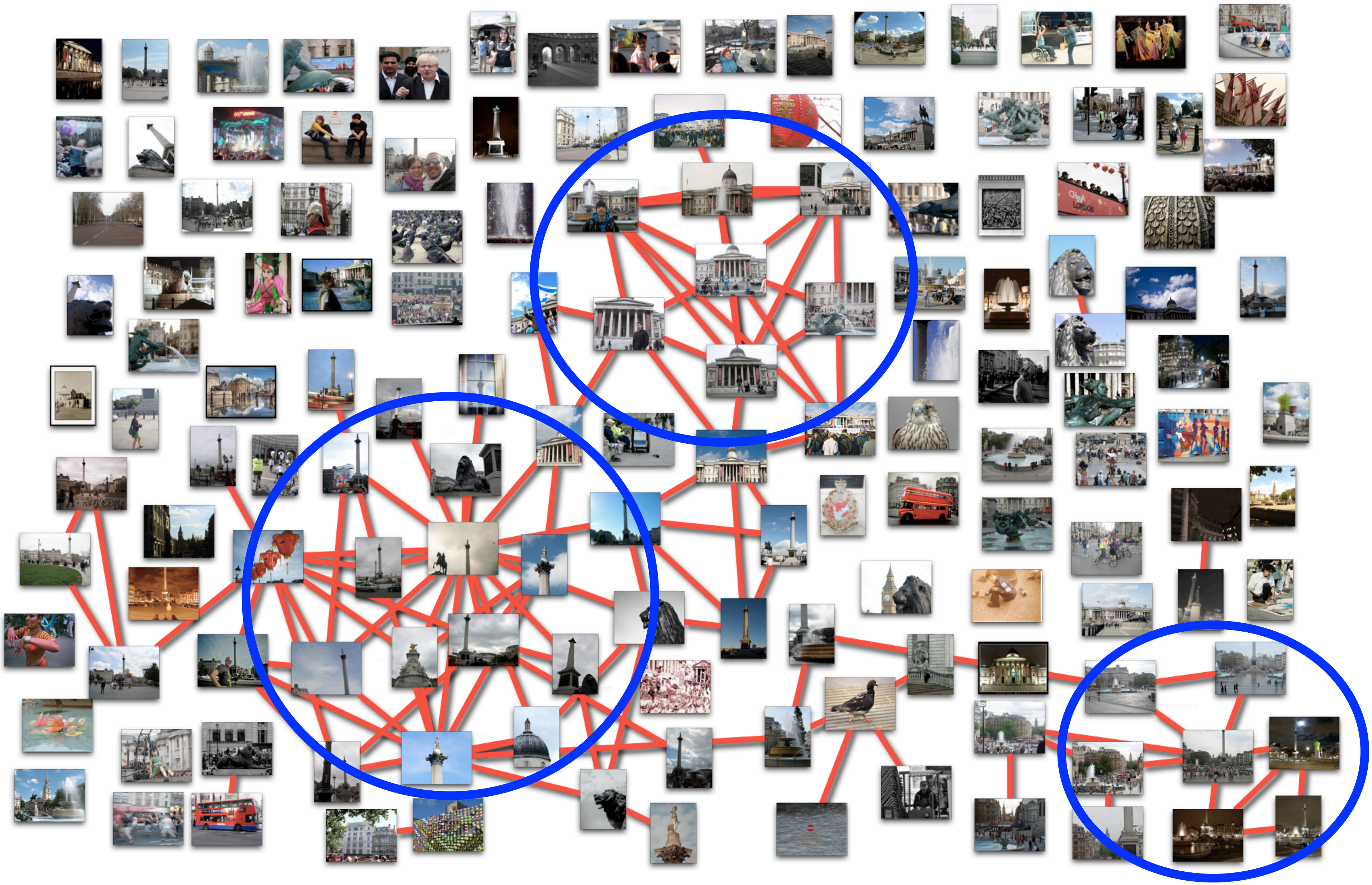
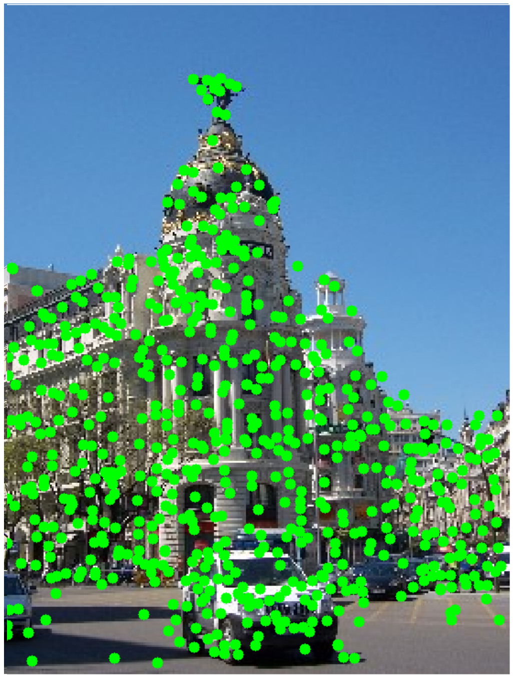


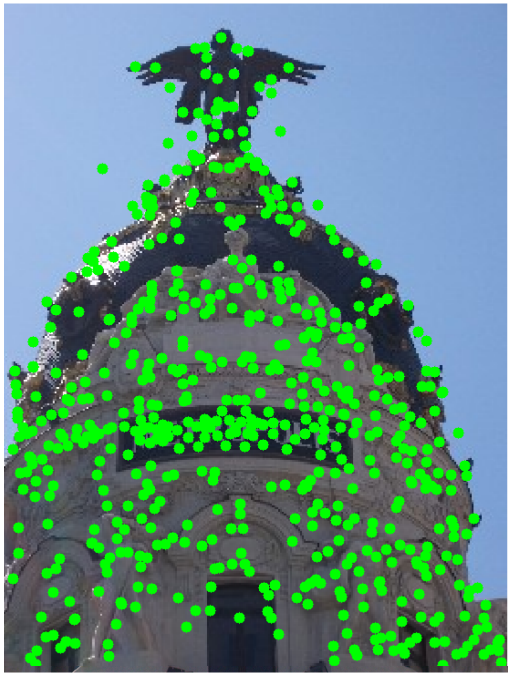
Image similarity graph: example

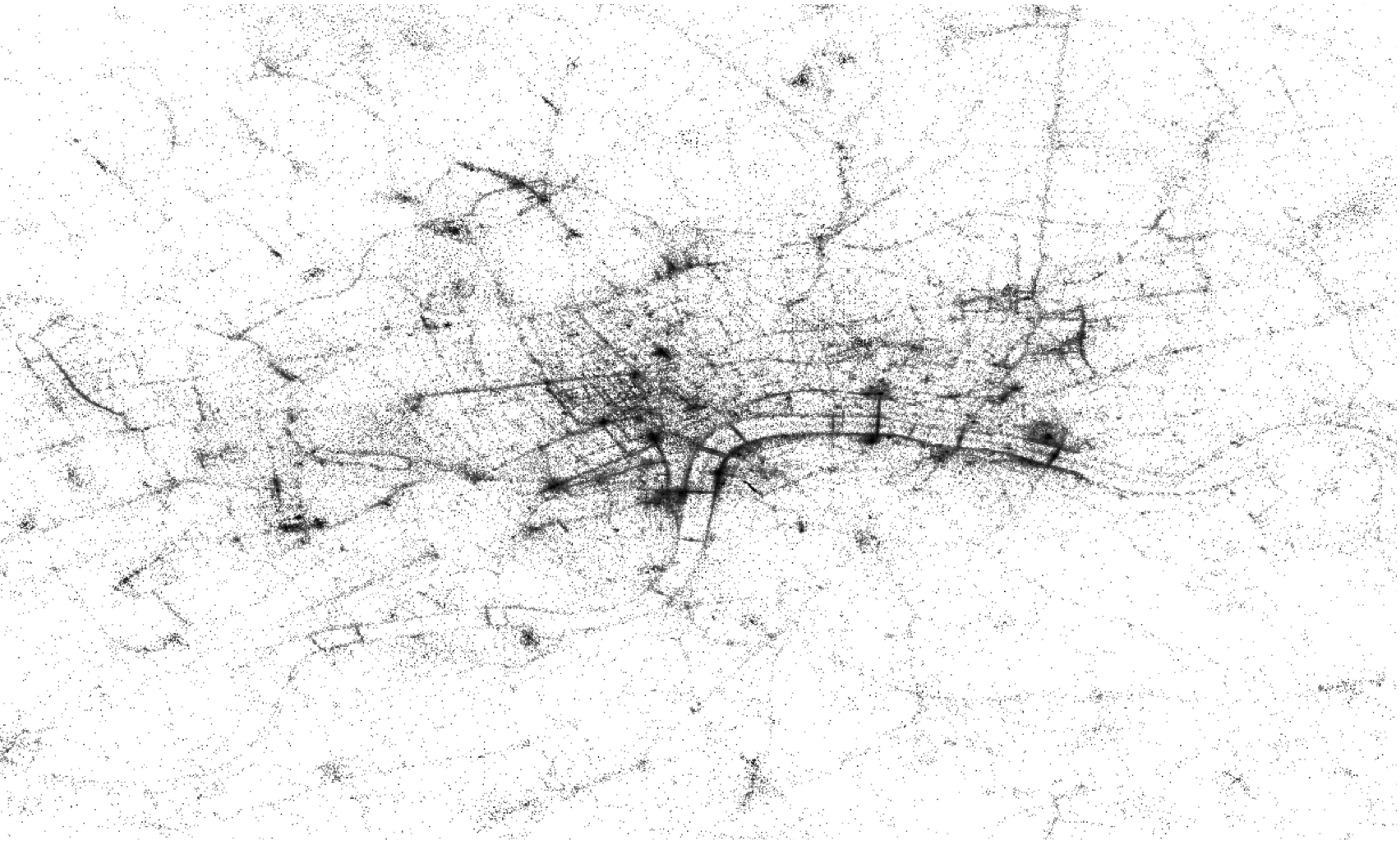


Measuring image similarity

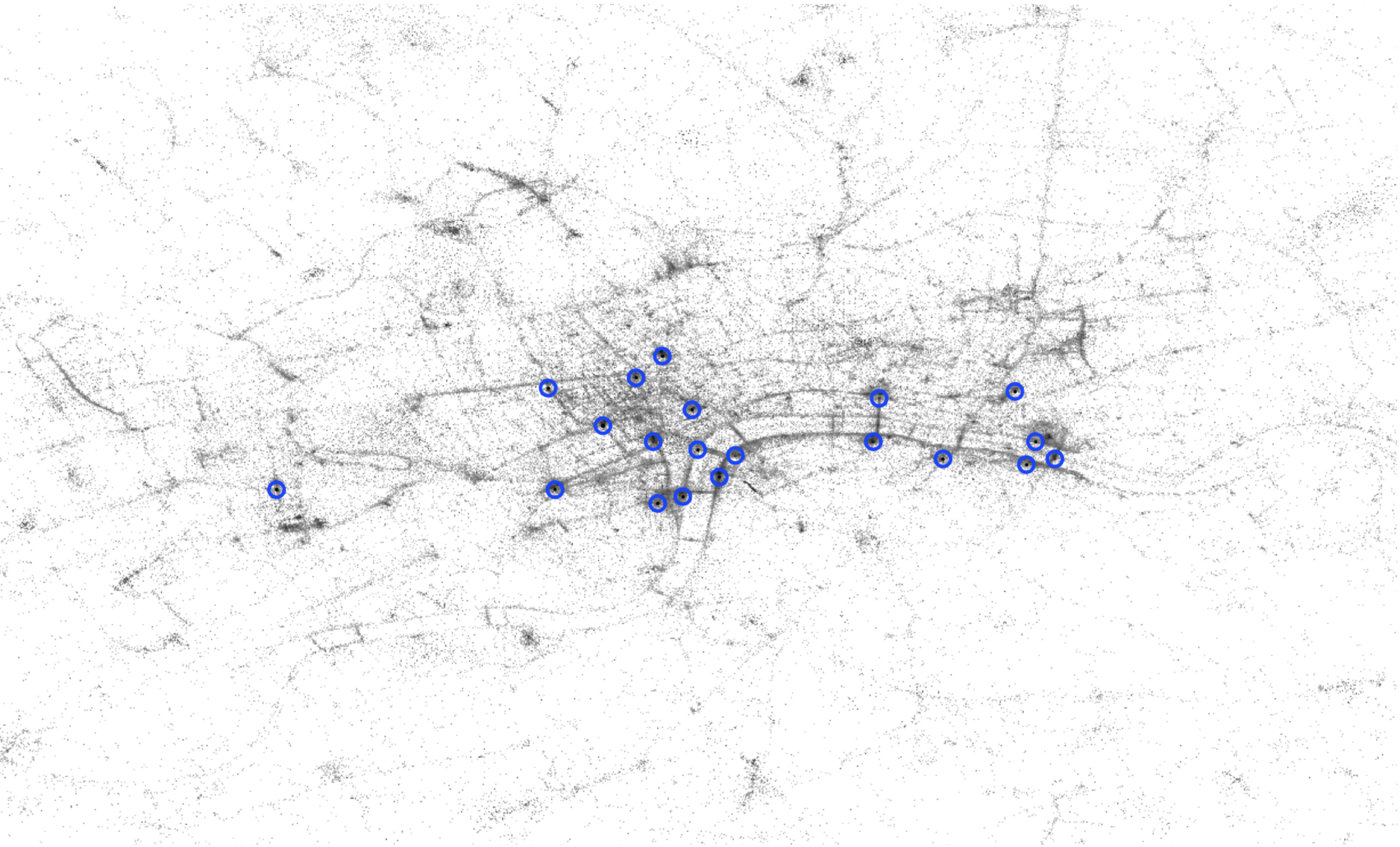
- We use SIFT to extract interest point descriptors [Lowe04]
  - Compute an invariant descriptor for each interest point
  - ~1000 interest points per image, 128-dimensional descriptors
  - To compare 2 images, count number of "matching" descriptors





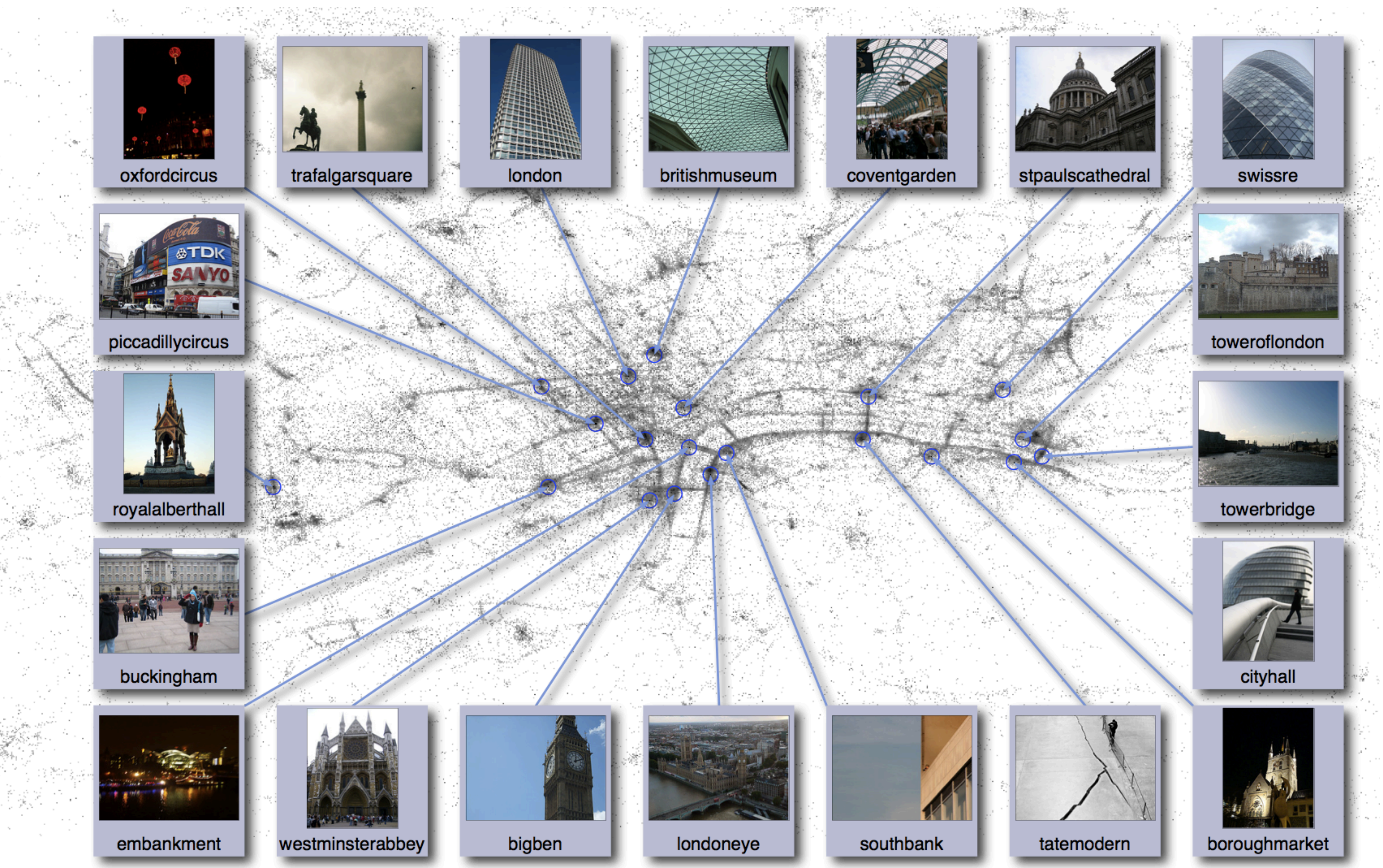


D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. “Mapping the World’s Photos,” *WWW* 2009.



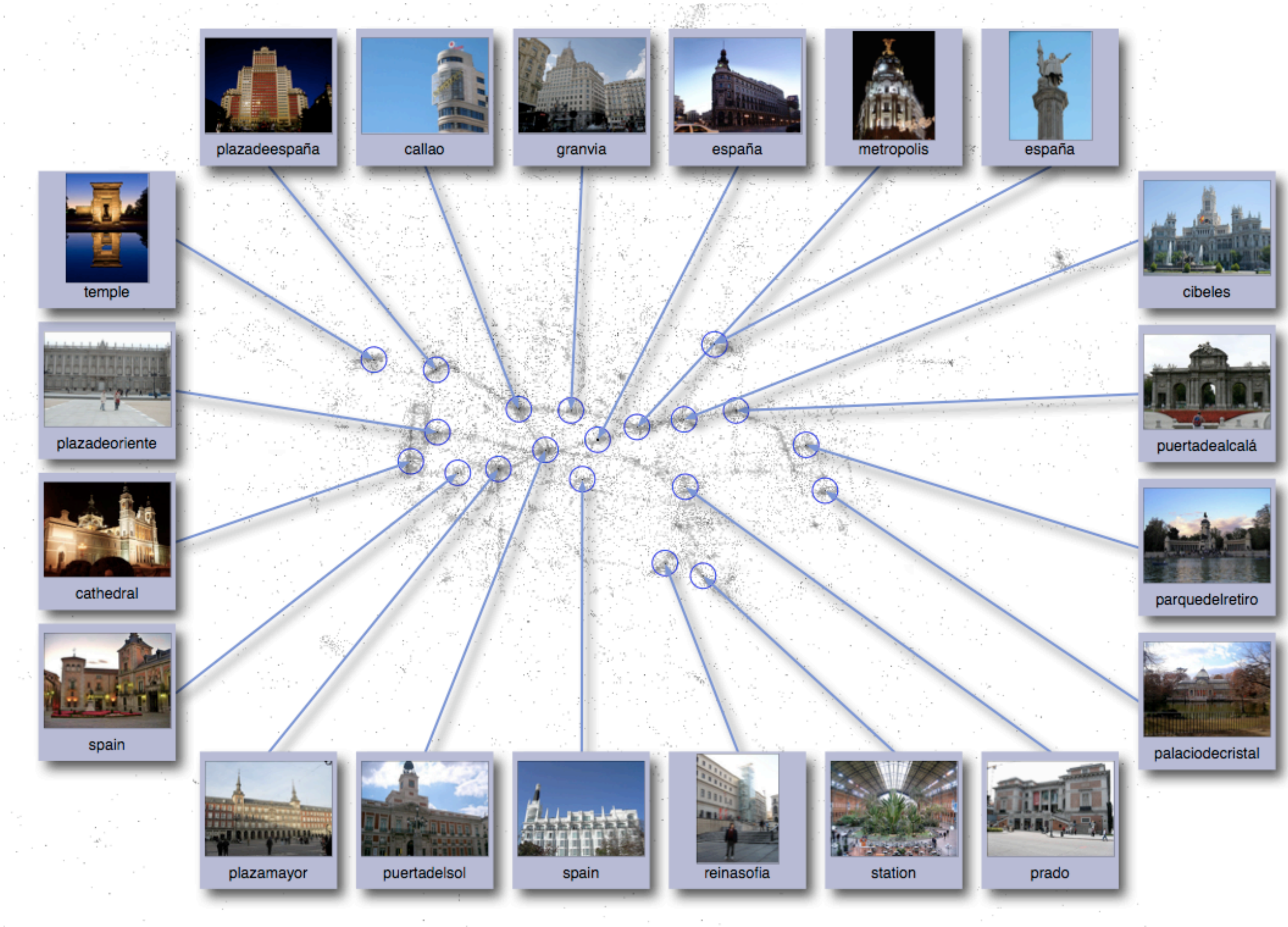
D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. "Mapping the World's Photos," *WWW* 2009.

D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. "Mapping the World's Photos," *WWW* 2009.

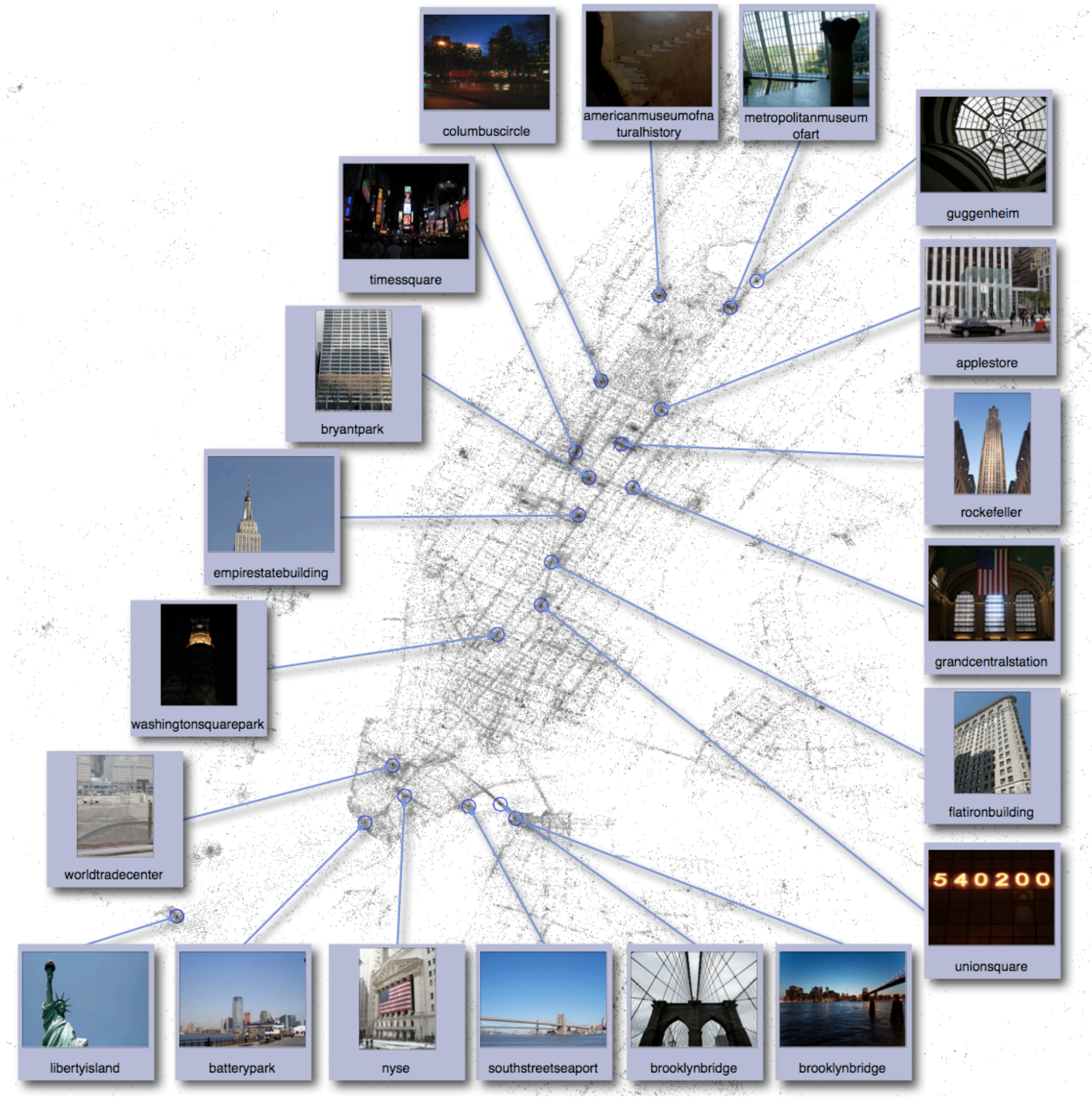


D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg. "Mapping the World's Photos," *WWW* 2009.

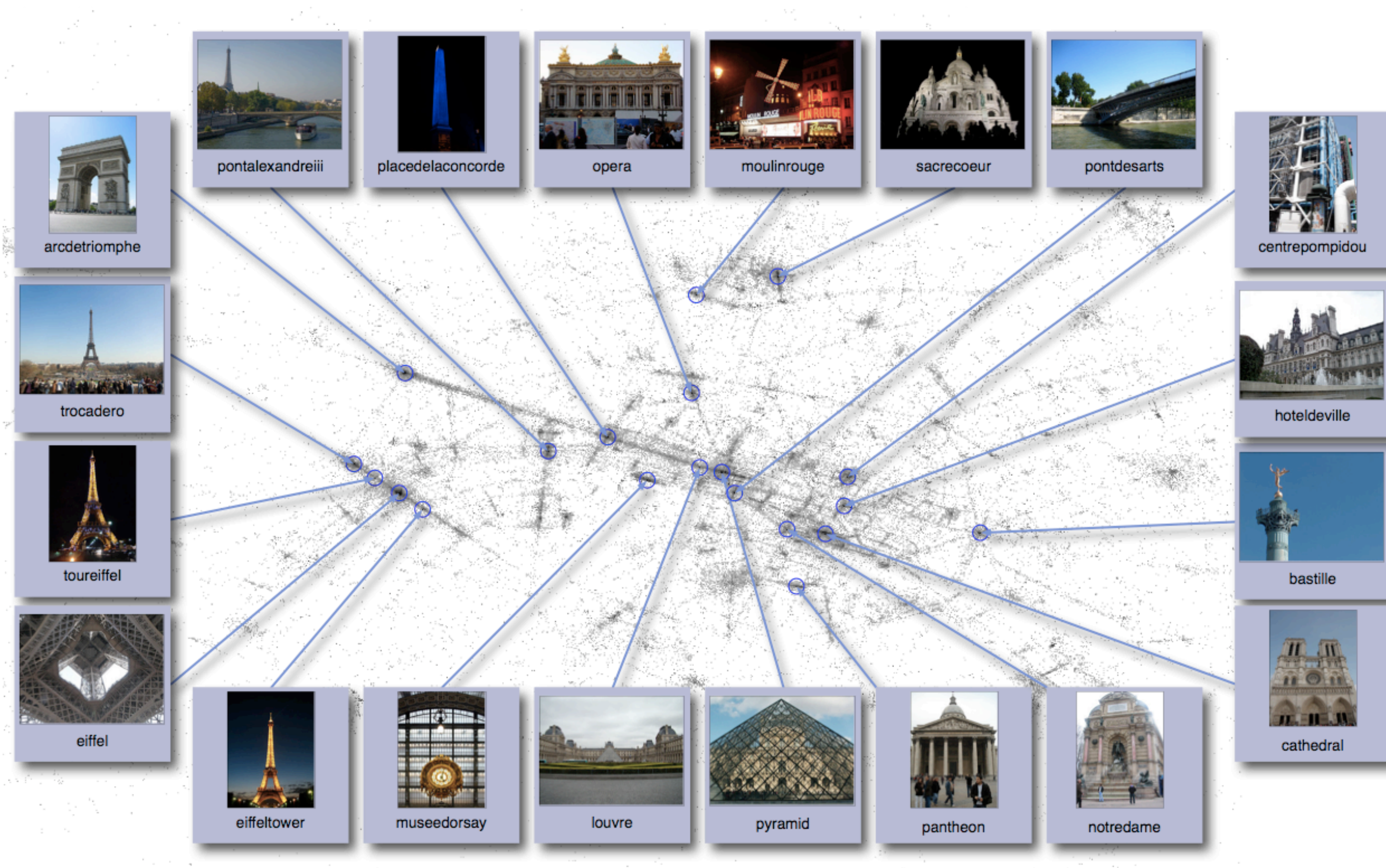
Example: Madrid



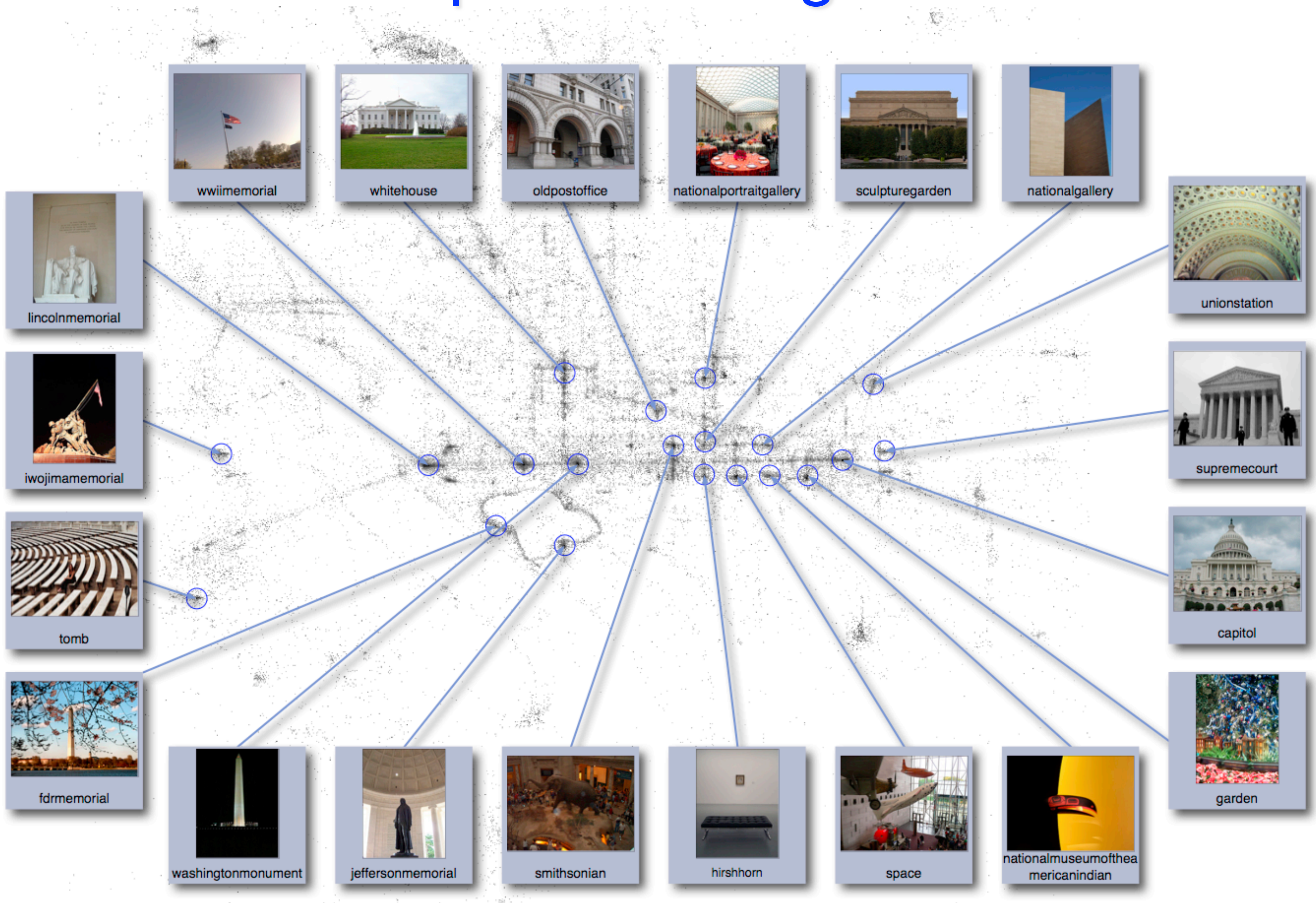
Example: Manhattan



Example: Paris

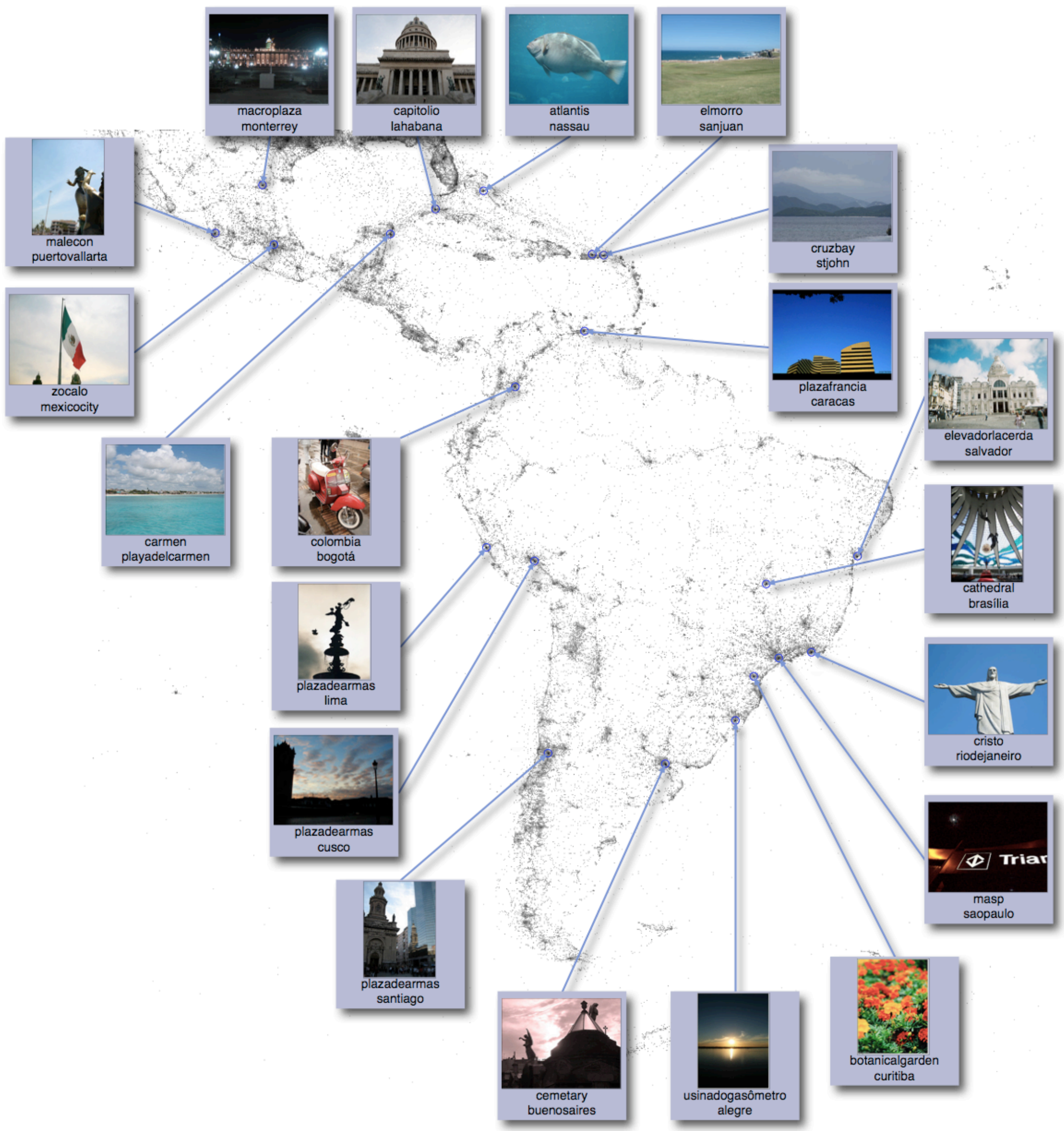


Example: Washington DC

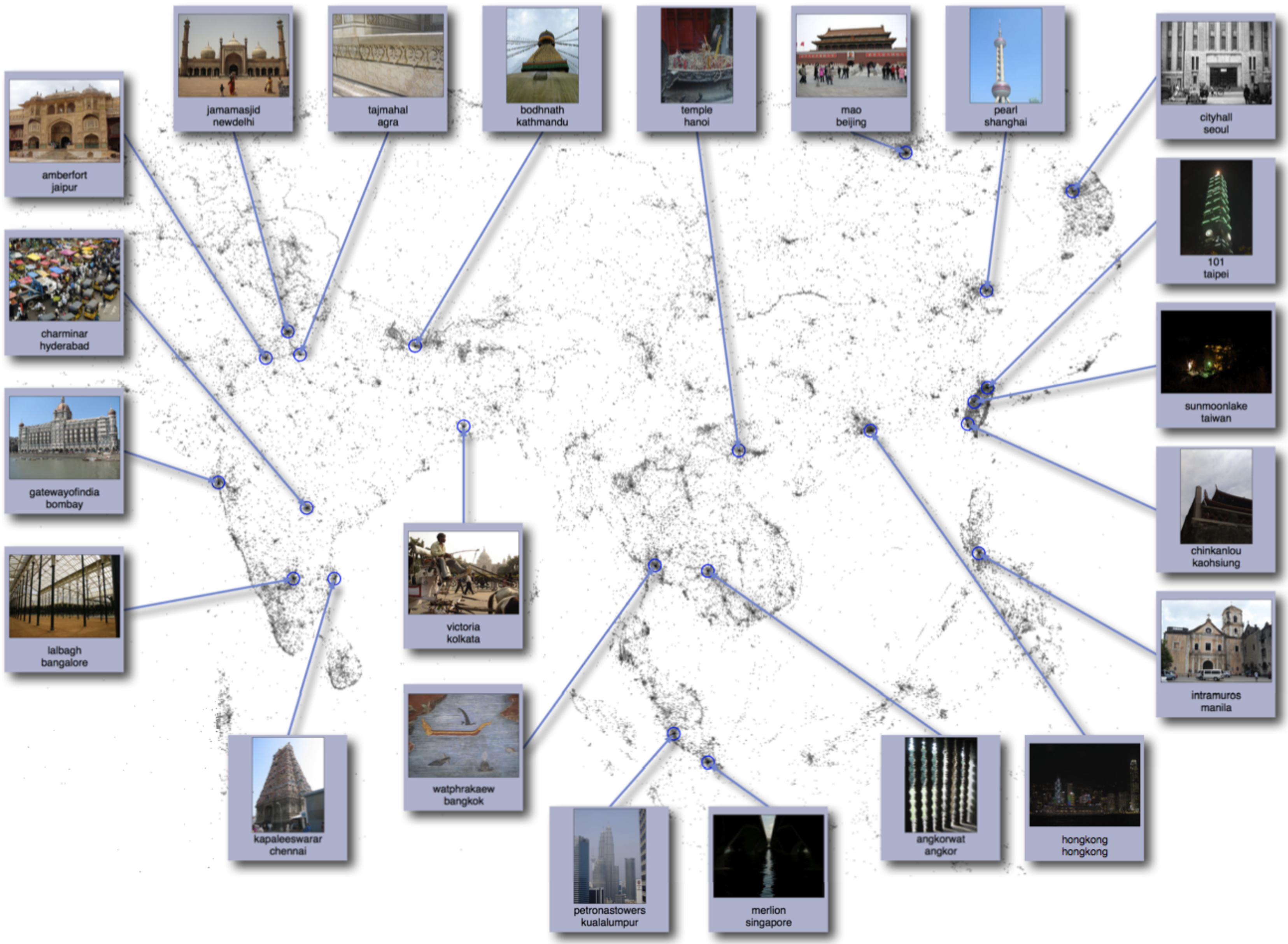


Example: North America

Example: South America



Example: Southeast Asia



Example: UK and Ireland

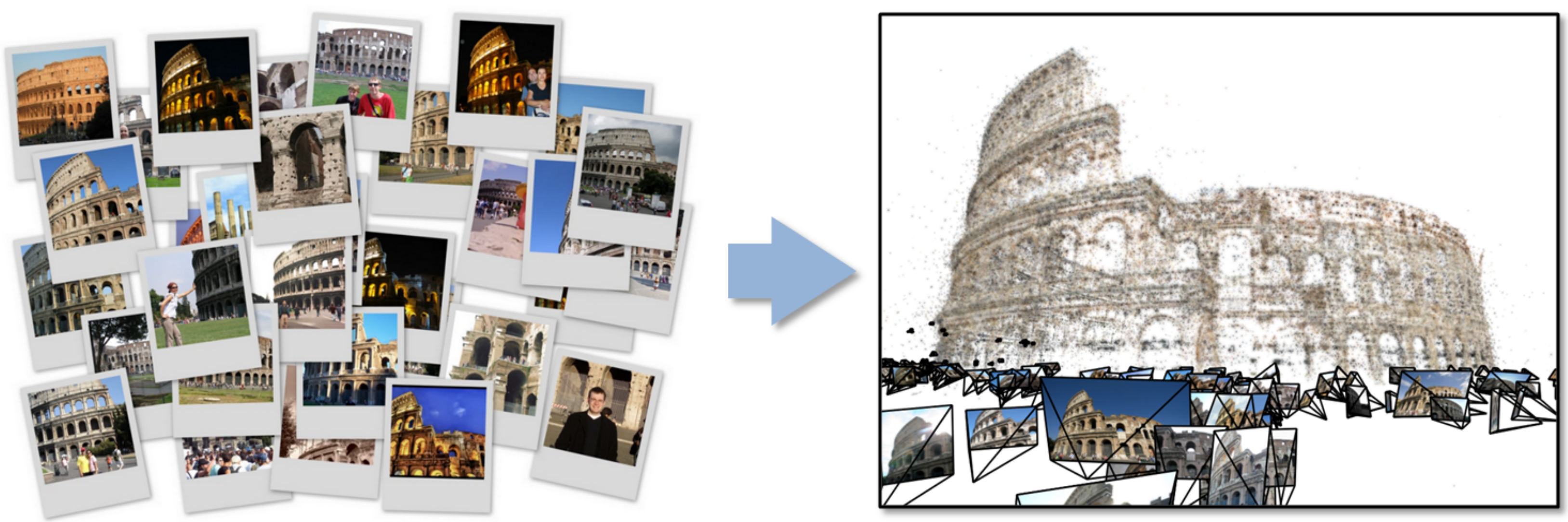
Example: Europe

Reconstructing 3D models of landmarks and cities

D. Crandall, A. Owens, N. Snavely, D. Huttenlocher, "Discrete-Continuous Optimization for Large-scale Structure from Motion," Proc. *CVPR* 2011.

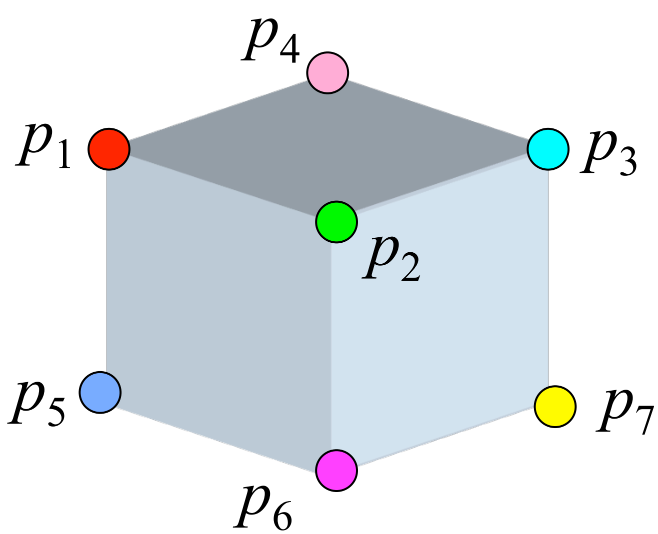
3D reconstruction

- From a collection of unstructured user-generated images, build a 3D model of a landmark
  - Related to the classical Structure from Motion (SfM) problem
  - e.g. [Snavely06], [Li08], [Agarwal09], [Frahm10], and commercial products (e.g. Microsoft PhotoSynth)

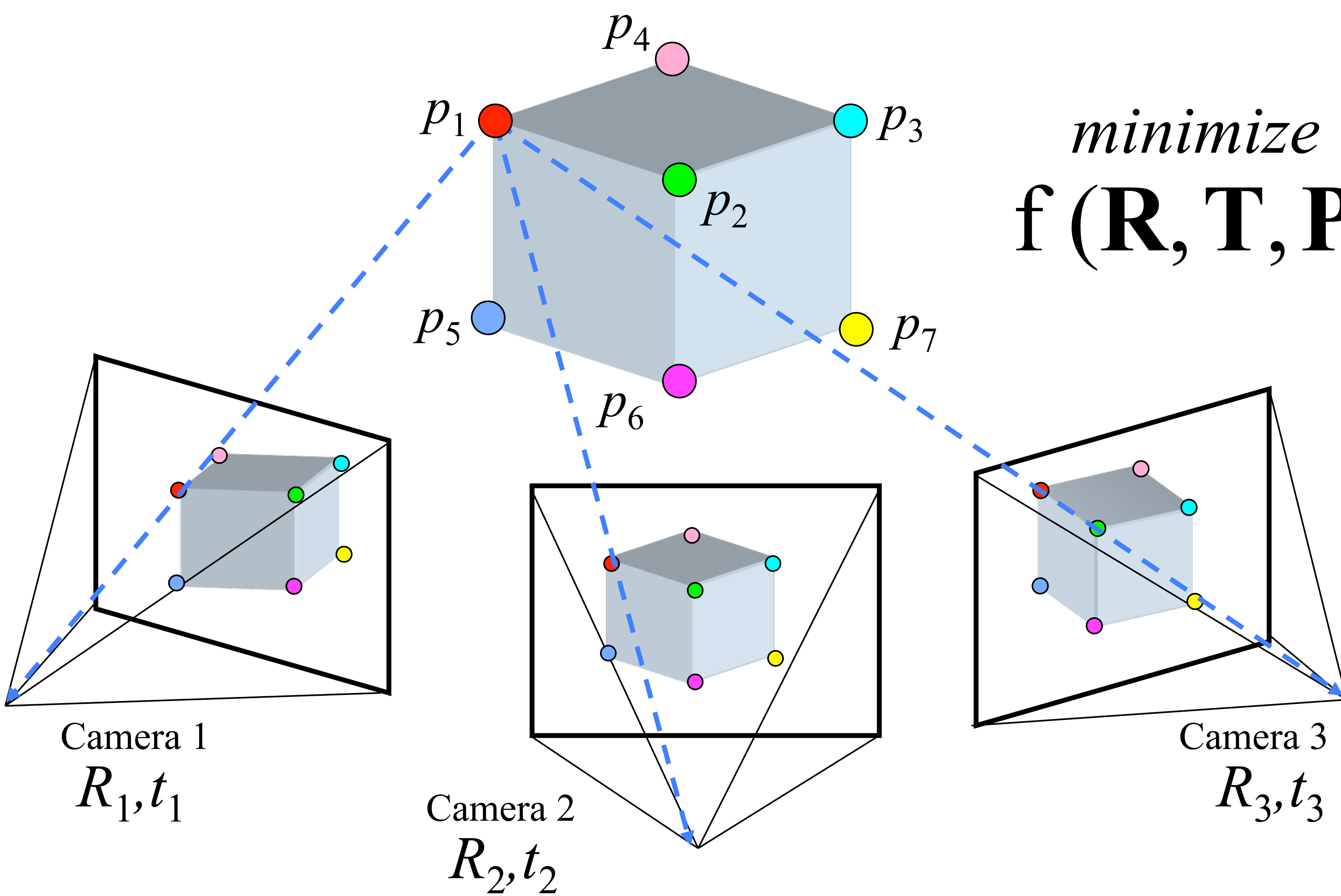


3d reconstruction

Solving for scene structure and camera poses



Solving for scene structure and camera poses



minimize

$$\mathrm{f}\,(\mathbf{R}, \mathbf{T}, \mathbf{P})$$

Existing algorithms

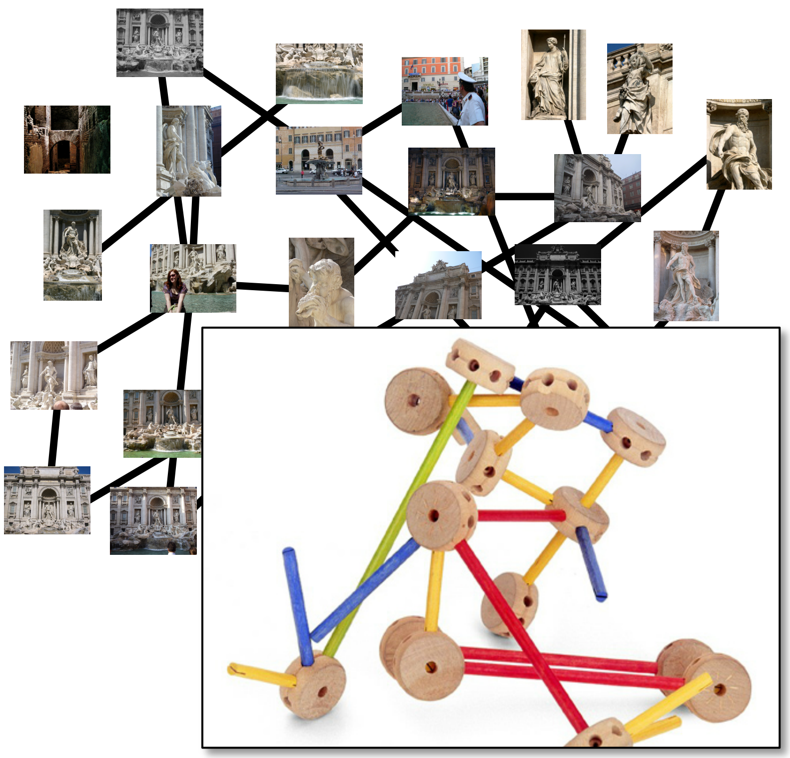
- Existing approaches use an incremental approach
  - Solve reconstruction problem for 2 images, then add a third and solve, then add a 4^{th} and solve, etc…

✓ Works very well for many scenes

✗ Poor scalability (running time quartic in # of images)
✗ Poor results if a bad seed image set is chosen
✗ Drift and bad local minima for some scenes

Our approach

- View as inference on a Markov Random Field (MRF) model, solving for all camera poses at once

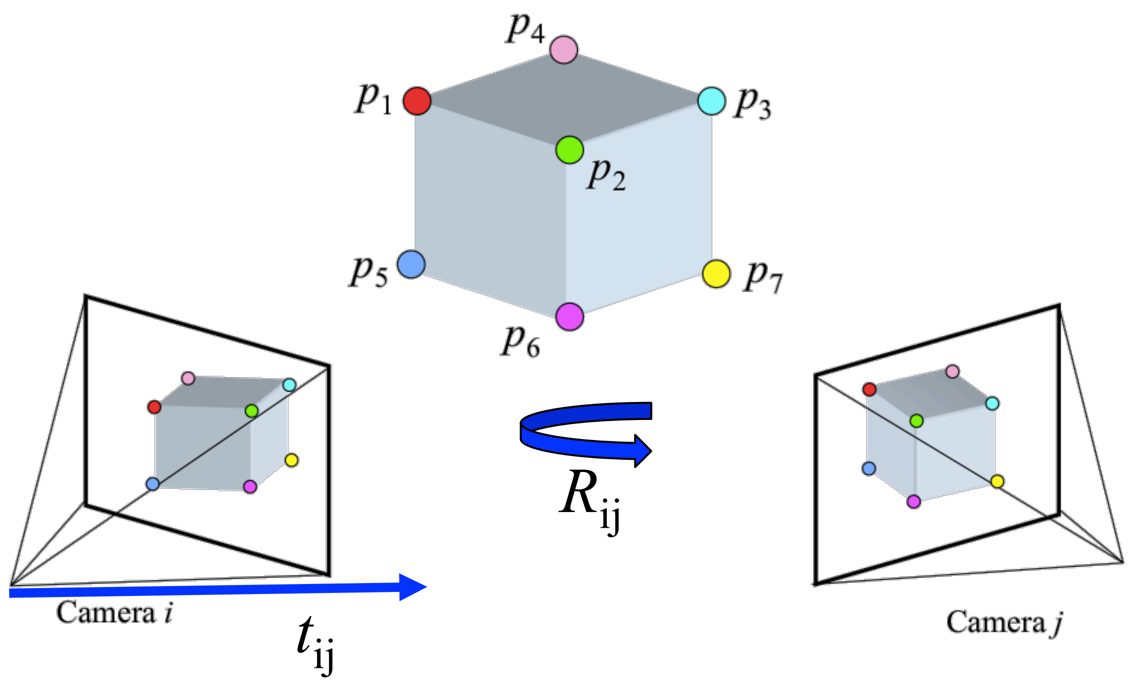


- **Vertices** are cameras (and/or points)
- Both **pairwise** and **unary** constraints
- **Inference problem:** label each image with a camera pose, such that constraints are satisfied

Constraints on camera pairs

- Compute relative pose between pairs of cameras, using 2-frame SfM

[Nister04]



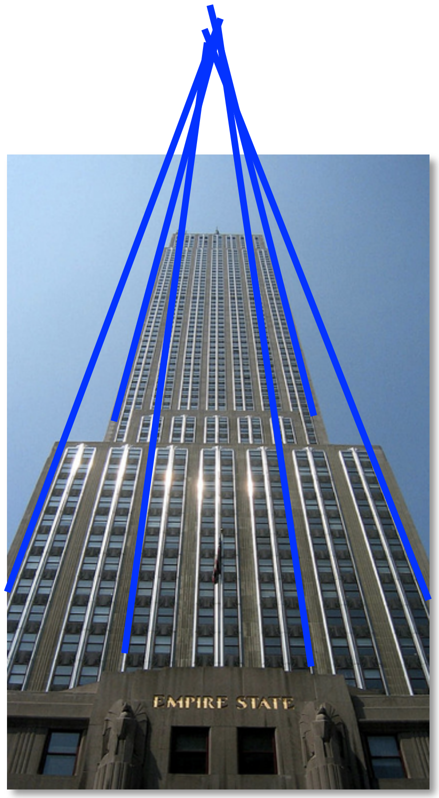
- Want to find absolute camera poses ($\boldsymbol{R}_i$, $\boldsymbol{t}_i$) and ($\boldsymbol{R}_j$, $\boldsymbol{t}_j$) such that:

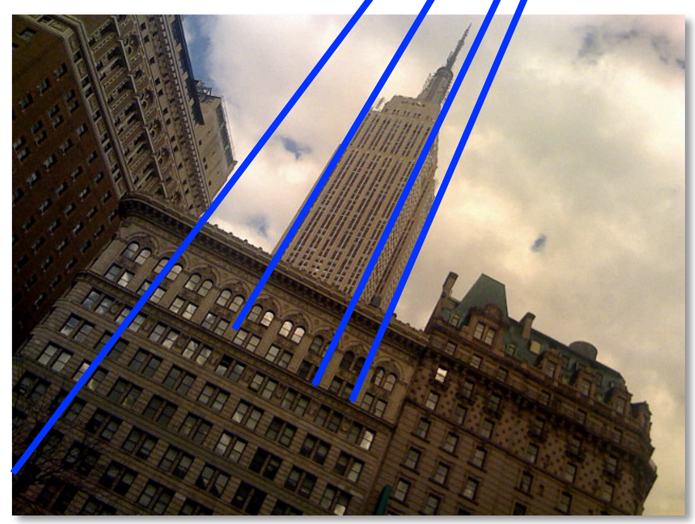
$$\mathbf{R}_{ij} = \mathbf{R}_i^\top \mathbf{R}_j$$

$$\lambda_{ij}\mathbf{t}_{ij} = \mathbf{R}_i^\top(\mathbf{t}_j - \mathbf{t}_i)$$

Prior pose information

- We may have <u>noisy</u> absolute pose estimates for some cameras
  - 2-d positions from geotags (GPS coordinates)
  - Orientations (tilt & twist angles) from vanishing point detection





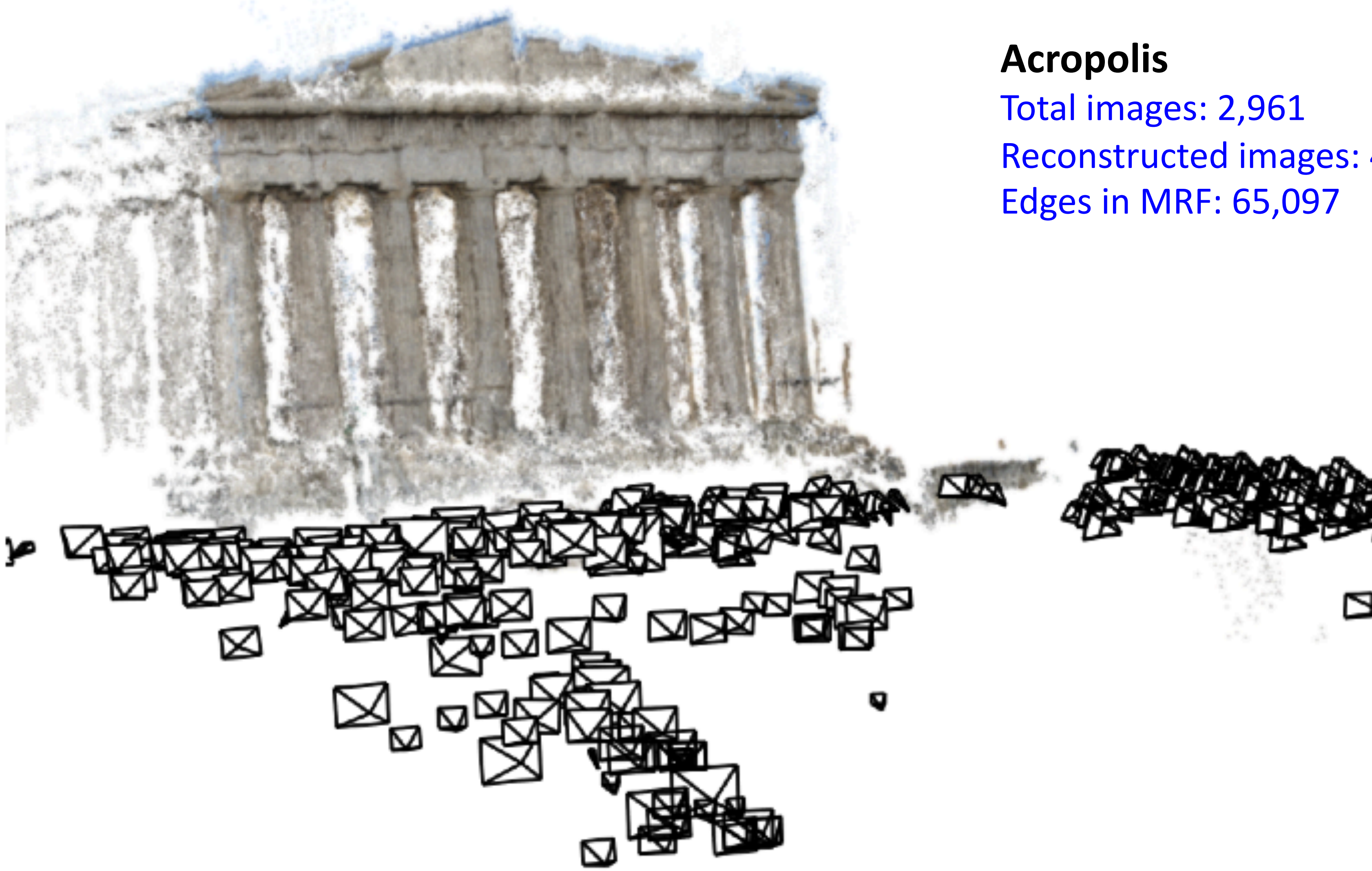
Overall optimization problem

- Given pairwise and absolute pose constraints, we want to solve for all absolute camera poses simultaneously
  - for n cameras, estimate $\mathcal{R}$=($\mathbf{R}_1$, $\mathbf{R}_2$, …, $\mathbf{R}_n$) and $\mathcal{T}$=($\mathbf{t}_1$, $\mathbf{t}_2$, …, $\mathbf{t}_n$) so as to minimize total error over the entire graph,

$$D^{\mathbf{R}}(\mathcal{R}) = \sum_{e_{ij} \in E_C} d^{\mathbf{R}}\left(\mathbf{R}_{ij}, \mathbf{R}_i^\top \mathbf{R}_j\right) + \alpha_1 \sum_{I_i \in \mathcal{I}} d_i^{\mathbf{O}}(\mathbf{R}_i)$$

$$D^{\mathbf{T}}(\mathcal{T}, \mathcal{R}) = \sum_{e_{ij} \in E_C} d^{\mathbf{T}}(\mathbf{t}_j - \mathbf{t}_i, \mathbf{R}_i \mathbf{t}_{ij}) + \alpha_2 \sum_{I_i \in \mathcal{I}} d_i^{\mathbf{G}}(\mathbf{t}_i)$$

- This is a Markov Random Field, which we solve approximately using Belief Propagation



Acropolis

Total images: 2,961

Reconstructed images: 454

Edges in MRF: 65,097

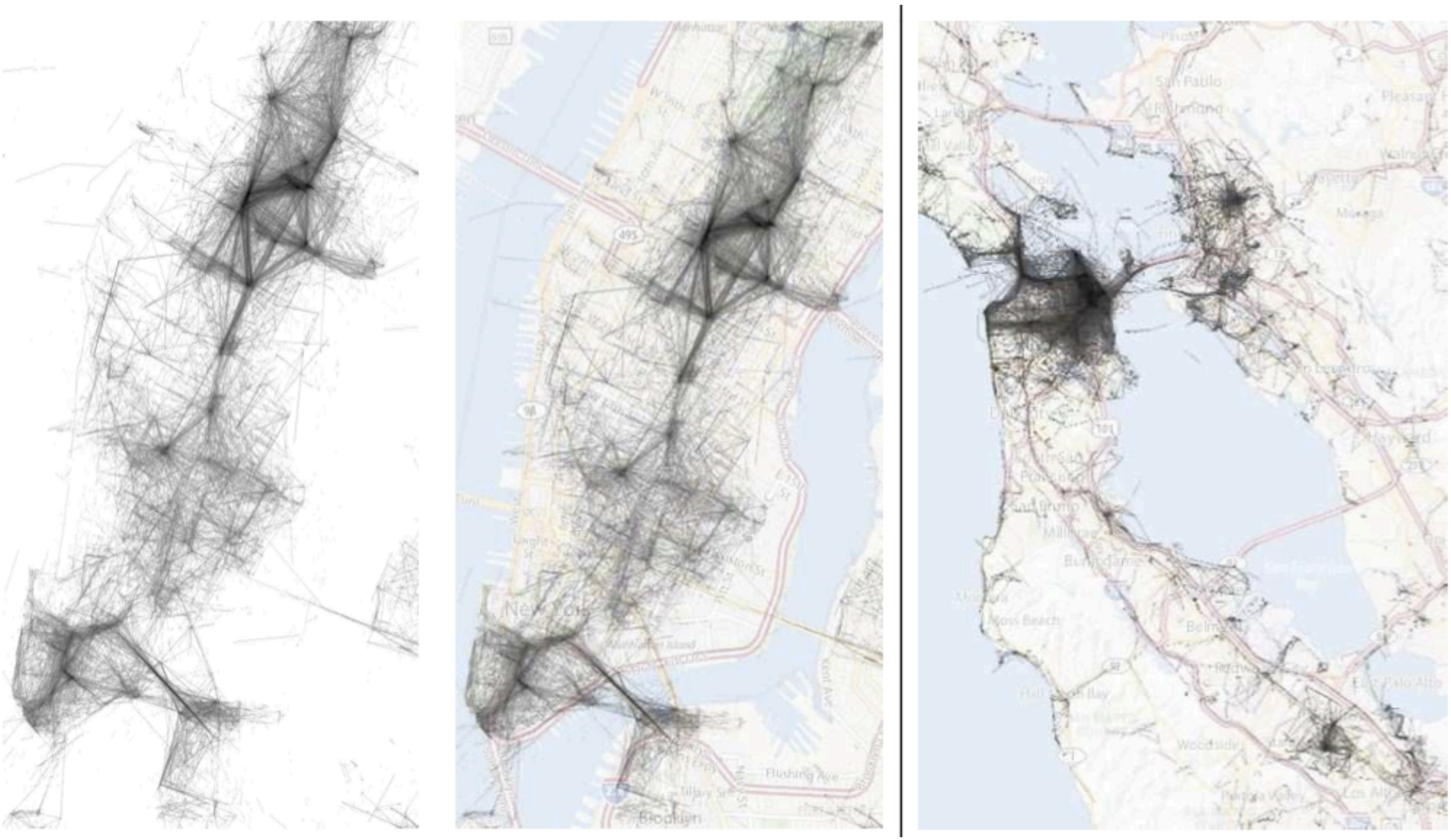
Videos

Inferring social ties from geographic coincidences

D. Crandall, L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, J. Kleinberg, Inferring Social Ties from Geographic Coincidences, *PNAS* 2010.

Geo-temporal trails of human travel

- From streams of geo-tagged, timestamped photos

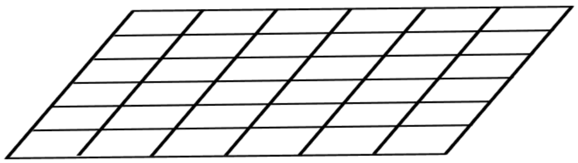


Mobility and the social network

- On *k* different occasions, you observe two people at about the same place. What are the chances that they know each other?
  - Or: what are the chances that these geo-temporal coincidences are just a coincidence?
- Related to much existing work on human mobility
  - [Brockmann06], [Gonzalez08], [Adrienko10]…
  - [Eagle09] predict social connections among small groups of people using dense geo-temporal data (cell phones)
- Also related
  - Inferring social network based on online activities: [Provost09], [Schifanella10] …
  - Studies of coincidences: [Diaconis89], [Griffiths01], [Guilt49] …

Geo-temporal coincidences on Flickr

- We discretize the world into $s° \times s°$ bins

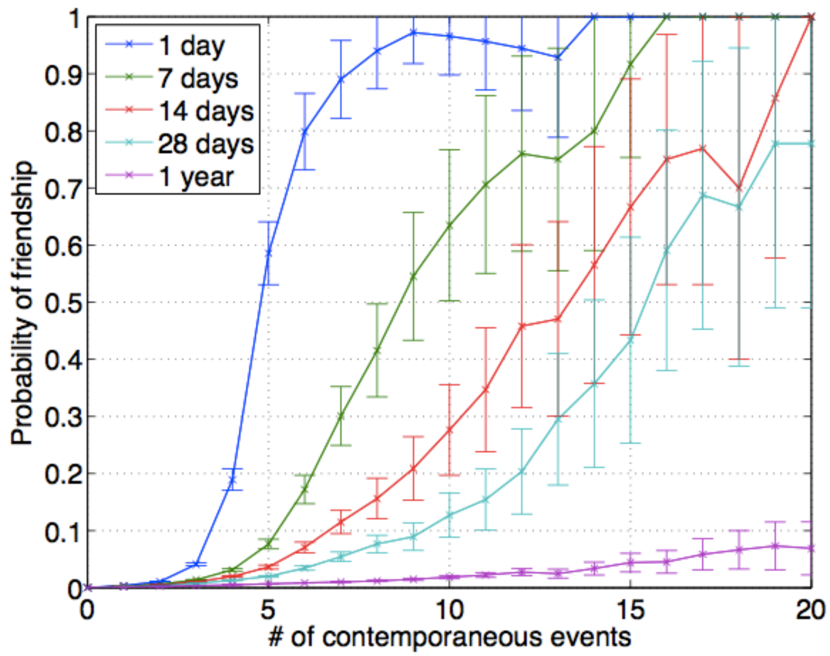


- For each pair of users, count the number of bins in which they have taken a photo within t days of one another

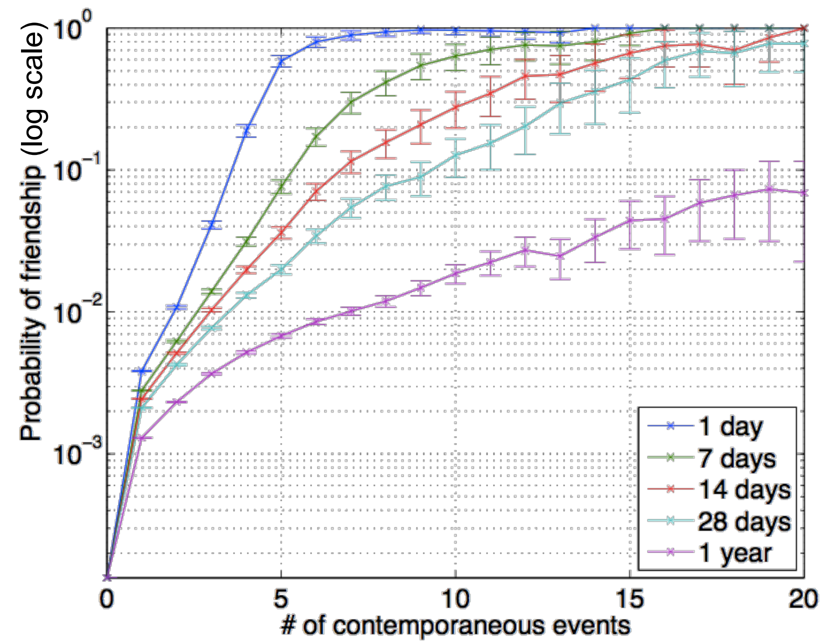
- Compute probability that the two people are "friends", given that we've observed k co-occurrences
  - where being contacts on Flickr is a proxy for friendship

Sample results

- After 5 co-occurrences with *s*=1° (roughly 100 km) and *t*=1 day, the probability that two people are contacts is ~60%
  - Even after 2 co-occurrences, probability is ~1% (~100 times baseline)

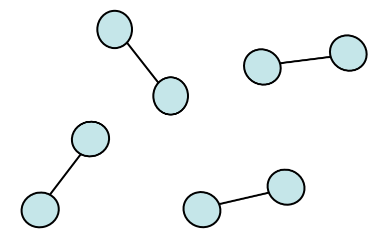


(D) $s = 1.0°$

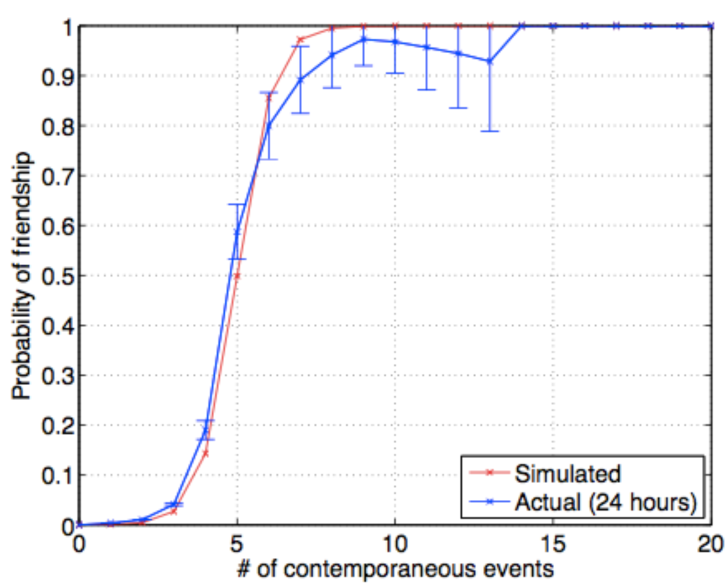


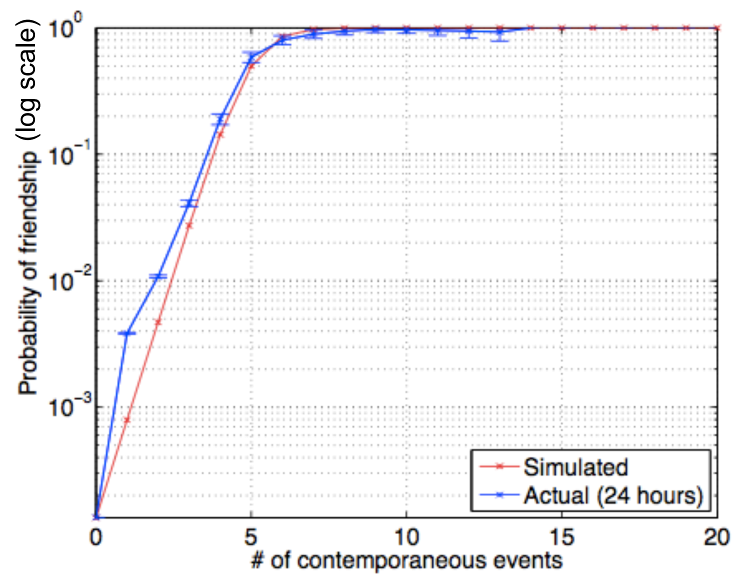
(D) $s = 1.0°$

A simple model



- The world consists of N spatial bins
  - There are M people, each with exactly one friend
  - Each day, each pair of friends chooses a random bin to visit either jointly (probability β) or independently (probability 1-β)
- Using Bayes Law, compute conditional probability of friendship as a function of *k*



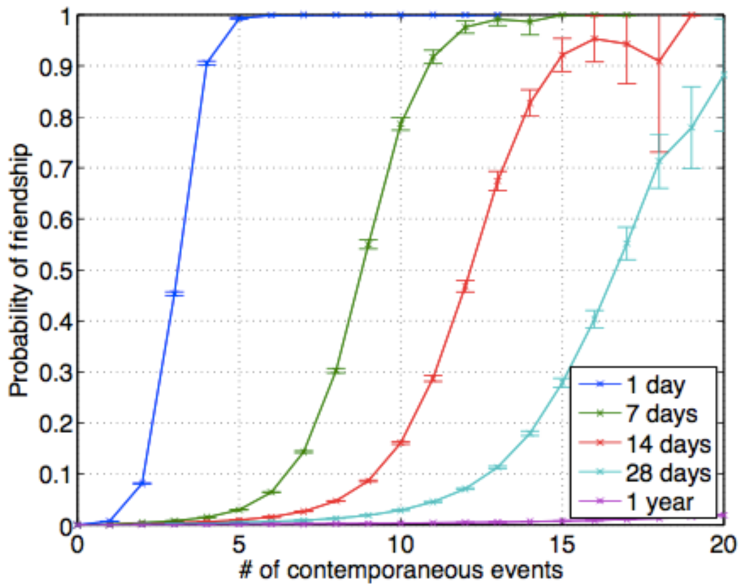


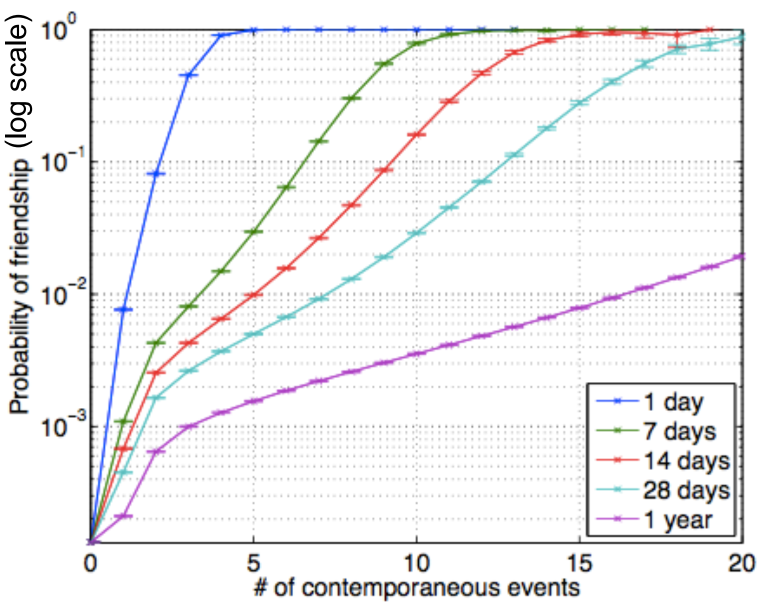
M=7500, N=100, β=0.05

A refined model

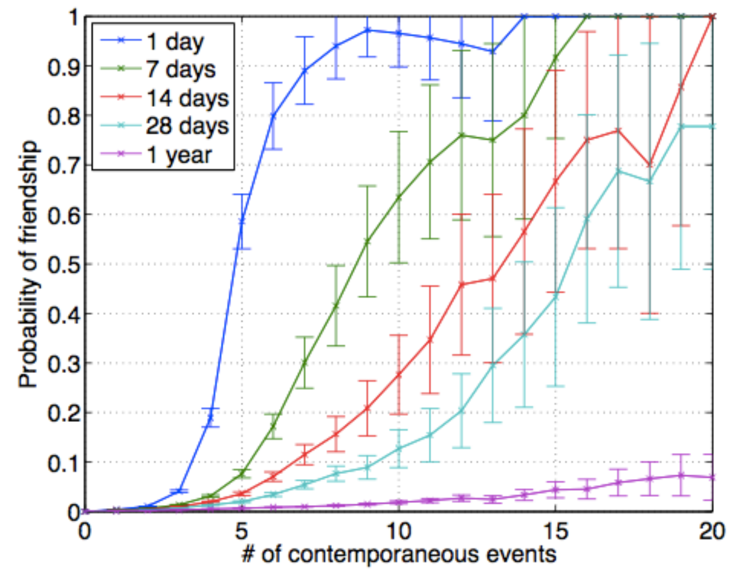
- Add notion of homophily to the model
  - Since friends are similar people, they'll choose similar destinations even when choosing independently
- Again, N spatial bins, M people, each with one friend
  - Each person has a "home cell"; friends have the same home cell
  - Sample bins to visit from a power law distribution centered at home cell [Brockmann06]
  - Choose bins jointly with probability β or independently (1-β)

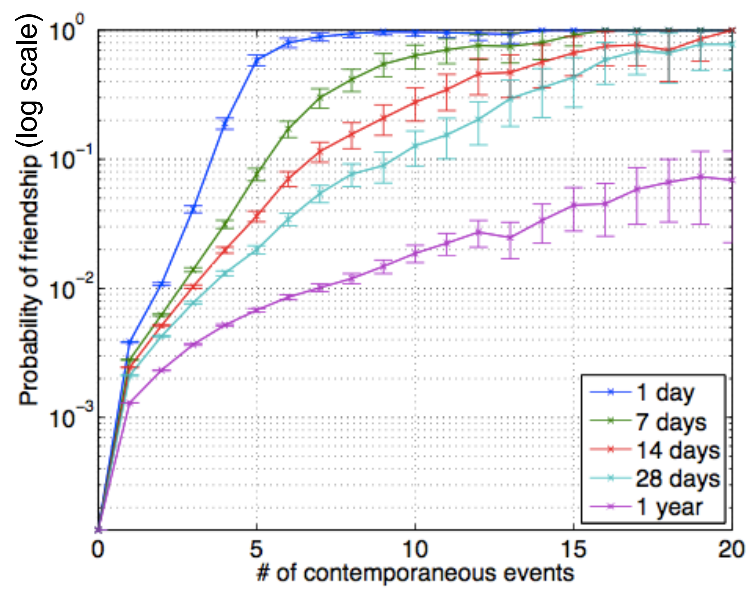
Simulated





Actual





Ongoing and future work

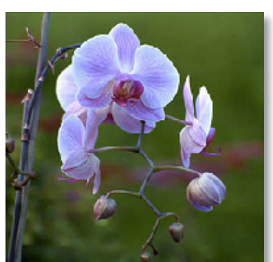
Broad research directions

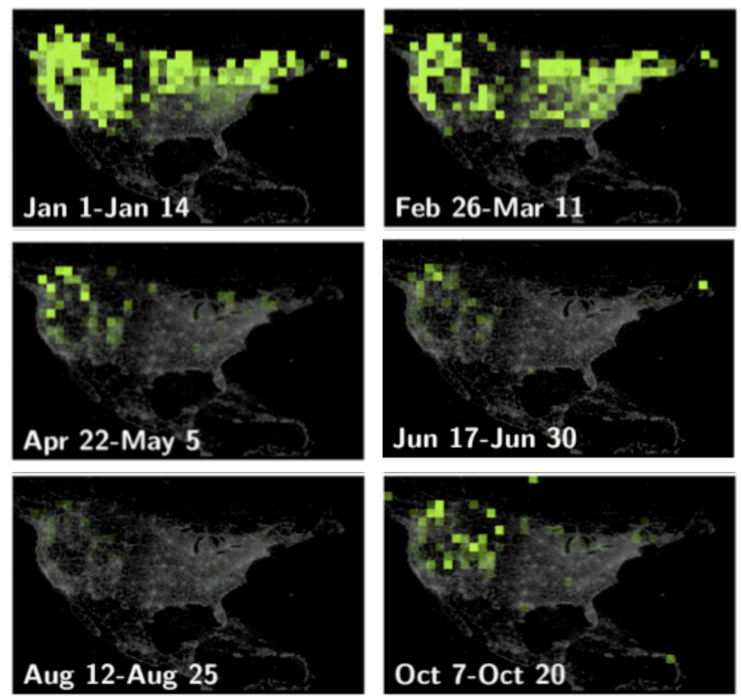
How do we use visual and non-visual analysis to:

1. organize vast collections of digital photos?
2. mine photos to study the world and its people?

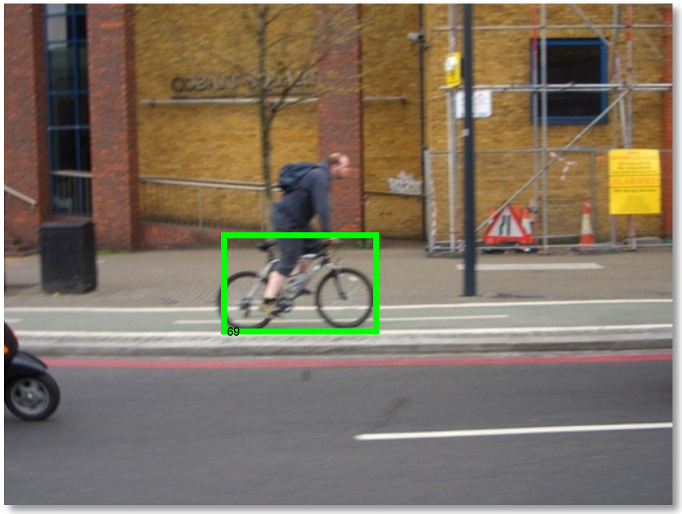
Mining Flickr for ecological phenomena

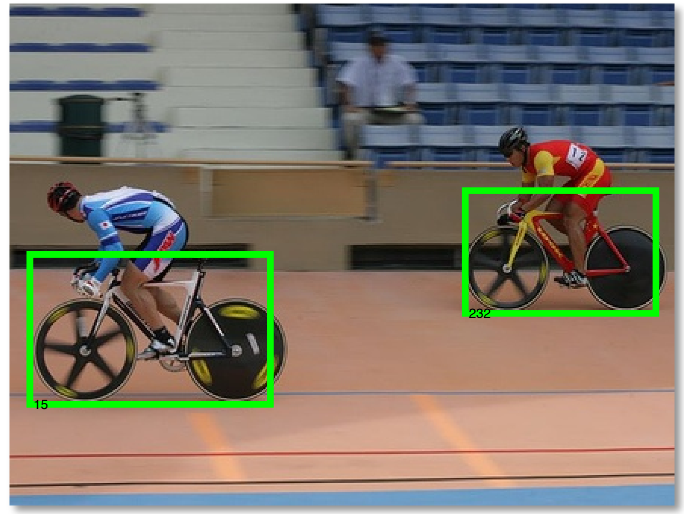
- Biologists & ecologists would like continental-scale observational data about flowers & wildlife
  - When & where are plants flowering? How are birds migrating?

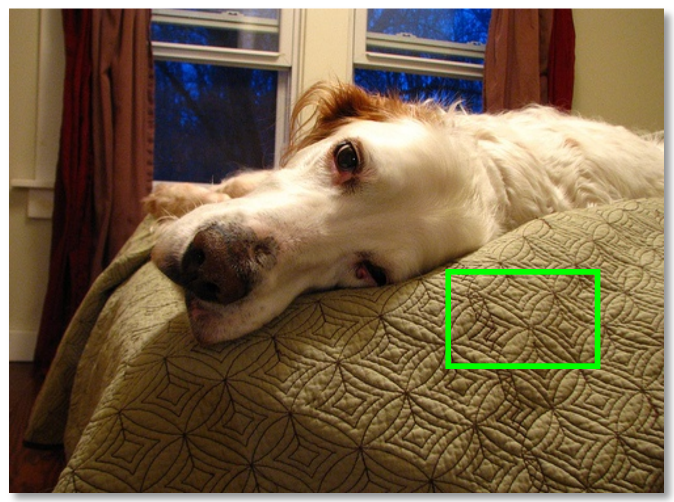




Object recognition

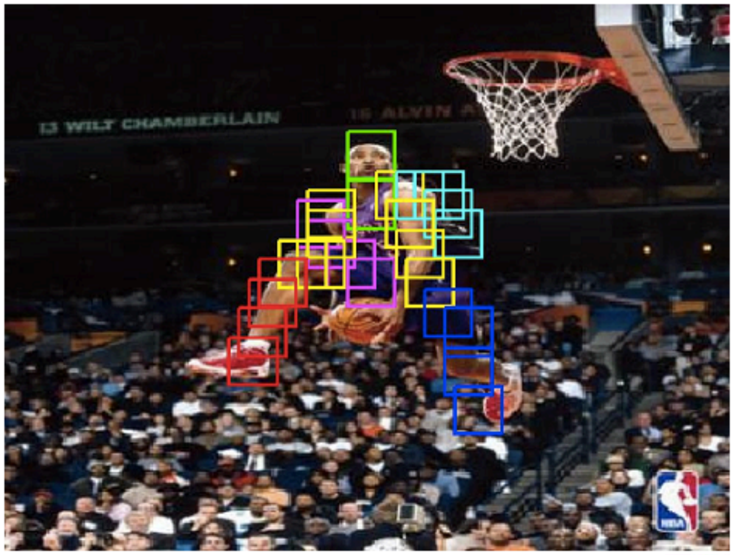


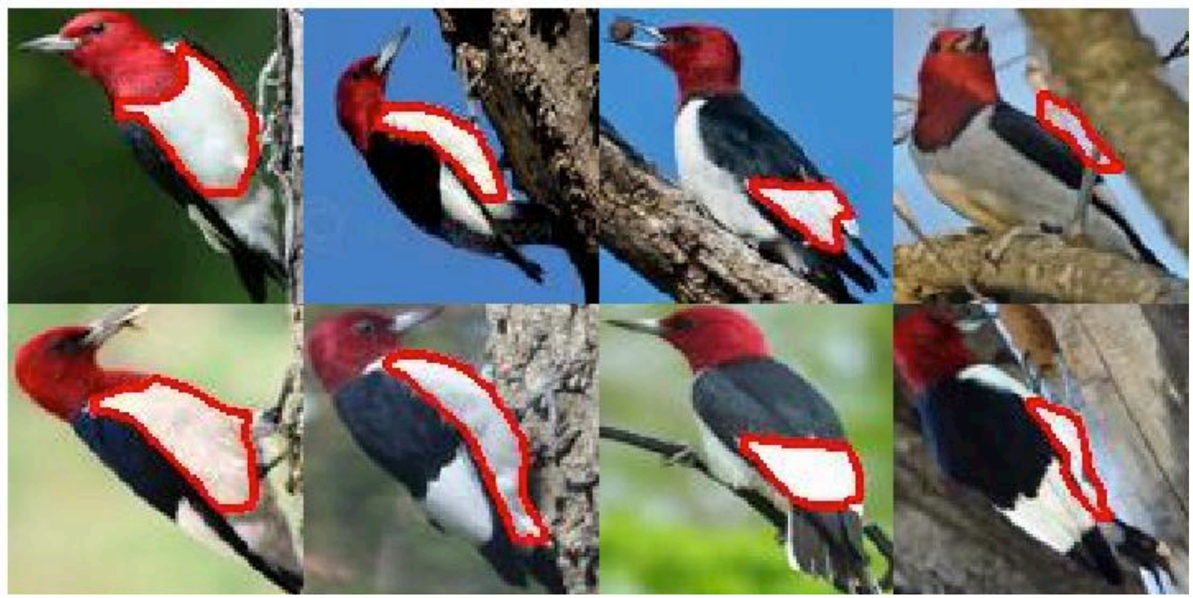




From D. Crandall, *Part-based Statistical Models for Visual Object Class Recognition*, 2008.

Person and animal recognition





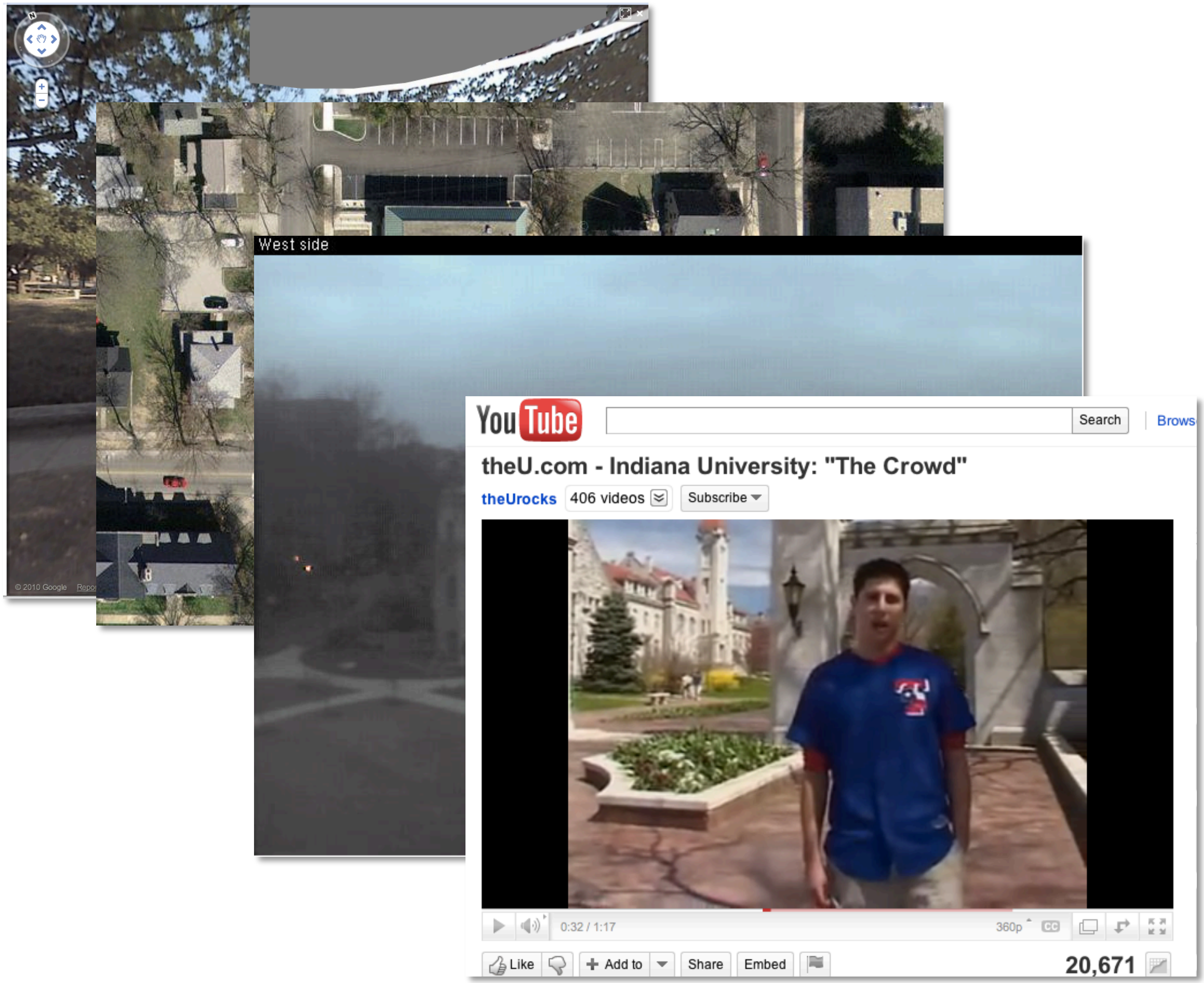
Automatic geo-location

(aka: Where were these photos taken?)




- Using visual, textual, temporal features, correct classification rate about 50% among top 500 landmarks

Y. Li, D. Crandall, D. Huttenlocher, "Landmark classification in large-scale image collections," *ICCV,* 2009.


West side
Search
theU.com - Indiana University: "The Crowd"
theUrocks
406 videos
Subscribe
0:32 / 1:17
360p
Like
Add to
Share
Embed
20,671

Thanks!

- *Funding:* IU Data to Insight Center, Lilly Foundation, Intel, NSF
- *Cyberinfrastructure:* Indiana University supercomputing facilities, Cornell Center for Advanced Computing
- *Collaborators:* Lars Backstrom, Dan Cosley, Dan Huttenlocher, Jon Kleinberg, Noah Snavely, Sid Suri
- *Students:* Kun Duan, Mohammed Korayem, Andrew Owens, Haipeng Zhang

For more information about these projects, please visit:

http://www.cs.indiana.edu/~djcran/