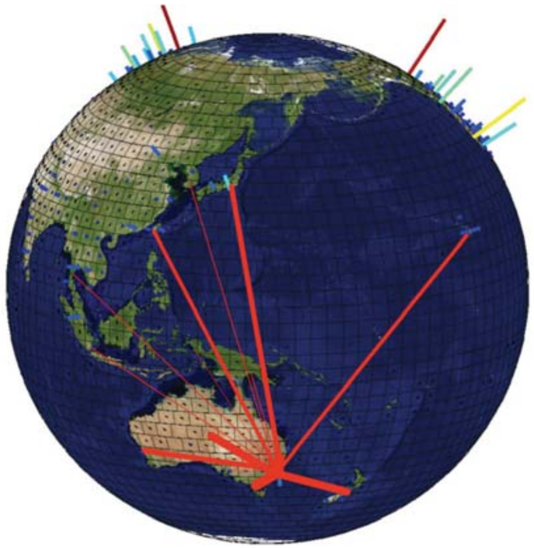
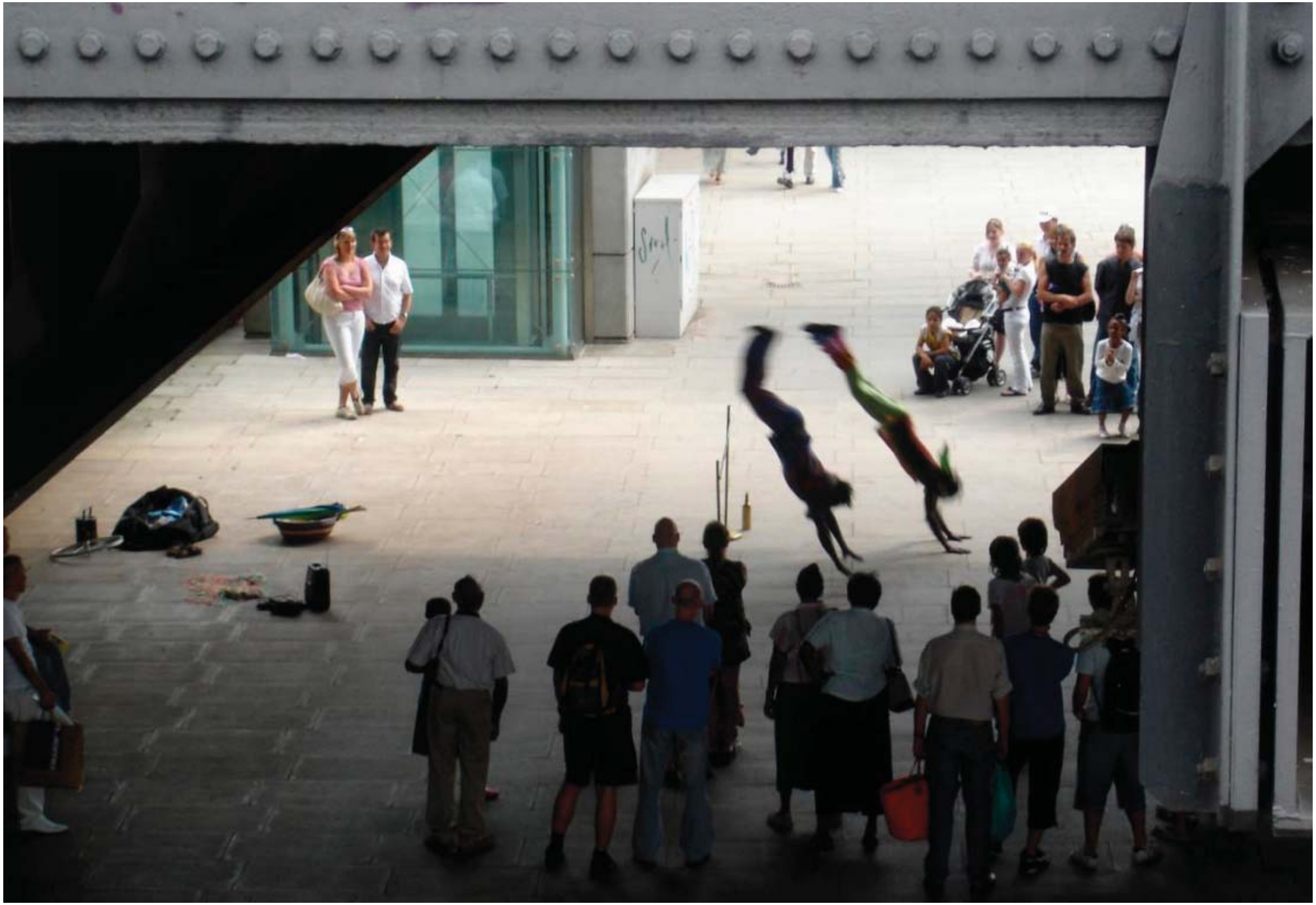


Image Sequence Geolocation with Human Travel Priors

Evangelos Kalogerakis*, Olga Vesselova*, James Hays+, Alexei A. Efros+, Aaron Hertzmann*

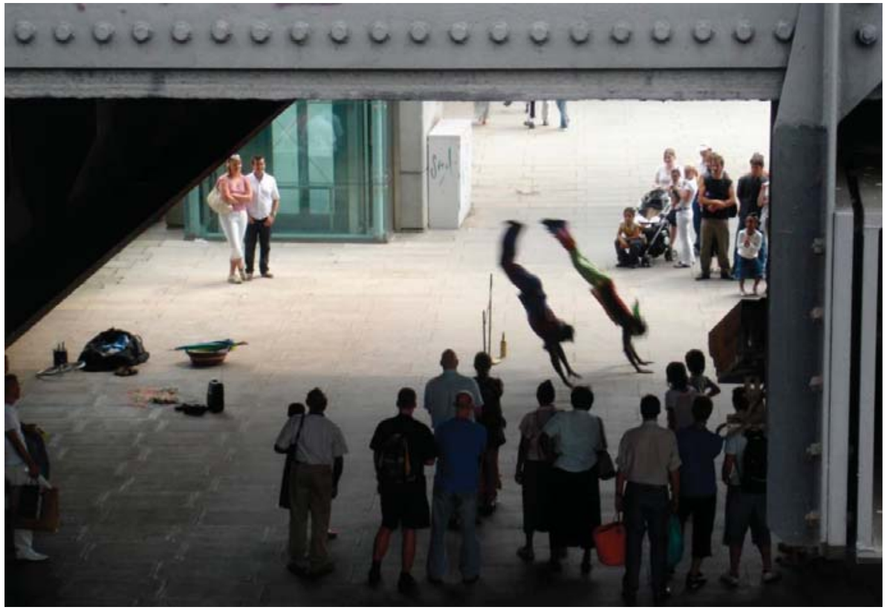
*University of Toronto, +Carnegie Mellon University

Where is this?



Where is this?

Where are these?



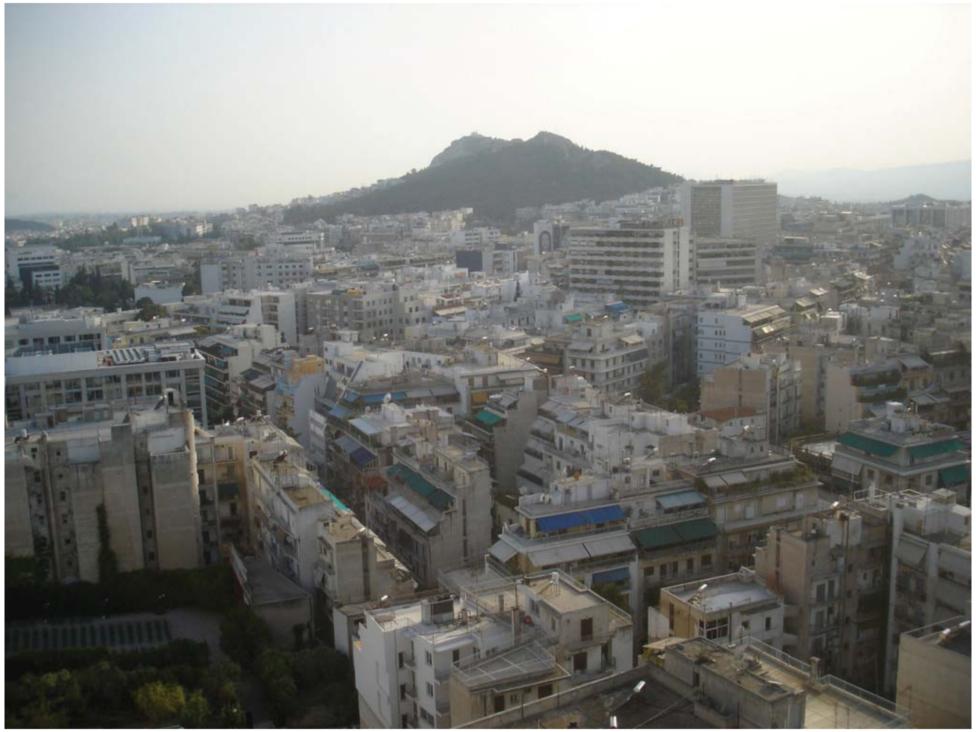
June 18, 2006, 15:45

June 18, 2006, 16:31

Where are these?

June 18, 2006, 15:45

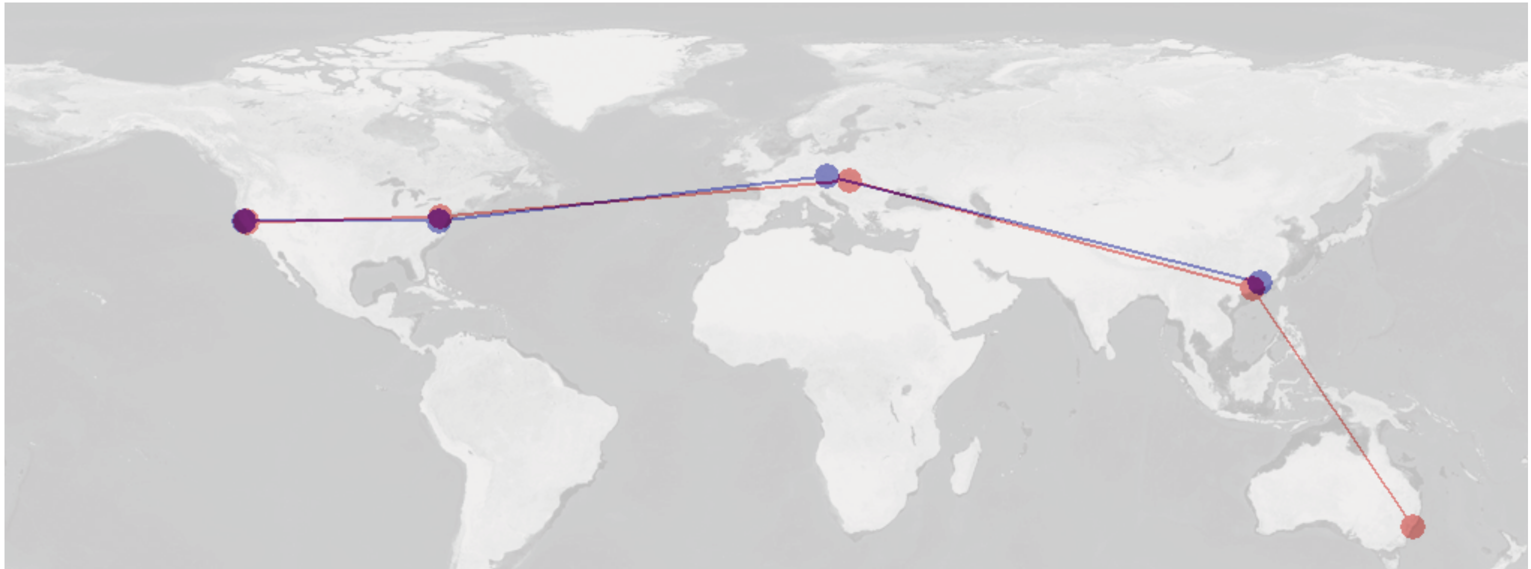
June 18, 2006, 16:31



June 19, 2006, 17:24

Problem statement

Want: geo-tags



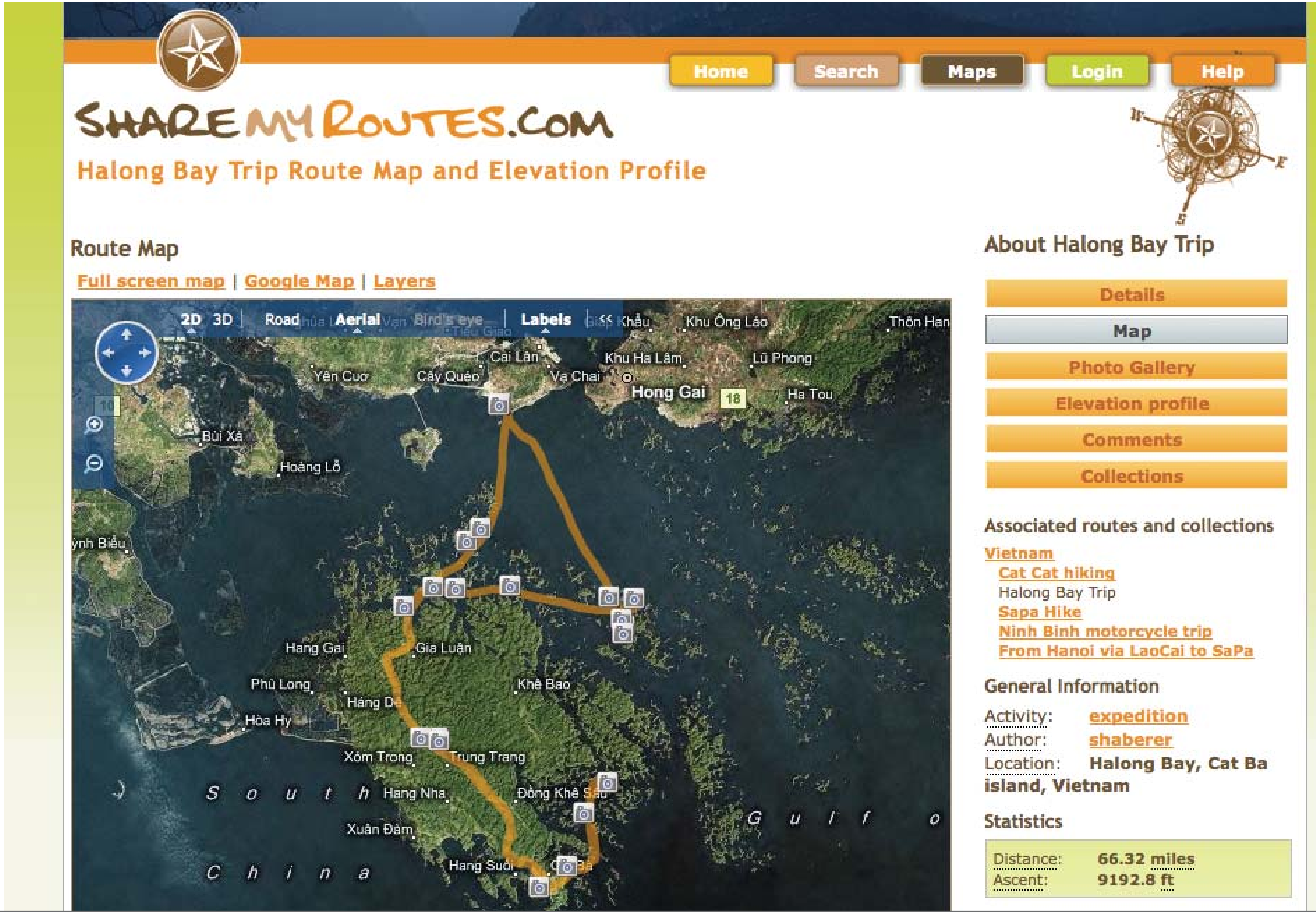
Key questions

How do we relate images to locations?

How do we model human travel?

Applications

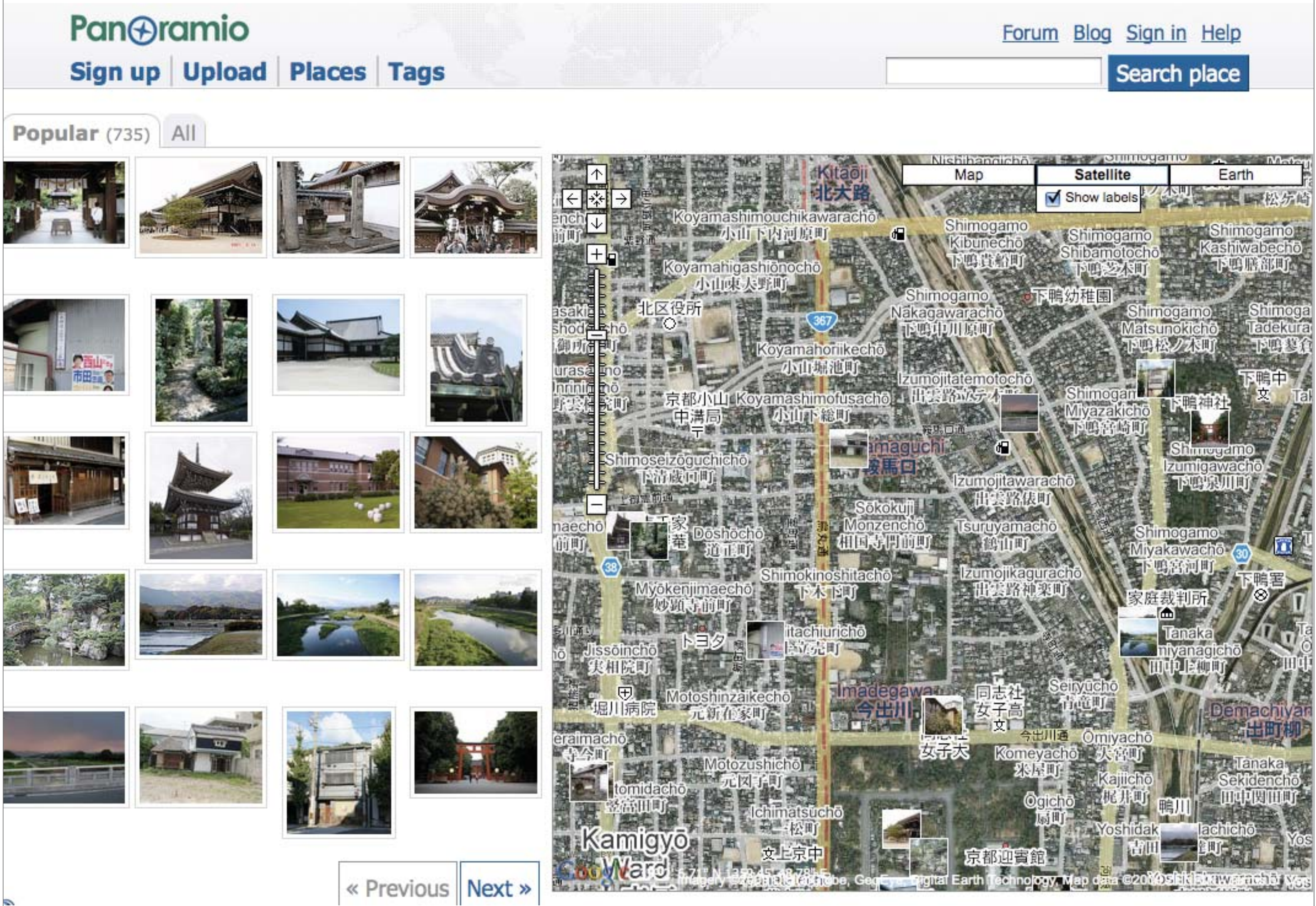
Geo-tagging your photos



Will all cameras have GPS?

This might not happen (cost; start-up time/ power consumption, urban/wilderness locations)

There are billions of existing images without good geotags


Panoramio
Forum Blog Sign in Help
Sign up | Upload | Places | Tags
Search place
Popular (735) All
« Previous Next »
Map Satellite Earth
Show labels
Kitaōji
北大路
Koyamashimouchikawarachō
小山下内河原町
Koyamahigashiōnochō
小山東大野町
Koyamahorikechō
小山堀池町
Koyamashimofusachō
小山下総町
Shimoseizōguchichō
下清蔵口町
Dōshōchō
道正町
Shimokinoshitachō
下木下町
Myōkenjimaechō
妙顕寺前町
Jissōinchō
実相院町
Motoshinzaikechō
元新在家町
Motozushichō
元図子町
Ichimatsuchō
一松町
Imadegawa
今出川
Shimogamo Kibunechō
下鴨貴船町
Shimogamo Shibamotochō
下鴨芝本町
Shimogamo Kashiwabechō
下鴨膳部町
Shimogamo Nakagawarachō
下鴨中川原町
Shimogamo Matsunokichō
下鴨松ノ木町
Izumojitatemotochō
Shimogamo Miyazakichō
下鴨宮崎町
Shimogamo Izumigawachō
下鴨泉川町
Izumojitawarachō
出雲路俵町
Sōkokuji Monzenchō
相国寺門前町
Tsuruyamachō
鶴山町
Shimogamo Miyakawachō
下鴨宮河町
Izumojikagurachō
出雲路神楽町
家庭裁判所
Tanaka
Seiryūchō
青竜町
Omiyachō
大宮町
Komeyachō
米屋町
Kajiichō
梶井町
Ōgichō
扇町
Tanaka Sekidenchō
田中関田町
Demachiyanagi
出町柳
下鴨神社
下鴨幼稚園
北区役所
京都小山中溝局
堀川病院
同志社女子高
同志社女子大
京都迎賓館
Kamigyō Ward
367
38
30

nature

Vol 439|26 January 2006|doi:10.1038/nature04292

LETTERS

The scaling laws of human travel

D. Brockmann[1,2], L. Hufnagel[3] & T. Geisel[1,2,4]

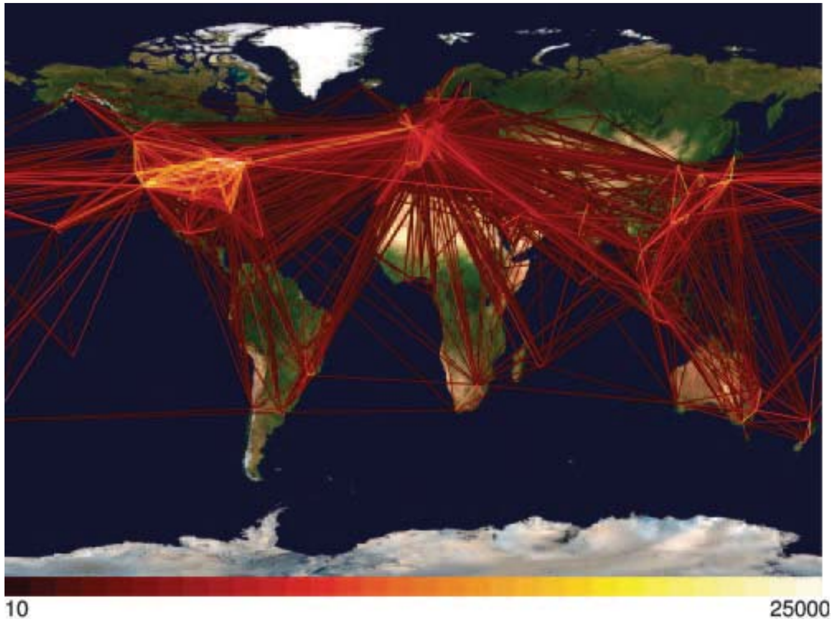
The dynamic spatial redistribution of individuals is a key driving force of various spatiotemporal phenomena on geographical scales. It can synchronize populations of interacting species, stabilize them, and diversify gene pools[1–3]. Human travel, for example, is responsible for the geographical spread of human infectious disease[4–9]. In the light of increasing international trade, intensified human mobility and the imminent threat of an influenza A epidemic[10], the knowledge of dynamical and statistical properties of human travel is of fundamental importance. Despite its crucial role, a quantitative assessment of these properties on geographical scales remains elusive, and the assumption that humans disperse diffusively still prevails in models. Here we report on a solid and quantitative assessment of human travelling statistics by analysing the circulation of bank notes in the United States. Using a comprehensive data set of over a million individual displacements, we find that dispersal is anomalous in two ways. First, the distribution of travelling distances decays as a power law, indicating that trajectories of bank notes are reminiscent of scale-free random walks known as Lévy flights. Second, the probability of remaining in a small, spatially confined region for a time T is dominated by algebraically long tails that attenuate the super-diffusive spread. We show that human travelling behaviour can be described mathematically on many spatiotemporal scales by a two-parameter continuous-time random walk model to a surprising accuracy, and conclude that human travel on geographical

quantitative assessment of human movements, however, is difficult, and a statistically reliable estimate of human dispersal comprising all spatial scales does not exist. The central aim of this work is to use data collected at online bill-tracking websites (which monitor the worldwide dispersal of large numbers of individual bank notes) to infer the statistical properties of human dispersal with very high spatiotemporal precision. Our analysis of human movement is based on the trajectories of 464,670 dollar bills obtained from the bill-tracking system www.wheresgeorge.com. We analysed the dispersal of bank notes in the United States, excluding Alaska and Hawaii. The core data consists of 1,033,095 reports to the bill-tracking website. From these reports we calculated the geographical displacements $r = |\mathbf{x}_2 - \mathbf{x}_1|$ between a first ($\mathbf{x}_1$) and secondary ($\mathbf{x}_2$) report location of a bank note and the elapsed time T between successive reports.

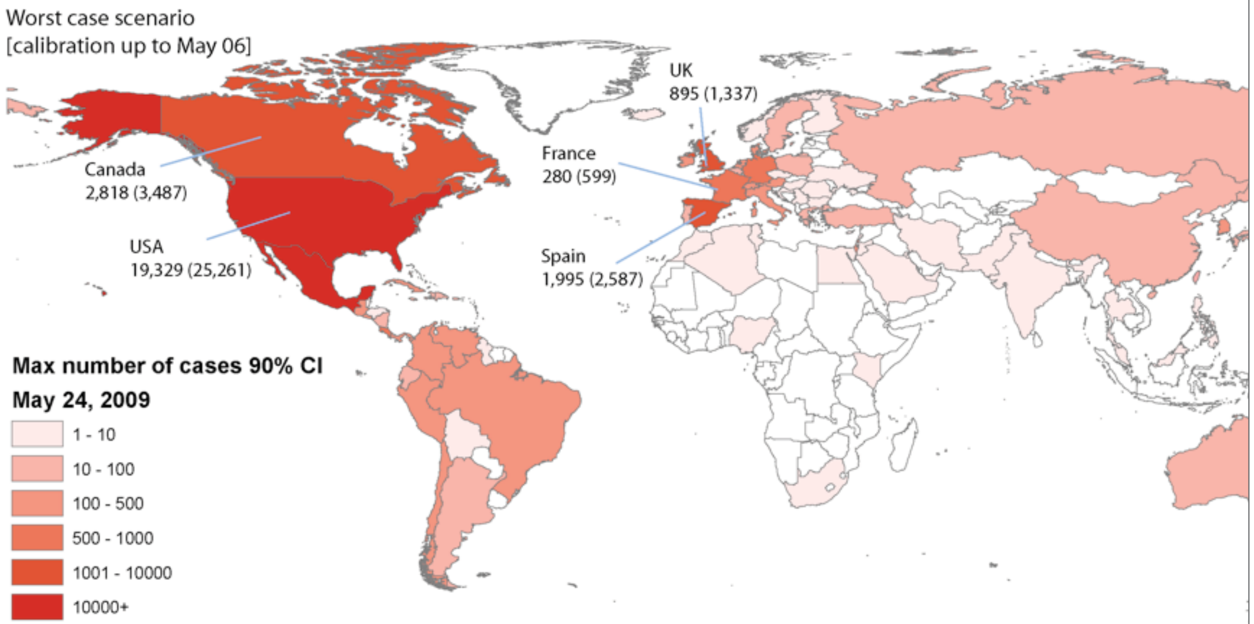
In order to illustrate qualitative features of bank note trajectories, Fig. 1b depicts short-time trajectories ($T < 14$ days) originating from three major US cities: Seattle, New York and Jacksonville. After their initial entry into the tracking system, most bank notes are next reported in the vicinity of the initial entry location, that is $|\mathbf{x}_2 - \mathbf{x}_1| \leq 10$ km (Seattle, 52.7%; New York, 57.7%; Jacksonville, 71.4%). However, a small but considerable fraction is reported beyond a distance of 800 km (Seattle, 7.8%; New York, 7.4%; Jacksonville, 2.9%).

From a total of 20,540 short-time trajectories originating across the United States, we measured the probability $P(r)$ of traversing a

Epidemic forecasting



World aviation network

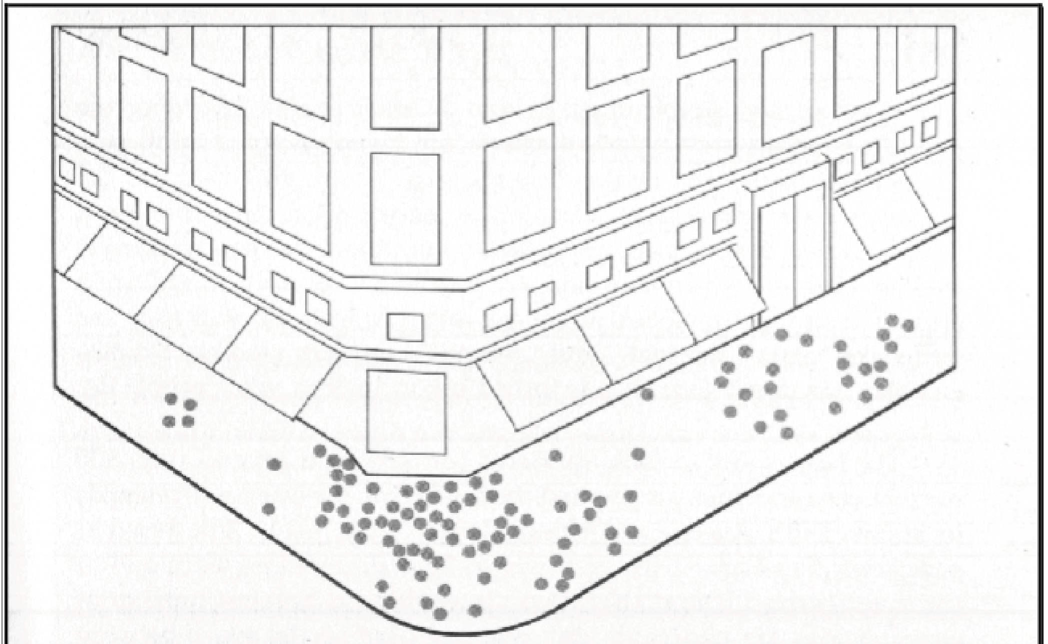


Swine flu projection for May 24

(Indiana University, http://www.gleamviz.org)

(Hufnagel 2004, Colizza 2007)

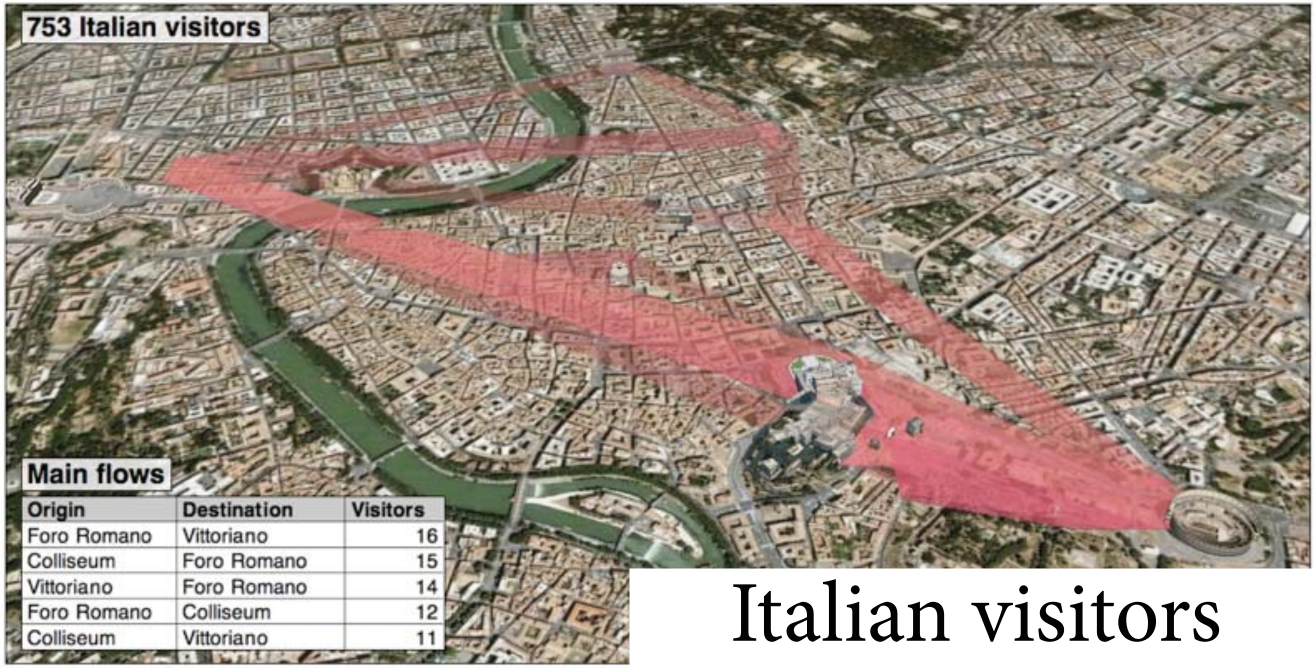
Urban planning



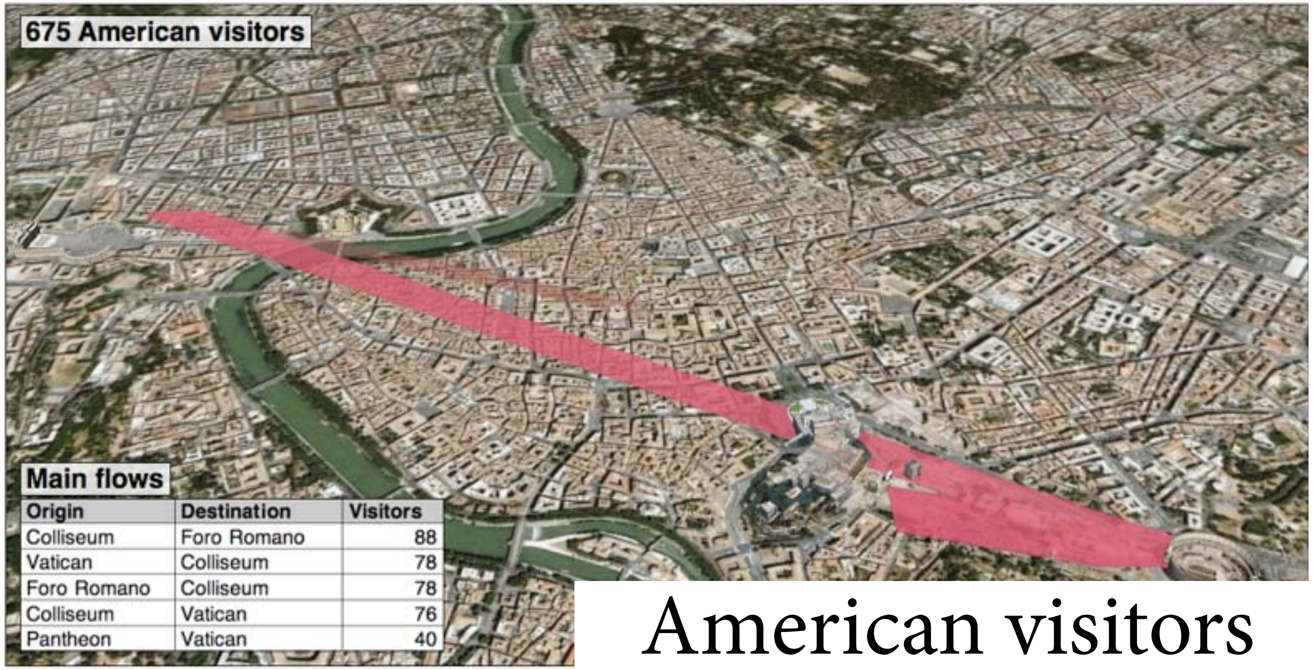
Location of street conversations lasting two minutes or more at Saks Fifth Avenue and Fiftieth Street. Cumulative for five days in June. Note main concentration at corner, secondary one outside entrance.

(Whyte, 1971)

2009



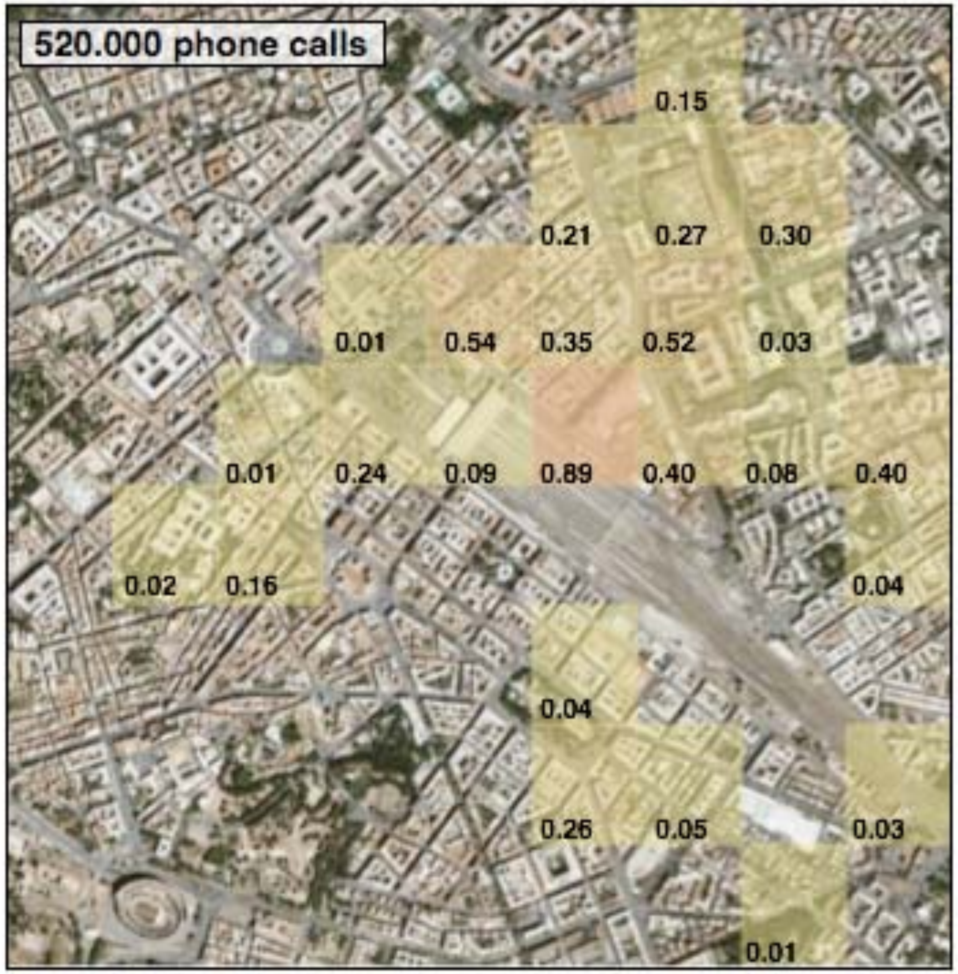
Italian visitors



American visitors

(Girardin et al., *Pervasive* 2008)

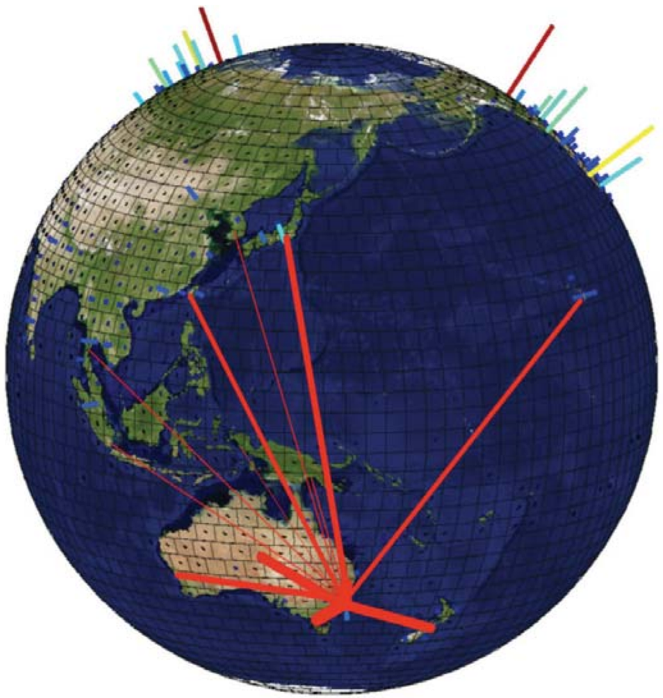
Photographs



Phone calls

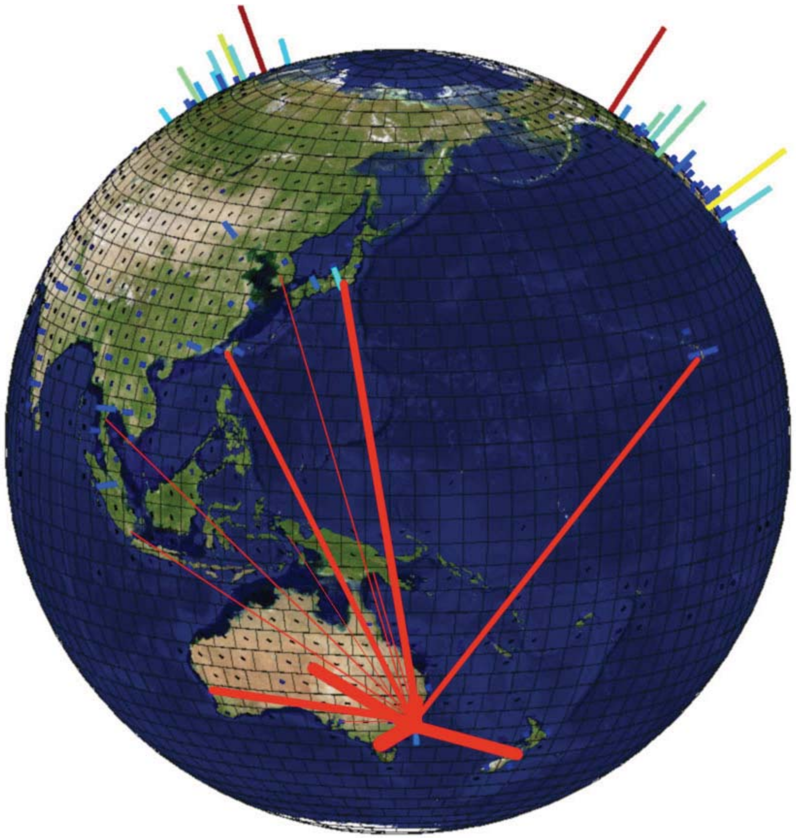
(Girardin et al., *Pervasive* 2008)

Human travel distributions

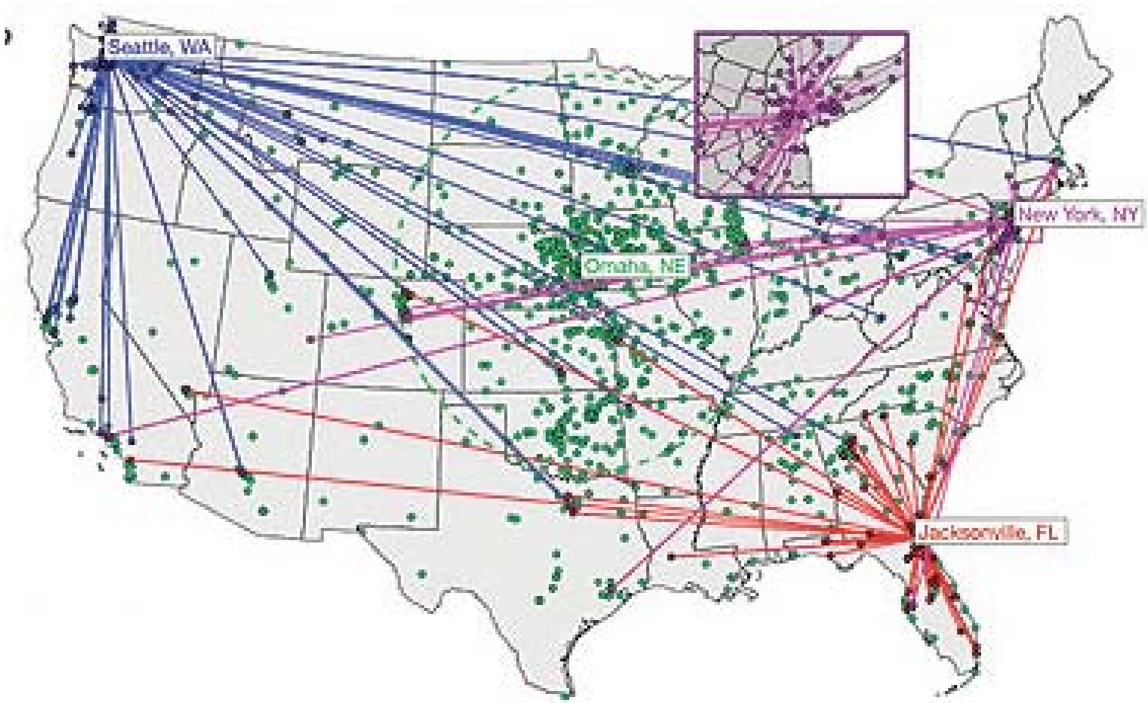


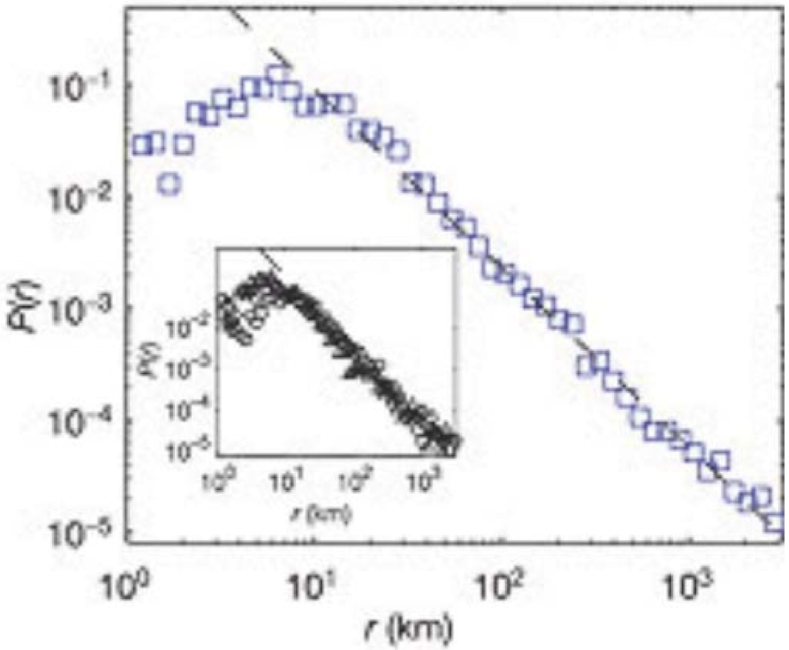
How likely are you to travel from one place to another in a fixed amount of time?

Need: $P(L_{t+1} = i | L_t = j, \Delta T_t)$



Related work



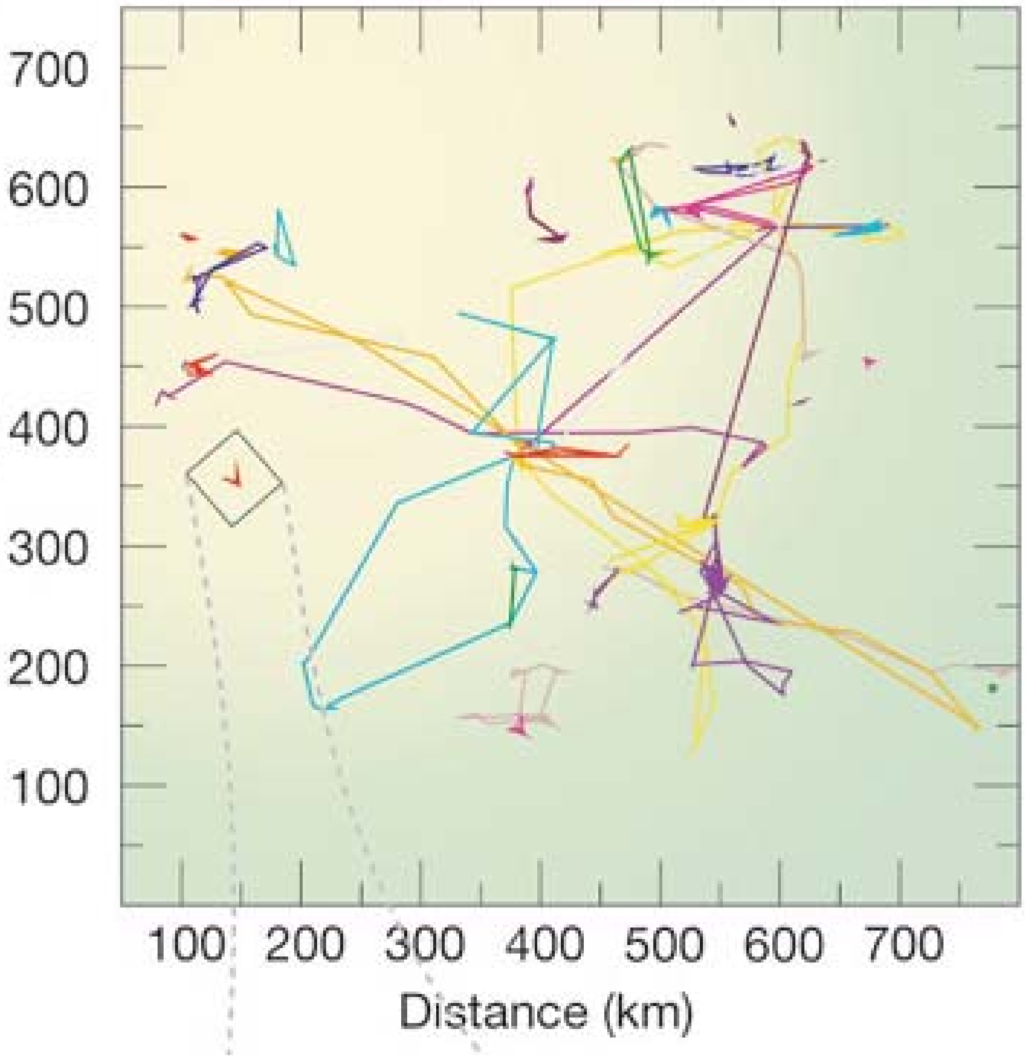


Data from **wheresgeorge.com**

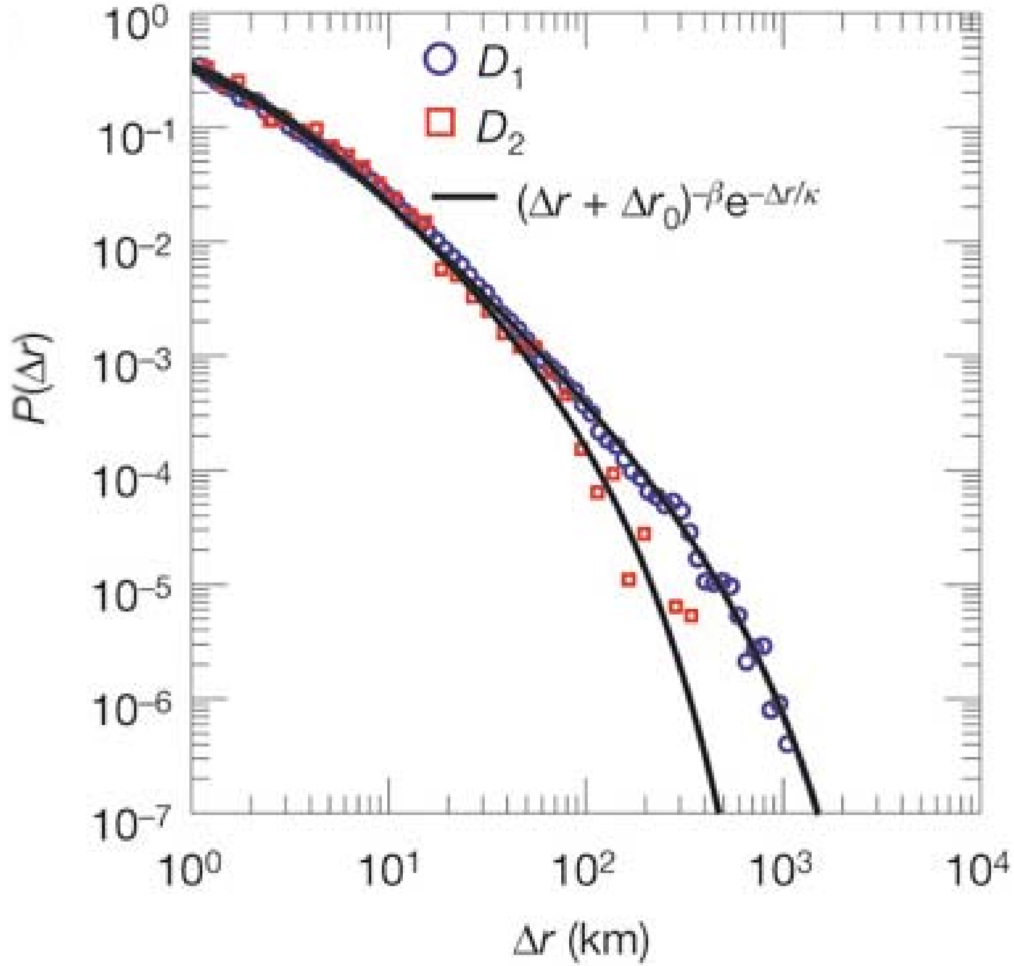
Lévy flight (power law):

$$r \sim r^{-\beta}$$

(Brockmann et al., *Nature* 2006)



Mobile phone traces



Power-law with cutoff

(González et al., *Nature* 2008)

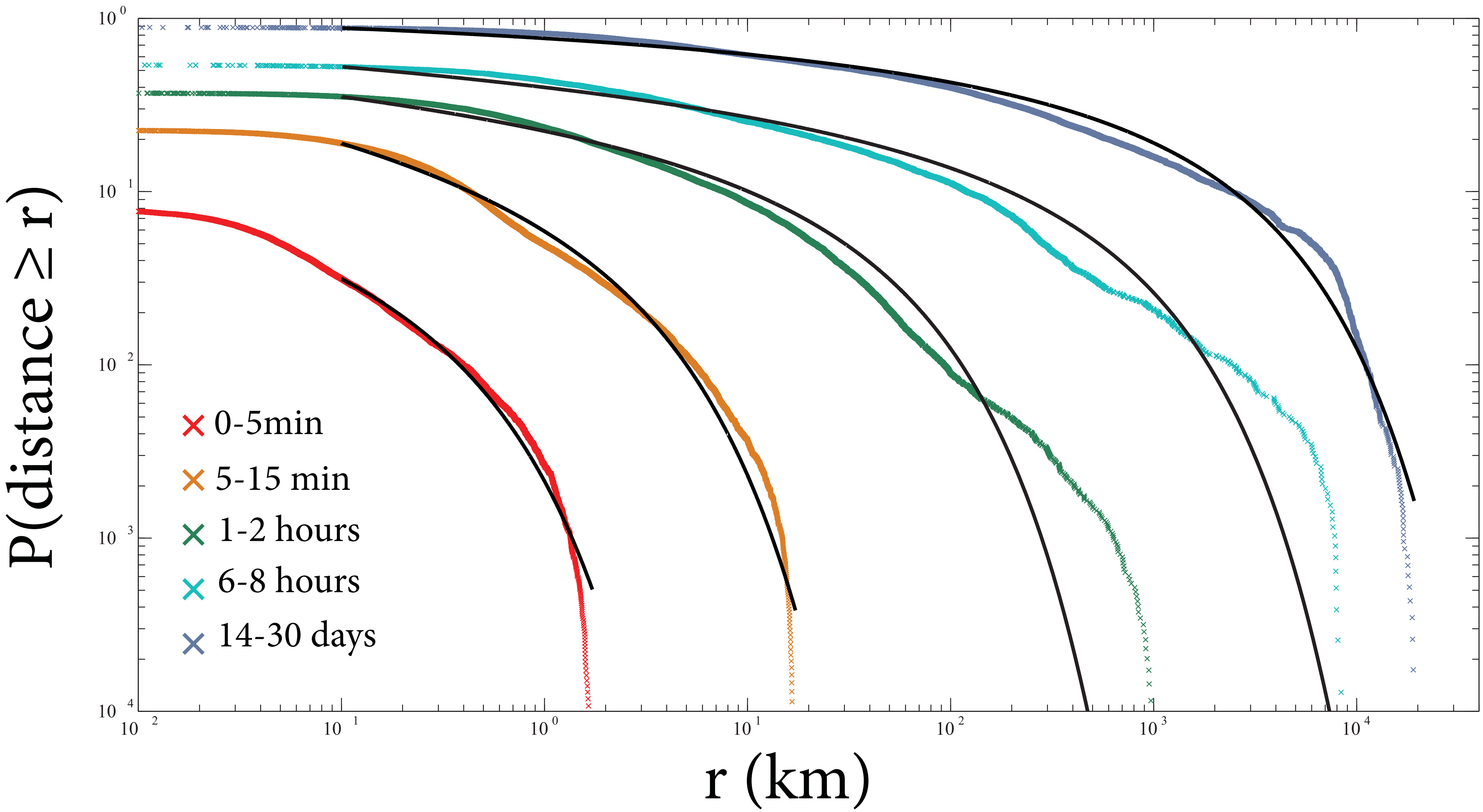
Photo travel database

6 million geotagged images downloaded from Flickr, through Nov 2007

Removed images based on tags (e.g., "birthday," "concert," "abstract," "cameraphone," etc.)

Removed users with no travel, implausible travel (e.g., 100 km in under 45 minutes) or obviously incorrect geotags (e.g., picture of Vancouver geotagged in Siberia)

Flickr distance histogram



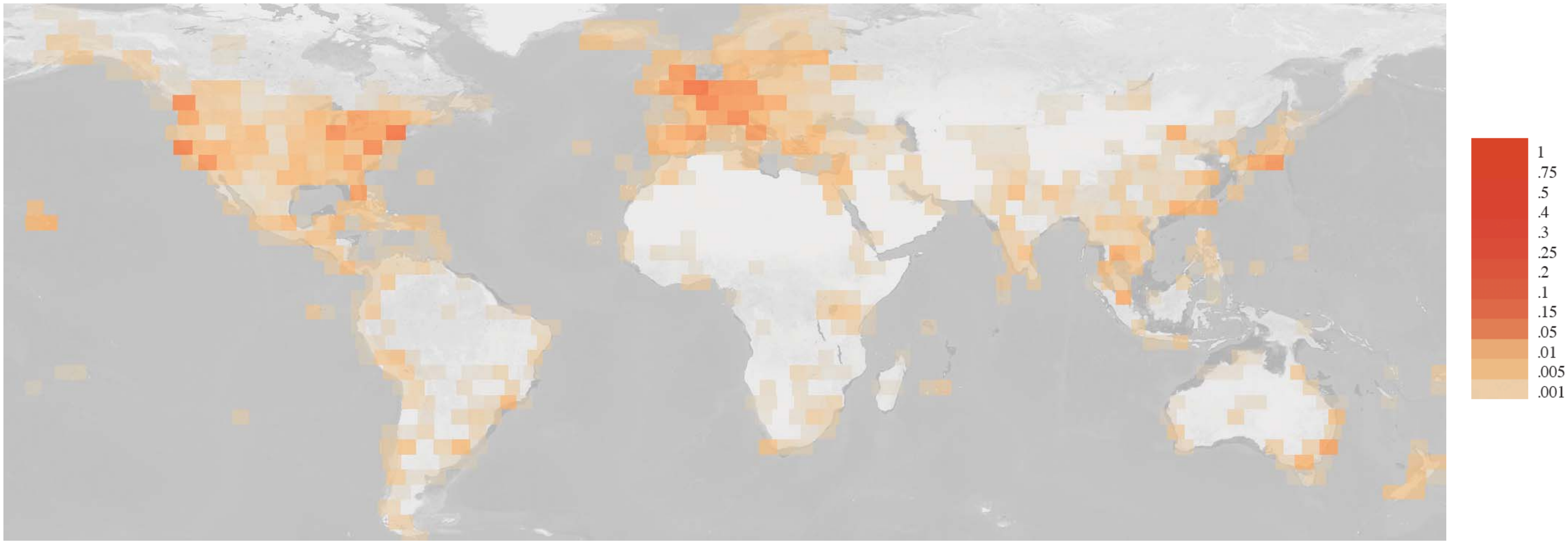
Discretization

400 km x 400 km, 3186 bins L_i

L

Empirical distribution

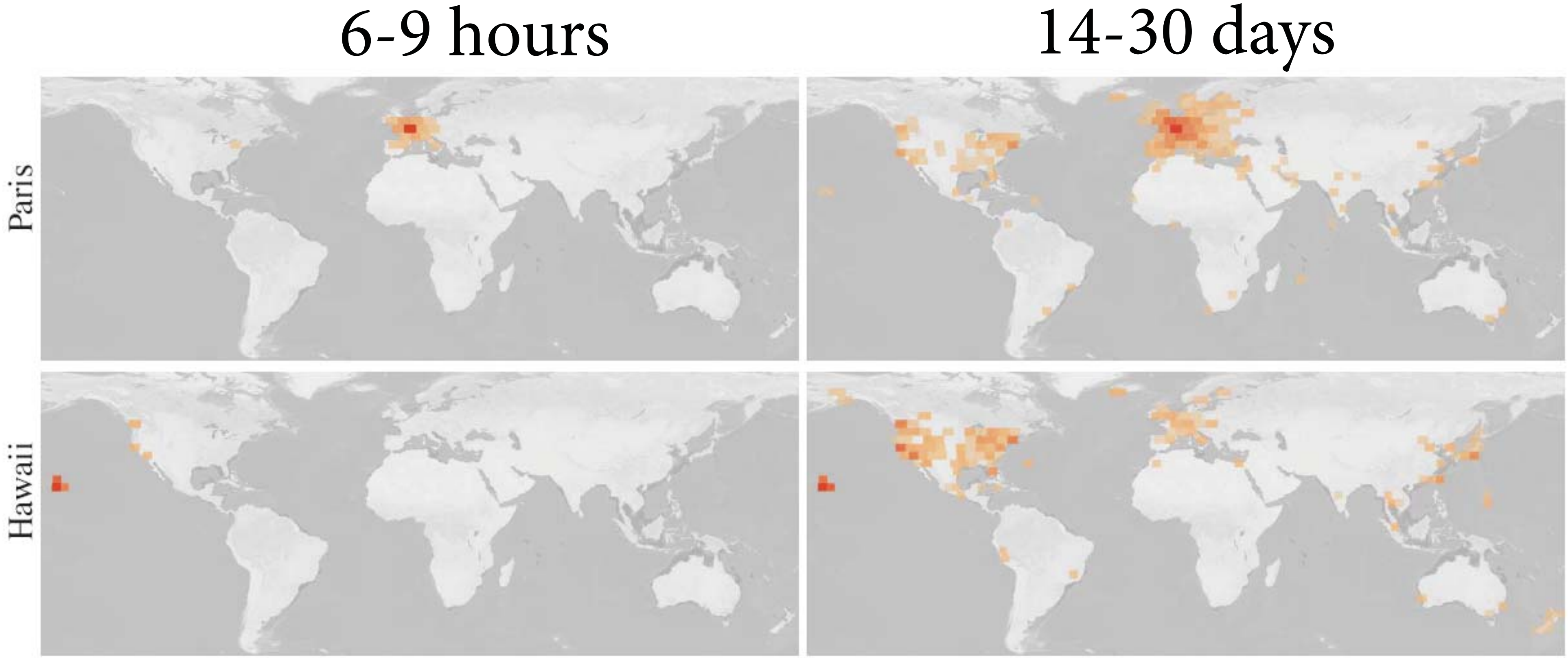
6 million geo-tagged images from Flickr.com



Spatially-varying distribution

$$P(L_{t+1} = j | L_t = i, \Delta T = k) = \frac{N_{ijk}}{\sum_i N_{ijk}}$$

Spatially-varying distribution



Single-image geolocation

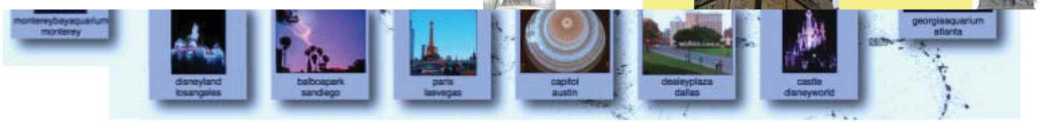
Related work

Urban (Zhang 2006, Schindler 2008)

Regional (Cristani 2008)

Global (Hays 2008)

Landmarks (Crandall 2009, Zheng 2009)



Location likelihood

Test image I

$$w_m = \frac{e^{-\lambda_m D(I, I_m)}}{\sum_{\ell=1}^{M} e^{-\lambda_m D(I, I_\ell)}}$$

$$P(L = i | I) \propto \left(\sum_m w_m \right) + \lambda_C$$

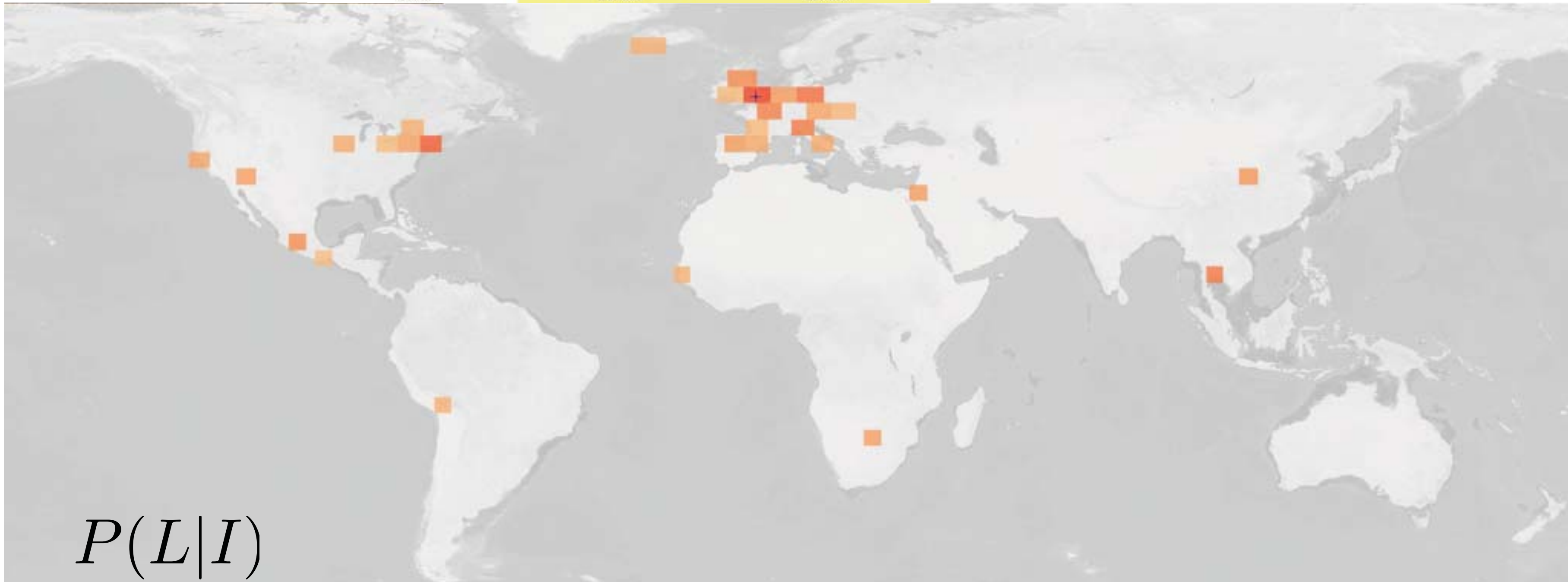
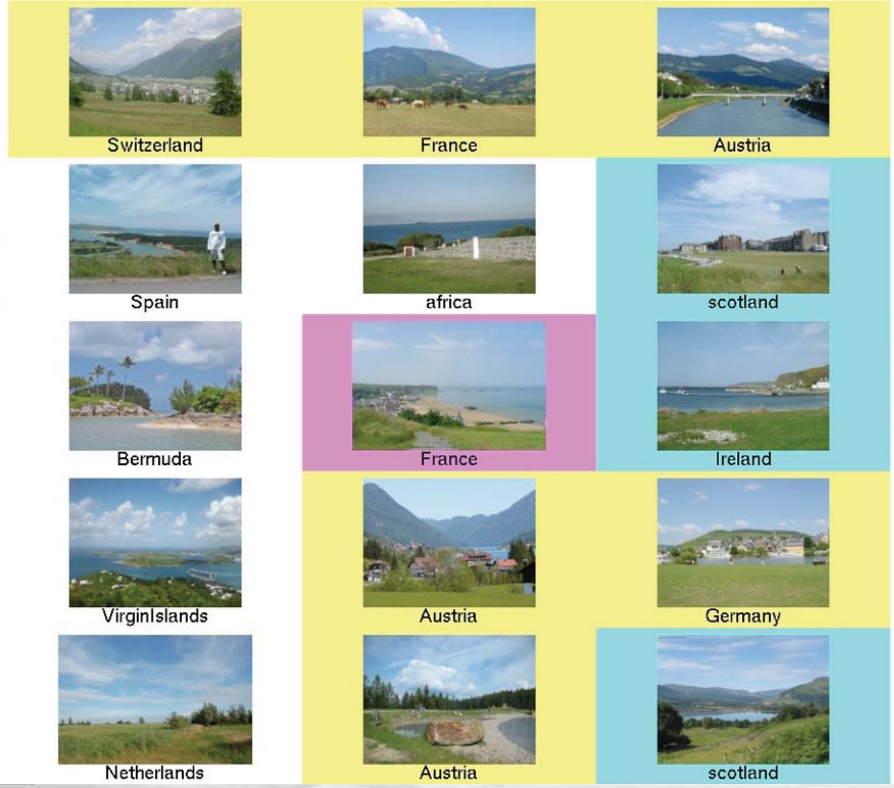
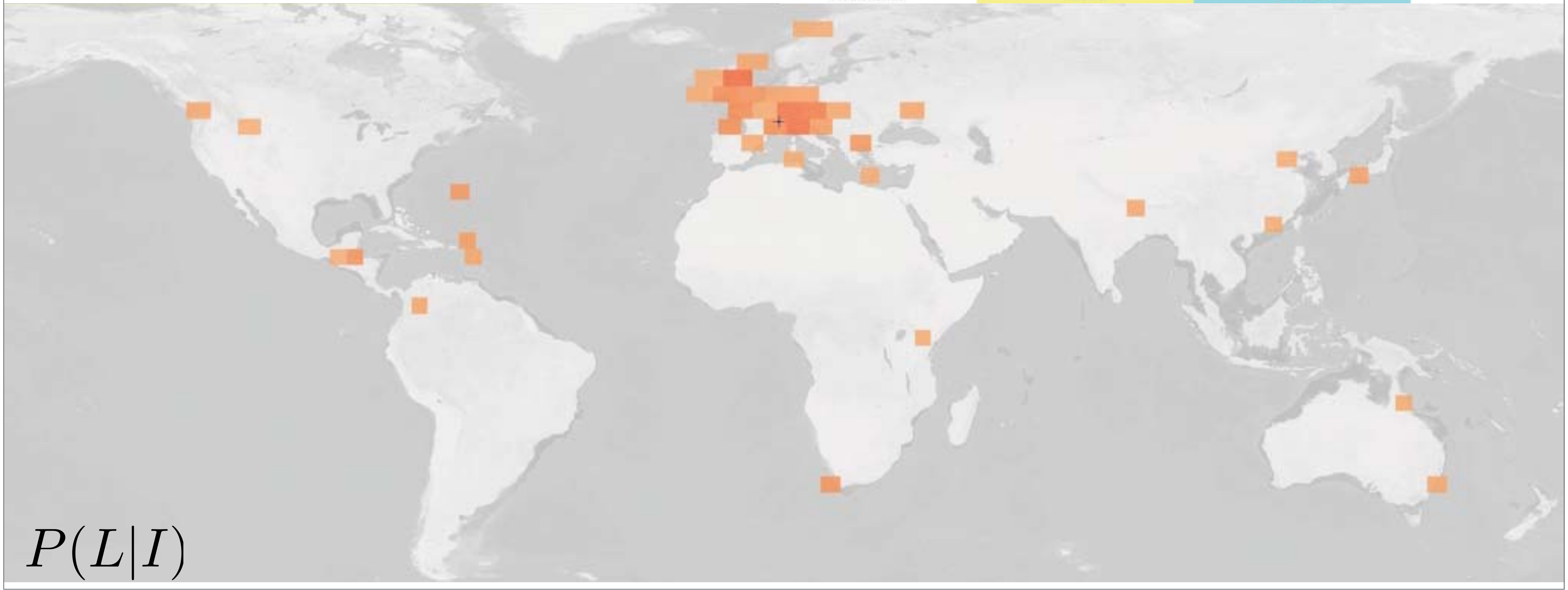


Image similarity score

Distance $D(I, I_m)$ between images is L_2 distance of:

- **Gist** descriptor (Oliva and Torralba 2006)
- **Color histograms**: L*A*B* 4x14x14 bins
- **Texton histograms**: 512 entry, filter-bank
- **Line histogram**


Switzerland
France
Austria
Spain
africa
scotland
Bermuda
France
Ireland
VirginIslands
Austria
Germany
Netherlands
Austria
scotland



$P(L|I)$


Indonesia
SriLanka
Philippines
Sydney
Rhode
Namibia
Vietnam
Australia
Italy
Australia
Italy
Argentina
Portugal
Norway
england

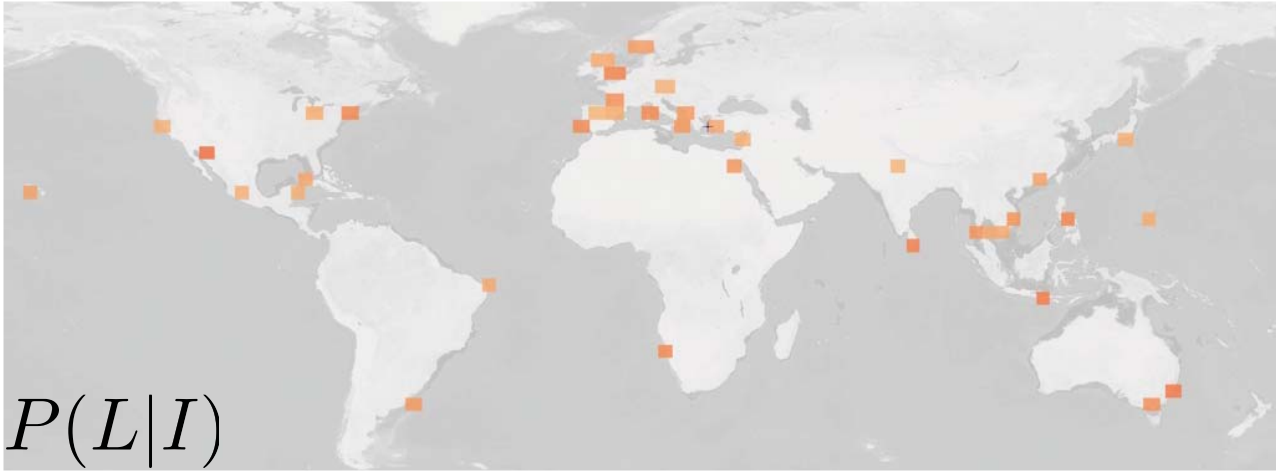
$P(L|I)$

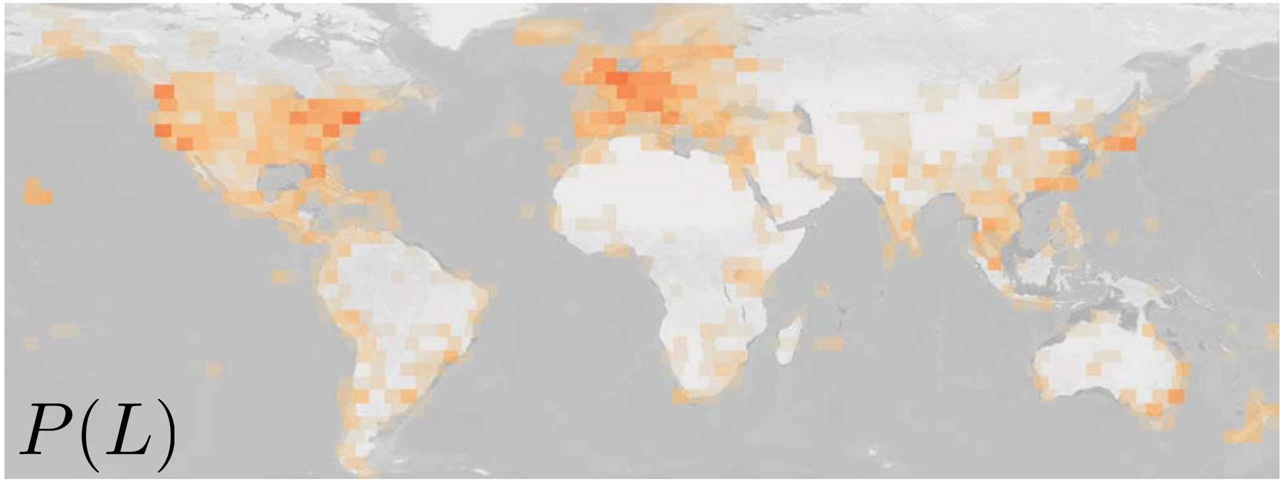
A loose continuum

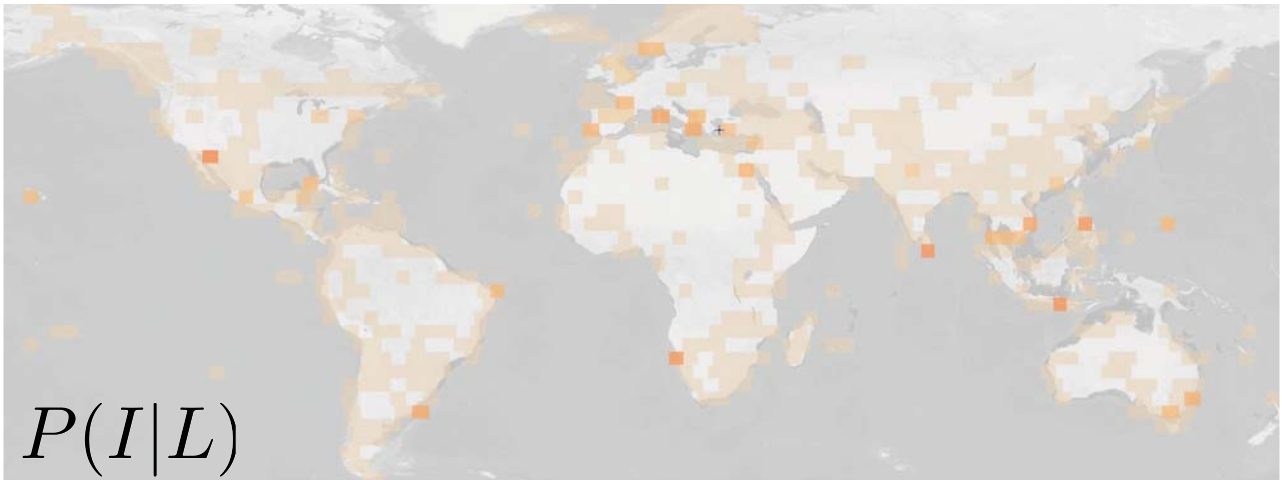
1. Distinctive
(e.g., landmarks)

2. Vague
(e.g., regional,
terrain/type)

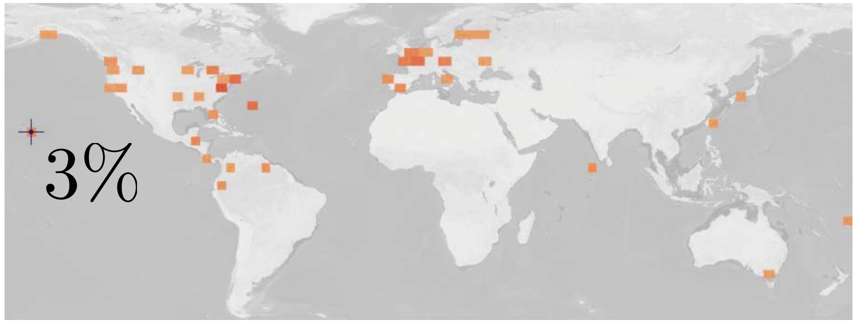
3. Nearly
uninformative


P(L|I)

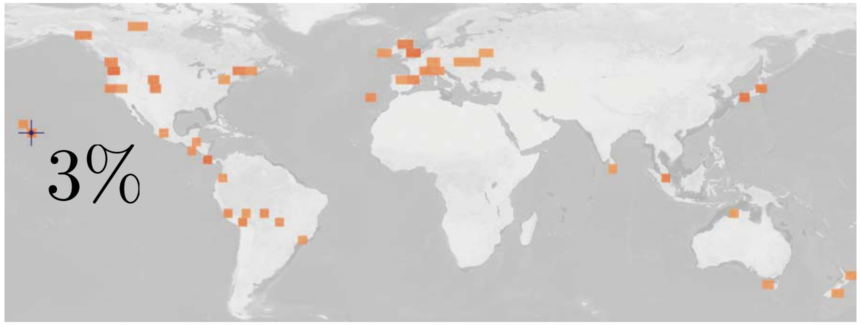

P(L)


P(I|L)

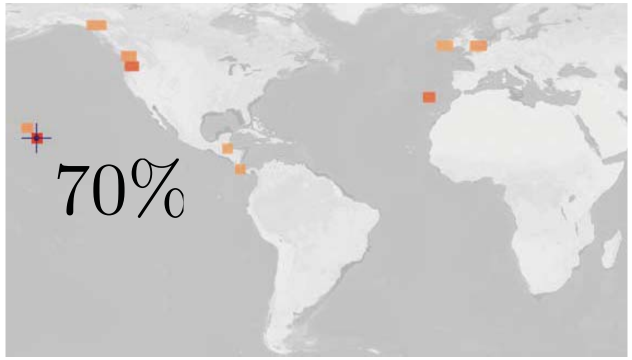
Combining “vague” results



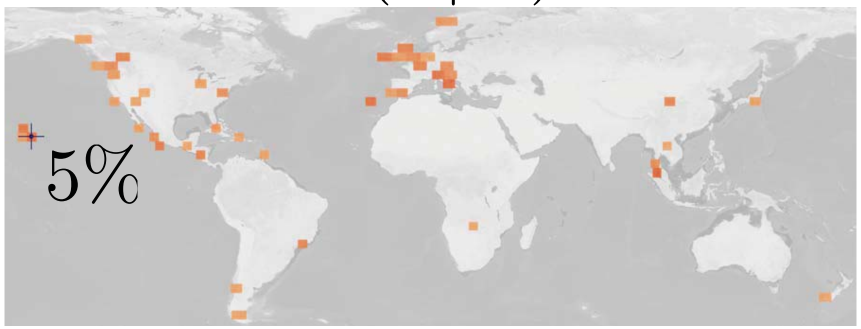
$P(L|I_1)$



$P(L|I_2)$

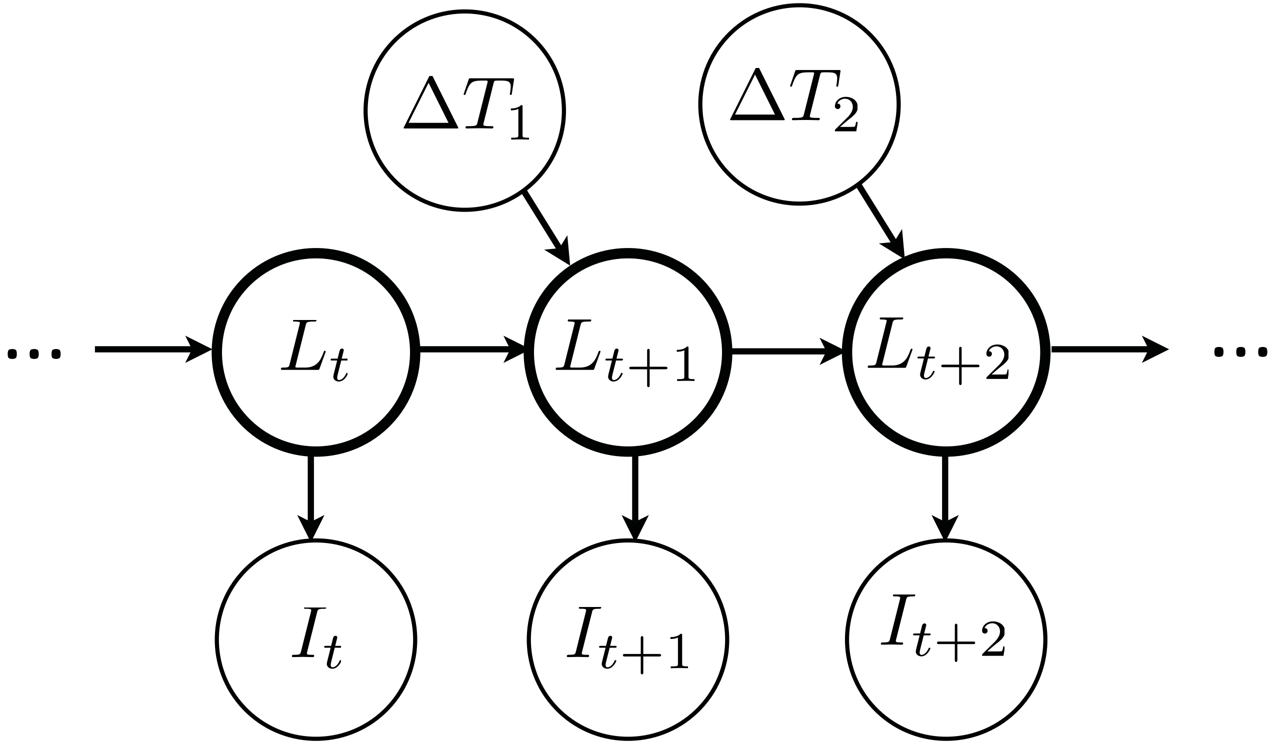


$P(L|I_1, I_2, I_3)$



$P(L|I_3)$

Hidden Markov Model



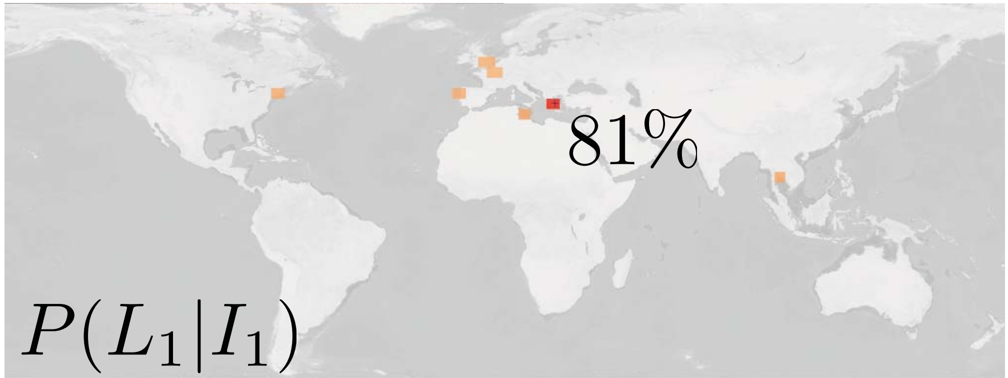
Forward-Backward algorithm computes

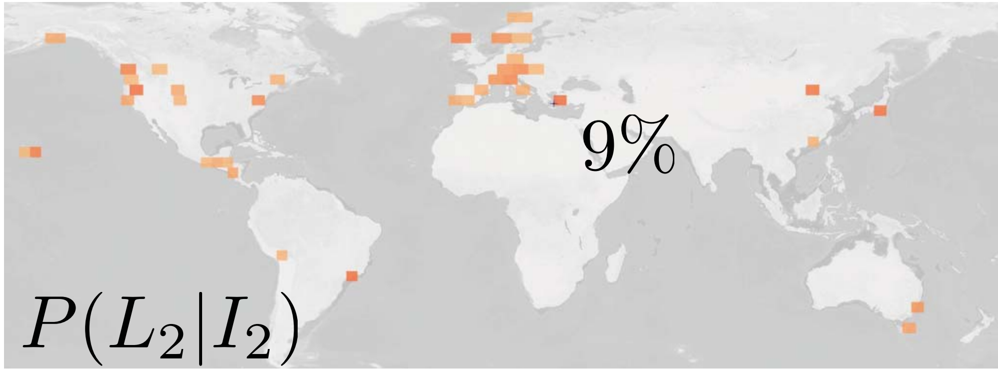
$$\gamma_{it} \equiv P(L_t = i | I_{1:N}, \Delta T_{1:N})$$

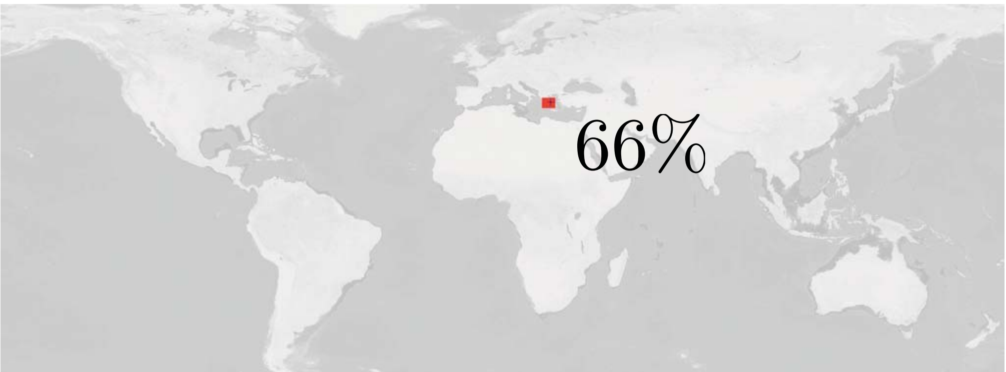
Given loss function, output a location estimate

Toy example

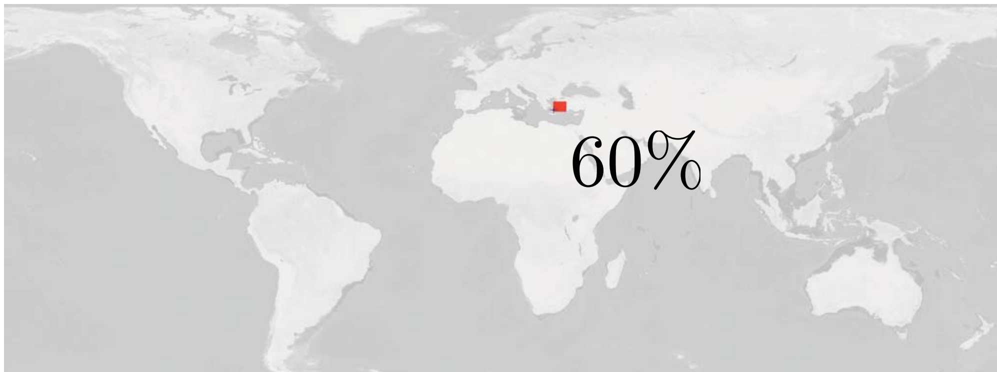
$\Delta T = 2$ hours







$P(L_1|I_1, I_2, \Delta T)$



$P(L_2|I_1, I_2, \Delta T)$

User-specific learning

User's results added to their training data

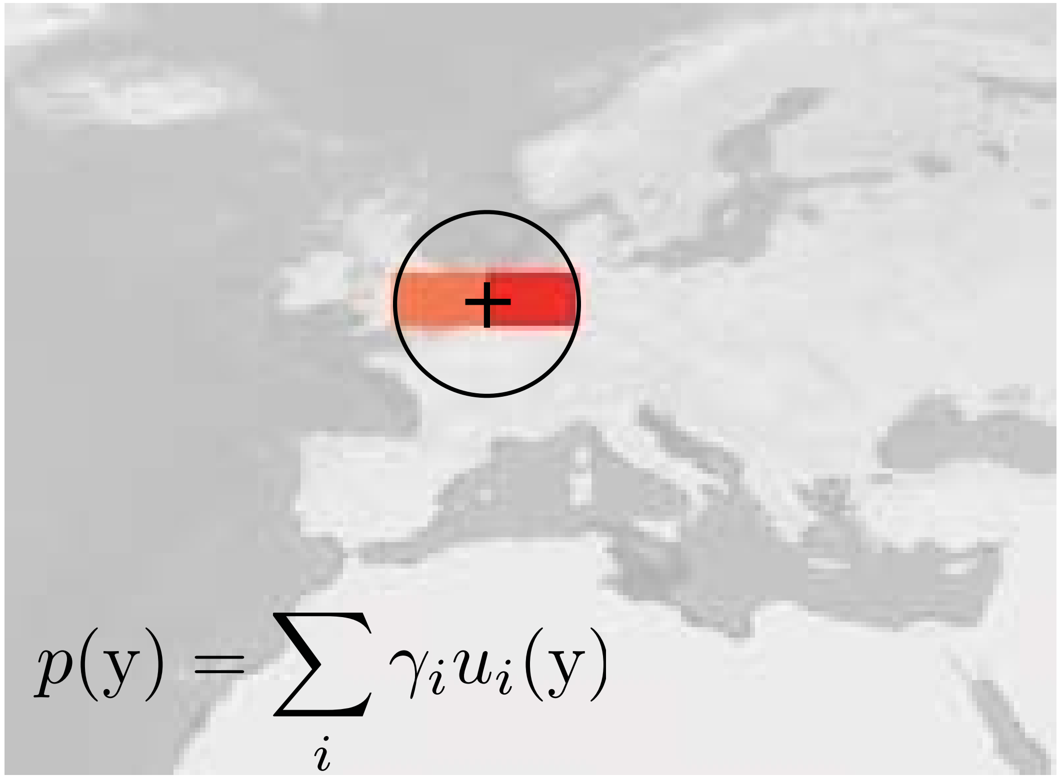
EM-like algorithm

New likelihood:

$$P(L = i|I) \propto \left(\sum_m w_m\right) + \left(\sum_n \gamma_{ni} w_n\right) + \lambda_C$$

Location estimation

Task: correct estimation with 400 km



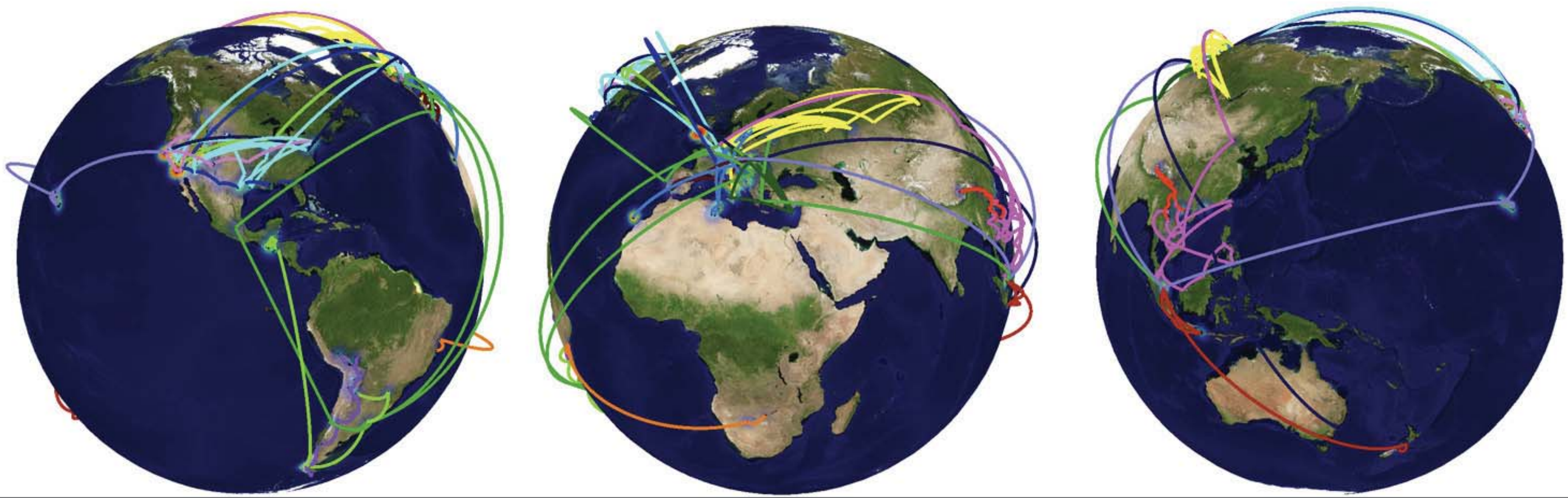
$$p(\mathrm{y}) = \sum_i \gamma_i u_i(\mathrm{y})$$

$$\mathrm{x}^* = \arg\max_{\mathrm{x}} \int_{||\mathrm{x}-\mathrm{y}||} p(\mathrm{y})dy$$

Evaluation

Validation set (6 users, 2005 photos):

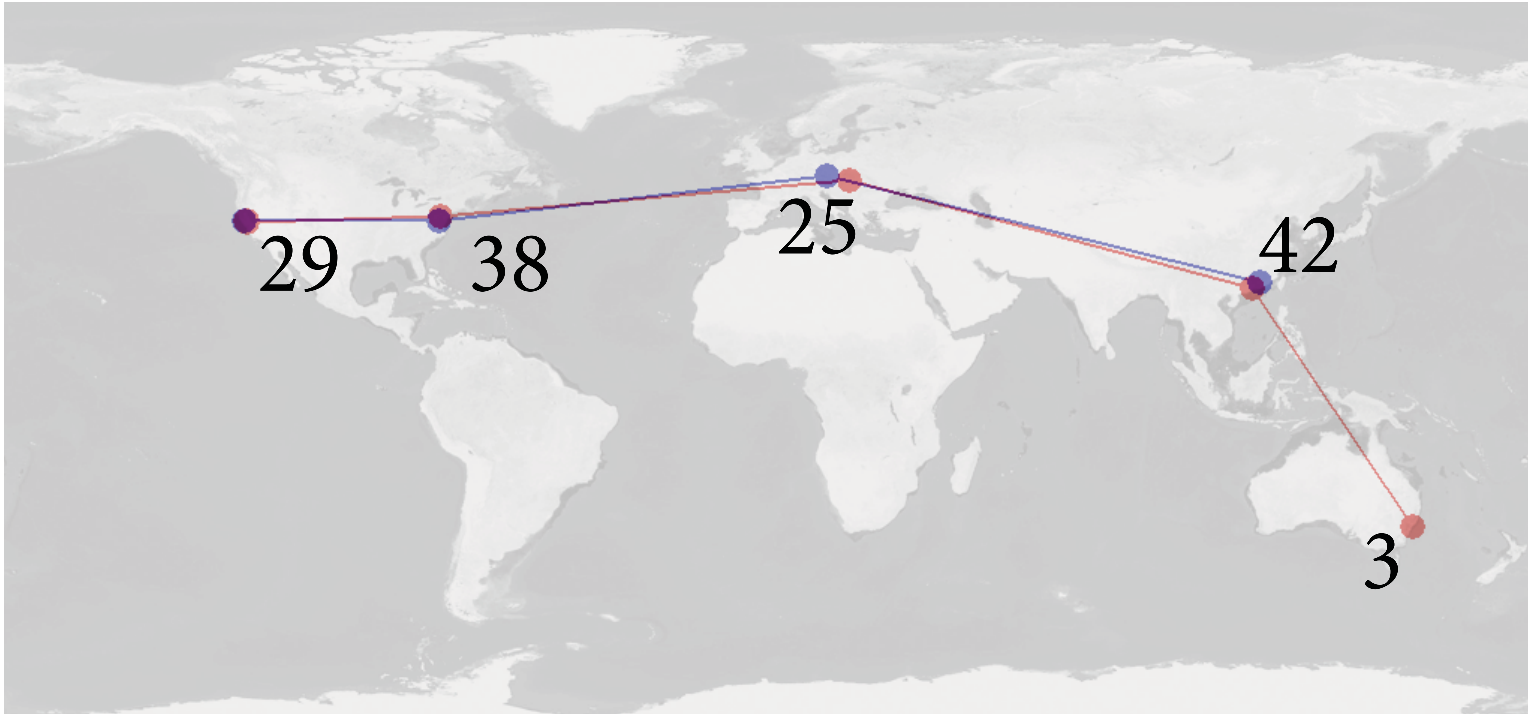
Test set (20 users, 4117 photos):



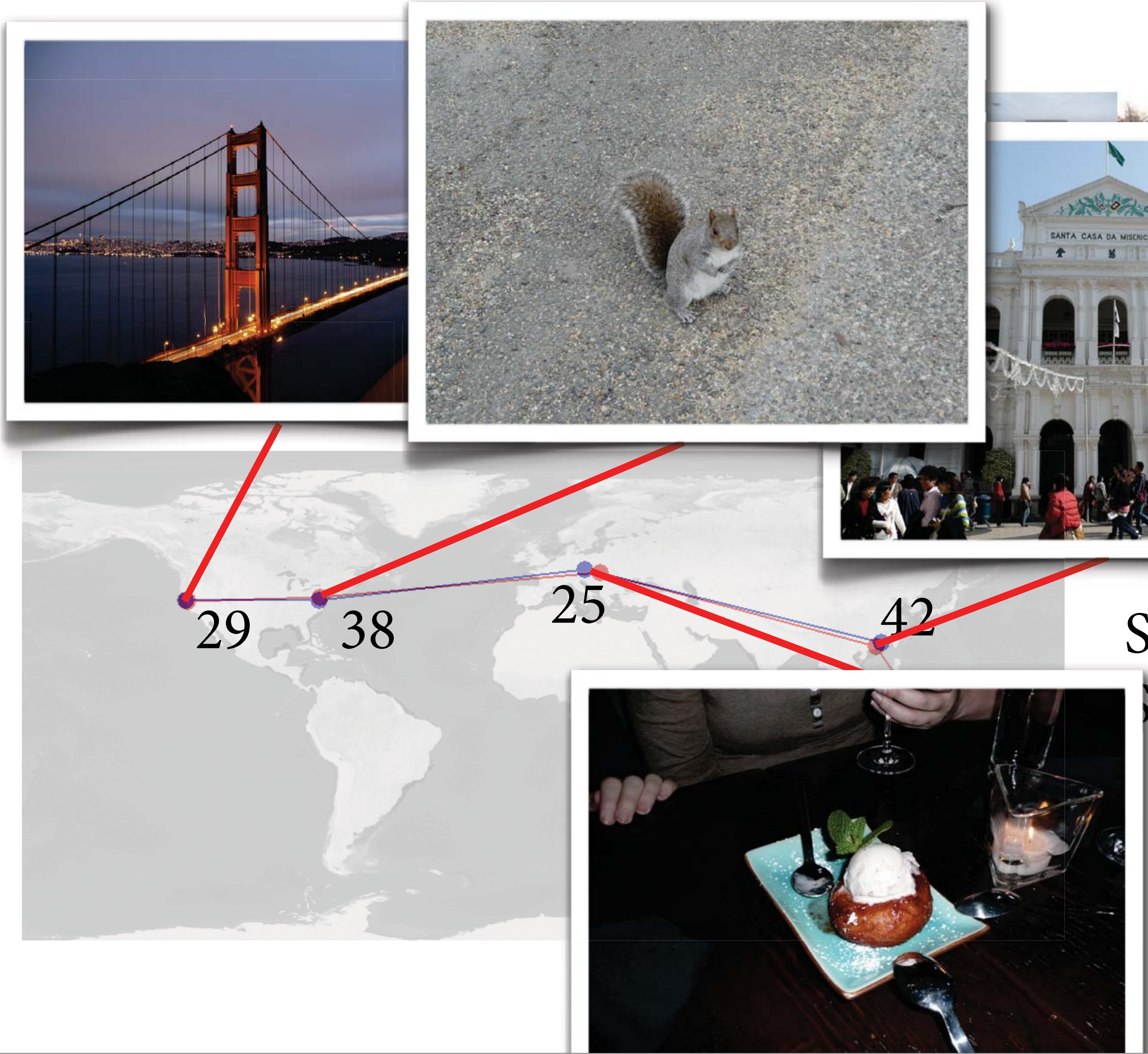
Results (correct within 400km) for test set:

London always	3%
IM2GPS (Hayes and Efros 2008)	10%
Sequence	58%

137 photos



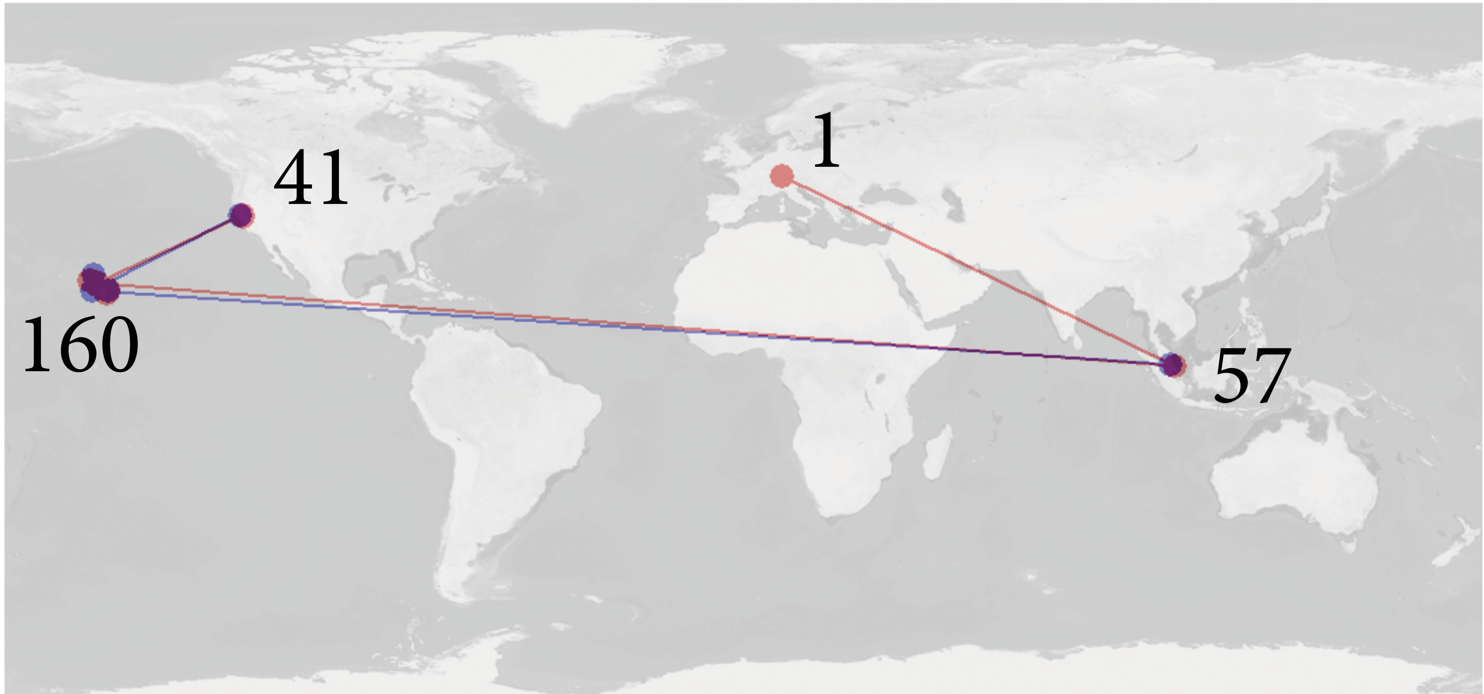
SIG: 37.7%
SEQ: 97.8%



SIG: 37.7%

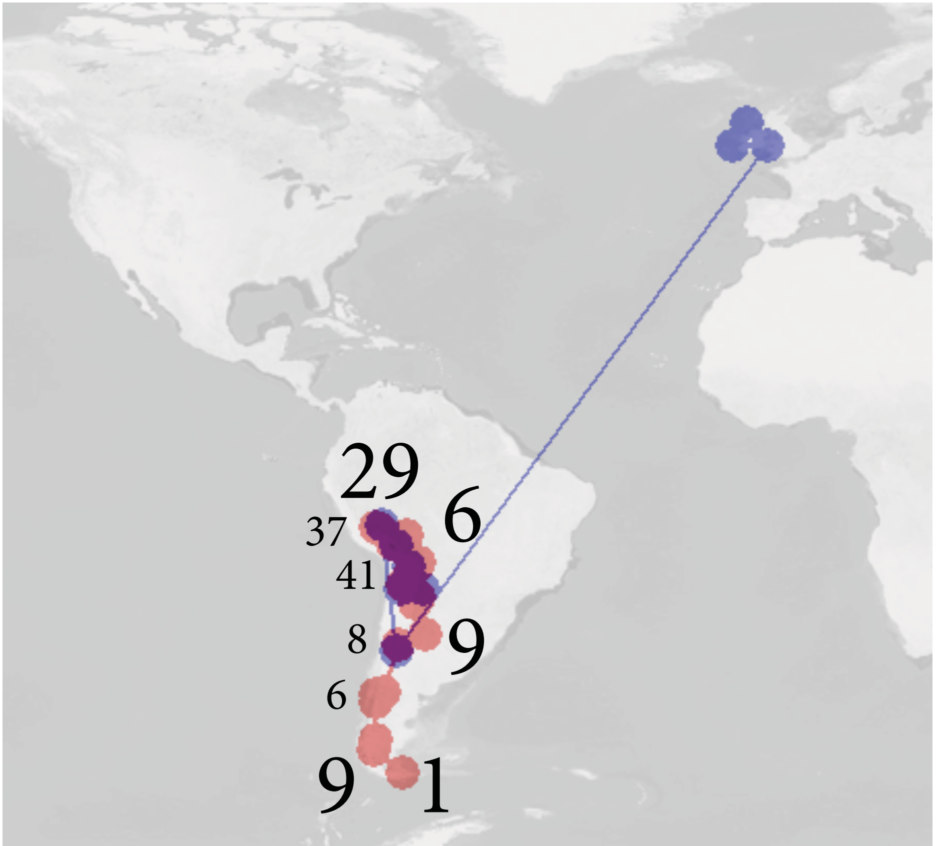
SEQ: 97.8%

259 photos



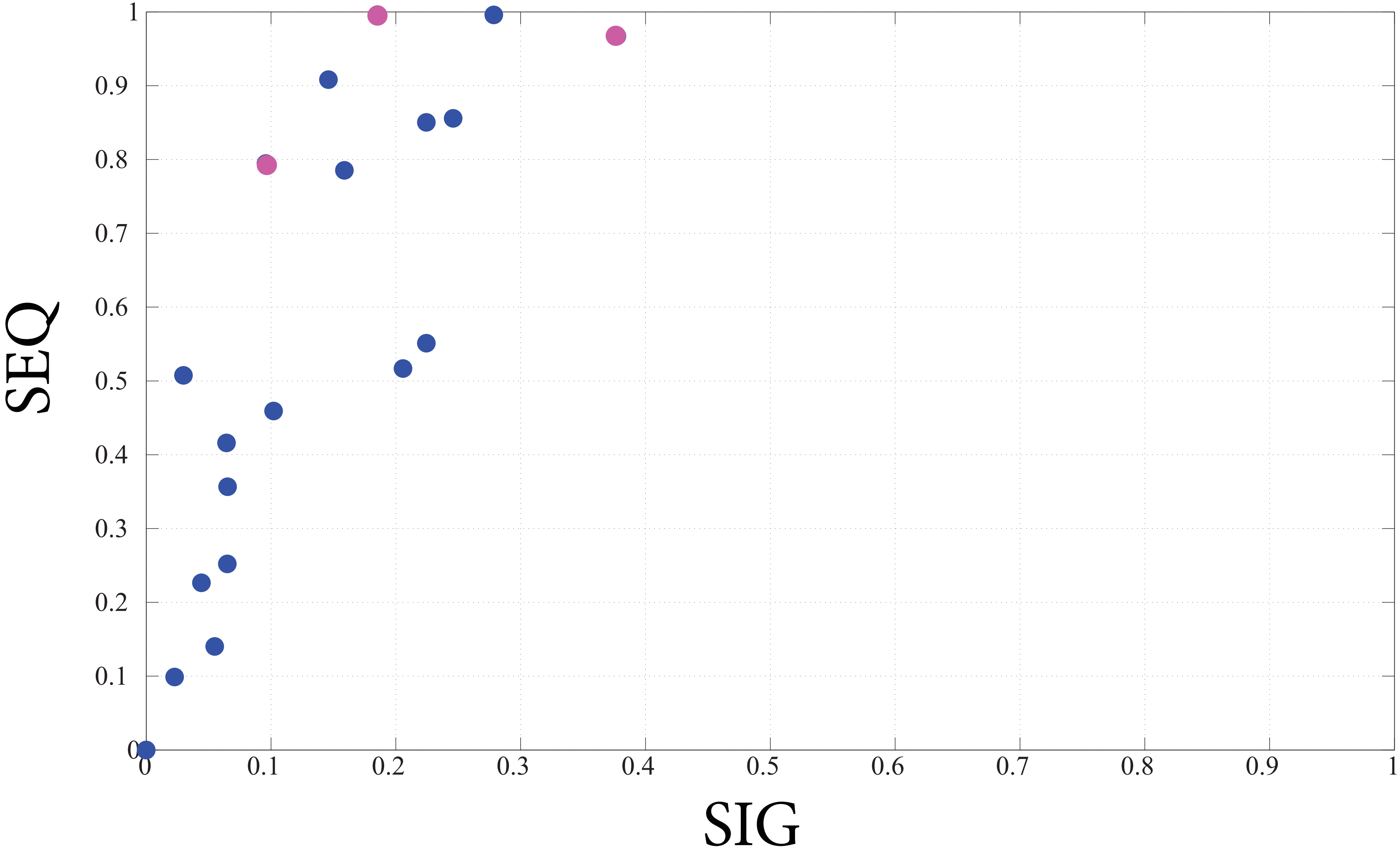
SIG: 18.5%
SEQ: 99.6%

146 photos



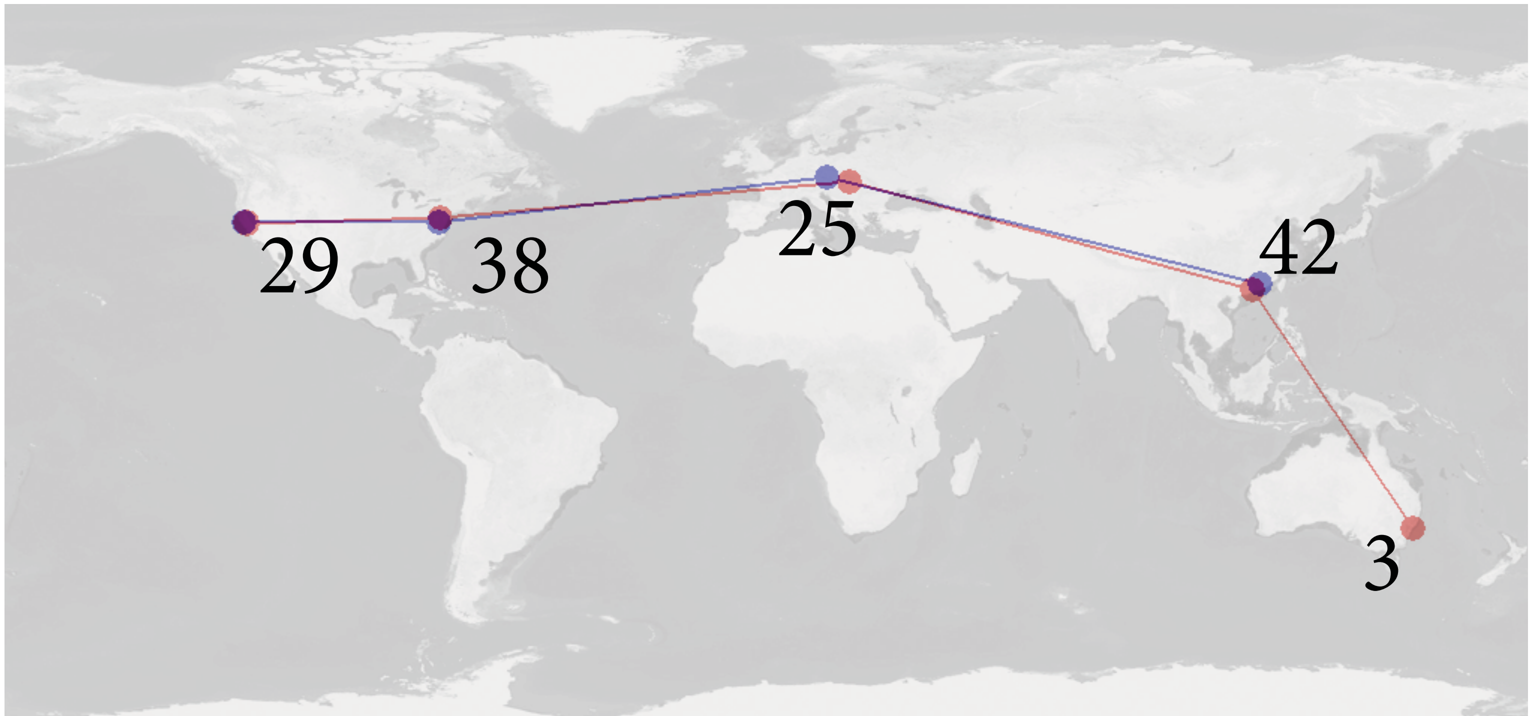
SIG: 10%

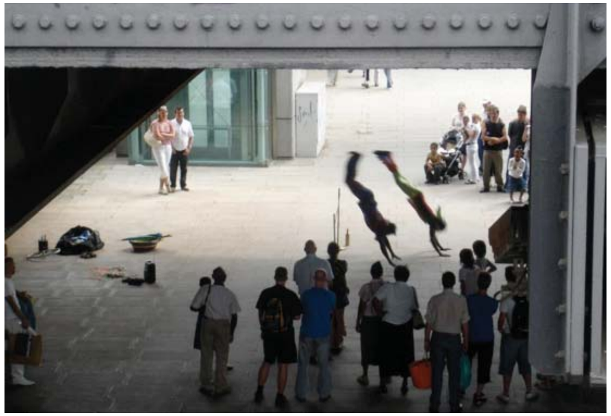
SEQ: 79%



SEQ
SIG
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Is it just landmark matching?





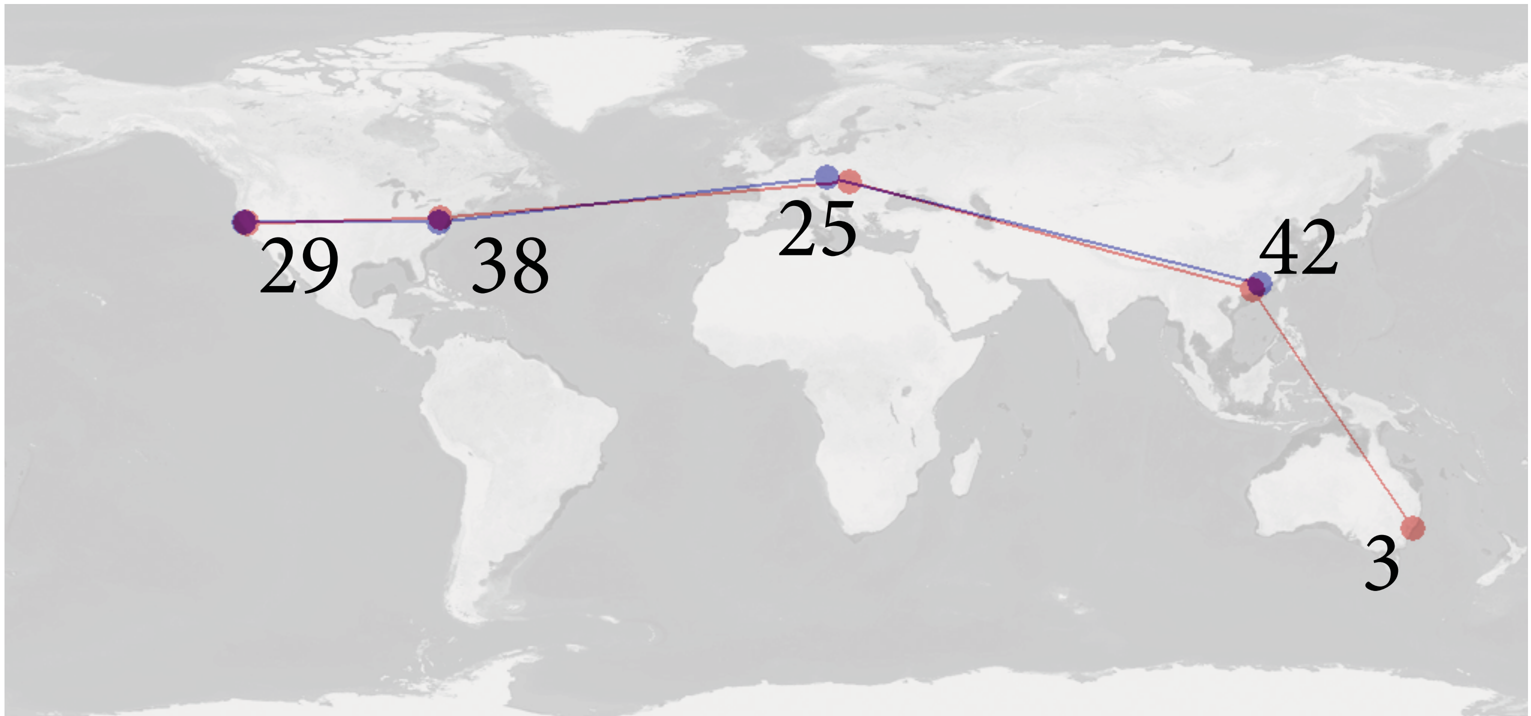
“Distinctive”

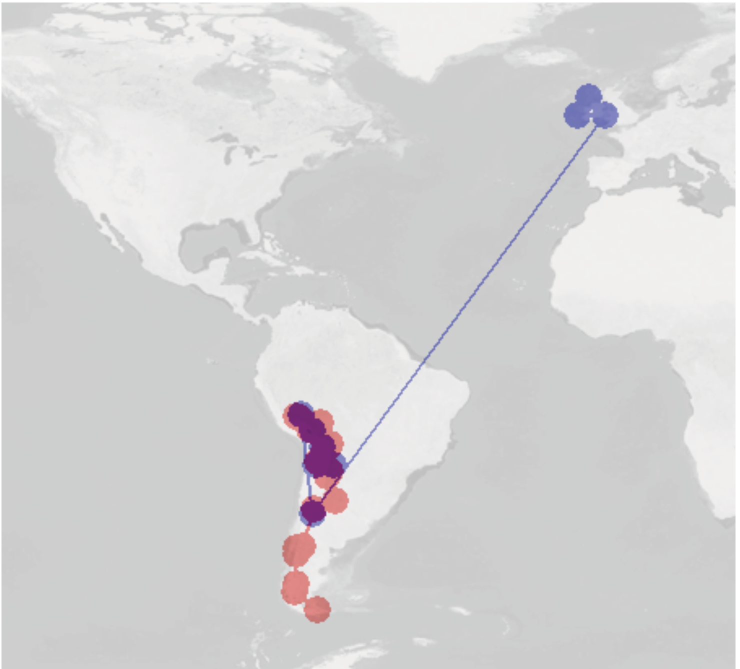
“Non-distinctive”

“Distinctive”

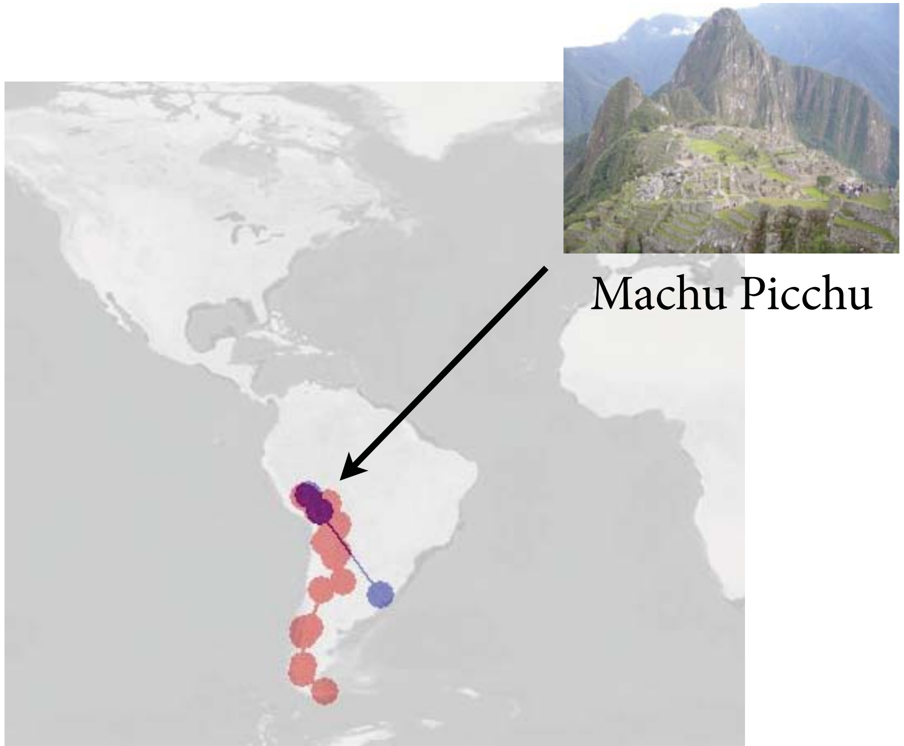
Landmark-only: 41%

Sequence: 58%

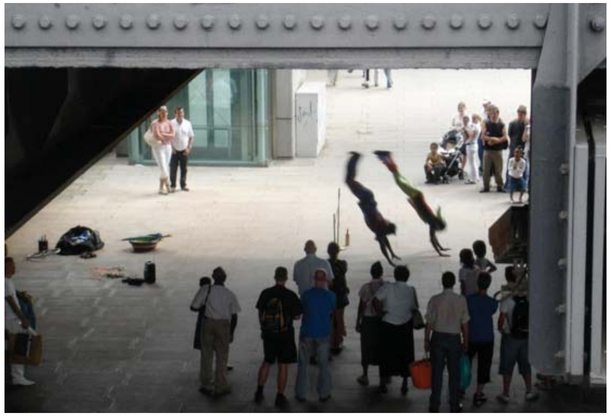

29
38
25
42
3

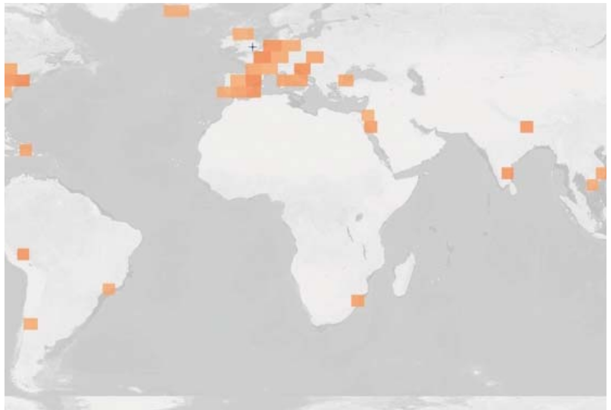


SEQ: 79%



Landmark SEQ: 55%



SIG: 0%

Landmark-less SEQ: 19.3%

Many possible improvements

Better binning

Better image matching

More general models (image meta-data, Flickr tags, user types, image types, weather, economy, transportation, etc).

... and so on

Conclusions

There is a wealth of travel data to explore and exploit

Given **images and timestamps**, we get much more information than from images alone

New application areas for computer vision