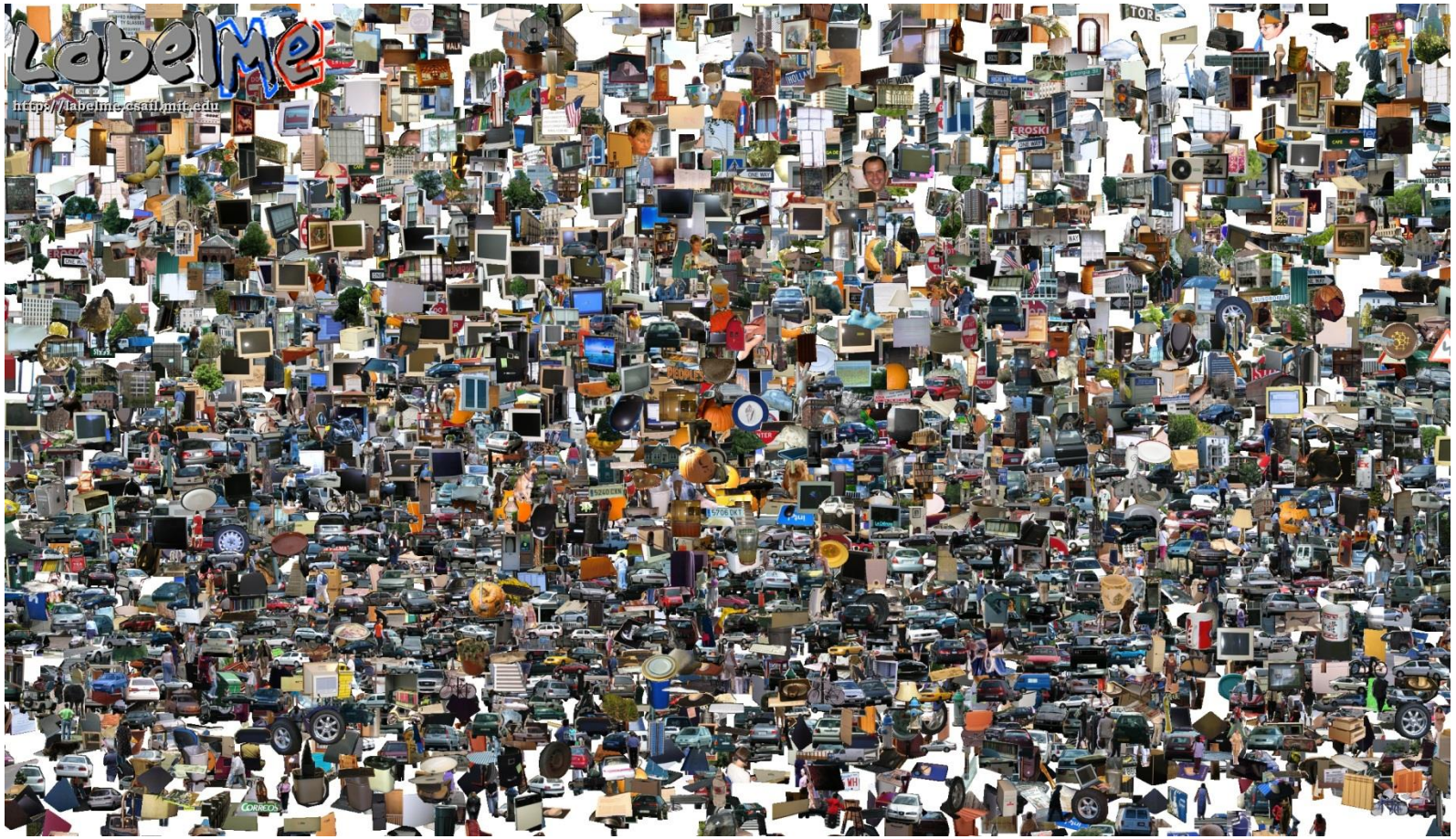
Opportunities of Scale



Computer Vision

James Hays, Brown

Many slides from James Hays, Alyosha Efros, and Derek Hoiem

Graphic from Antonio Torralba

Opportunities of Scale: Data-driven methods

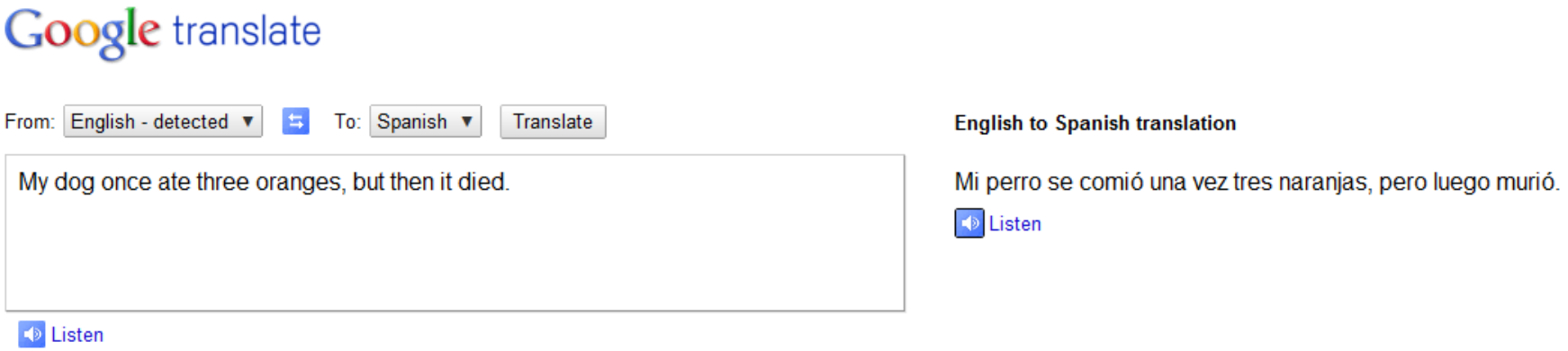
- Today's class
  - Scene completion
  - Im2gps
- Next class
  - Recognition via Tiny Images
  - More recognition by association

Google and massive data-driven algorithms

A.I. for the postmodern world:

- all questions have already been answered…many times, in many ways
- Google is dumb, the "intelligence" is <u>in the data</u>

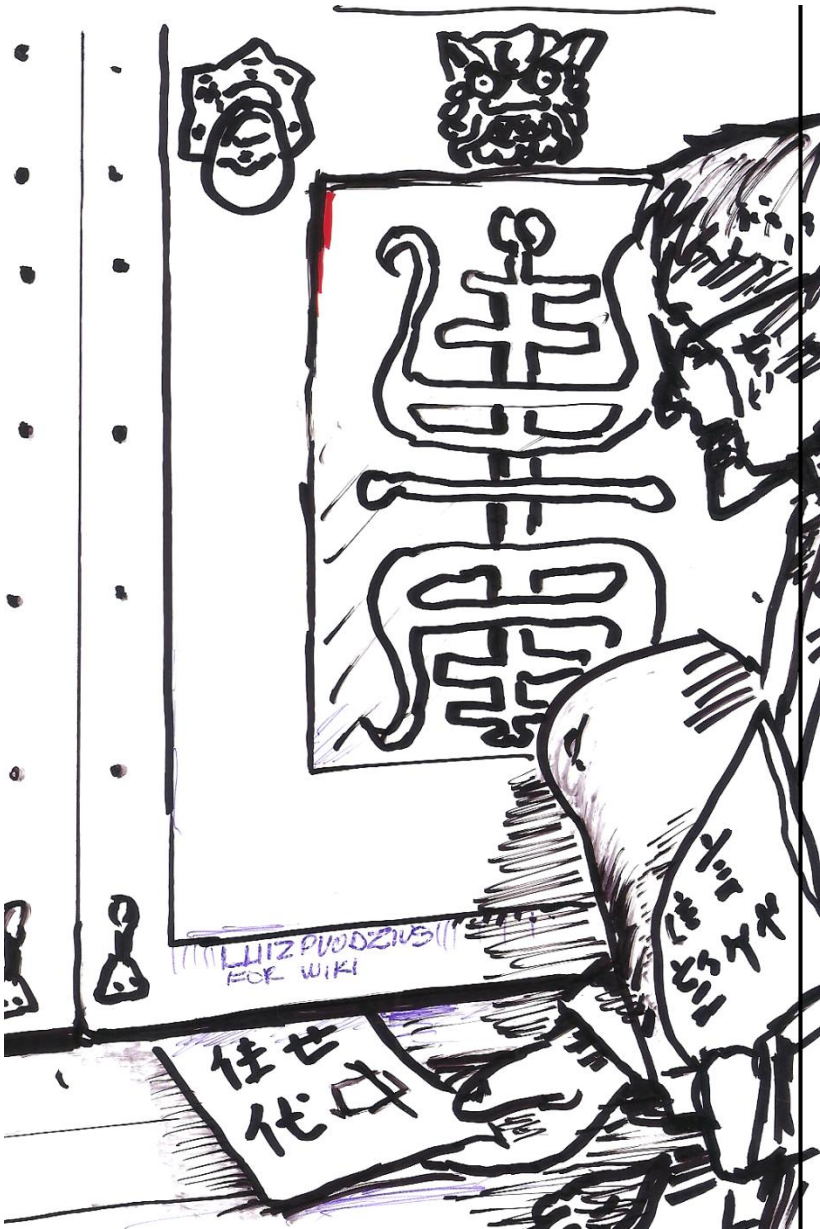
Google Translate



Chinese Room, John Searle (1980)

If a machine can convincingly simulate an intelligent conversation, does it necessarily understand? In the experiment, Searle imagines himself in a room, acting as a computer by manually executing a program that convincingly simulates the behavior of a native Chinese speaker.

Most of the discussion consists of attempts to refute it. "The overwhelming majority," notes *BBS* editor Stevan Harnad,“ still think that the Chinese Room Argument is dead wrong." The sheer volume of the literature that has grown up around it inspired Pat Hayes to quip that the field of cognitive science ought to be redefined as "the ongoing research program of showing Searle's Chinese Room Argument to be false.



Big Idea

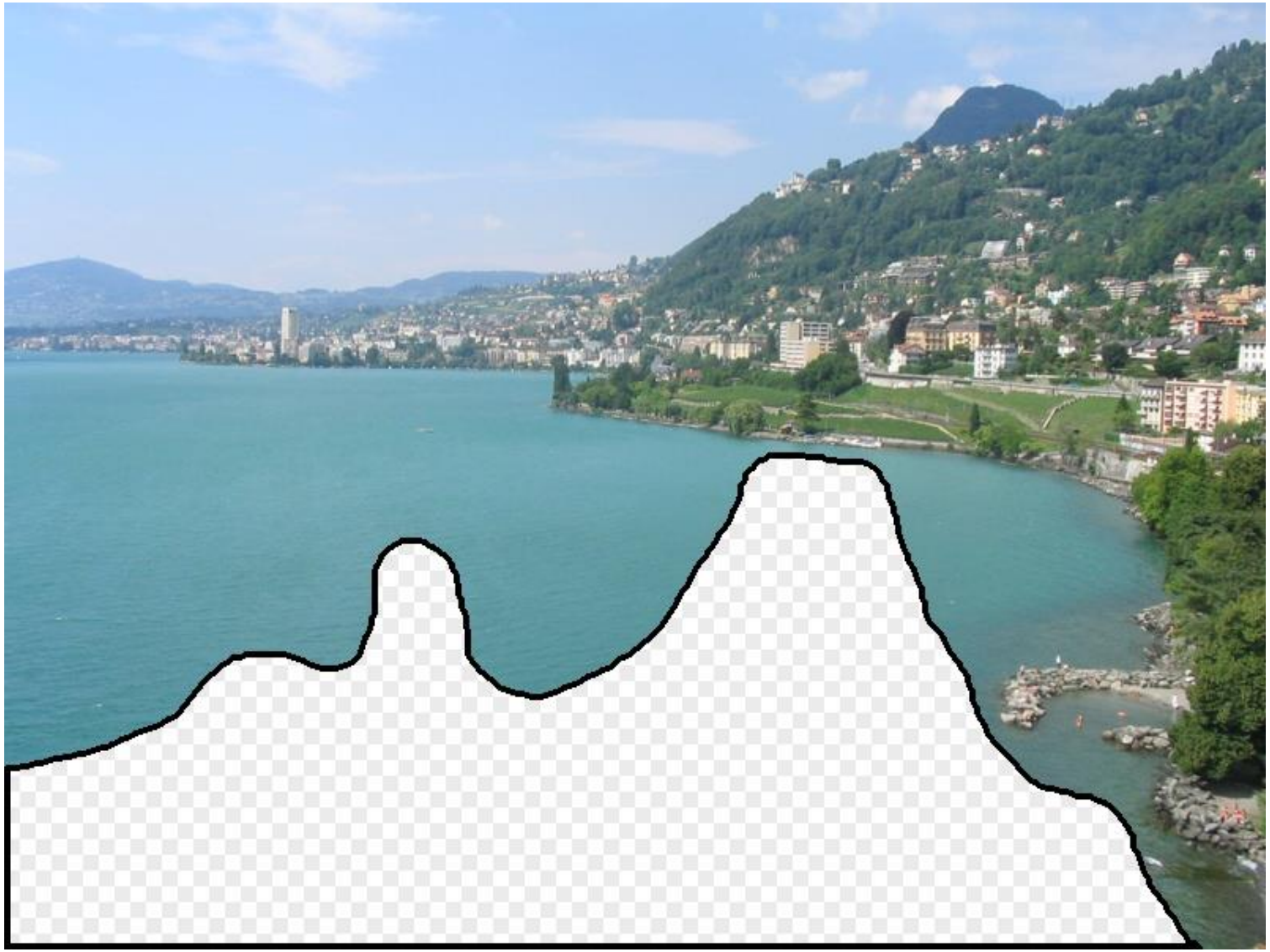
- What if invariance / generalization isn't actually the core difficulty of computer vision?
- What if we can perform high level reasoning with brute-force, data-driven algorithms?

Image Completion Example

[Hays and Efros. Scene Completion Using Millions of Photographs. SIGGRAPH 2007 and CACM October 2008.]

http://graphics.cs.cmu.edu/projects/scene-completion/

What should the missing region contain?



Which is the original?

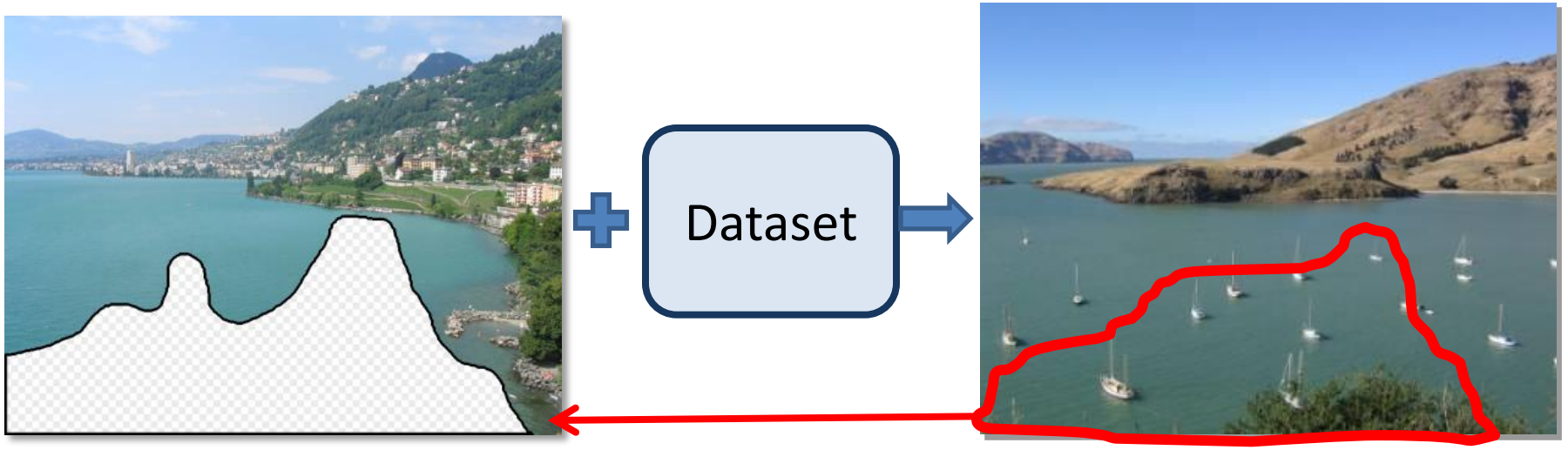
(a)

(b)

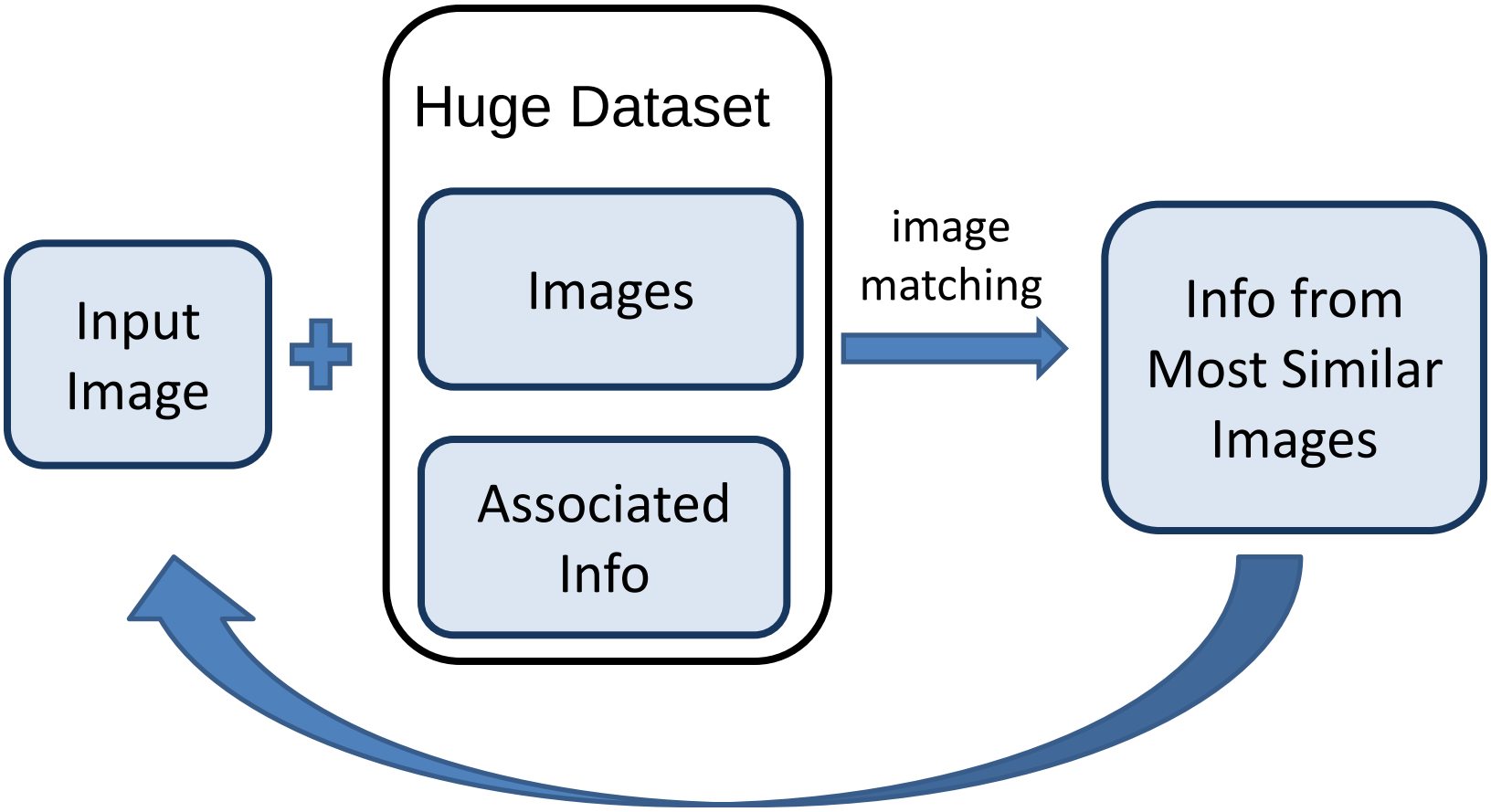
(c)

How it works

- Find a similar image from a large dataset
- Blend a region from that image into the hole

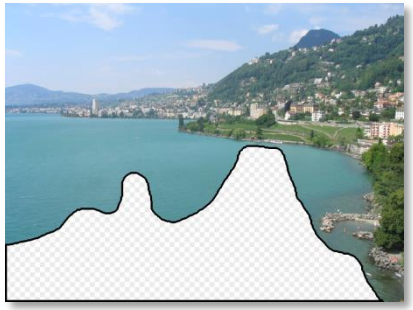


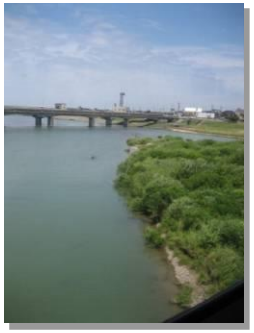
General Principal



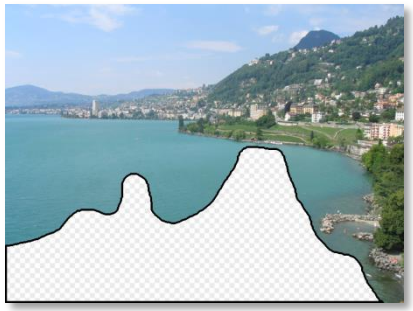
Hopefully, If you have enough images, the dataset will contain very similar images that you can find with simple matching methods.

How many images is enough?



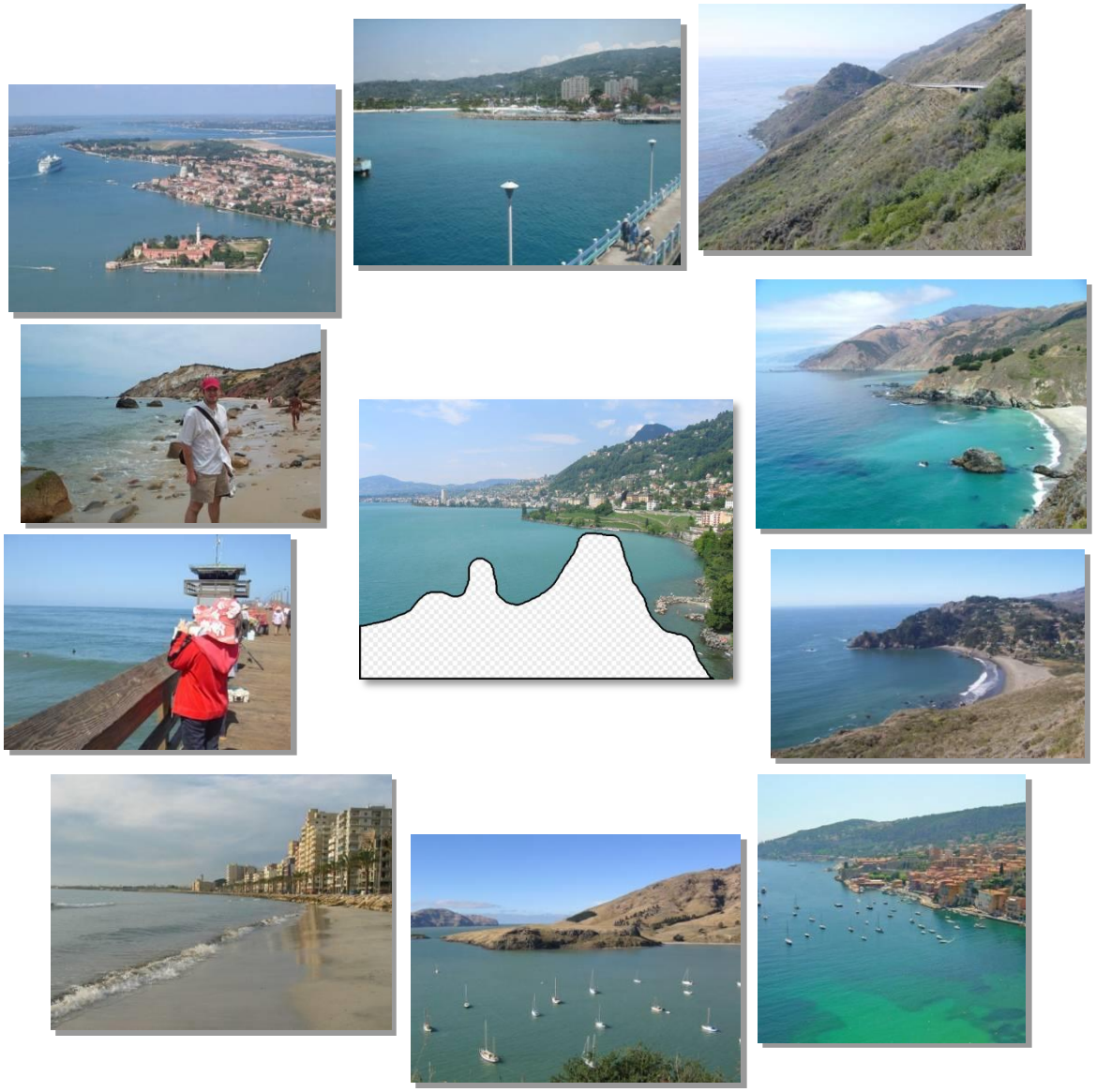







Nearest neighbors from a
collection of 20 thousand images



Nearest neighbors from a
collection of 2 million images

Image Data on the Internet

- Flickr (as of Sept. 19th, 2010)
  - 5 billion photographs
  - 100+ million geotagged images
- Facebook (as of 2009)
  - 15 billion

http://royal.pingdom.com/2010/01/22/internet-2009-in-numbers/

Image Data on the Internet

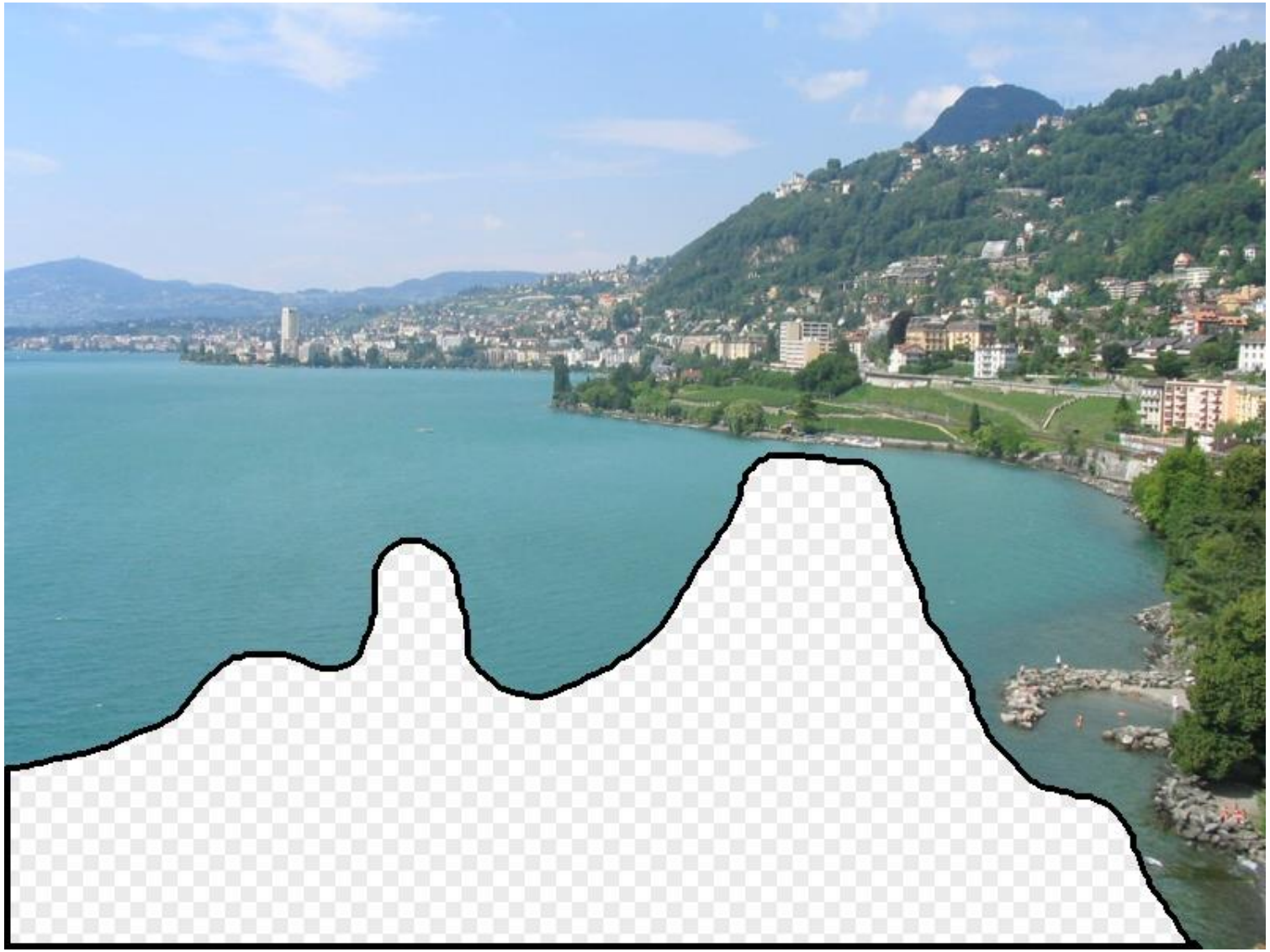
- Flickr (as of Nov 2013)
  - 10 billion photographs
  - 100+ million geotagged images
  - 3.5 million a day
- Facebook (as of Sept 2013)
  - 250 billion+
  - 300 million a day
- Instagram
  - 55 million a day

Image completion: how it works

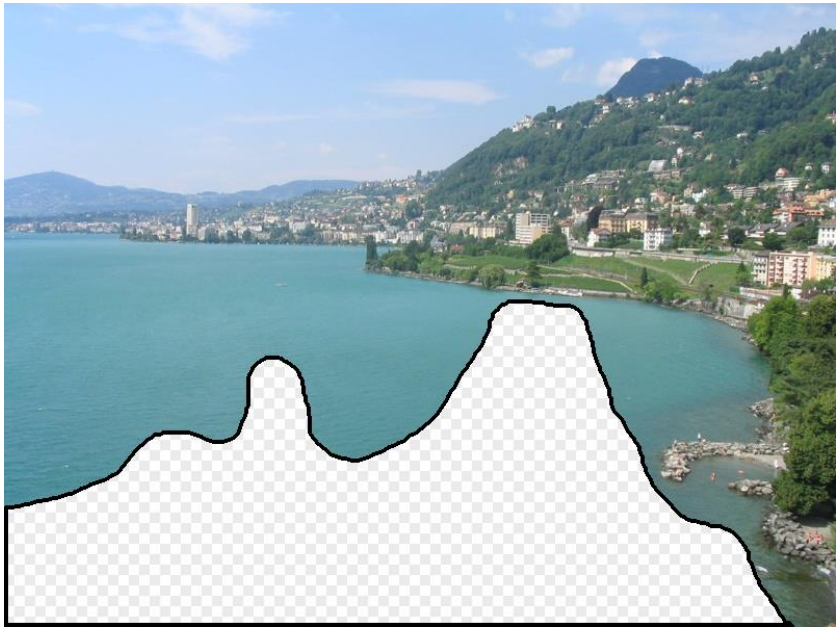
[Hays and Efros. Scene Completion Using Millions of Photographs. SIGGRAPH 2007 and CACM October 2008.]

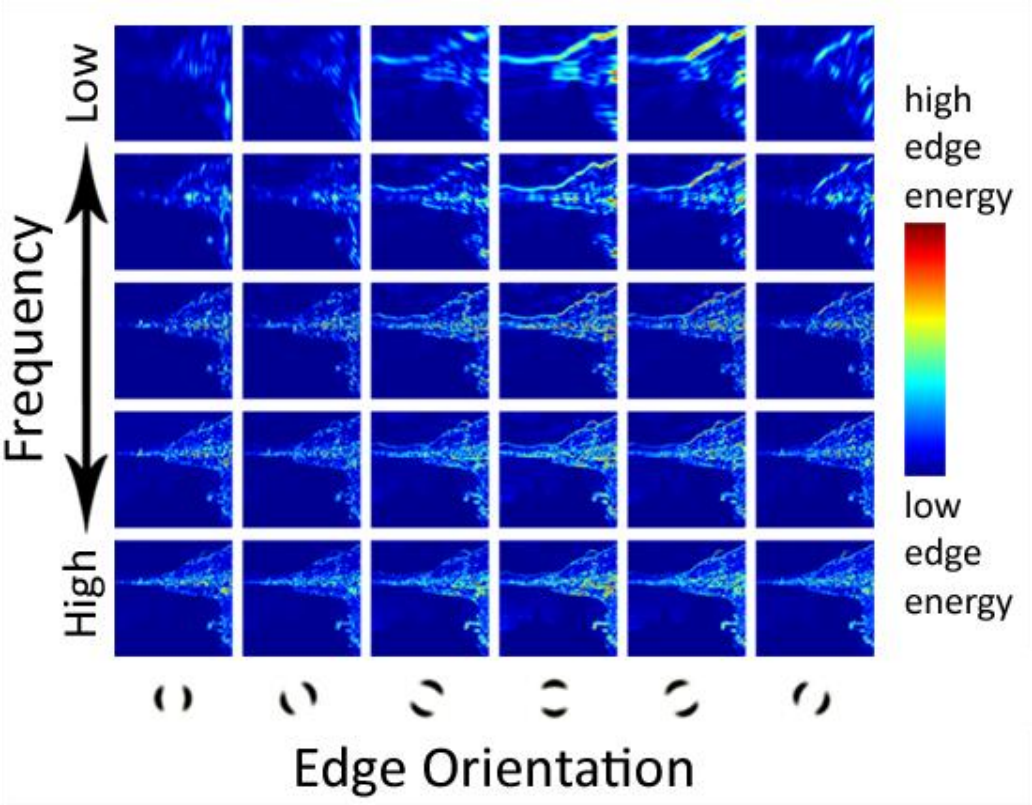
The Algorithm

Scene Matching

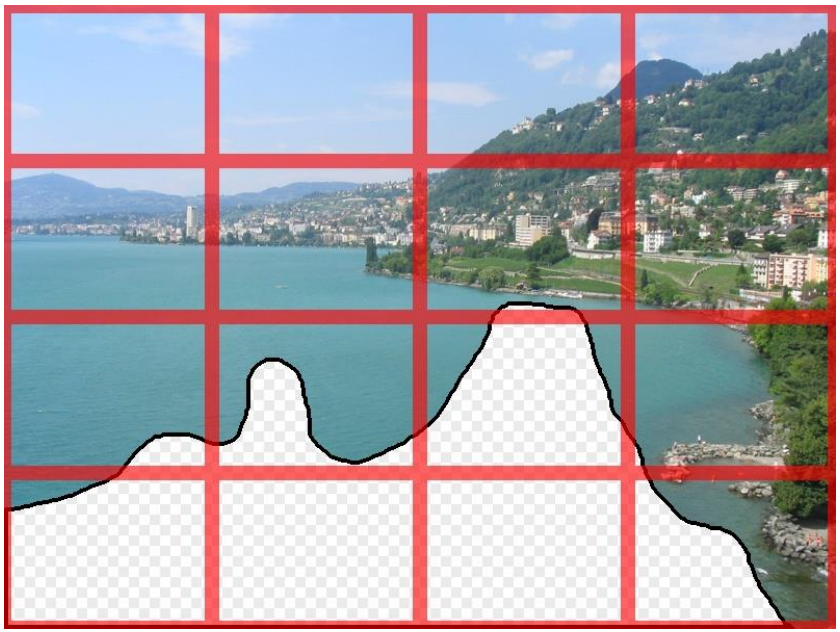


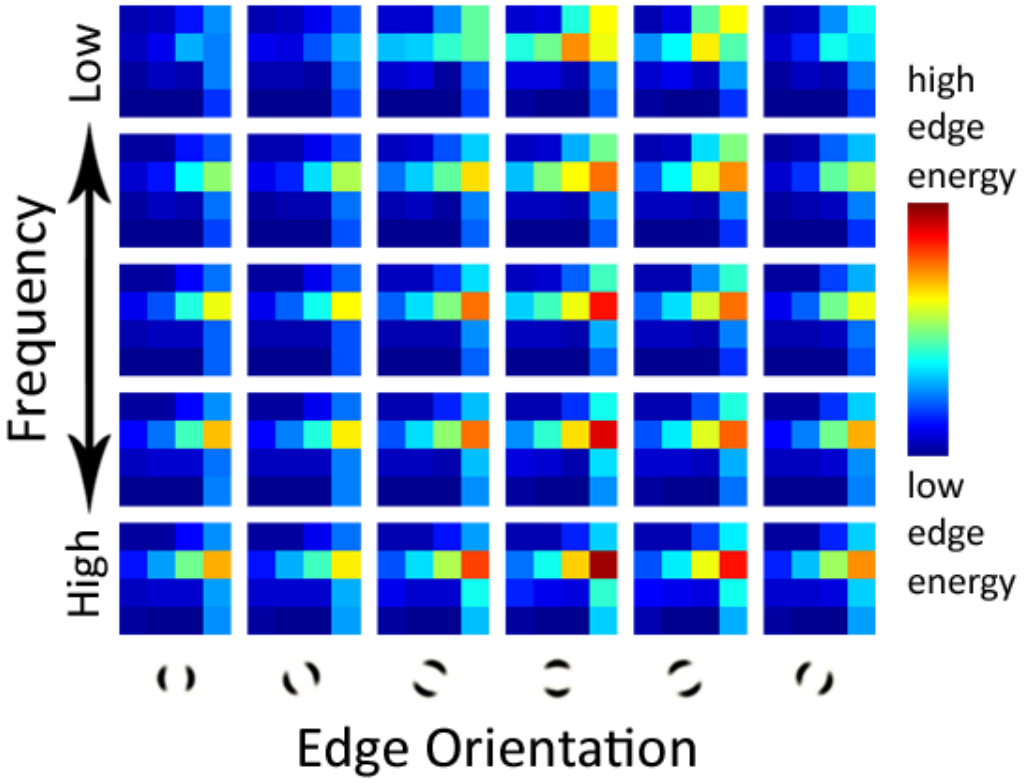
Scene Descriptor





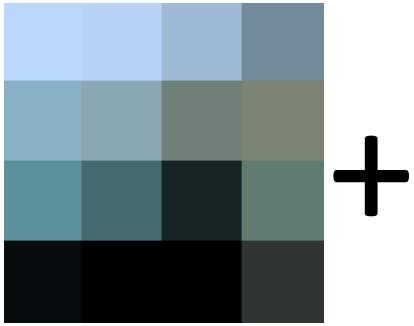
Scene Descriptor

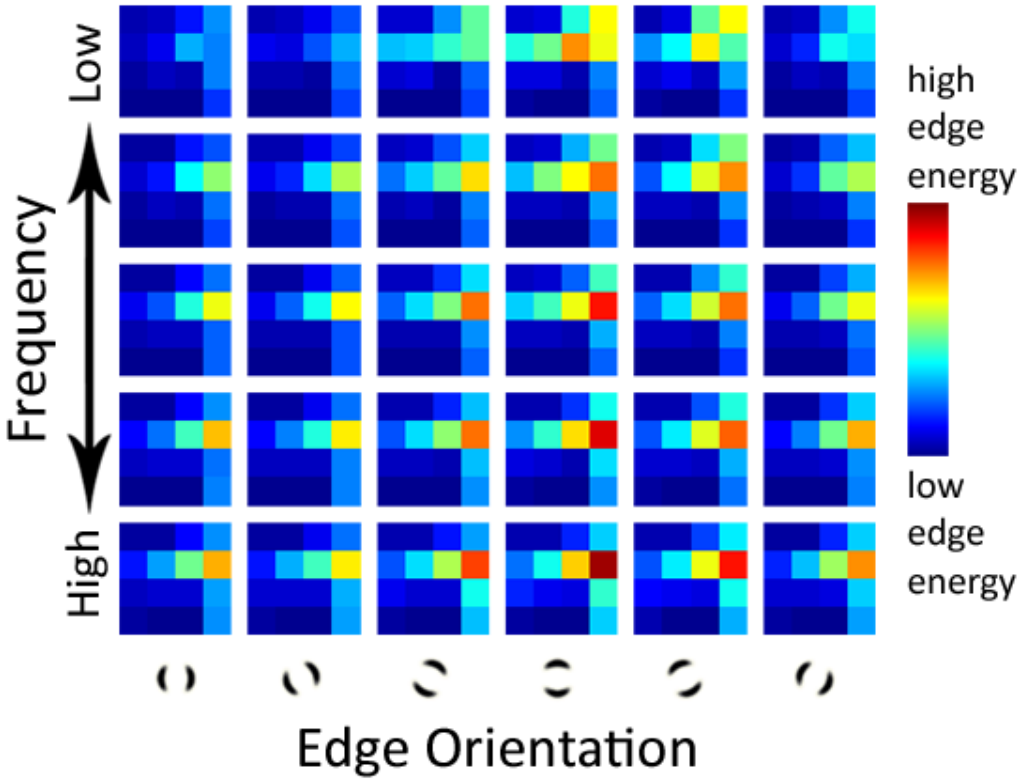




Scene Gist Descriptor
(Oliva and Torralba 2001)

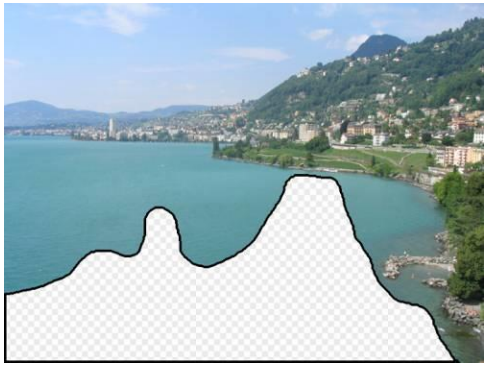
Scene Descriptor

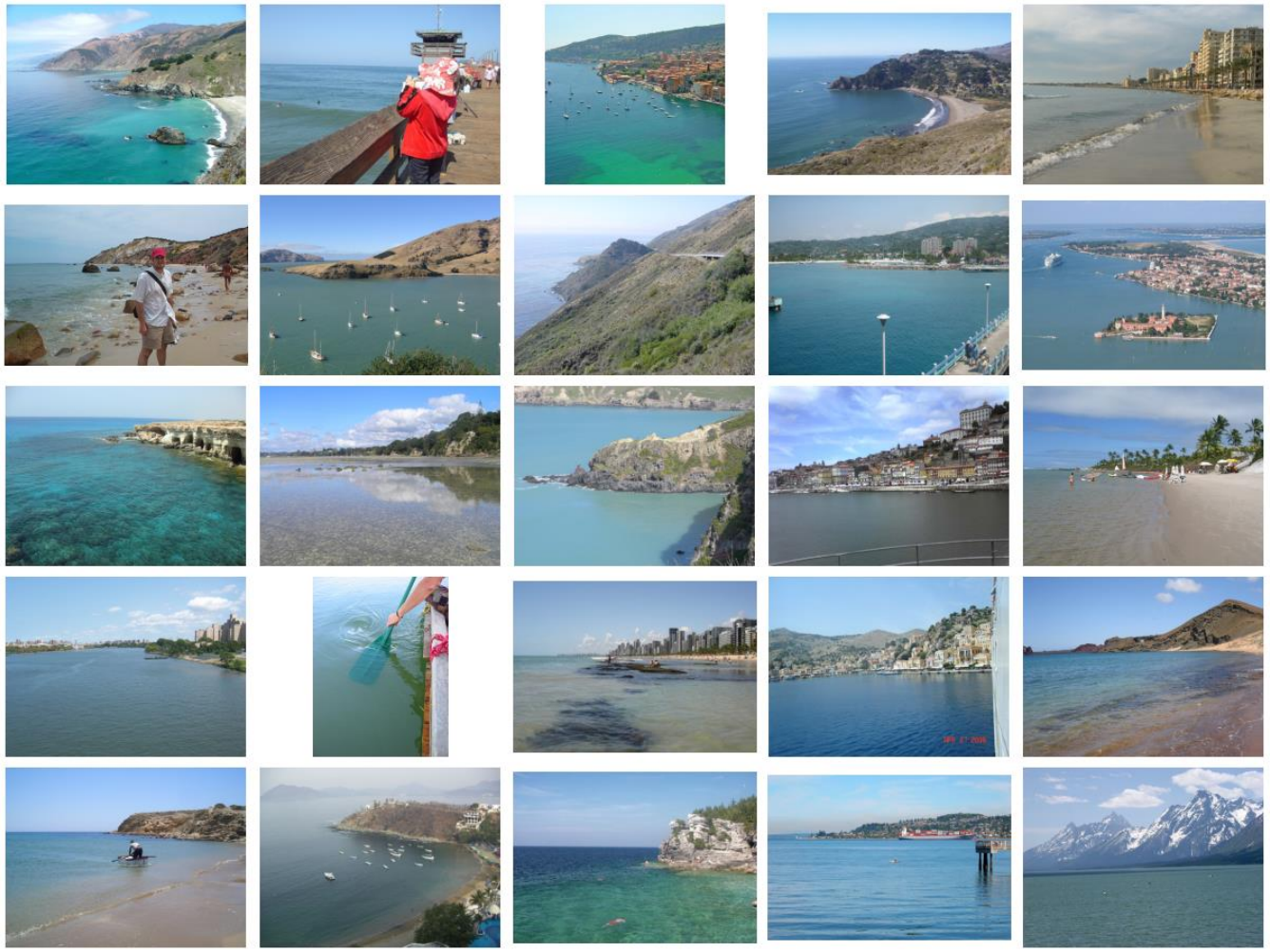




Scene Gist Descriptor
(Oliva and Torralba 2001)

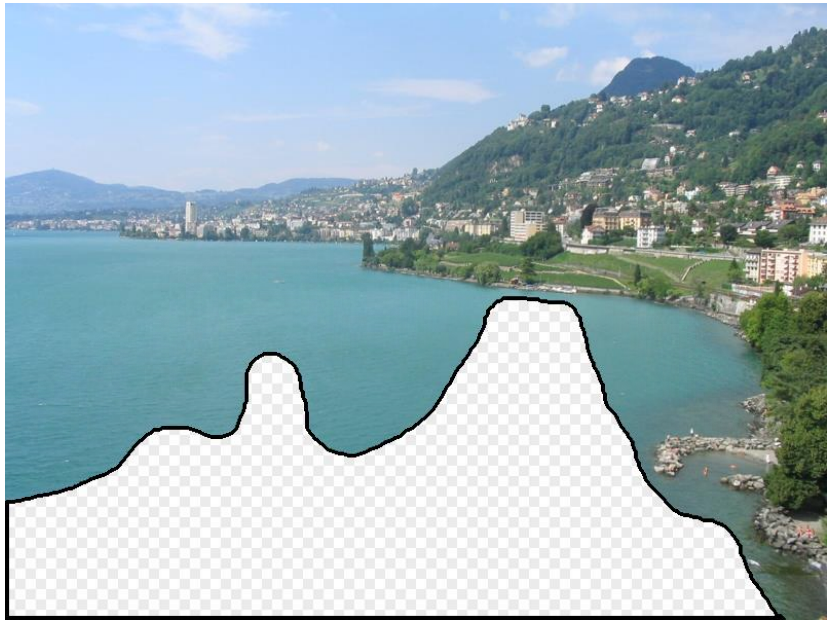

2 Million Flickr Images





… 200 total

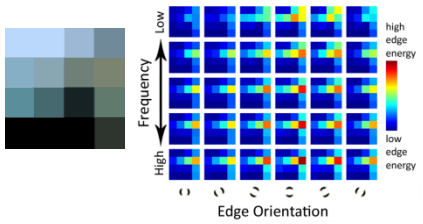
Context Matching




Graph cut + Poisson blending

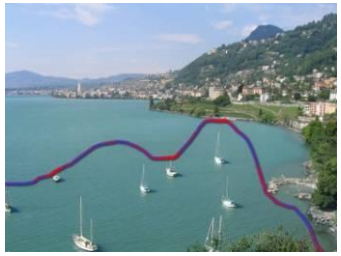
Result Ranking

We assign each of the 200 results a score which is the sum of:

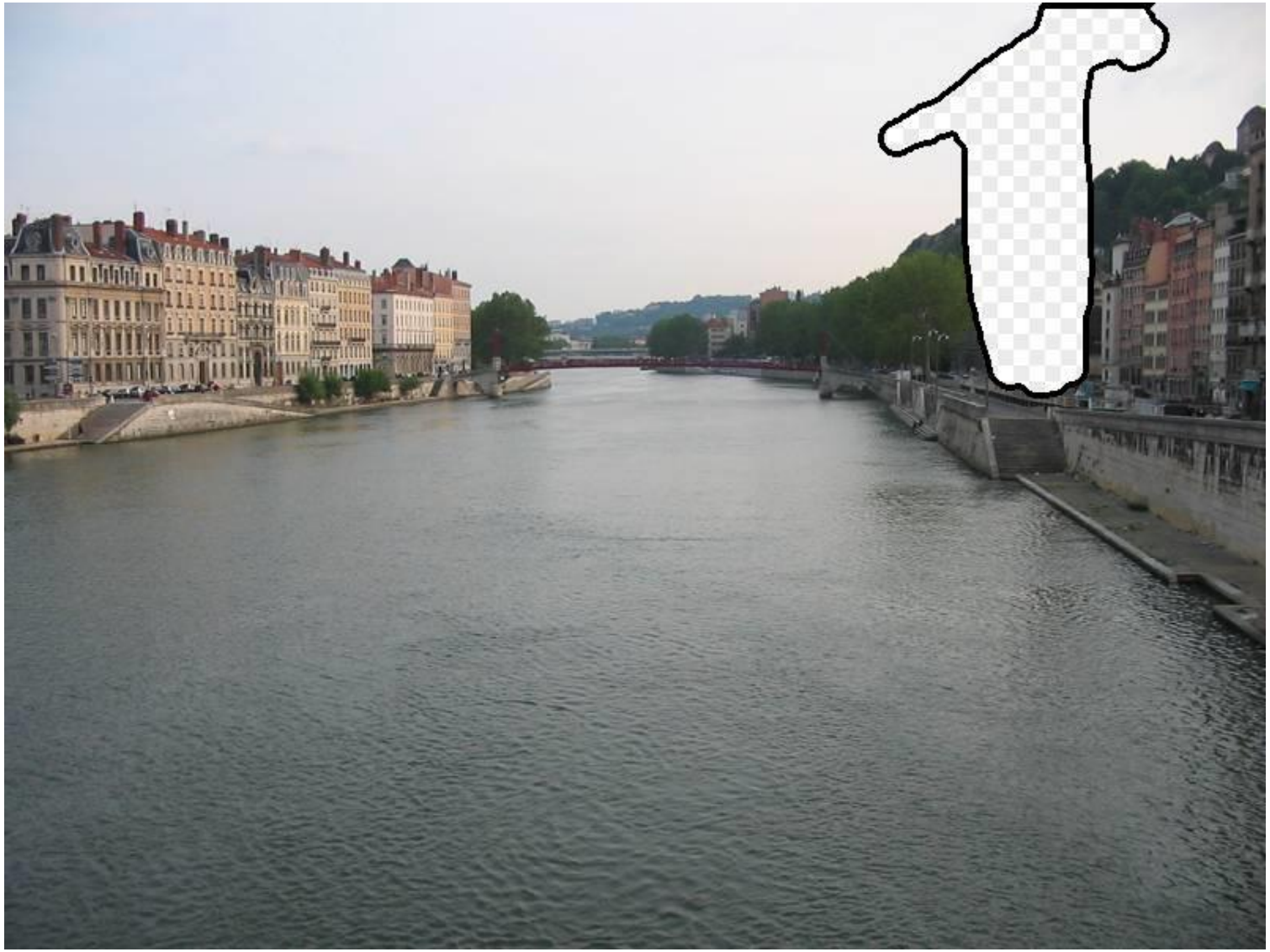


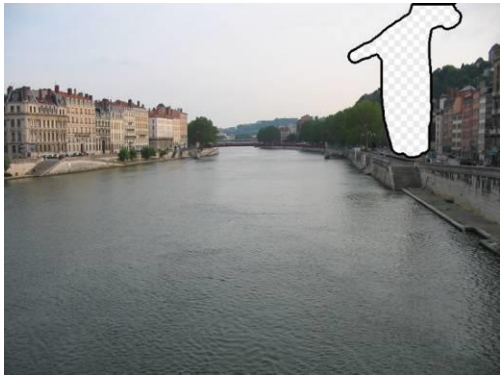
The scene matching distance

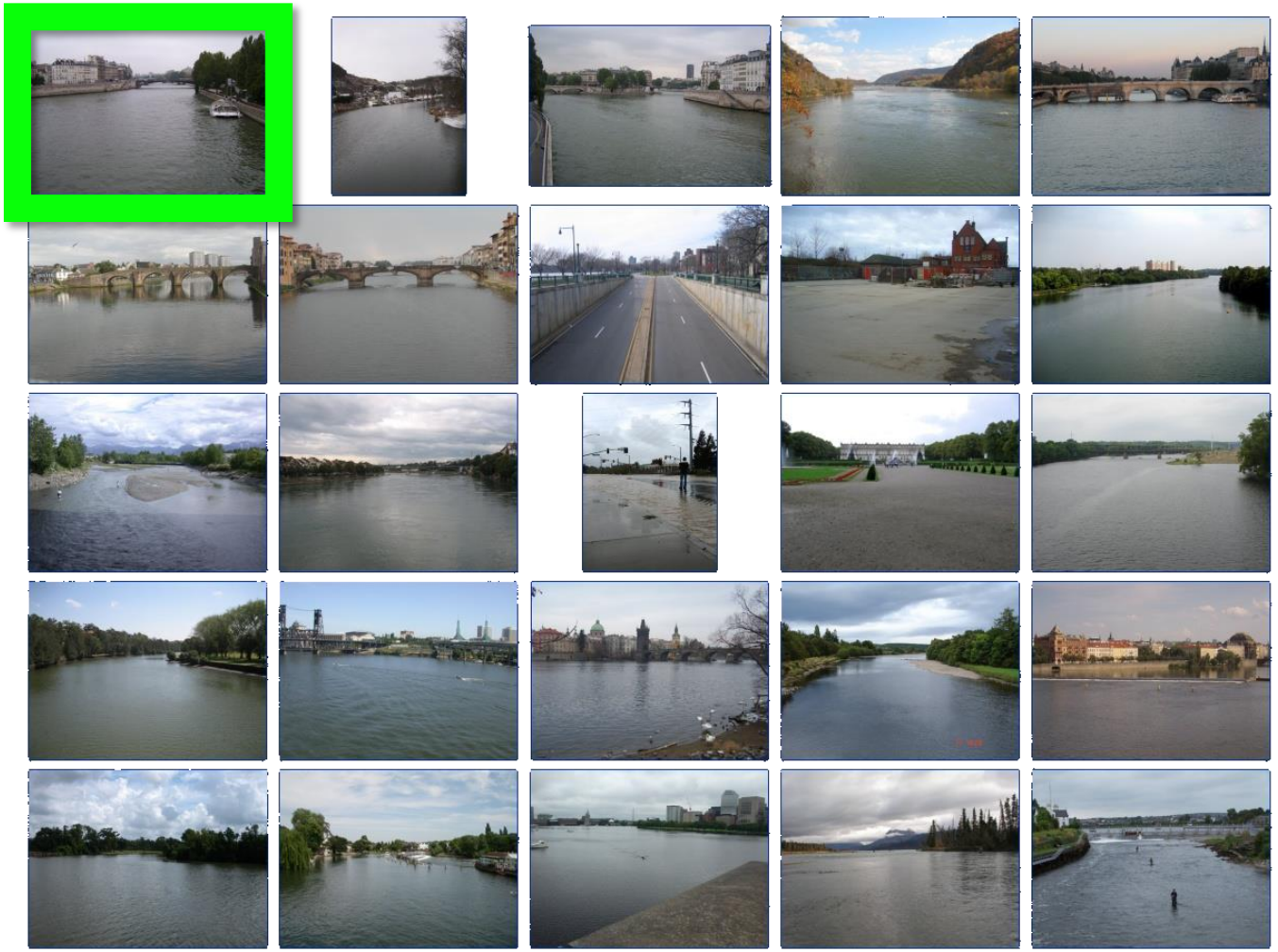
The context matching distance (color + texture)



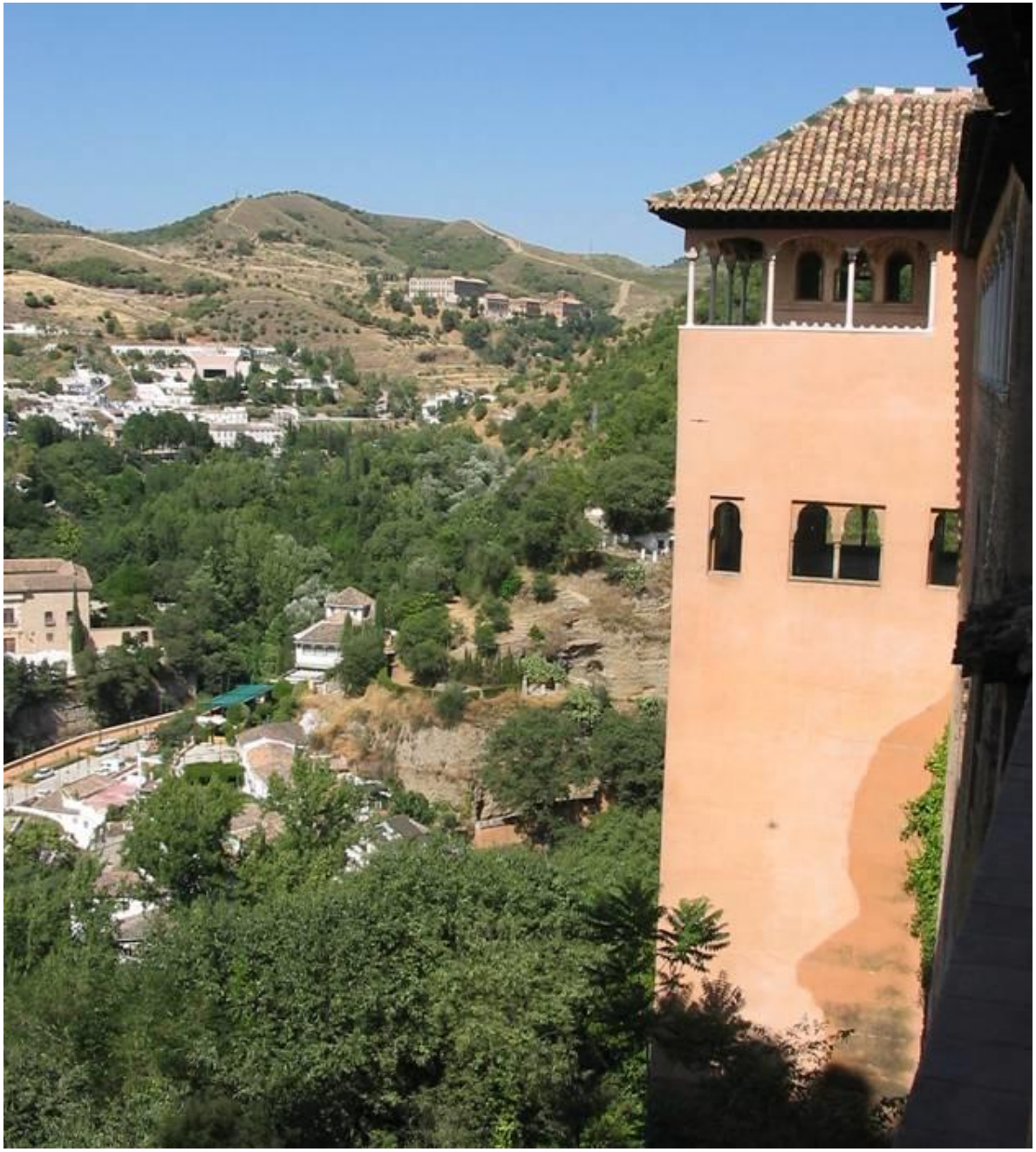
The graph cut cost

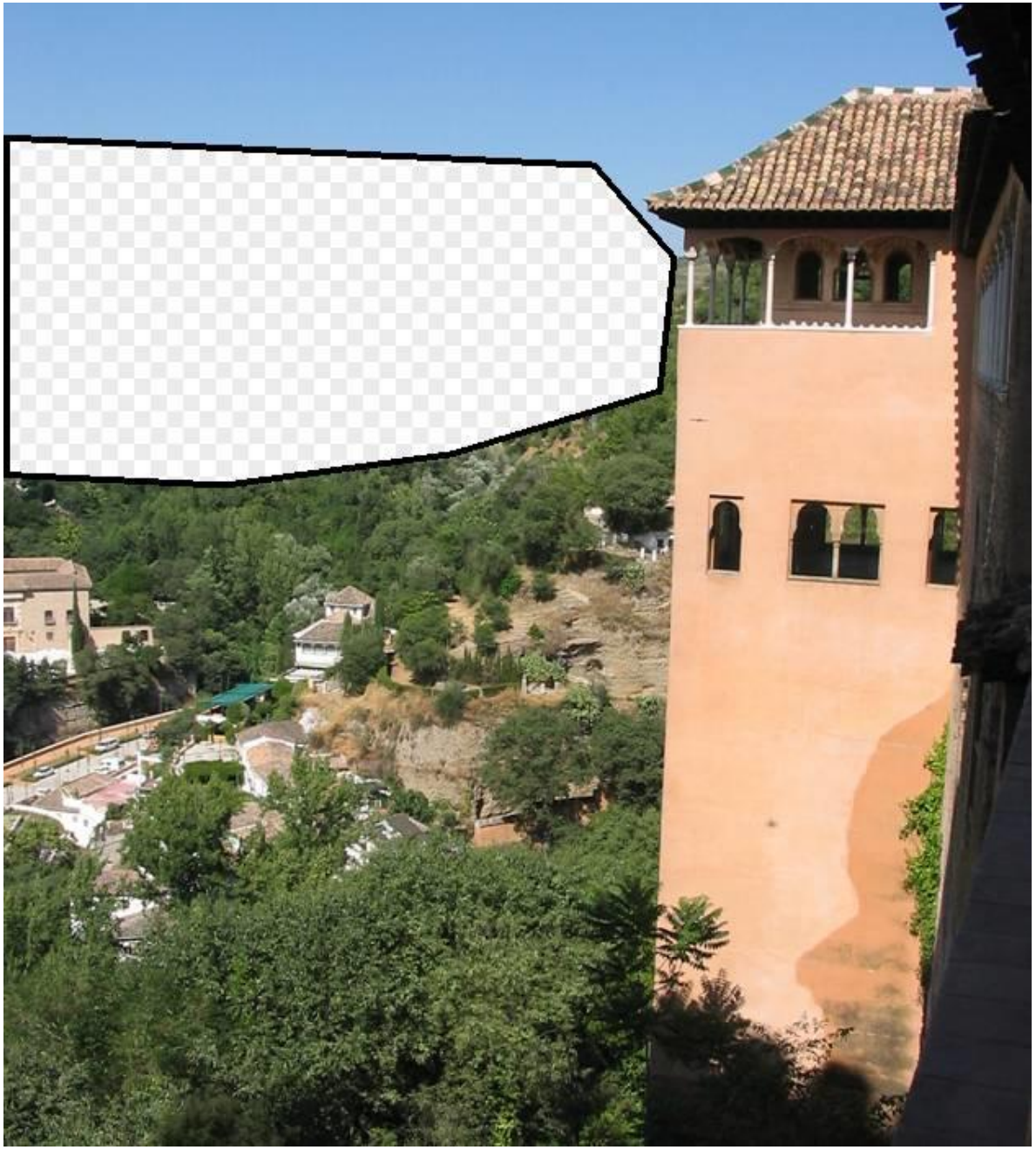






… 200 scene matches





Which is the original?

im2gps (Hays & Efros, CVPR 2008)

6 million geo-tagged Flickr images

http://graphics.cs.cmu.edu/projects/im2gps/

How much can an image tell about its geographic location?


Paris


Paris


Paris


Paris


Paris


Paris


Paris


Madrid


Rome


Paris


Cuba


Paris

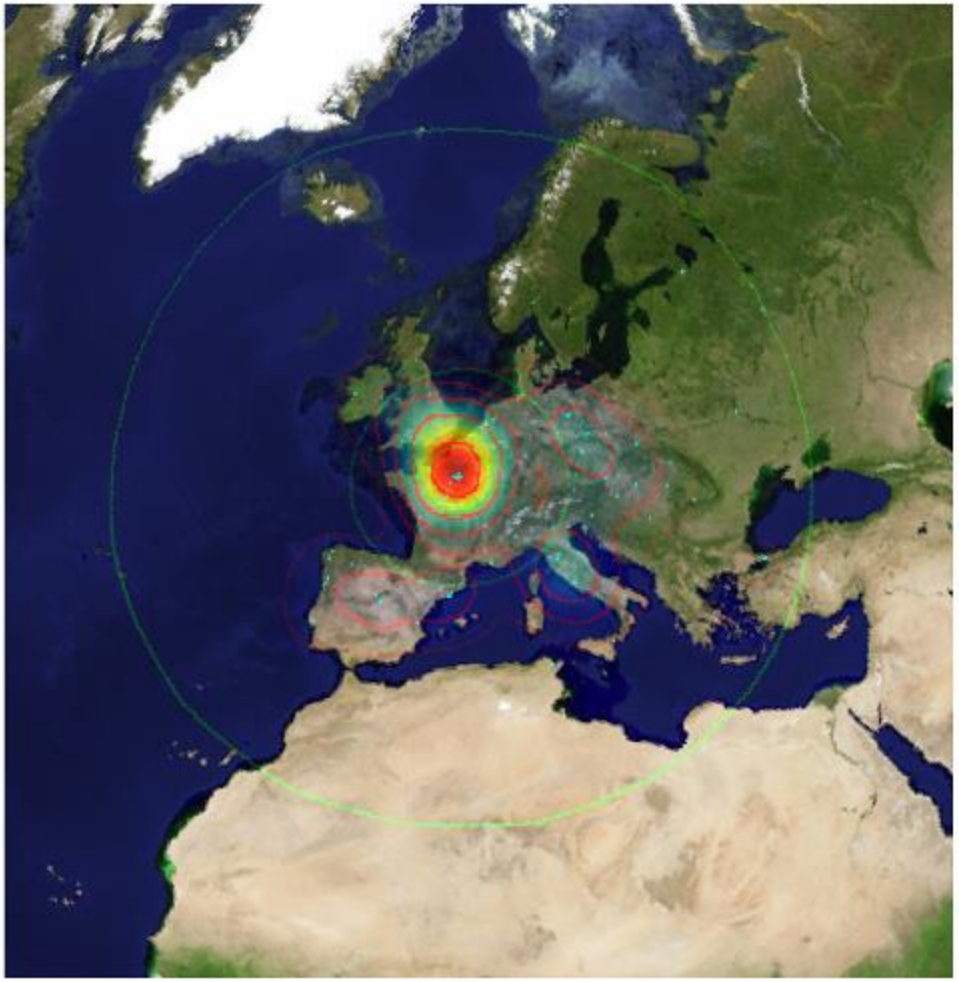

Paris


Poland


Paris


Paris

Nearest Neighbors according to gist + bag of SIFT + color histogram + a few others



Im2gps

Example Scene Matches


Madrid


england


France


Paris


Croatia


heidelberg


Macau


Malta


Cairo


Italy


Italy


Italy

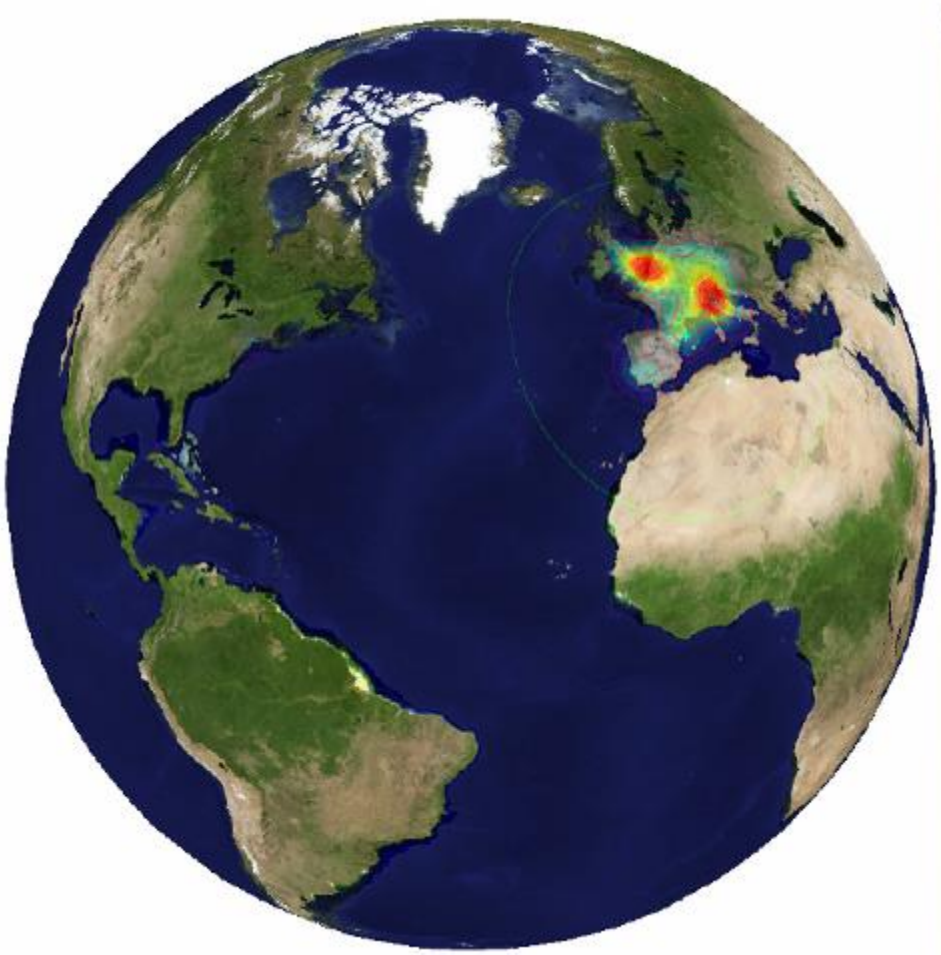

Latvia

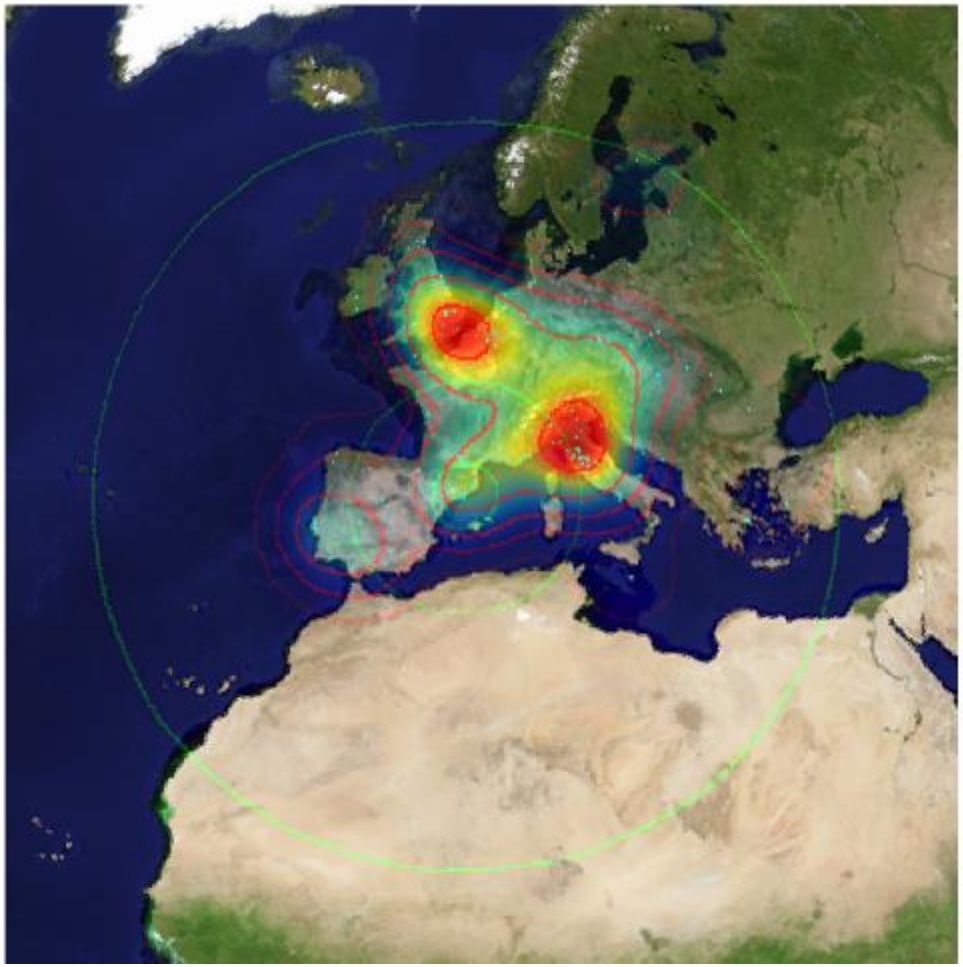

europe


Barcelona


Austria

Voting Scheme





im2gps

Philippines

Houston

Thailand

Houston

Maldives

Philippines

NewZealand

Bermuda

Palau

Mexico2

Brazil

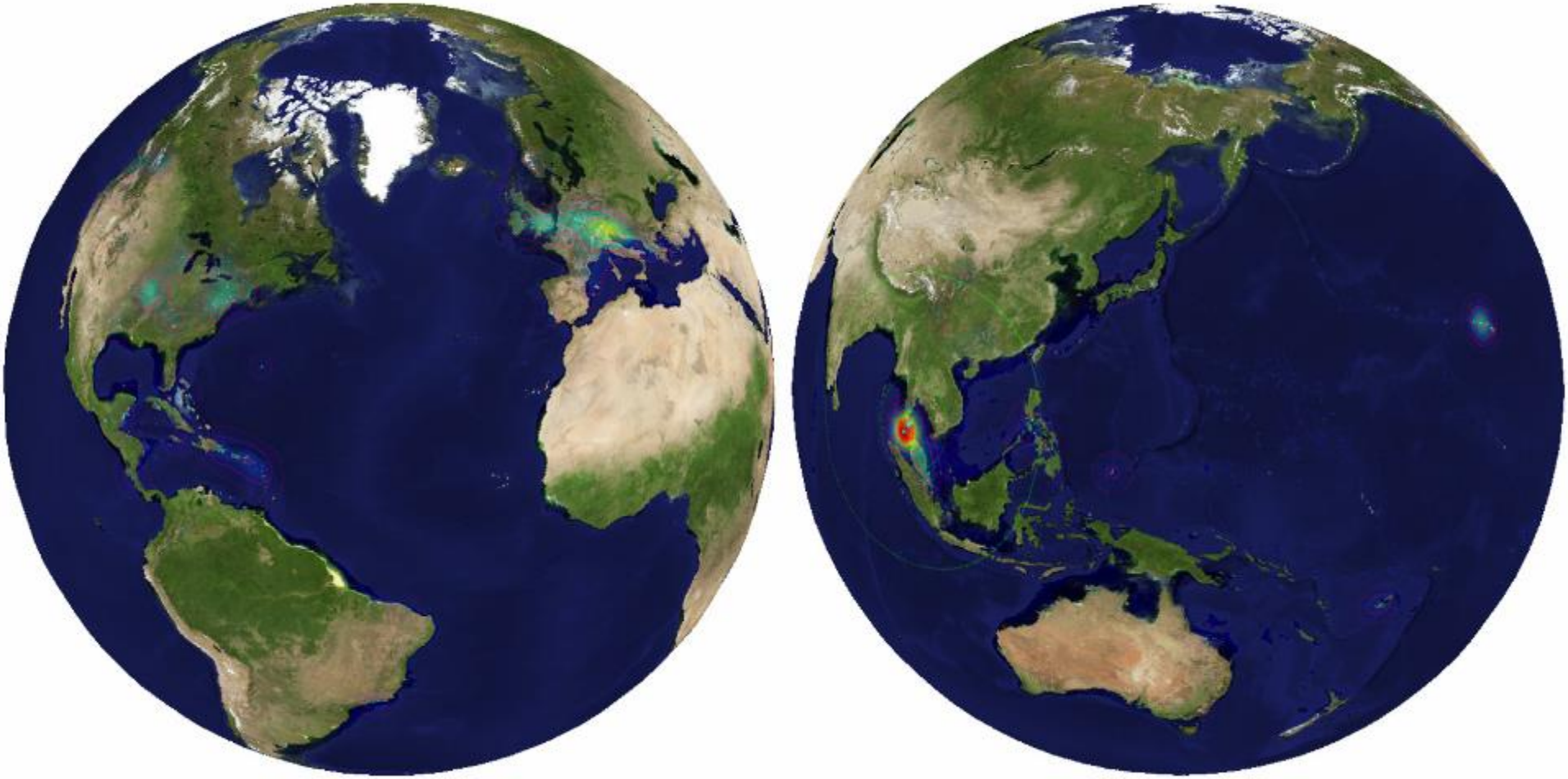
Mendoza

Brazil

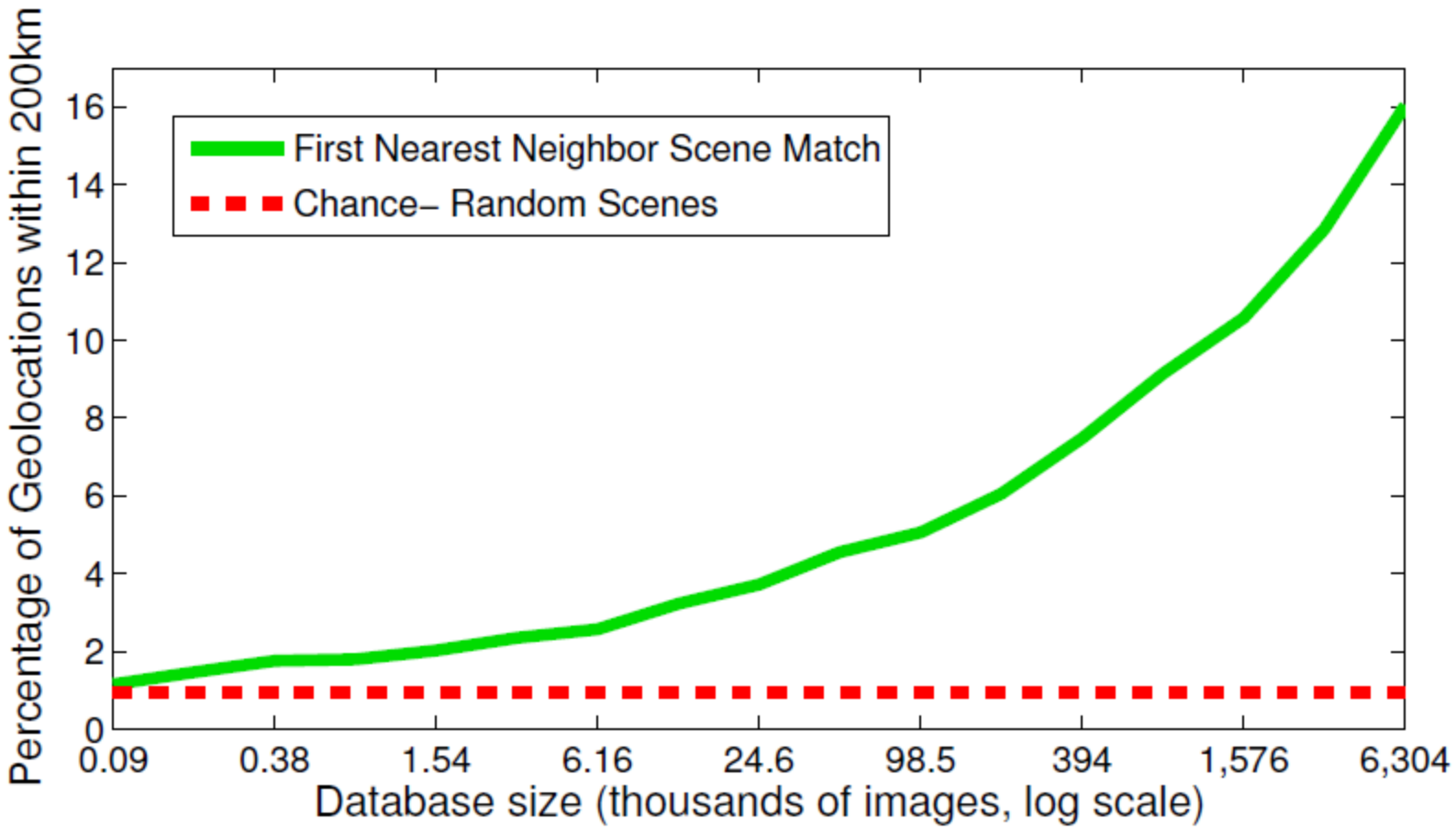
Thailand

Arkansas

Hawaii



Effect of Dataset Size



Population density ranking

High Predicted Density

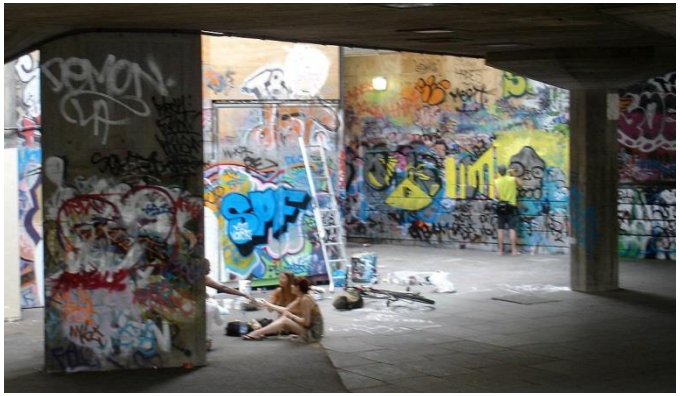
Low Predicted Density

Where is This?

[Olga Vesselova, Vangelis Kalogerakis, Aaron Hertzmann, James Hays, Alexei A. Efros. Image Sequence Geolocation. ICCV'09]

Where is This?

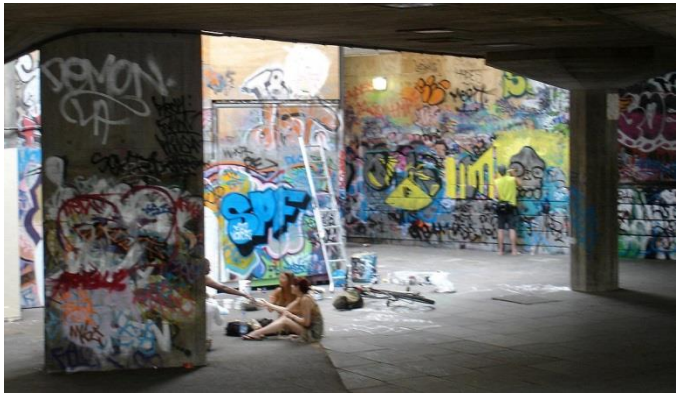
Where are These?



15:14,
June 18th, 2006

16:31,
June 18th, 2006

Where are These?



15:14,
June 18th, 2006

16:31,
June 18th, 2006

17:24,
June 19th, 2006

Results

- im2gps – 10% (geo-loc within 400 km)
- temporal im2gps – 56%