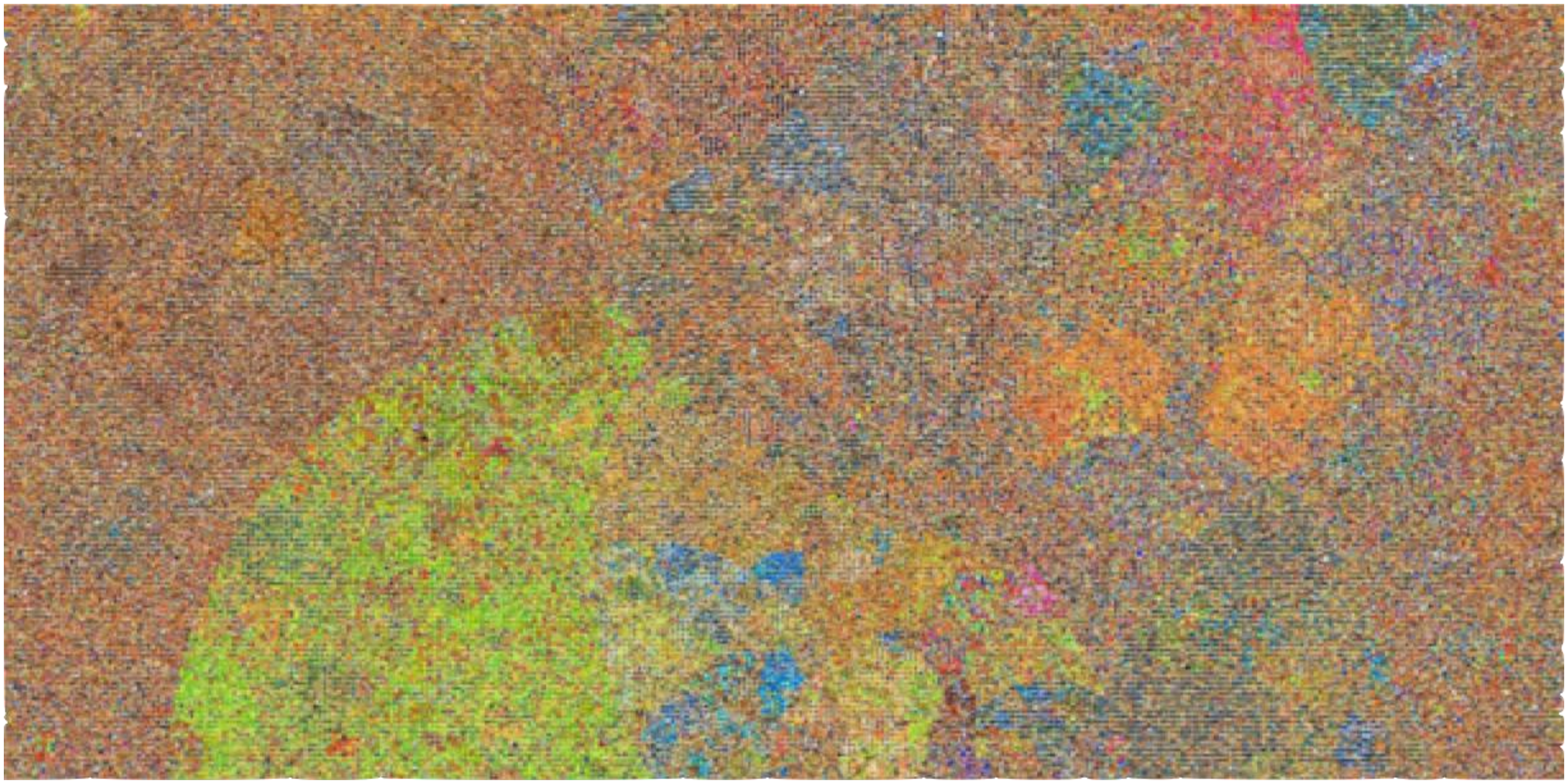
Visual Data on the Internet - Part 2

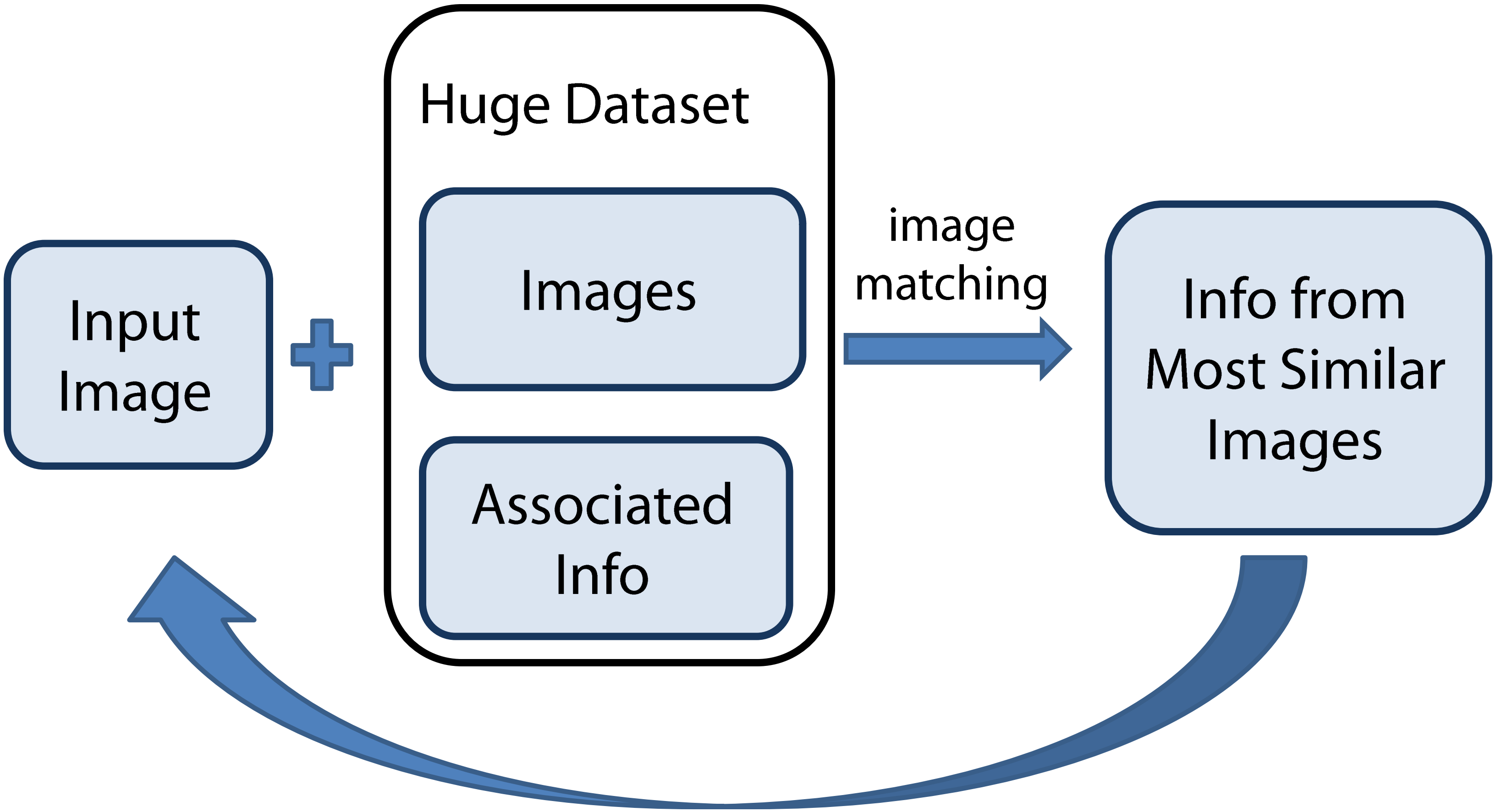


Visualization of 53,464 english nouns, credit: A. Torralba,
http://groups.csail.mit.edu/vision/TinyImages/

CS 129: Computational Photography
James Hays, Brown, Fall 2012

With slides from Alexei Efros, Jean Francois Lalonde,
Derek Hoiem, and Antonio Torralba

Recap: Using lots of data!

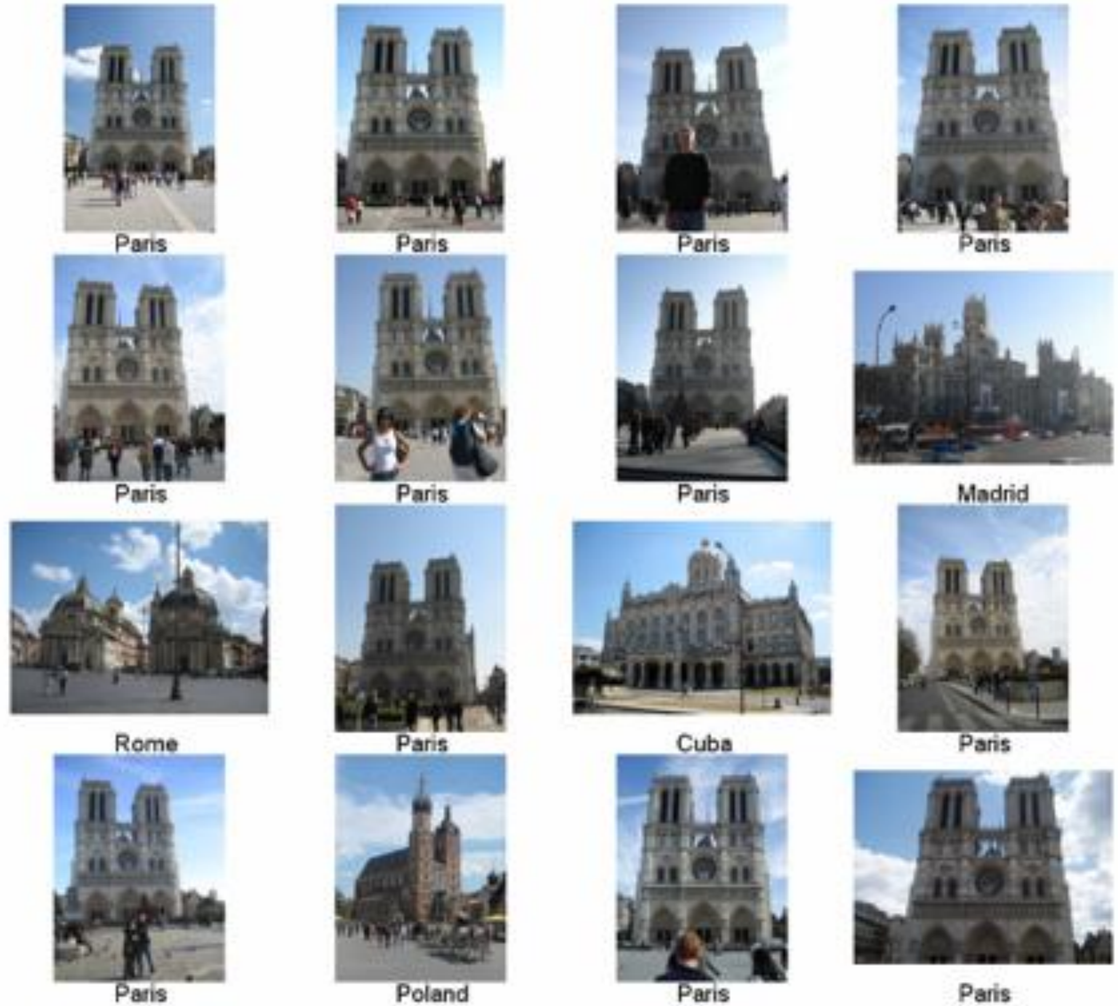


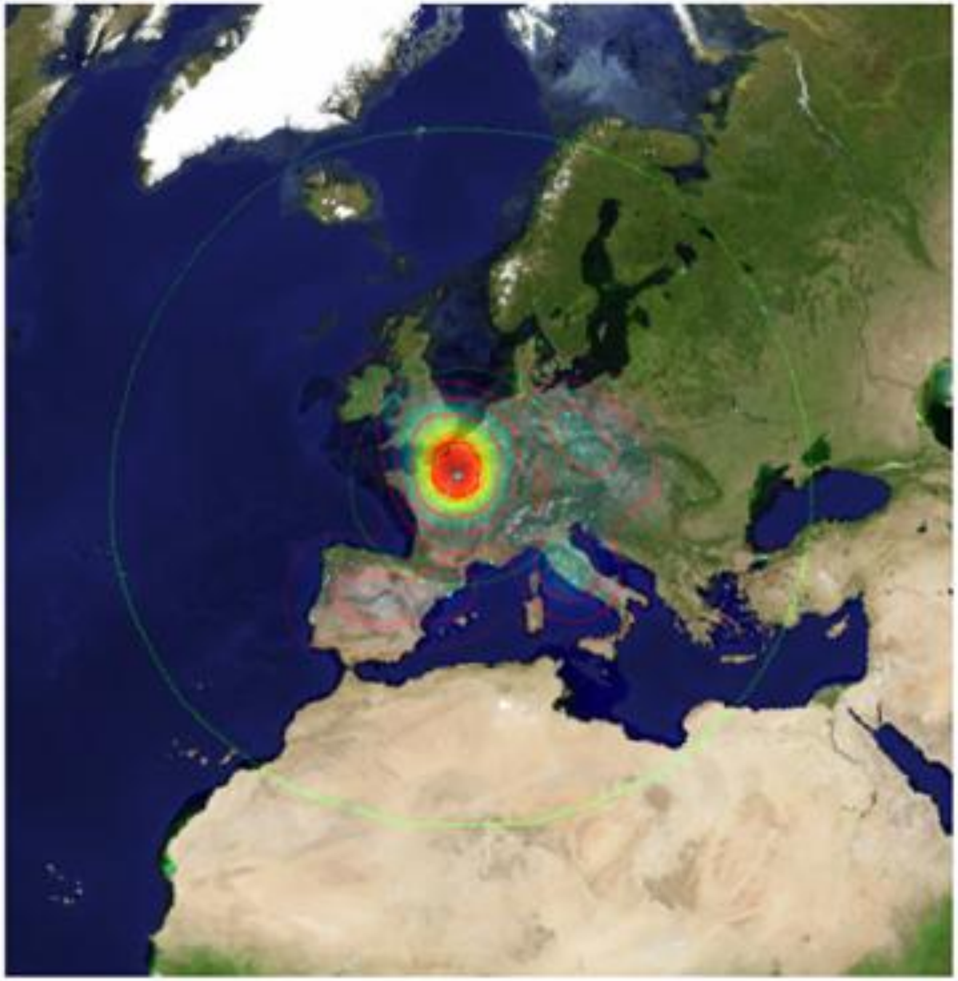
Trick: If you have enough images, the dataset will contain very similar images that you can find with simple matching methods.

im2gps (Hays & Efros, CVPR 2008)

6 million geo-tagged Flickr images

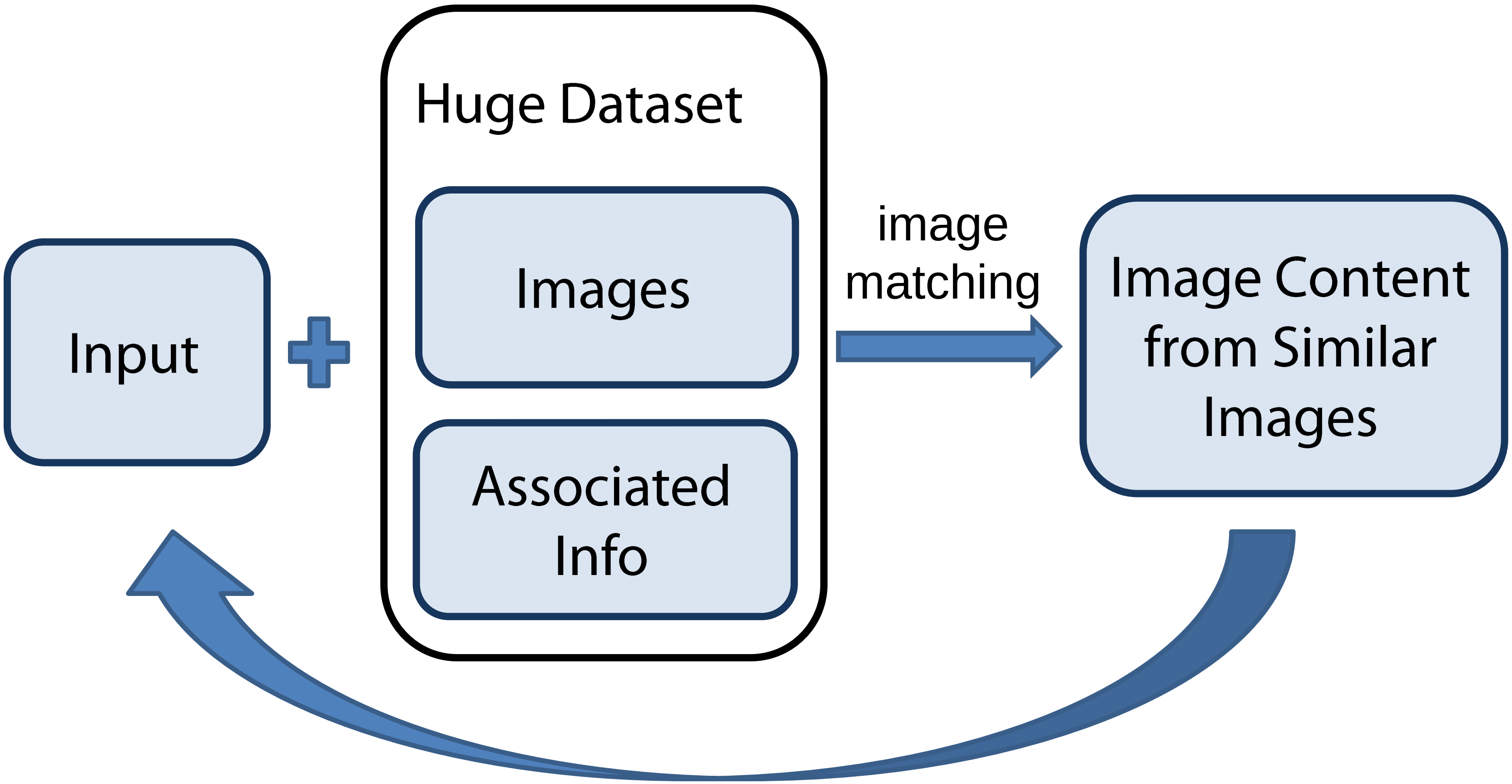
How much can an image tell about its geographic location?


Paris
Paris
Paris
Paris
Paris
Paris
Paris
Madrid
Rome
Paris
Cuba
Paris
Paris
Poland
Paris
Paris

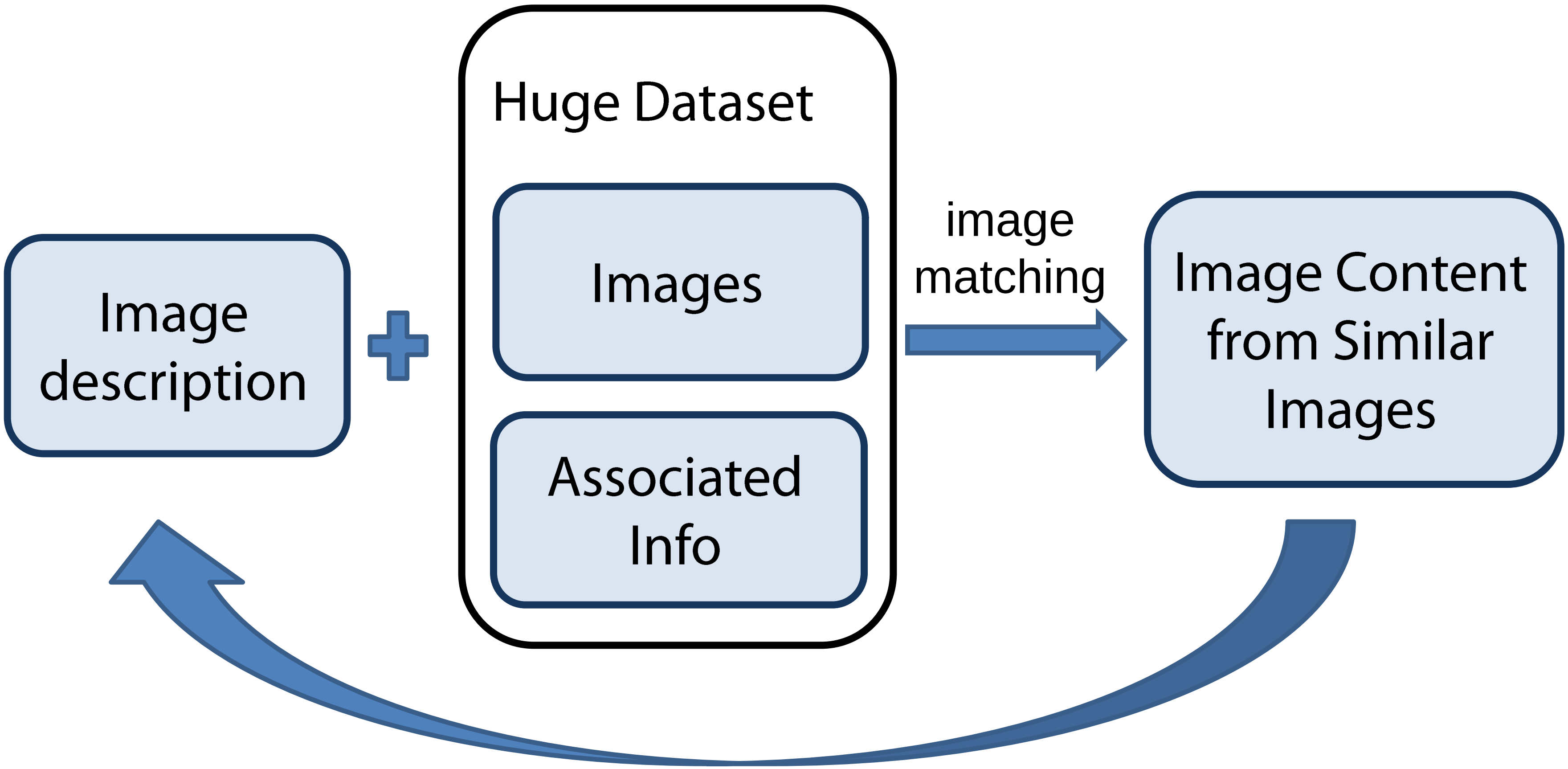



Today

Using lots of data to create new images

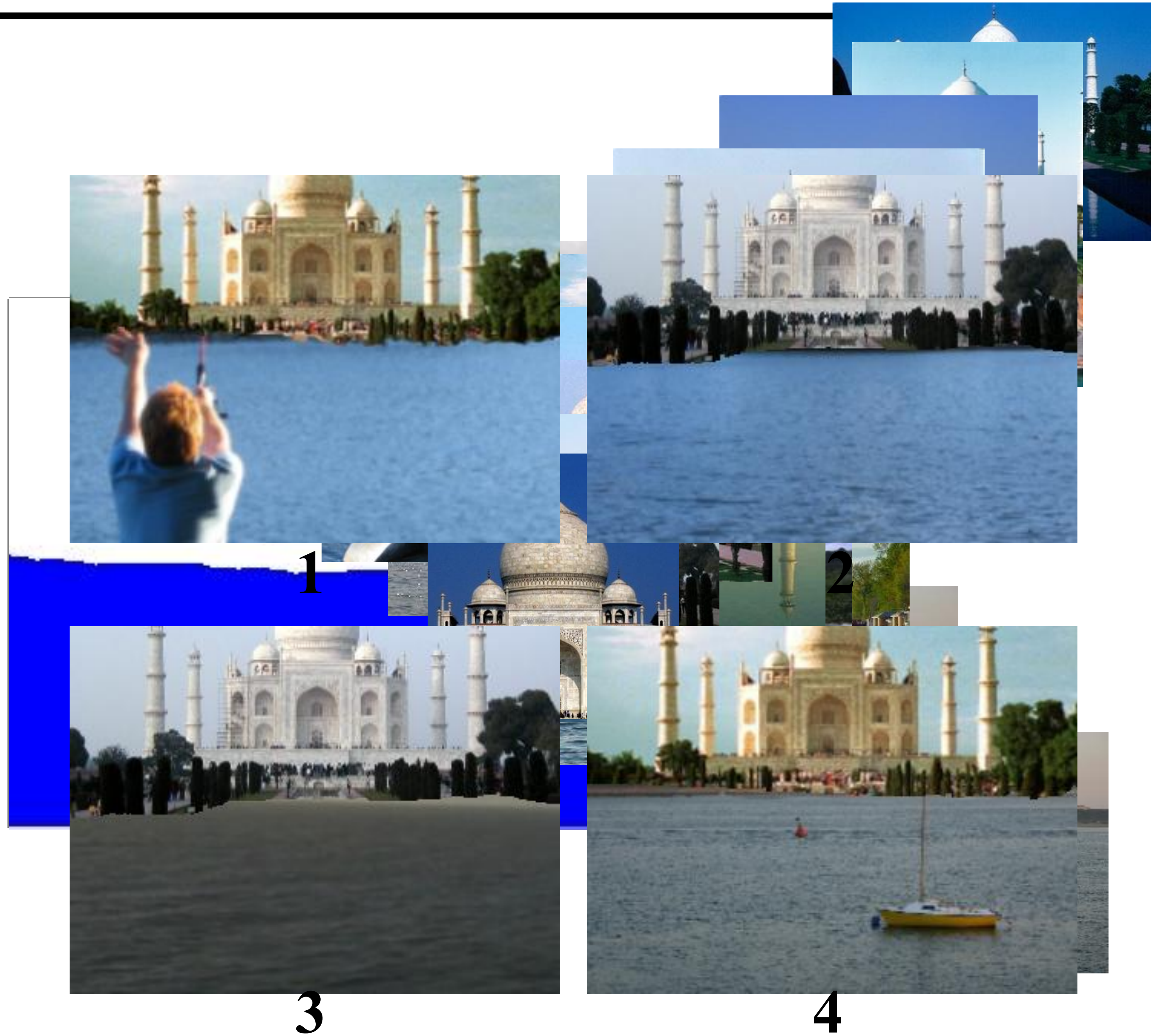


Semantic Photo Synthesis

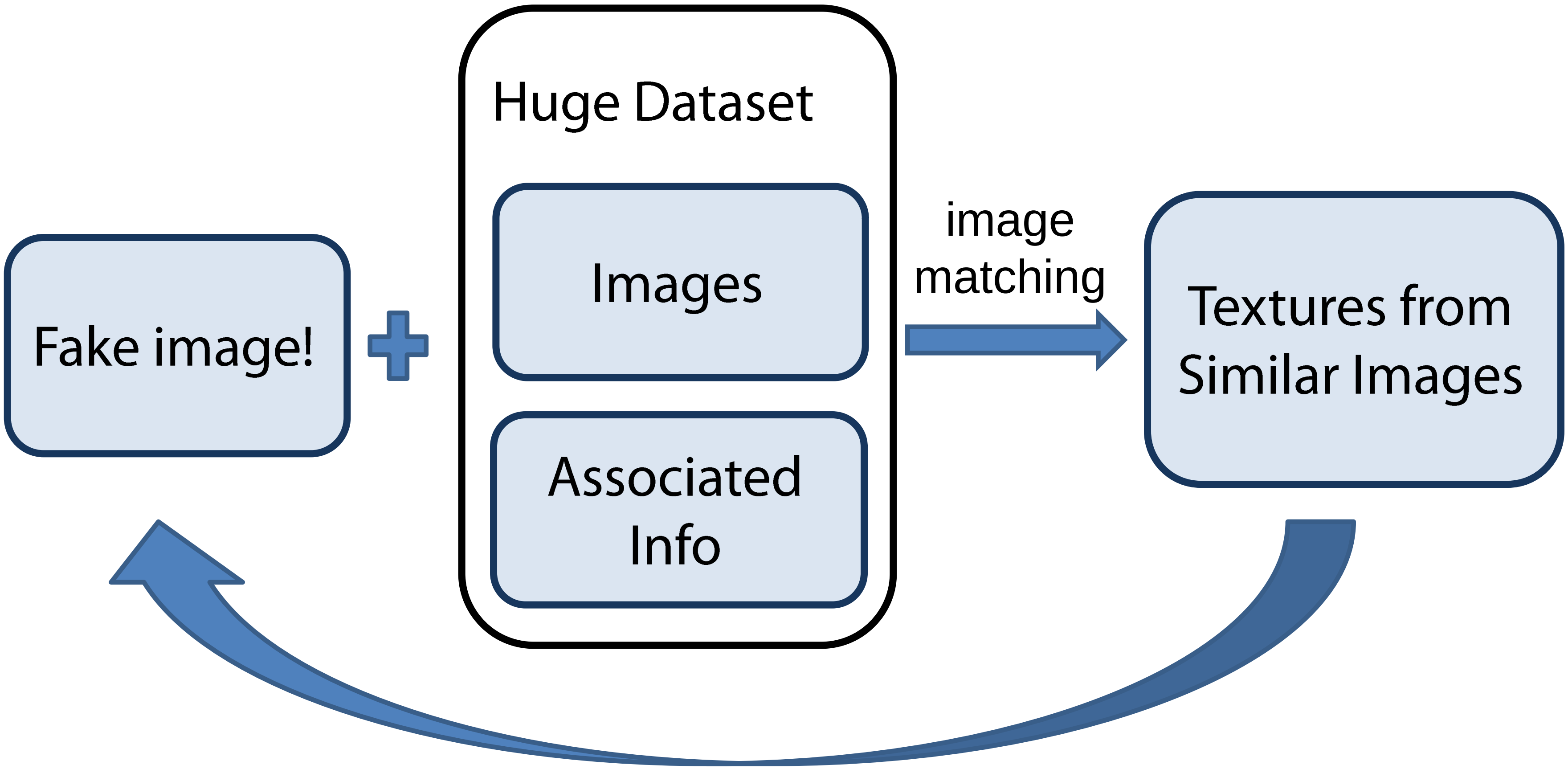


M. Johnson, G. Brostow, J. Shotton, O. A. c, and R. Cipolla, “Semantic Photo Synthesis,” Computer Graphics Forum Journal (Eurographics 2006), vol. 25, no. 3, 2006.

Semantic Photo Synthesis

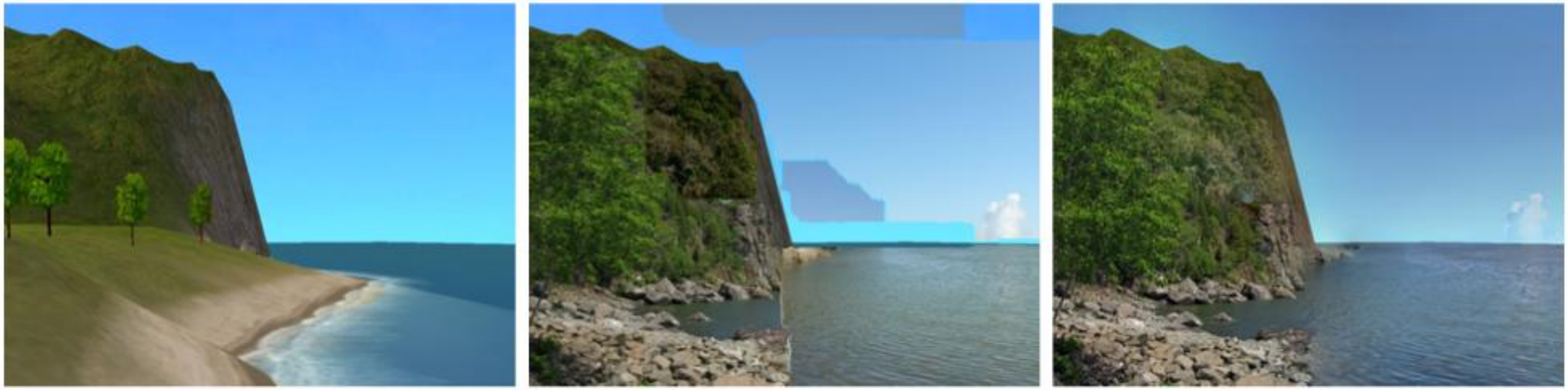


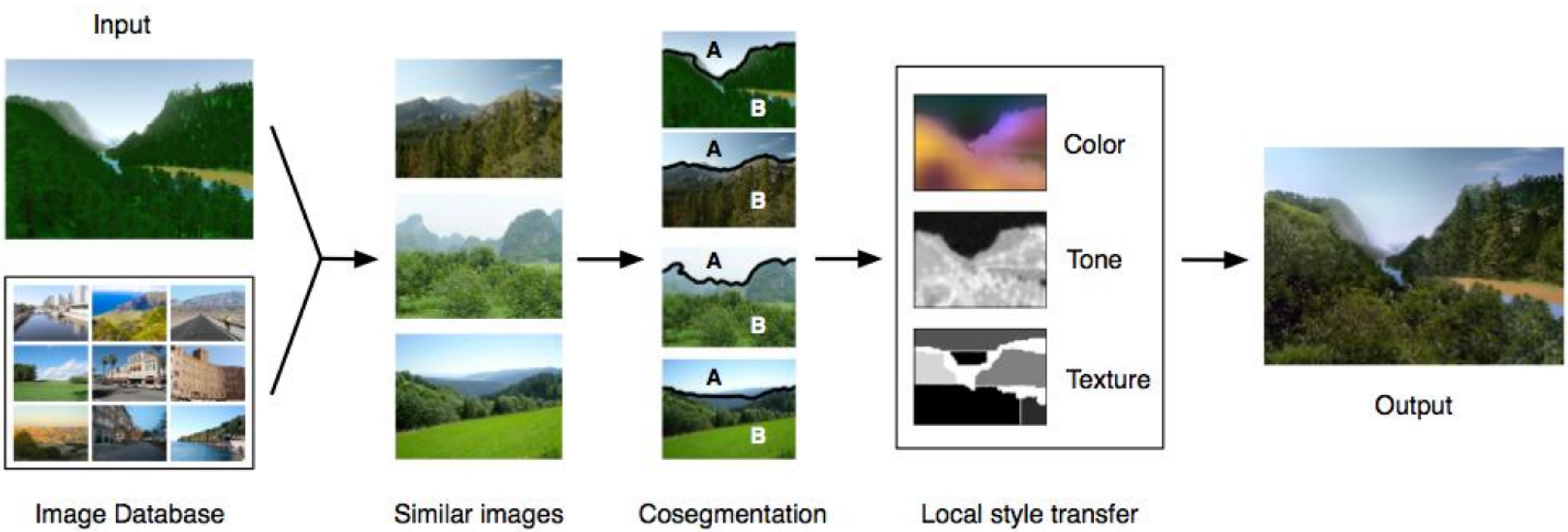
CG2Real



M. K. Johnson, K. Dale, S. Avidan, H. Pfister, W. T. Freeman, and W. Matusik, "CG2Real: Improving the realism of computer generated images using a large collection of photographs," IEEE Transactions on Visualization and Computer Graphics, 2010.

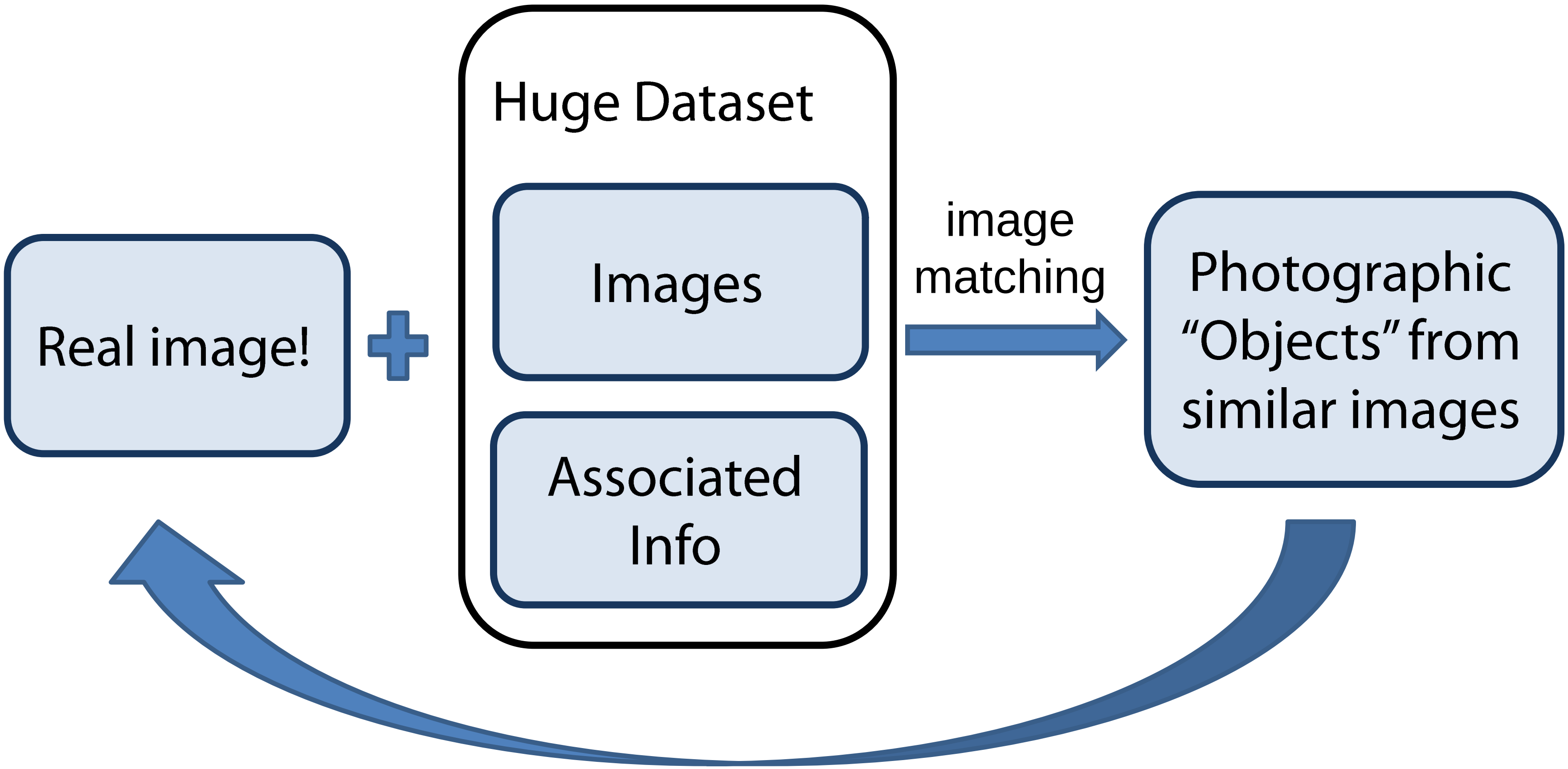
CG2Real





M. K. Johnson, K. Dale, S. Avidan, H. Pfister, W. T. Freeman, and W. Matusik, "CG2Real: Improving the realism of computer generated images using a large collection of photographs," IEEE Transactions on Visualization and Computer Graphics, 2010.

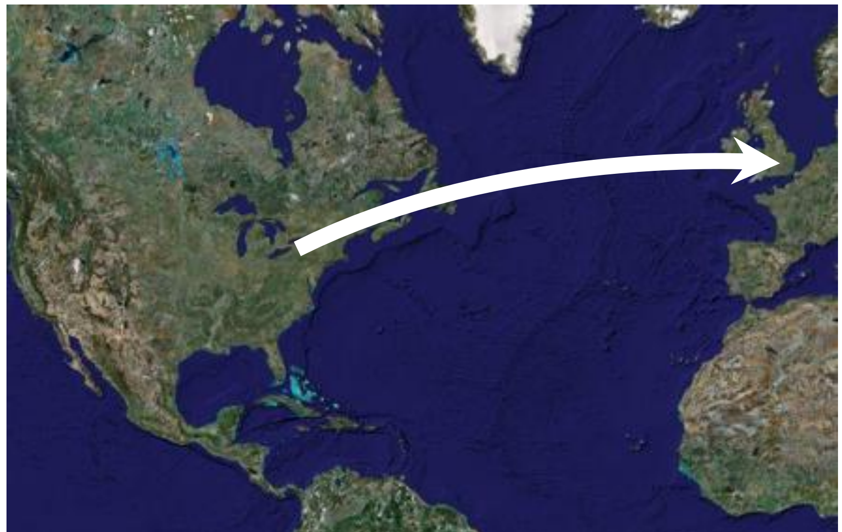
Photo Clip Art



J.-F. Lalonde, D. Hoiem, A. A. Efros, C. Rother, J. Winn, and A. Criminisi, "Photo Clip Art," ACM Transactions on Graphics (SIGGRAPH 2007), vol. 26, no. 3, Aug. 2007.

Inserting objects into images

Inserting objects into images



[Debevec, '98]

Inserting objects in images

Highly detailed geometry
Highly detailed materials
Very expensive

Realistic renderings

Expensive and impractical

[Debevec, '98]

Alternative: Clip art

Easy, intuitive, cheap

Not realistic

Creating images (2-D)

Photo-realistic — Cartoon

Image-based rendering

Photo Clip Art

?

Clip Art

Expensive and impractical — Cheap and intuitive

“Photoshopping”

Composite by David Dewey

Inserting objects into images

Challenges

Insert THIS object: impossible!

object orientation
scene illumination

The use of data

Insert SOME object: much easier!

The Google model

Database

Query

Results

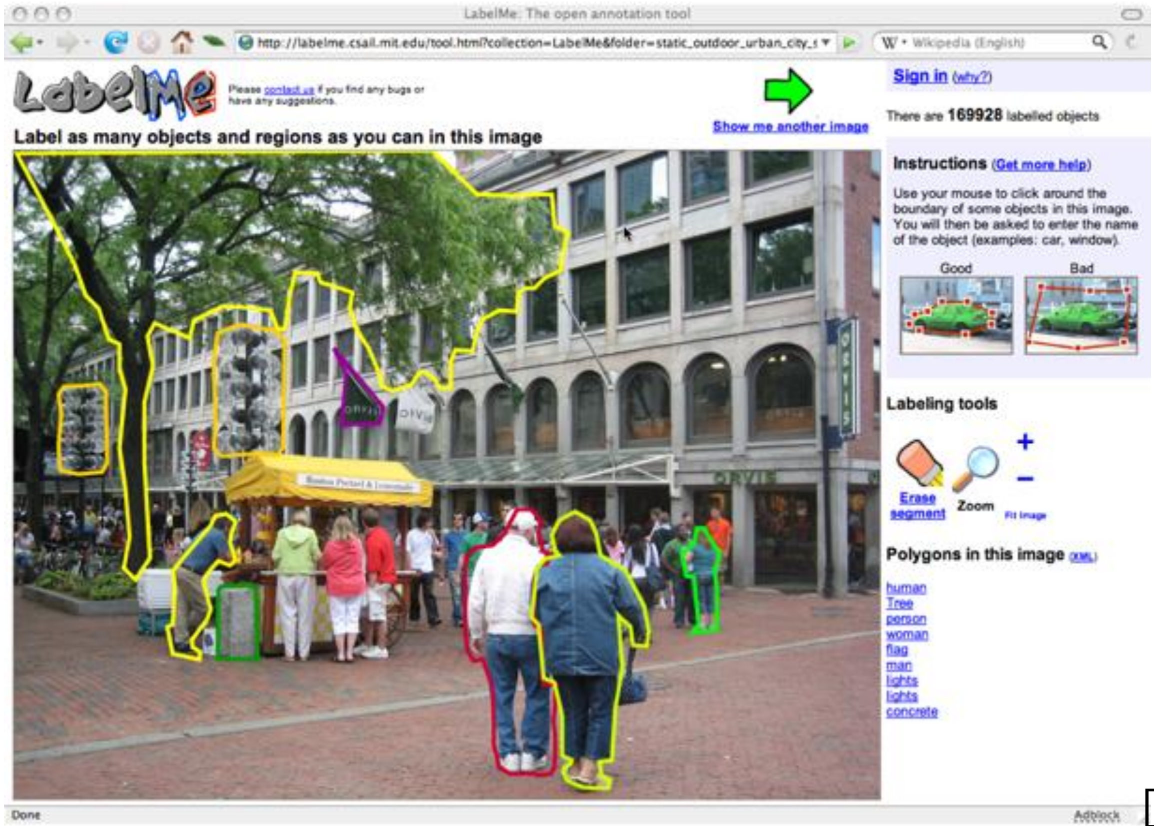
Sort the objects

Data source: LabelMe

Online (http://labelme.csail.mit.edu), user-contributed

170,000 objects in 40,000 images

Polygons and names



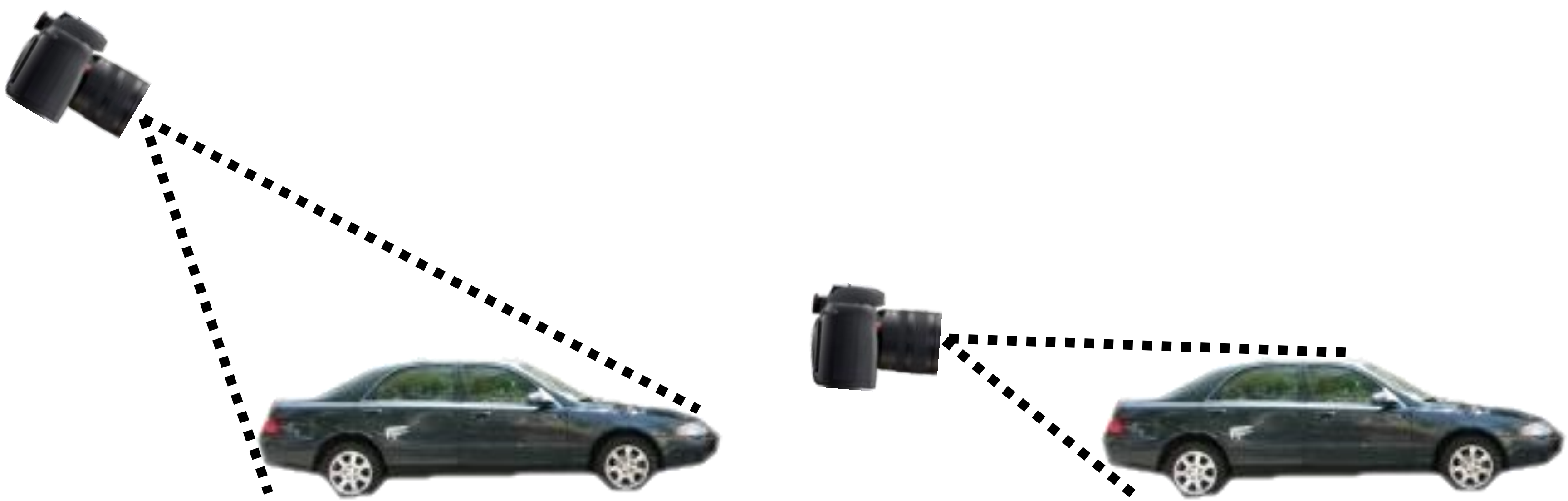
[Russell et al., 2005]

Data organization

Top-level categories (chosen manually, 16 total)

What should we match?

Camera parameters



Assume

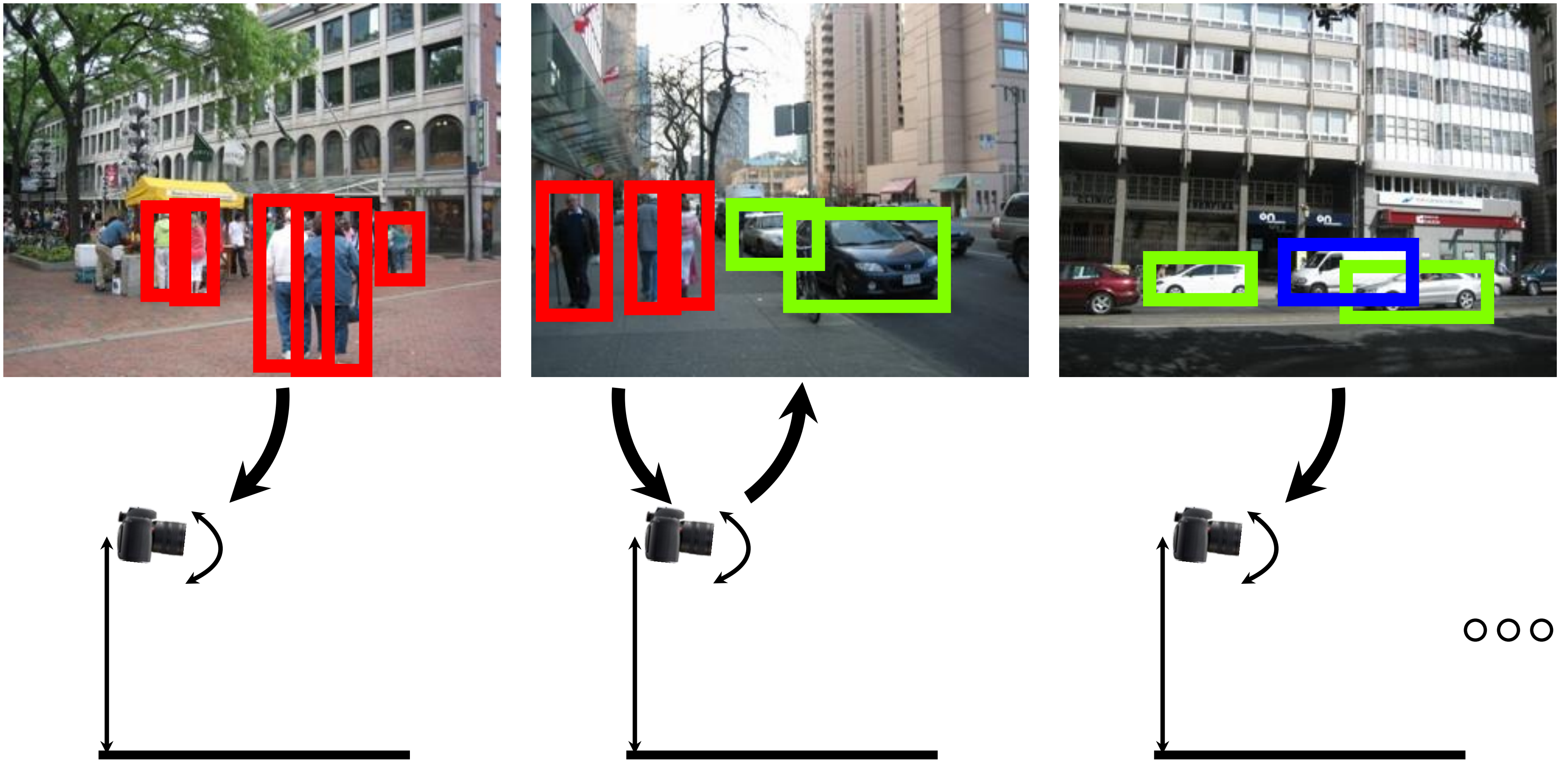
- flat ground plane
- all objects on ground
- camera roll is negligible (consider pitch only)

Camera parameters: height and orientation

Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

Car height distribution
1.5 +/- 0.19 m
(automatically learned)



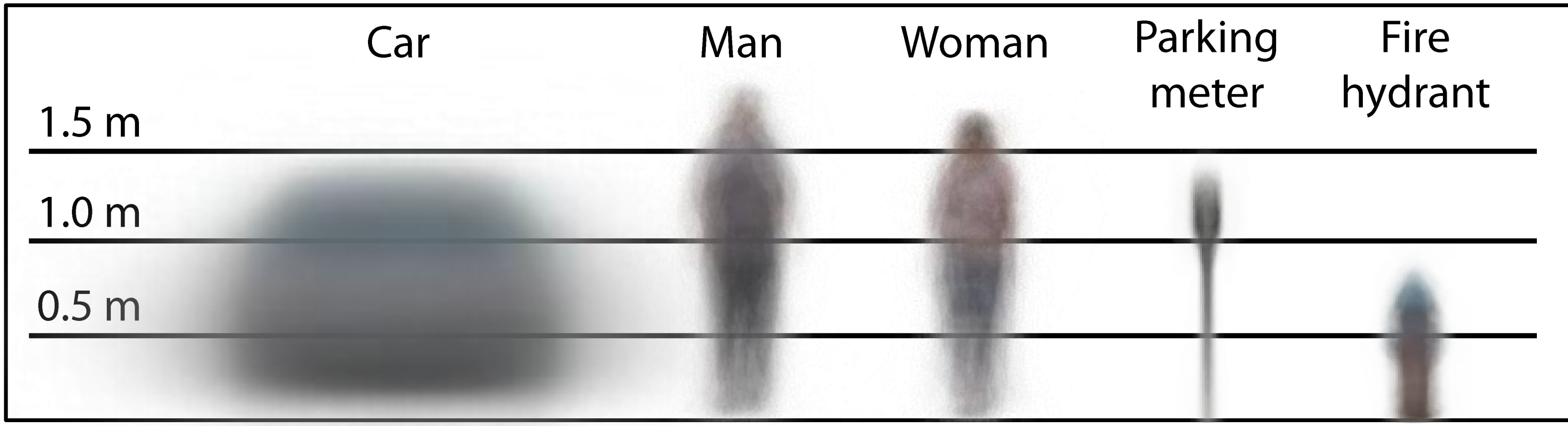
Object heights

Database image

Pixel heights

Real heights

Estimated object heights



Object	Estimated average height (m)
Car	1.51
Man	1.80
Woman	1.67
Parking meter	1.36
Fire hydrant	0.87

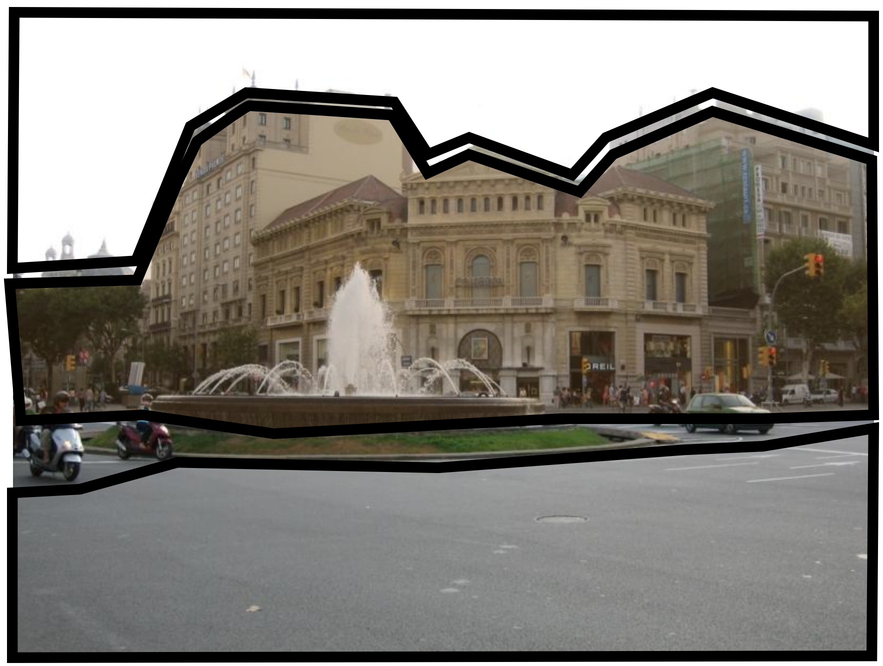
What should we match?

Geometry is not enough

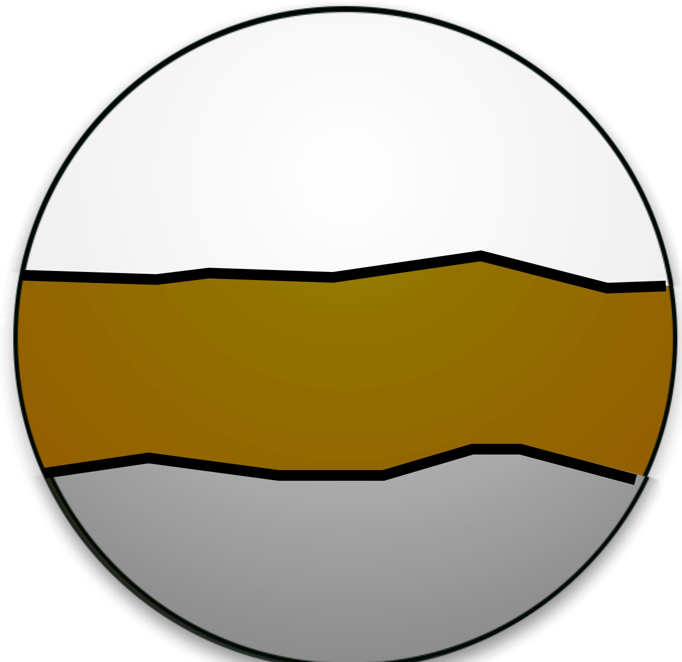
Illumination context

- Exact environment map is impossible
- Approximations [Khan et al., '06]

Database image

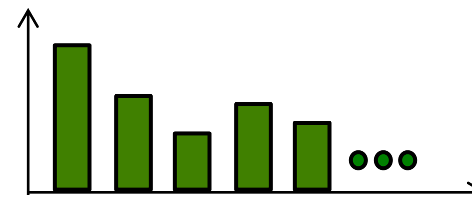


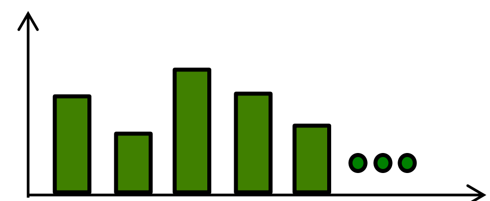
Environment map rough approximation



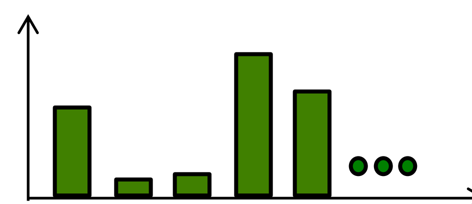
Illumination context

Database image | P(pixel|class) | CIE L*a*b* histograms

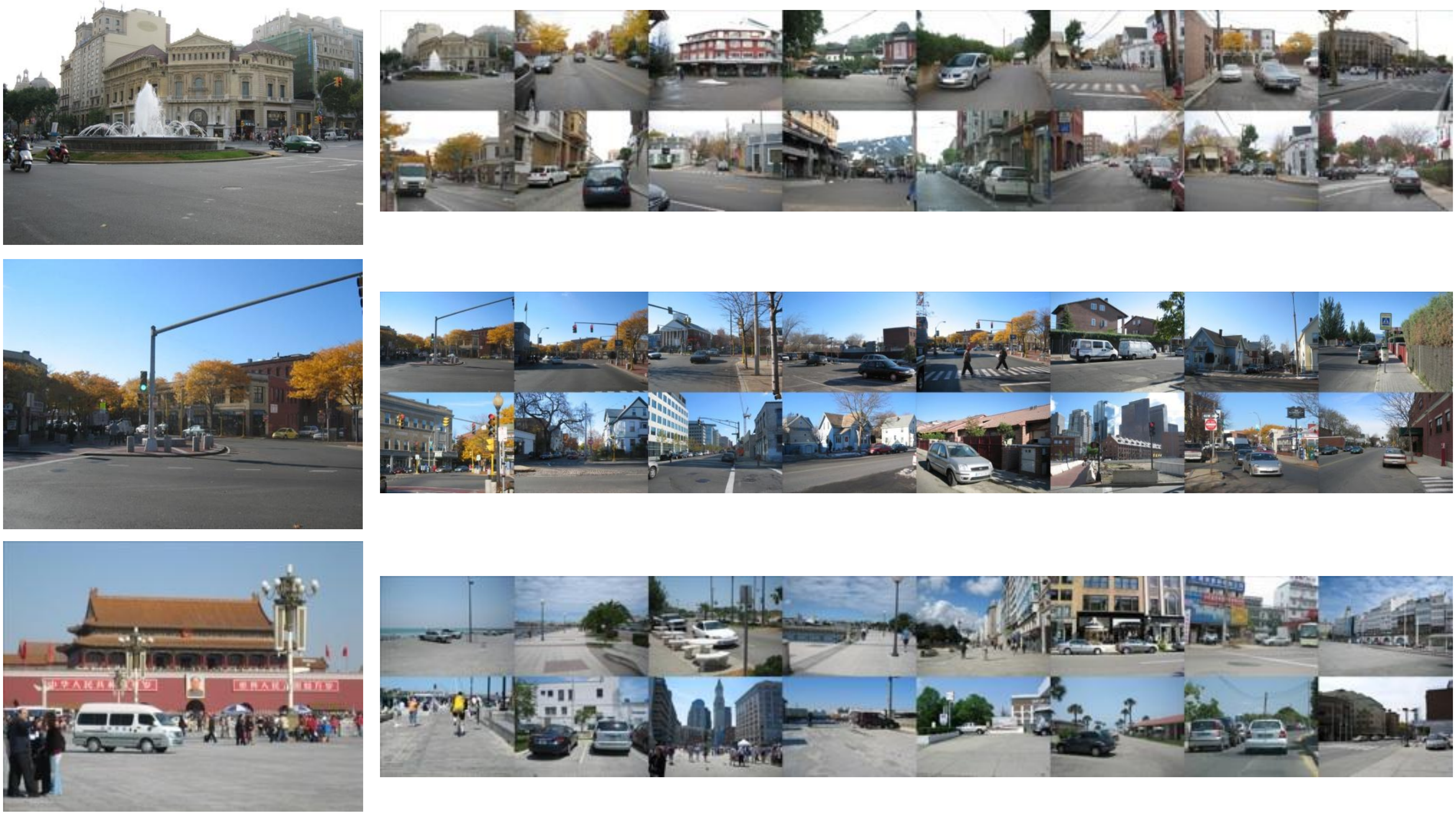




matic Photo PopupHoiem et al., SIGGRAPH '0



Illumination nearest-neighbors



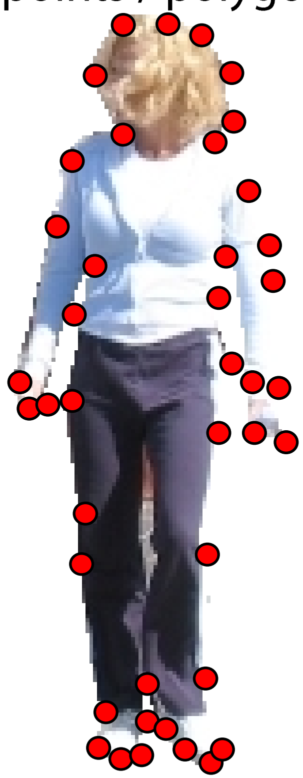
Other criteria: local context

Other criteria: segmentation

LabelMe contributors not always reliable

Segmentation quality

38 points / polygon



4 points / polygon

Other criteria: blur

Resolution: avoid up-sampling

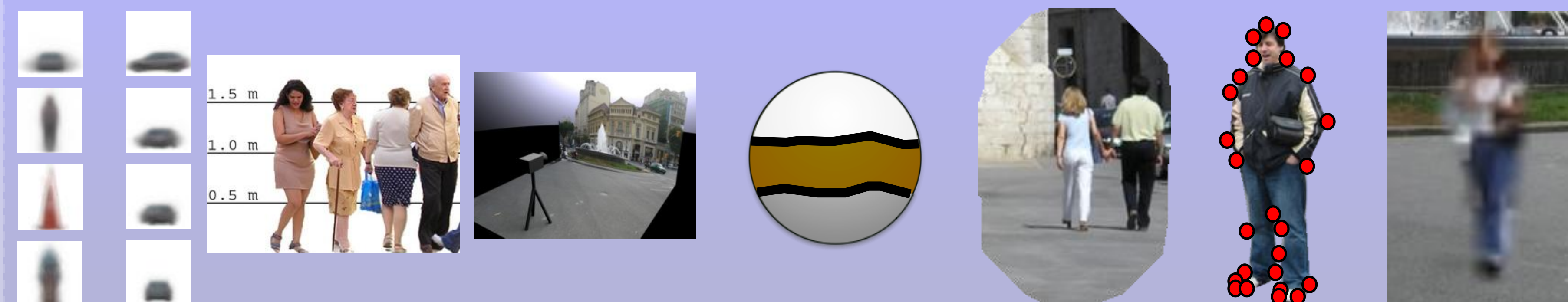
x3 up-sampling

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label | Cluster | 3-D height | Camera | Illumination context | Local context | Segmentation | Blur



Phase II: Object insertion

Let’s insert an object!

Poor user-provided segmentations

Noticeable seams

Seams

Input

Destination image

Result

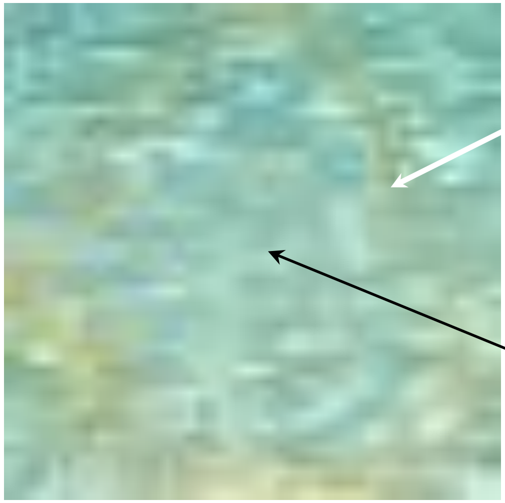
Visible seam!

[Perez et al., 2003]

Poisson blending: idea

Input

Destination



Enforce boundary color
(seamless result)

Enforce same gradient
than input

Result

[Perez et al., 2003]

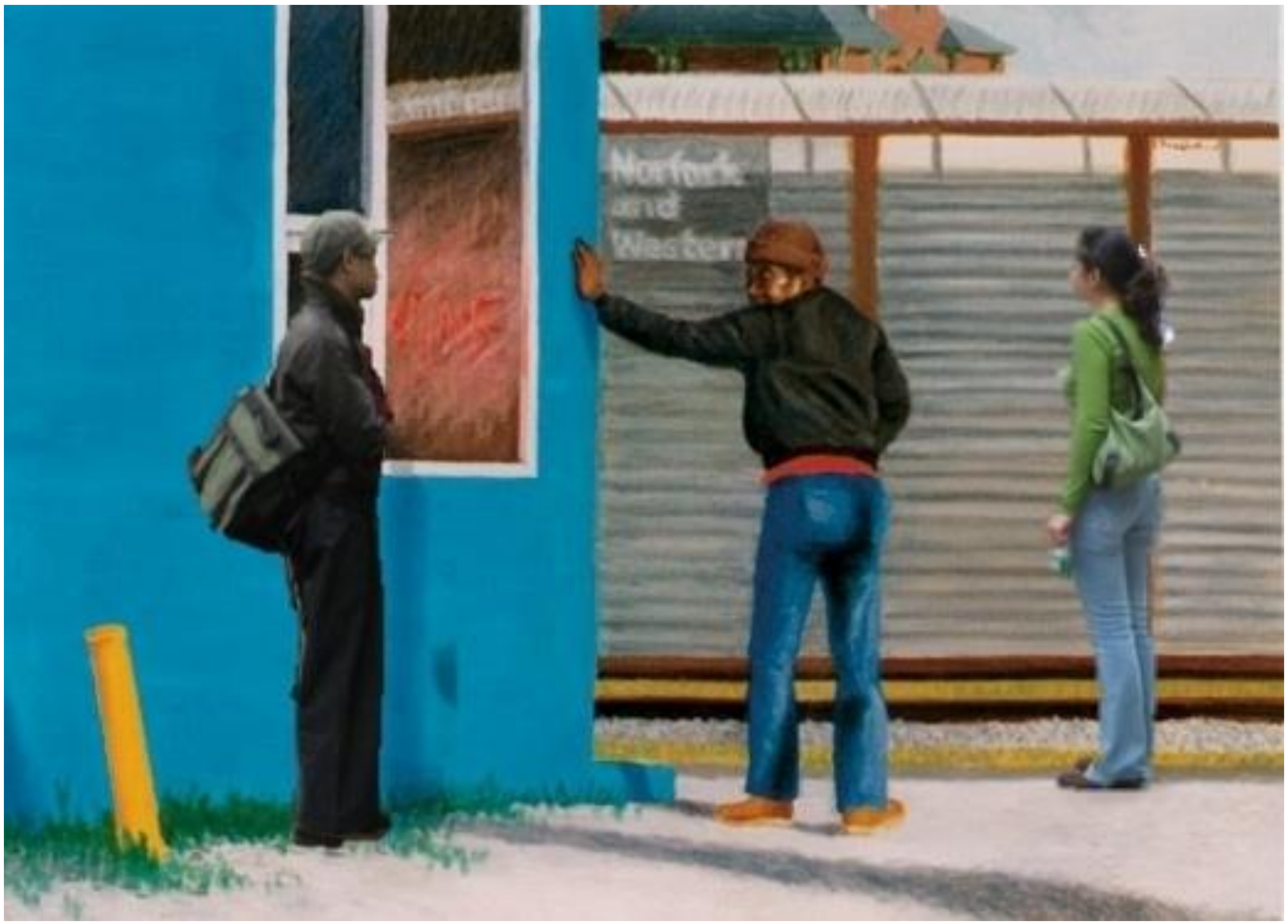
Still not right!

Not so sensitive to shadow direction [Cavanagh, 2005]

Street accident

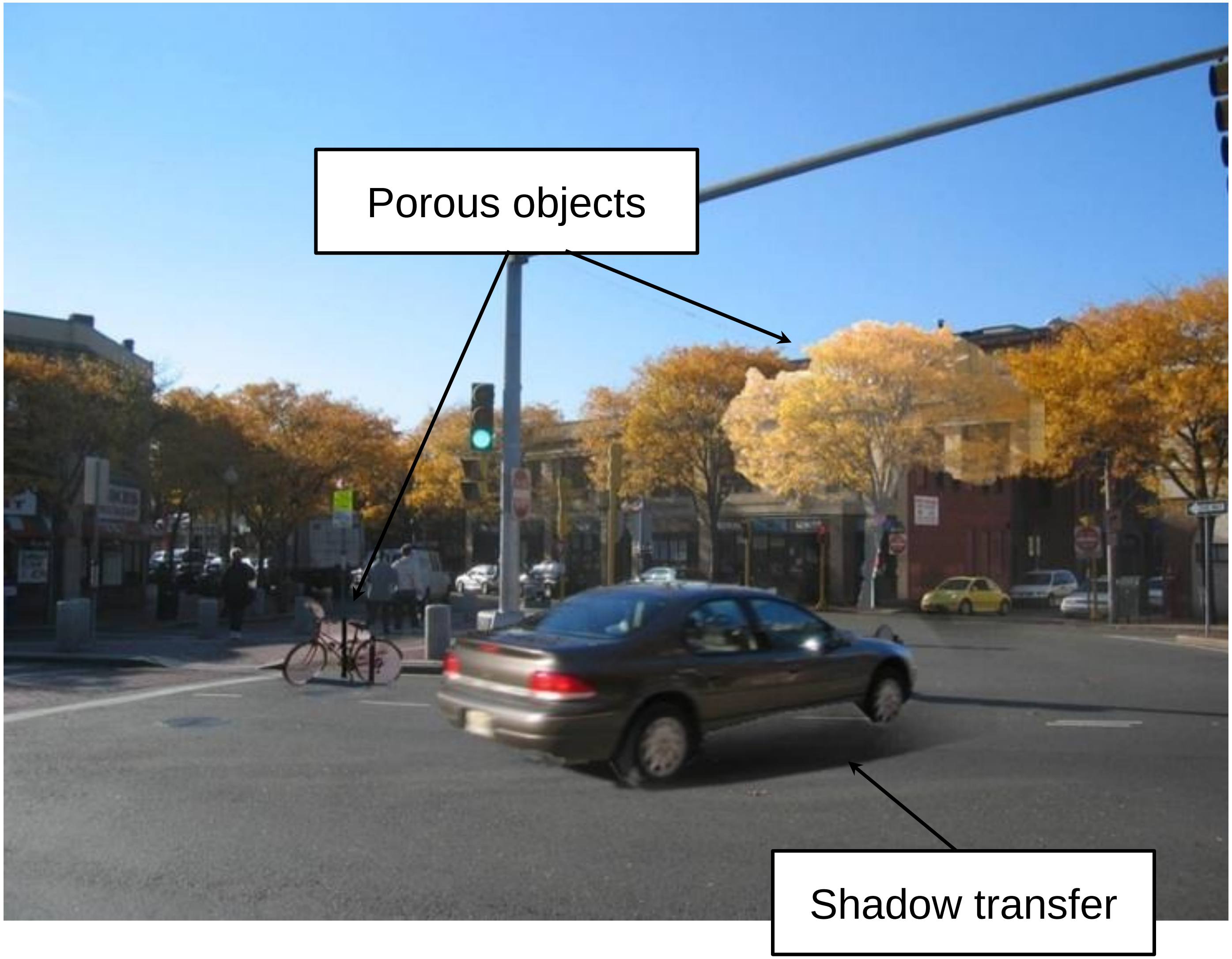
Bridge

Painting



Alley

Failure cases



Failure cases

The Dangers of Data

Bias

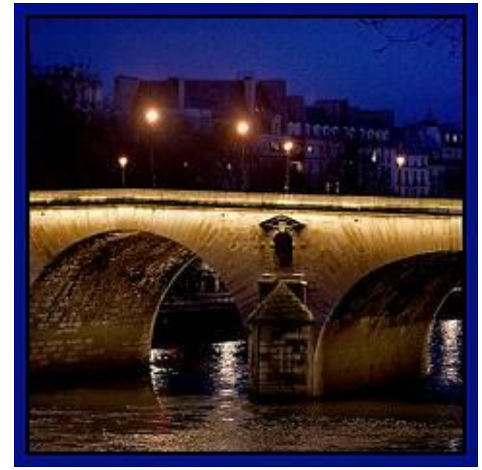
Internet is a tremendous repository of visual data (Flickr, YouTube, Picasa, etc.)

But it's not random samples of visual world

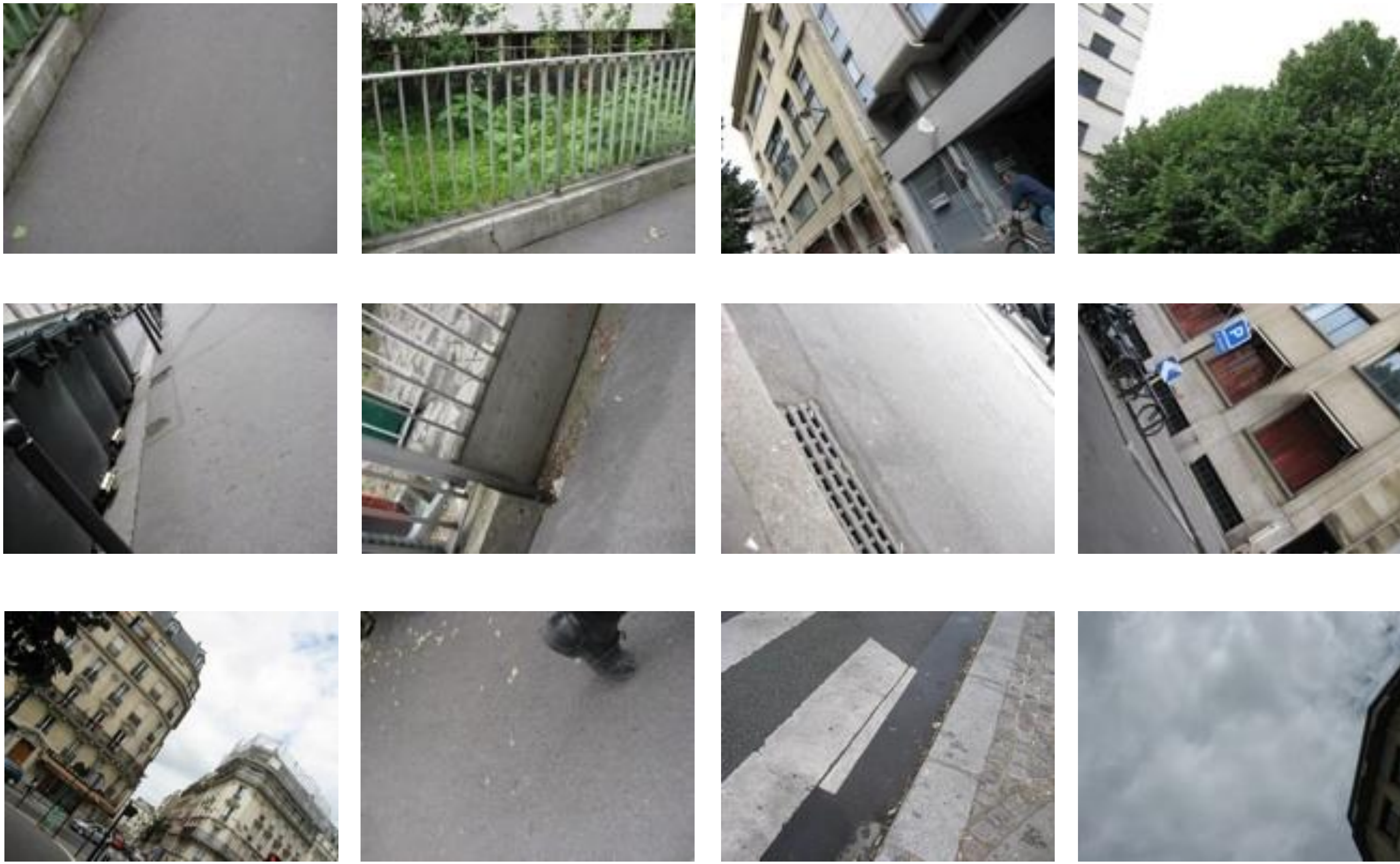
Many sources of bias:

- Sampling bias
- Photographer bias
- Social bias

Flickr Paris



Real Paris



Real Notre Dame



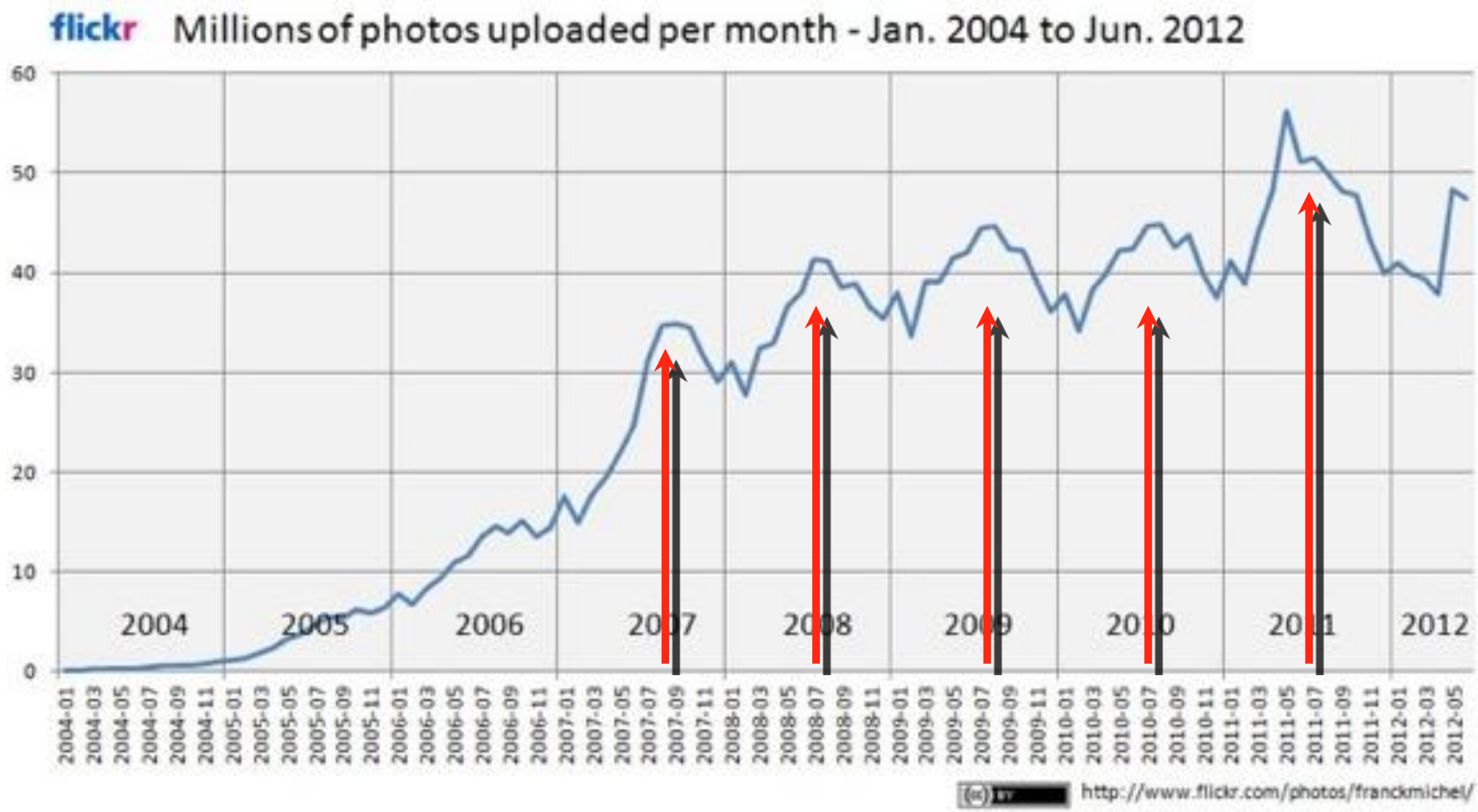



Sampling Bias

People like to take pictures on vacation

Sampling Bias

People like to take pictures on vacation



Photographer Bias

People want their pictures to be recognizable and/or interesting

vs.

Photographer Bias

People follow photographic conventions

vs.

Social Bias

Little Leaguer

Kids with Santa

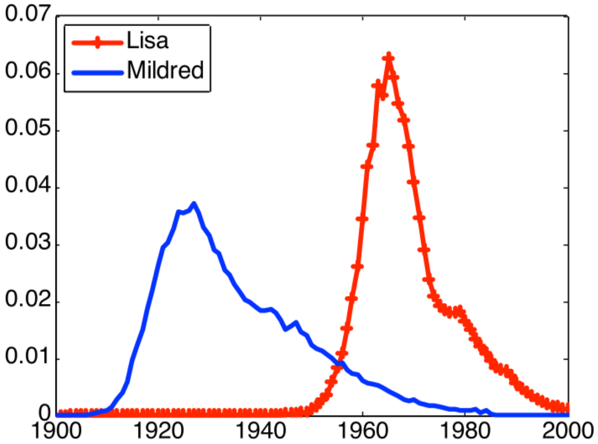
The Graduate

Newlyweds

“100 Special Moments” by Jason Salavon

Social Bias

Mildred and Lisa



Source: U.S. Social Security Administration

Gallagher et al CVPR 2008

Social Bias

Reducing / Changing Bias

Street side
Google StreetView

Satellite
google.com

Webcams

Autonomous capture methods can reduce / change bias

- But it won't go away completely

Sometimes you can just pick your data to suit your problem, but not always…

Overview

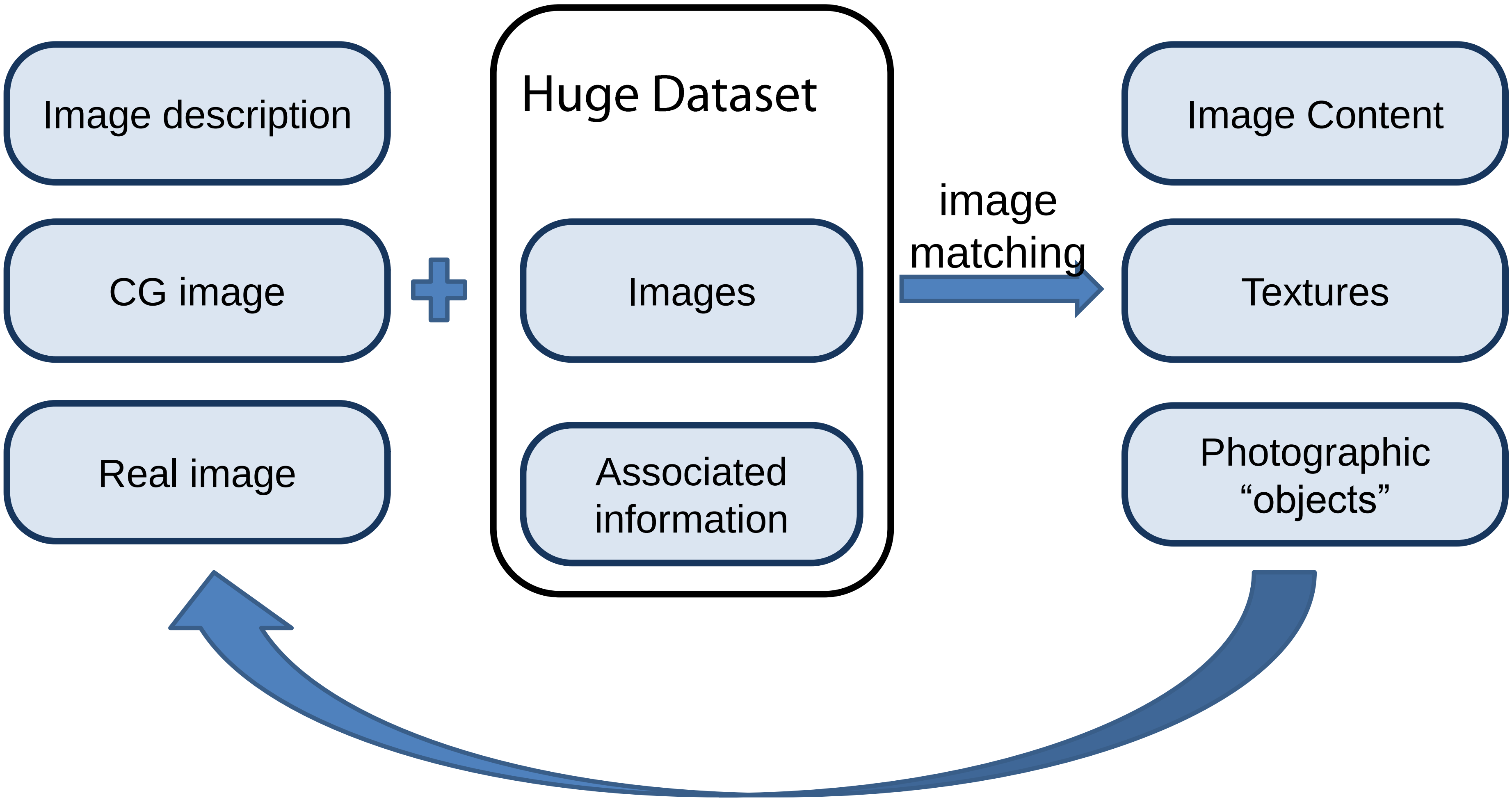


Image description
CG image
Real image
Huge Dataset
Images
Associated information
image matching
Image Content
Textures
Photographic “objects”