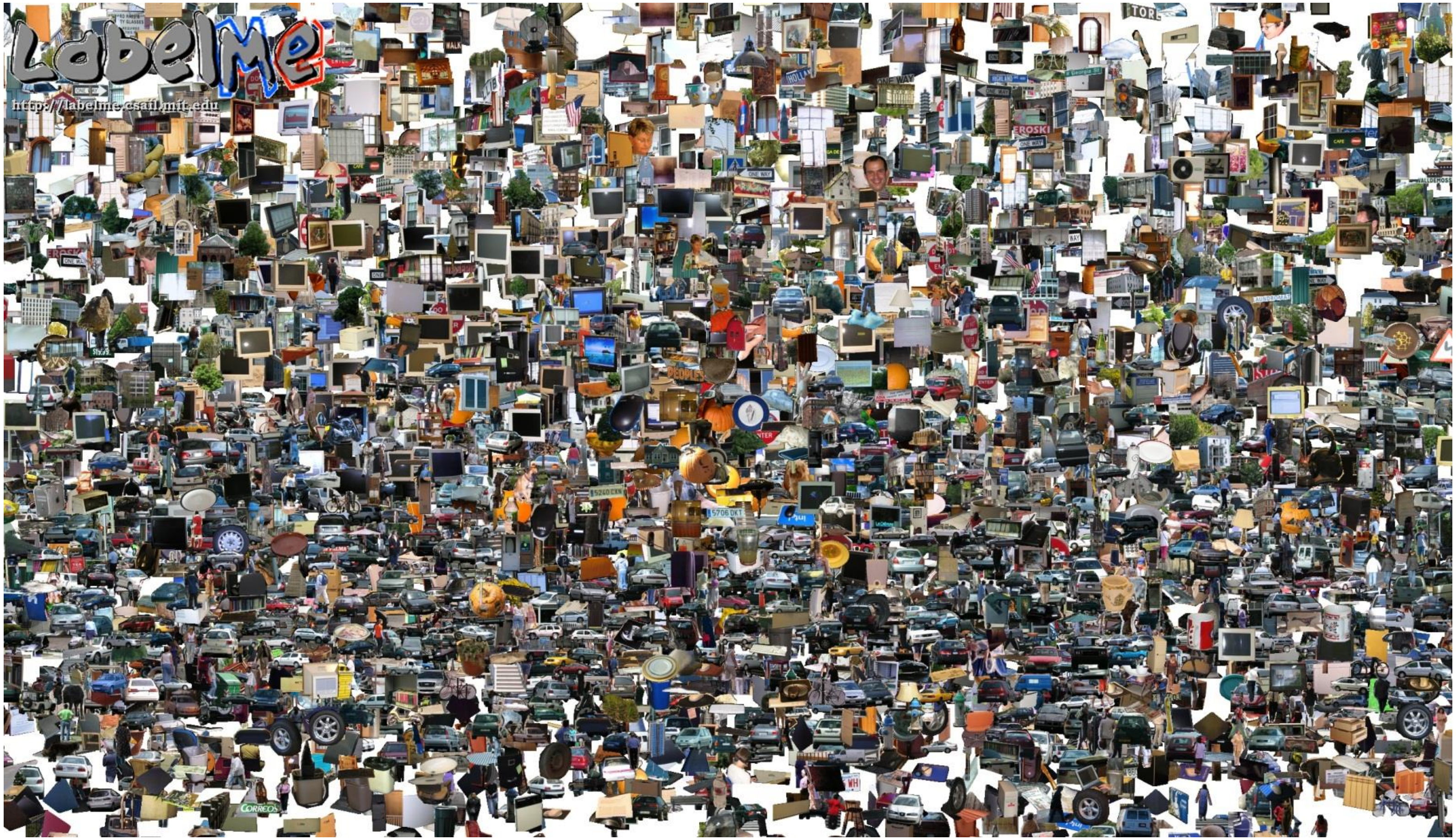
Opportunities of Scale, Part 2



Computer Vision

James Hays, Brown

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

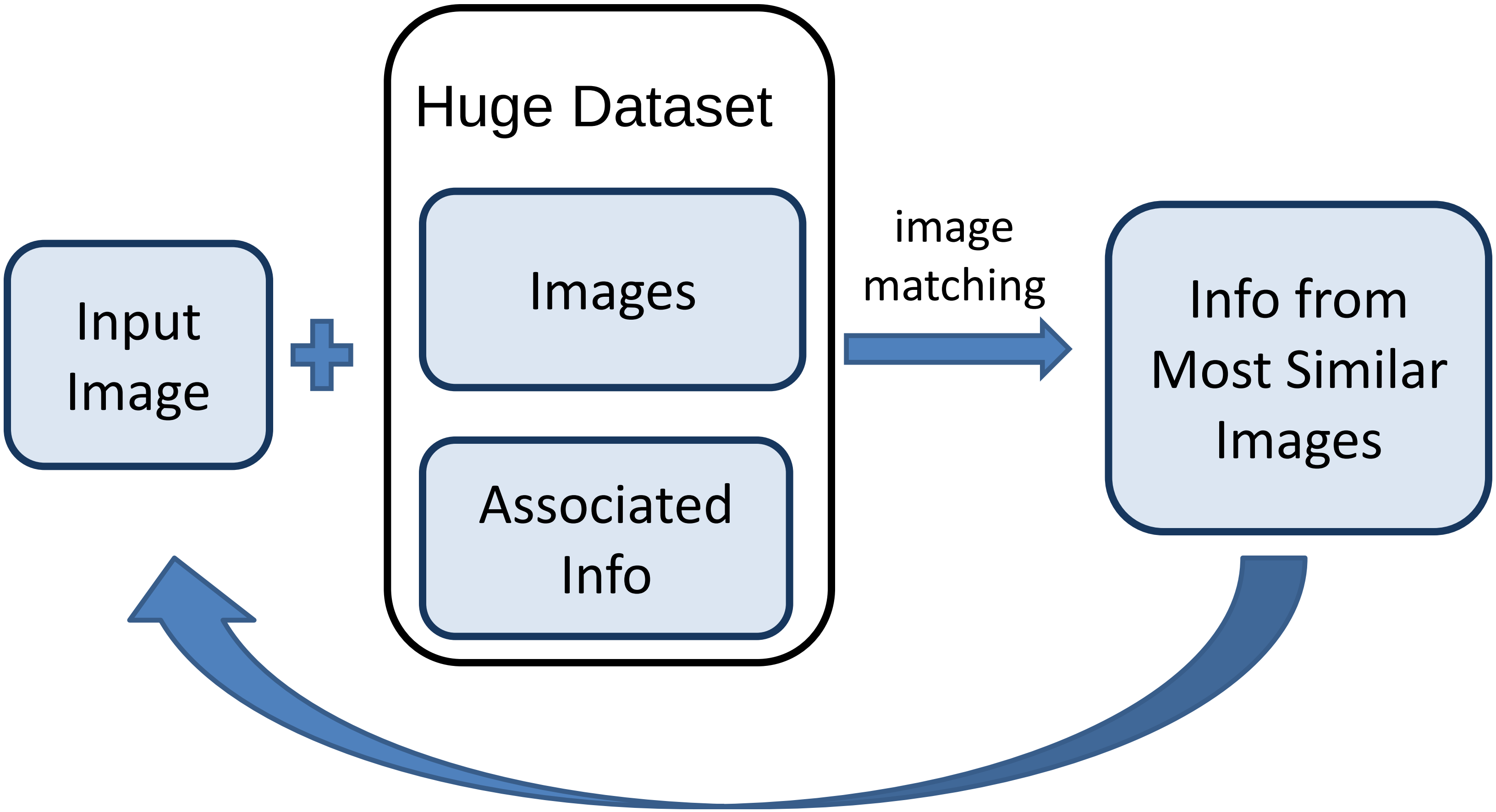
Graphic from Antonio Torralba

Recap

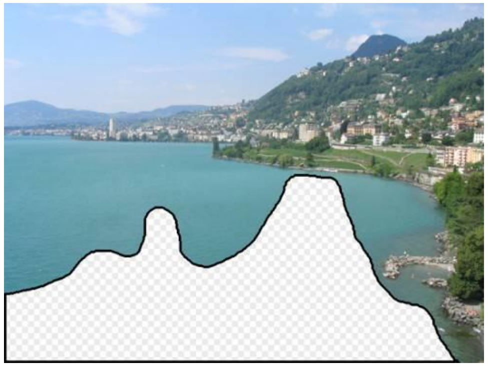
Opportunities of Scale: Data-driven methods

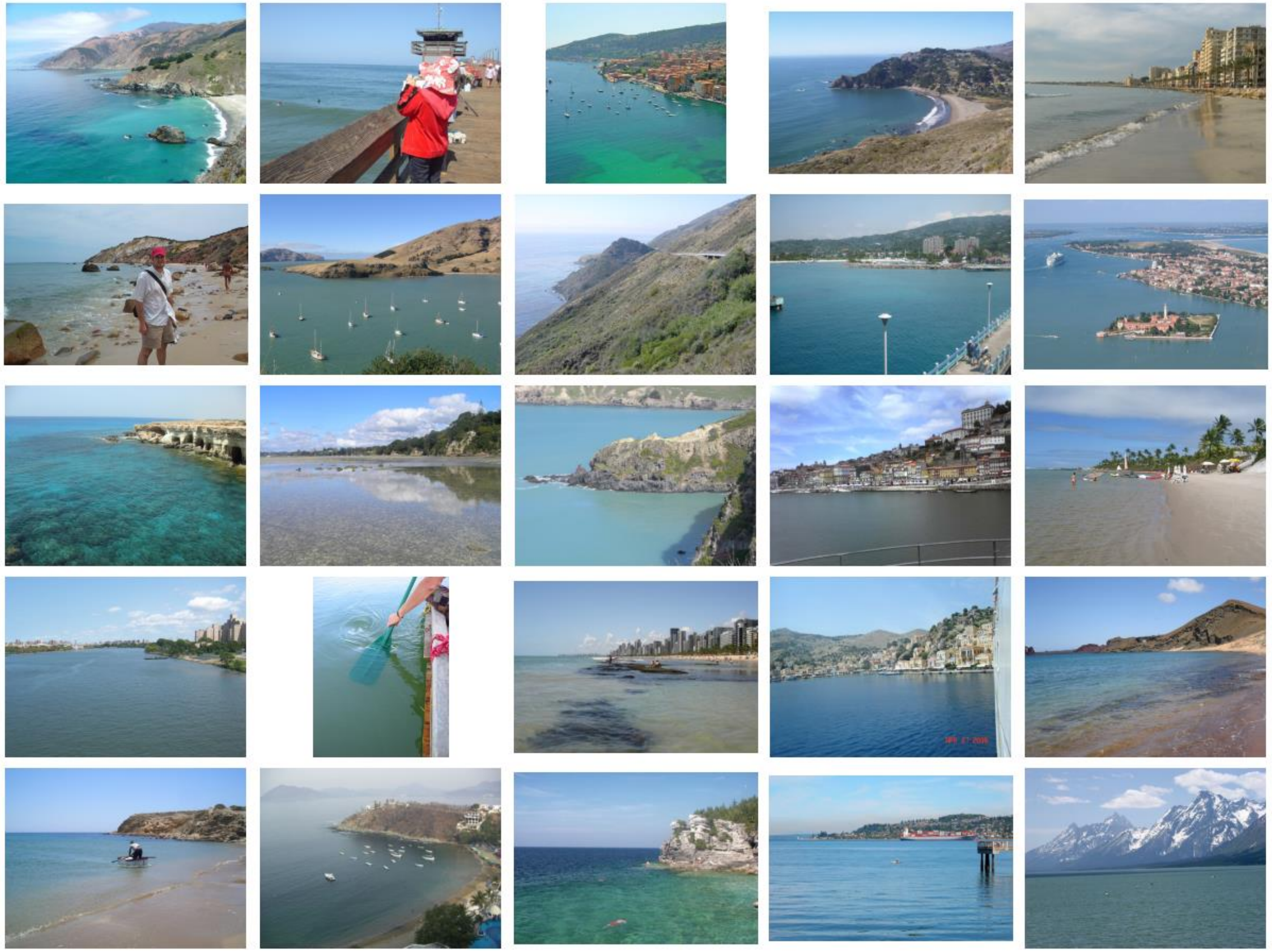
- Monday
  - Scene completion
  - Im2gps
- Today
  - Recognition via Tiny Images
  - More recognition by association

General Principal



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





... 200 total

Graph cut + Poisson blending

im2gps (Hays & Efros, CVPR 2008)

6 million geo-tagged Flickr images

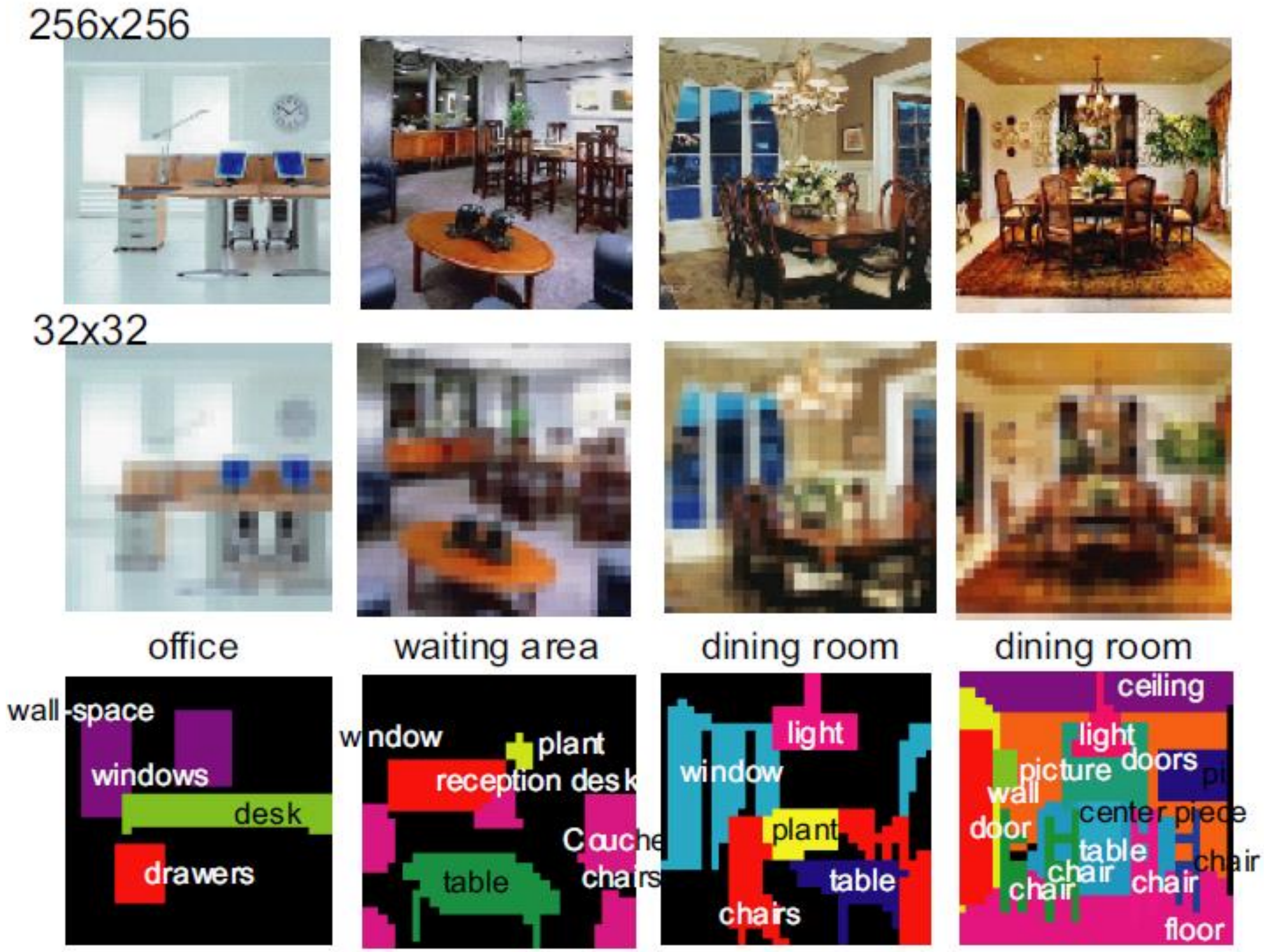
http://graphics.cs.cmu.edu/projects/im2gps/

Tiny Images

80 million tiny images: a large dataset for non-parametric object and scene recognition
Antonio Torralba, Rob Fergus and William T. Freeman. PAMI 2008.

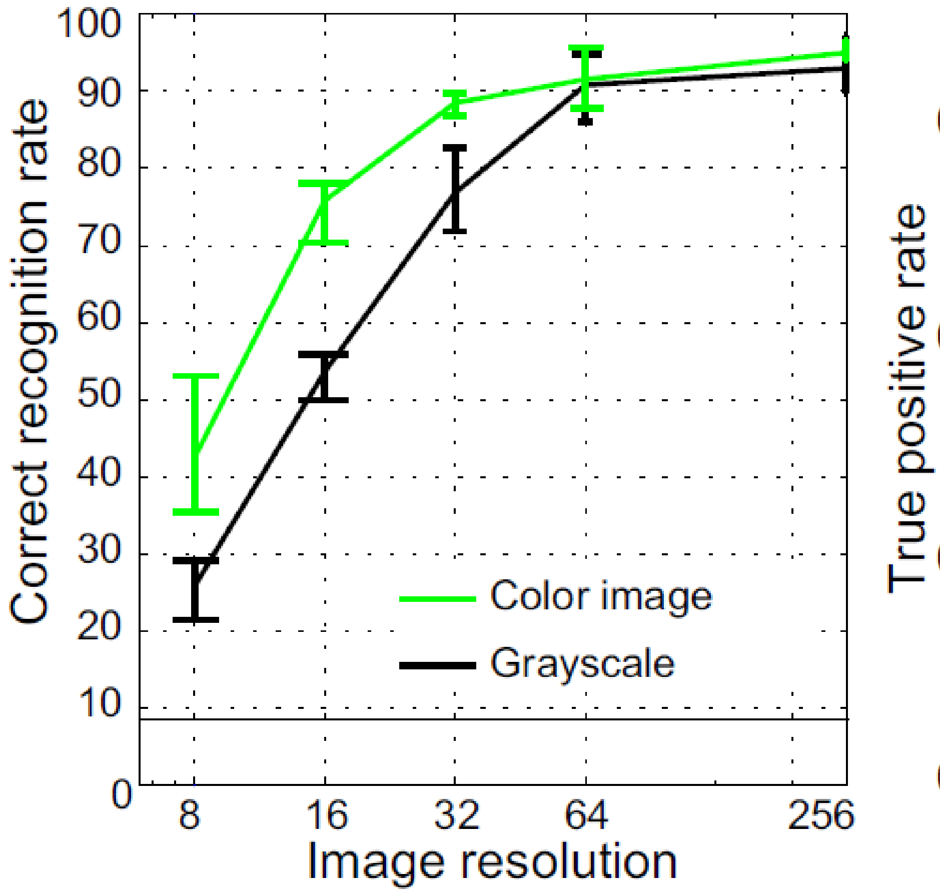
http://groups.csail.mit.edu/vision/TinyImages/

256x256



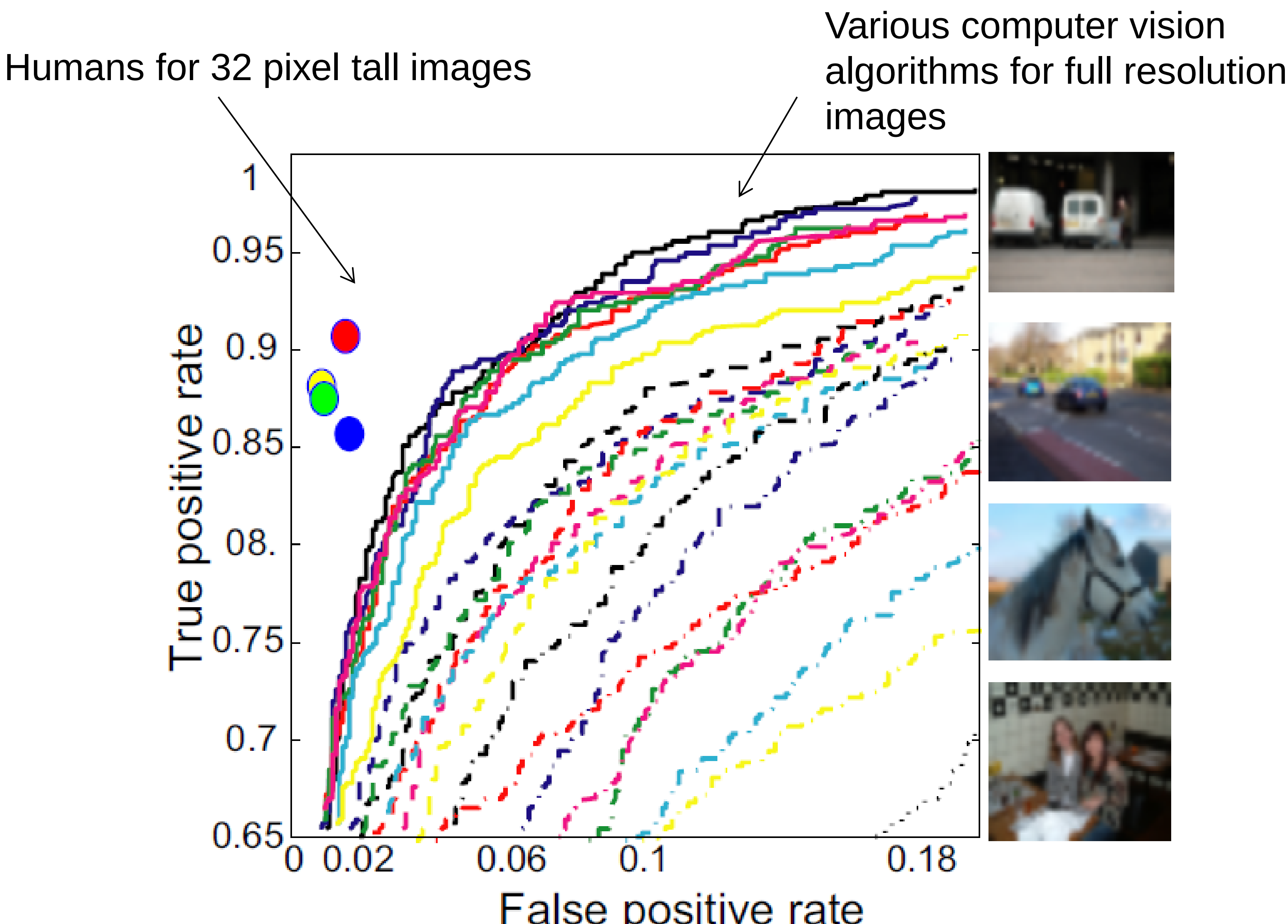
c) Segmentation of 32x32 images

Human Scene Recognition



a) Scene recognition

Humans vs. Computers: Car-Image Classification



Powers of 10

Number of images on my hard drive: 10^4

Number of images seen during my first 10 years: 10^8

(3 images/second * 60 * 60 * 16 * 365 * 10 = 630720000)

Number of images seen by all humanity: 10^{20}

106,456,367,669 humans[1] * 60 years * 3 images/second * 60 * 60 * 16 * 365 =

1 from http://www.prb.org/Articles/2002/HowManyPeopleHaveEverLivedonEarth.aspx

Number of photons in the universe: 10^{88}

Number of all 32x32 images: 10^{7373}

$256^{32*32*3} \sim 10^{7373}$

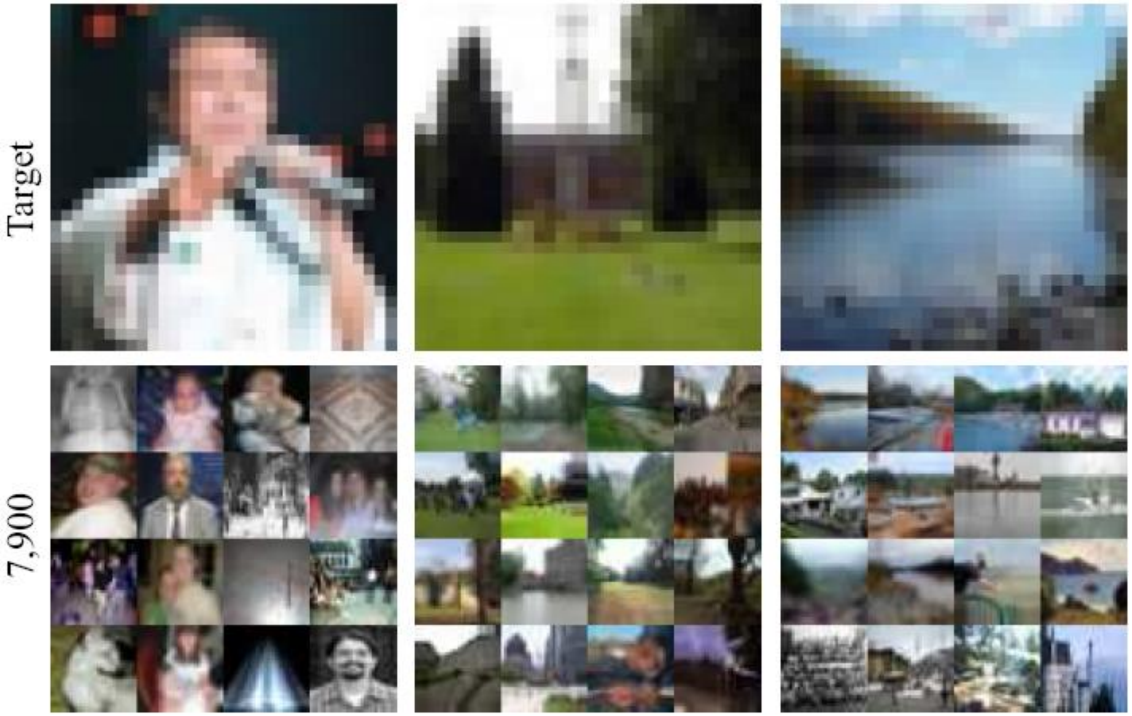
Scenes are unique

But not all scenes are so original

Lots

Of

Images

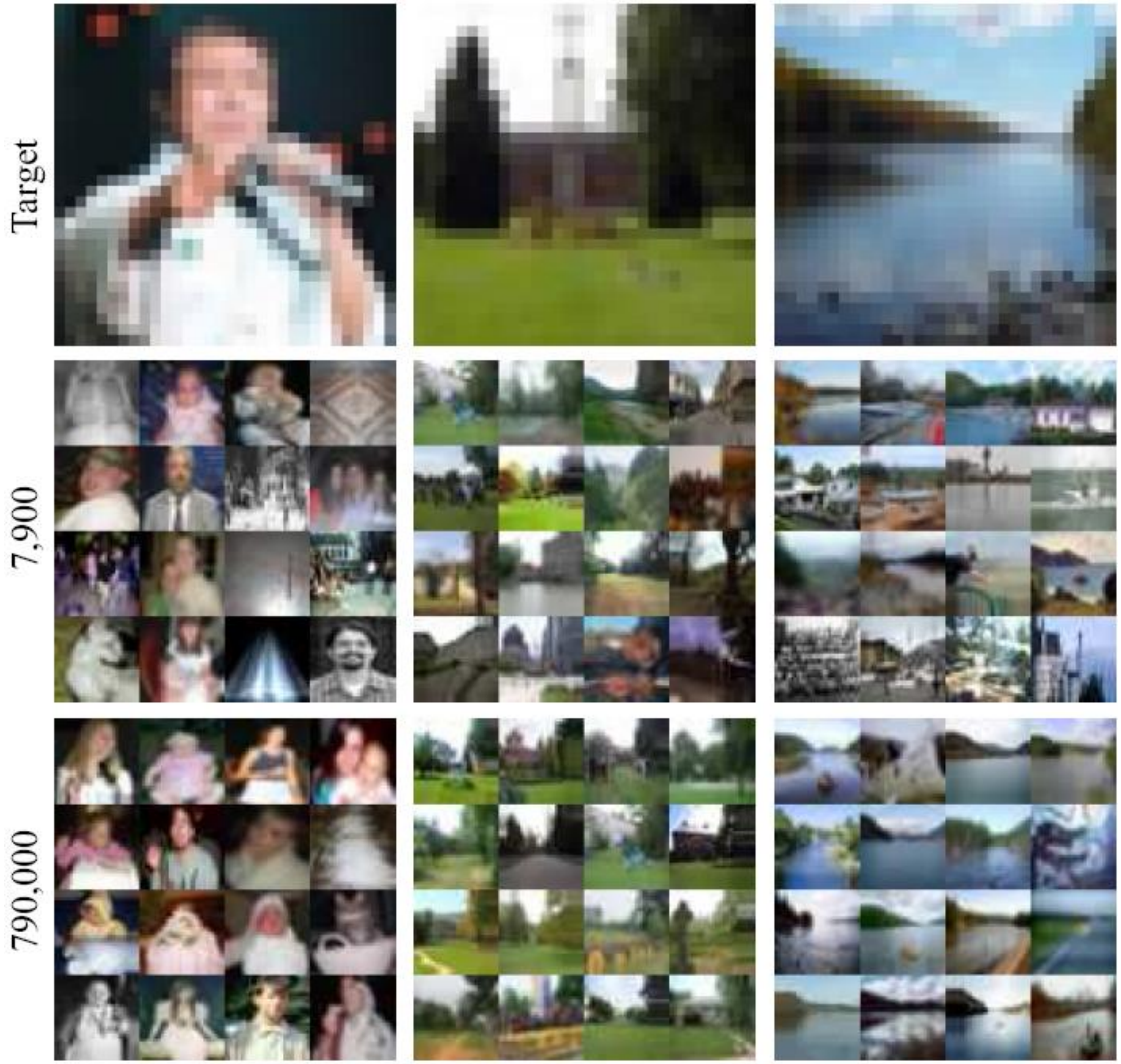


A. Torralba, R. Fergus, W.T.Freeman. PAMI 2008

Lots

Of

Images

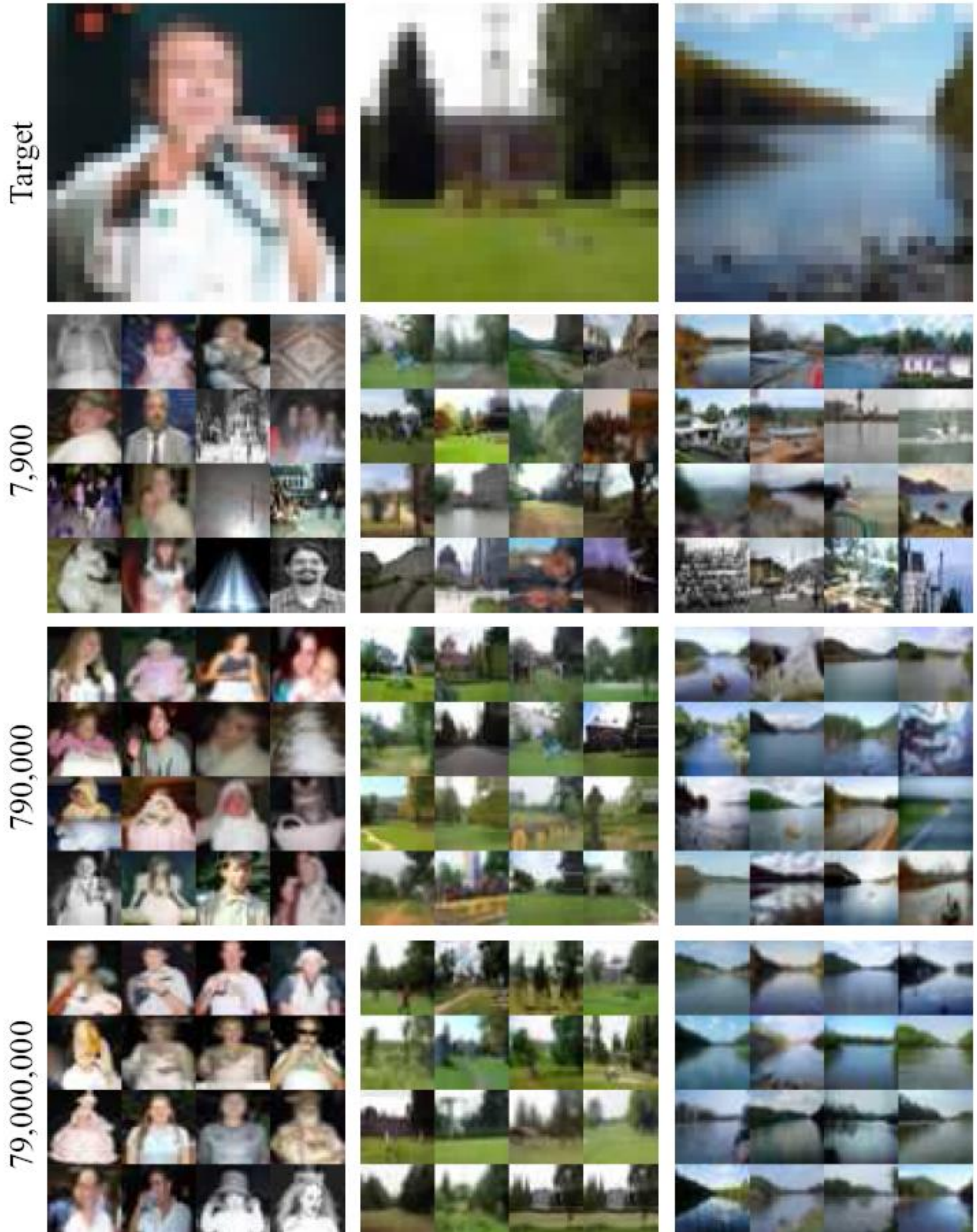


A. Torralba, R. Fergus, W.T.Freeman. PAMI 2008

Lots

Of

Images



Application: Automatic Colorization

Input

Color Transfer

Color Transfer

Matches (gray)

Matches (w/ color)

Avg Color of Match

Application: Automatic Colorization

Input

Color Transfer

Color Transfer

Matches (gray)

Matches (w/ color)

Avg Color of Match

Recognition by Association

Rather than categorizing objects, associate them with stored examples of objects and transfer the associated labels.

Malisiewicz and Efros (CVPR 2008)

Training procedure

- Learn a region similarity measure from hand-segmented objects in LabelMe

- Similarity features
  - Shape: region mask, pixel area, bounding box size
  - Texture: normalized texton histogram
  - Color: mean RGB, std RGB, color histogram
  - Position: coarse 8x8 image mask, coords of top/bottom pixels

Training procedure

- Learn a distance/similarity measure *for each region*
  - Minimize distance to K most similar examples from same category
  - Maximize distance to examples from other categories

distance weights

distance measures

$$\{\mathbf{w}^*, \boldsymbol{\alpha}^*\} = \operatorname*{argmin}_{\mathbf{w}, \boldsymbol{\alpha}} f(\mathbf{w}, \boldsymbol{\alpha})$$

$$f(\mathbf{w}, \boldsymbol{\alpha}) = \sum_{i \in C} \alpha_i L(-\mathbf{w} \cdot \mathbf{d}_i) + \sum_{i \notin C} L(\mathbf{w} \cdot \mathbf{d}_i)$$

$$\mathbf{w} \geq 0, \ \alpha_j \in \{0, 1\}$$

$$\sum_j \alpha_j = K$$

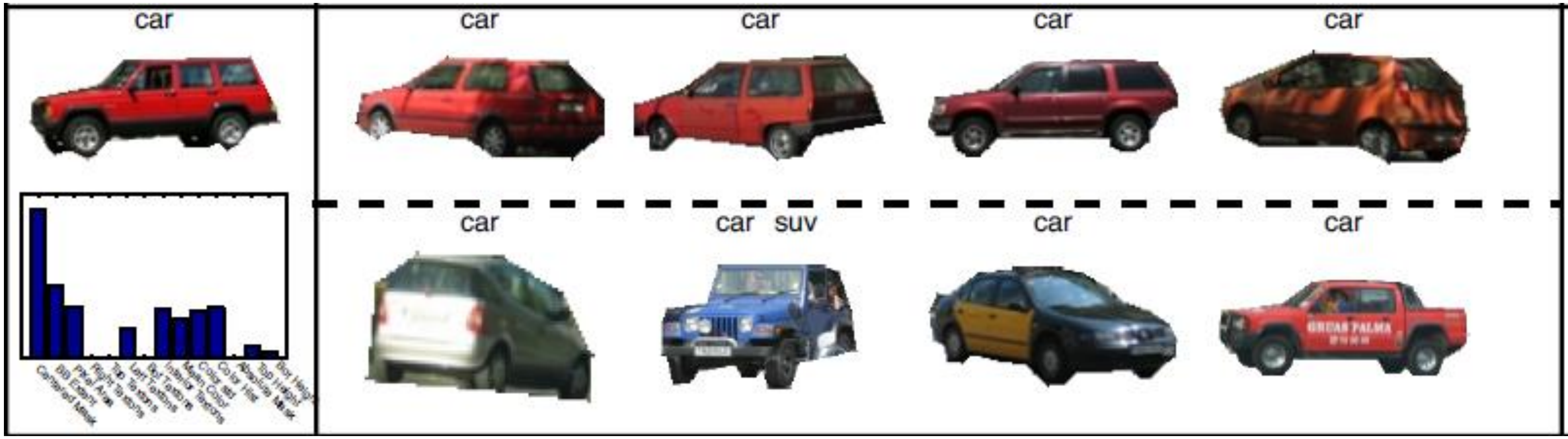
Set to 1 for K nearest examples

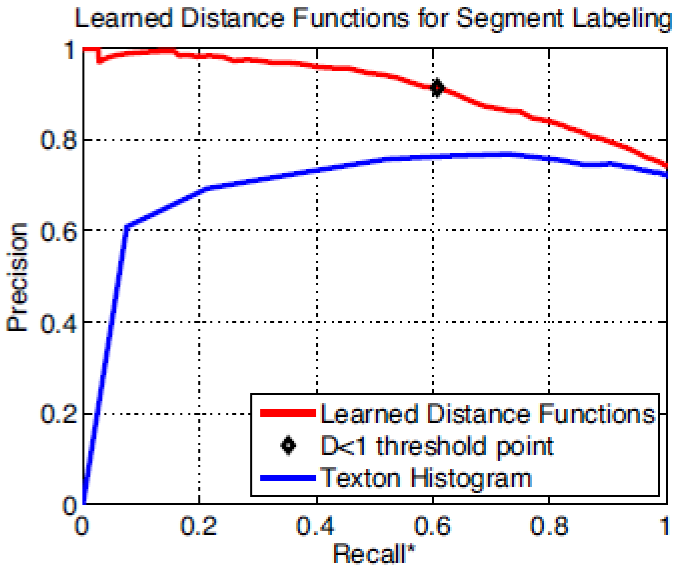
Hinge Loss

Learned Similarity Measure

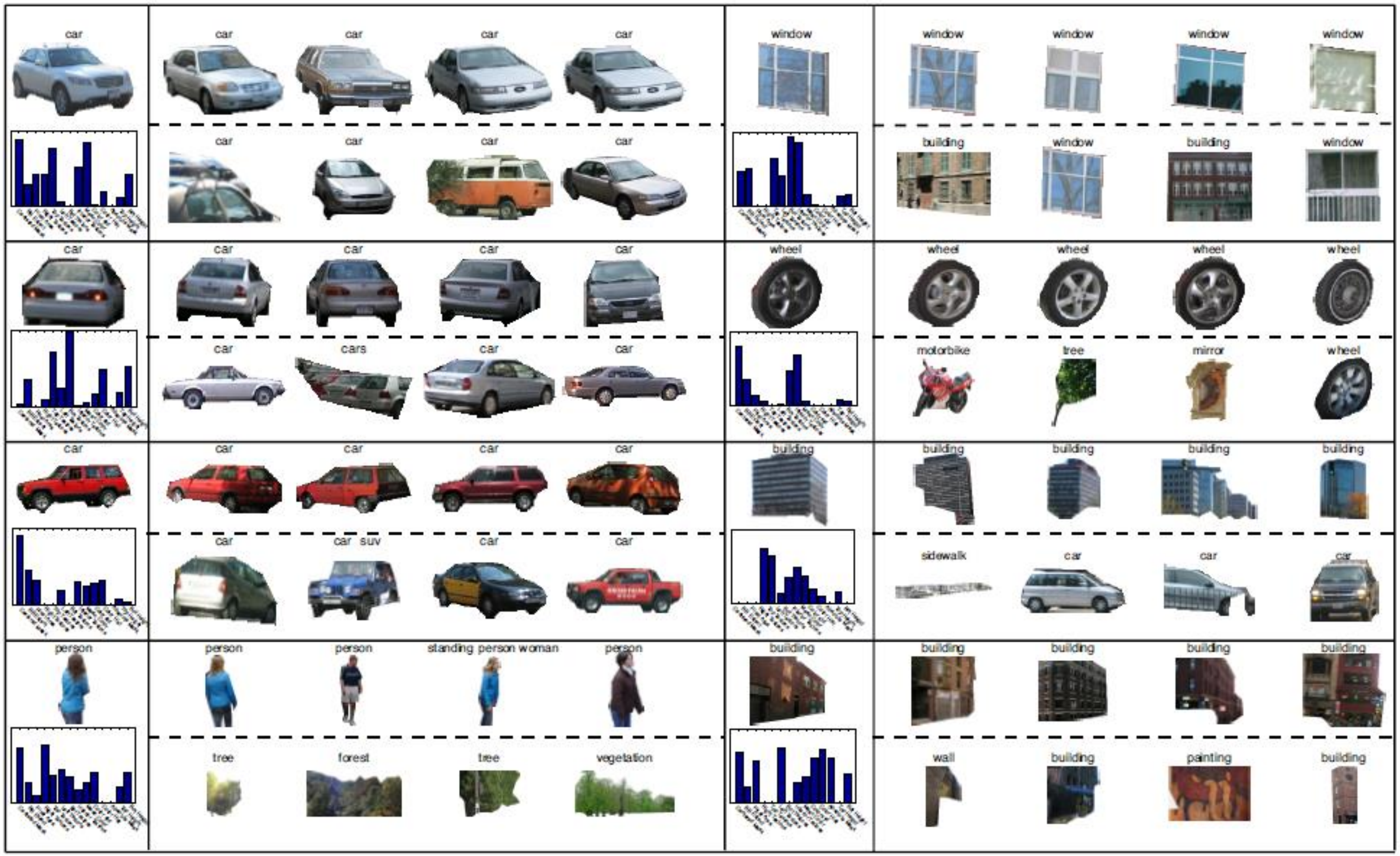
Learned Distance

Texton Distance



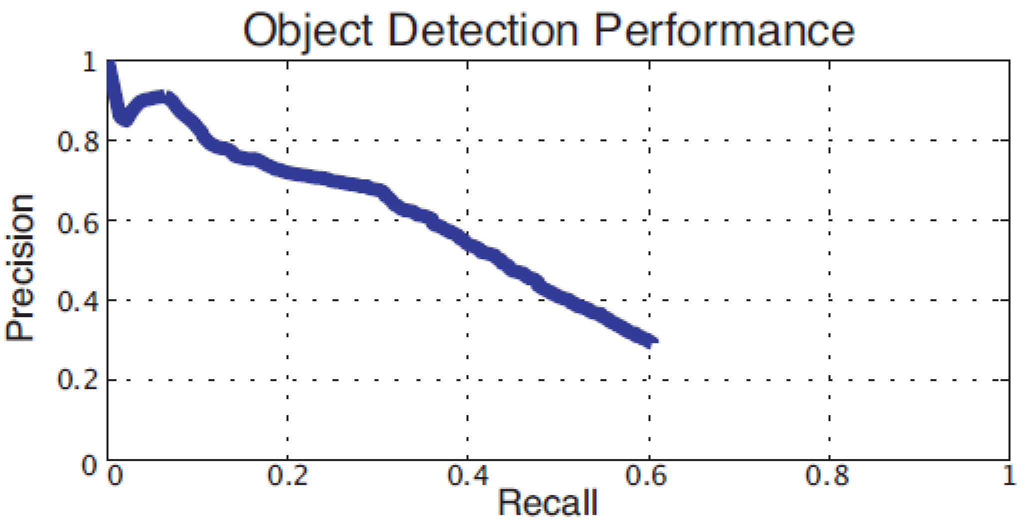


Learned Similarity Measure

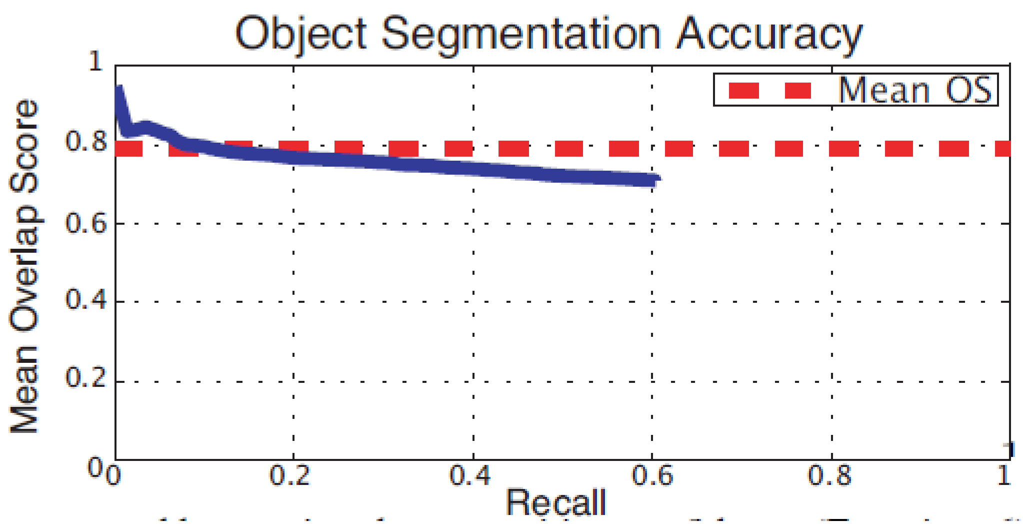


Testing procedure

- Create multiple segmentations (MeanShift + Ncuts)
- Find similar object regions in training set; each votes for the object label
- What about bad segments?
  - Most of the time, they don’t match any objects in the training set
  - Consider only associations with distance < 1



Object Detection Performance
Precision
Recall
0
0.2
0.4
0.6
0.8
1



Object Segmentation Accuracy
Mean OS
Mean Overlap Score
Recall
0
0.2
0.4
0.6
0.8
1

Automatic Parses

Summary

- With billions of images on the web, it's often possible to find a close nearest neighbor

- In such cases, we can shortcut hard problems by "looking up" the answer, stealing the labels from our nearest neighbor

- For example, simple (or learned) associations can be used to synthesize background regions, colorize, or recognize objects