Depth from a single image based on the amount of blur

Estimate the amount of blur using and recover a sharp image by deconvolution with a sparse gradient prior.
Visual Data on the Internet


With slides from Alexei Efros, James Hays, Antonio Torralba, Jean-Francois Lalonde, and Frederic Heger

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

What is out there on the Internet?
How do we get it?
What can we do with it?
Subject-specific Data

Photos of Coliseum (Snavely et al.)

Portraits of Bill Clinton
Much of Captured World is “Generic”
Generic Data

street scenes

Food plates

faces

pedestrians
The Internet as a Data Source
How big is Flickr?

100M photos updated *daily*
6B photos as of August 2011!
• ~3B public photos

Credit: Franck_Michel (http://www.flickr.com/photos/franckmichel/)
How Annotated is Flickr? (tag search)

Party – 23,416,126
Paris – 11,163,625
Pittsburgh – 1,152,829
Chair – 1,893,203
Violin – 233,661
Trashcan – 31,200
“Trashcan” Results

• http://www.flickr.com/search/?q=trashcan+NOT+party&m=tags&z=t&page=5
Big Issues

What is out there on the Internet?
How do we get it?
What can we do with it?

• Let’s see a motivating example...
Scene Completion
Scene Matching
Scene Descriptor
Scene Descriptor

Scene Gist Descriptor
(Oliva and Torralba 2001)
Scene Descriptor

Scene Gist Descriptor
(Oliva and Torralba 2001)
Context Matching
Graph cut + Poisson blending
Nearest neighbors from a collection of 20 thousand images
Nearest neighbors from a collection of 2 million images
“Unreasonable Effectiveness of Data”

Parts of our world can be explained by elegant mathematics
  - physics, chemistry, astronomy, etc.

But much cannot
  - psychology, economics, genetics, etc.

Enter The Data!
  - Great advances in several fields:
    - e.g. speech recognition, machine translation
    - Case study: Google

[Halevy, Norvig, Pereira 2009]
A.I. for the postmodern world:

• all questions have already been answered…many times, in many ways

• Google is dumb, the “intelligence” is in the data
How about *visual* data?

Text is simple:
- clean, segmented, compact, 1D

Visual data is much harder:
- Noisy, unsegmented, high entropy, 2D/3D

Quick Overview
- Comparing Images
- Uses of Visual Data
- The Dangers of Data
Distance Metrics

- Euclidian distance of 5 units

- Gray value distance of 50 values

- ?
SSD says these are not similar
Tiny Images

Image Segmentation (by humans)
Image Segmentation (by humans)
Human Scene Recognition

The graph shows the relationship between image resolution and correct recognition rate. The correct recognition rate is higher for color images compared to grayscale images. The true positive rate also increases with image resolution, but the difference between color and grayscale images is less pronounced at higher resolutions.
Tiny Images Project Page

http://groups.csail.mit.edu/vision/TinyImages/
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 \(\times 60 \times 60 \times 16 \times 365 \times 10 = 630720000\)

Number of images seen by all humanity: \(10^{20}\)
106,456,367,669 humans\(^1\) \(\times 60\) years \(\times 3\) images/second \(\times 60 \times 60 \times 16 \times 365 =\)

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

Number of all 8-bits 32x32 images: \(10^{7373}\)
\(256^{32 \times 32} \approx 10^{7373}\)
Scenes are unique
But not all scenes are so original
But not all scenes are so original
Automatic Colorization Result

Grayscale input High resolution

Colorization of input using average

A. Torralba, R. Fergus, W.T.Freeman. 2008
Automatic Orientation

Many images have ambiguous orientation
Look at top 25% by confidence

- correlation score

Examples of high and low confidence images
Automatic Orientation Examples

A. Torralba, R. Fergus, W.T.Freeman. 2008
Tiny Images Discussion

Why SSD on color images?
Can we build a better image descriptor?
Image Representations: Histograms

**global histogram**

- Represent distribution of features
  - Color, texture, depth, …

Images from Dave Kauchak
Image Representations: Histograms

Joint histogram
- Requires lots of data
- Loss of resolution to avoid empty bins

Marginal histogram
- Requires independent features
- More data/bin than joint histogram
Image Representations: Histograms

Adaptive binning

- Better data/bin distribution, fewer empty bins
- Can adapt available resolution to relative feature importance
Clustering / Signatures

- “super-adaptive” binning
- Does not require discretization along any fixed axis
Issue: How to Compare Histograms?

Bin-by-bin comparison
Sensitive to bin size.
Could use wider bins ...
... but at a loss of resolution

Cross-bin comparison
How much cross-bin influence is necessary/sufficient?
Red Car Retrievals (Color histograms)

\[ \chi^2(h_i, h_j) = \frac{1}{2} \sum_{m=1}^{K} \frac{[h_i(m) - h_j(m)]^2}{h_i(m) + h_j(m)} \]

Histogram matching distance
Capturing the “essence” of texture

...for real images

We don’t want an actual texture realization, we want a texture invariant

What are the tools for capturing statistical properties of some signal?
Multi-scale filter decomposition

Filter bank

Input image
Filter response histograms
Start with a noise image as output

Main loop:

• Match *pixel* histogram of output image to input
• Decompose input and output images using multi-scale filter bank (Steerable Pyramid)
• Match sub-band histograms of input and output pyramids
• Reconstruct input and output images (collapse the pyramids)
SIFT local feature descriptor

Based on 16*16 patches
4*4 subregions
8 bins in each subregion
4*4*8=128 dimensions in total
Image Descriptors

• Blur + SSD
• Gist descriptor (average edge response in a coarse spatial grid)
• Color histograms
• Filter response histograms
• “Bag of Visual Words” – histograms of quantized SIFT or HOG features.
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?
Example Scene Matches
The Importance of Data

![Graph showing the percentage of geolocations within 200km versus database size. The graph compares 'First Nearest Neighbor Scene Match' in green with 'Chance: Random Scenes' in red. The x-axis represents the database size (thousands of images, log scale) ranging from 0.09 to 6,304, and the y-axis represents the percentage of geolocations within 200km ranging from 0 to 16.]
Where is This?

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 (geo-loc within 400 km)

im2gps – 10%
temporal im2gps – 56%