Are Categories Necessary?

Alexei (Alyosha) Efros
CMU

L'ecipe n'est pas une pipe.
Before We Begin…

• Not an AI/learning person

• My work is in Computer Vision and Computer Graphics

• But I want to utilize lots of data
A Confession

I have a problem…

…I am a nearest-neighbor addict!
Classical Texture Synthesis

Sample texture

Synthesis

Parametric Texture Model

Analysis

Novel texture

Sample texture

This is hard!
Non-parametric Approach

Synthesis

Analysis

Sample texture

Novel texture
Non-parametric Synthesis [Efros&Leung’99]

- Assuming Markov property, compute $P(p|N(p))$
  - Building explicit probability tables infeasible
  - Instead, we search the input image for all similar neighbourhoods — that’s our distribution for $p$
  - To sample from distribution, just pick one match at random
Hole Filling
input image

Portilla & Simoncelli

Our algorithm
Portilla & Simoncelli

Our algorithm

input image

Portilla & Simoncelli

Our algorithm

input image
2 Million Flickr Images
Two Kinds of Things in the World

Navier-Stokes Equation

\[ \frac{\partial u}{\partial t} = - (u \cdot \nabla) u + v \nabla^2 u - \frac{1}{d} \nabla p + f \]

+ weather
+ location
+ ...
Lots of data available
“Unreasonable Effectiveness of Data”  
[Halevy, Norvig, Pereira 2009]

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

• But much cannot:
  – psychology, genetics, economics, etc.

• Enter: The Magic of **Big Data**
  – Great advances in several fields:
    • e.g. speech recognition, machine translation, Google
• 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
Are Categories Necessary?

Leci n’est pas une pipe.

Alexei (Alyosha) Efros
CMU

Joint work with
Tomasz Malisiewicz
Acknowledgements

Talks by Moshe Bar; writings of Shimon Edelman

Many great discussions with many colleagues, especially Tomasz Malisiewicz, James Hays, and Derek Hoiem
Understanding an Image

slide by Fei Fei, Fergus & Torralba
Object naming -> Object categorization

- sky
- building
- flag
- banner
- face
- wall
- street lamp
- bus
- cars

Slide by Fei Fei, Fergus & Torralba
Object categorization

- sky
- building
- flag
- face
- banner
- street lamp
- wall
- bus
- cars
Why Categorize?

1. Knowledge Transfer
2. Communication
Classical View of Categories

• Dates back to Plato & Aristotle
  1. Categories are defined by a list of properties shared by all elements in a category
  2. Category membership is binary
  3. Every member in the category is equal
Problems with Classical View

• Humans don’t do this!
  – People don’t rely on abstract definitions / lists of shared properties (Wittgenstein 1953, Rosch 1973)
    • e.g. define the properties shared by all “games”
    • e.g. are curtains furniture? Are olives fruit?
  – Typicality
    • e.g. Chicken -> bird, but bird -> eagle, pigeon, etc.
  – Language-dependent
    • e.g. “Women, Fire, and Dangerous Things” category is Australian aboriginal language (Lakoff 1987)
  – Doesn’t work even in human-defined domains
    • e.g. Is Pluto a planet?
Problems with **Visual Categories**

- A lot of categories are functional
- World is too varied
- Categories are 3D, but images are 2D
Typical HOG car detector

Felzenszwalb et al, PASCAL 2007
Why not?
Solution: hierarchy?

Ontologies, hierarchies, levels of categories (Rosch), etc.
WordNet, ImageNet, etc etc

Diagram showing a Venn diagram with categories: cat, dog, leopard, tiger.
Still Problematic!

– Intransitivity
  • e.g. car seat is chair, chair is furniture, but ...
– Multiple category membership
  • it’s not a tree, it’s a forest!

Clay Shirky, “Ontologies are Overrated”
Fundamental Problem with Categorization

Making decisions too early!
We should only categorize at run-time, once we know the task!
categories are losing…

vs.

Yahoo! vs. Google
On-the-fly Categorization?

1. Knowledge Transfer
2. Communication
Association instead of categorization

Ask not “what is this?”, ask “what is this like”

– Moshe Bar

• Exemplar Theory (Medin & Schaffer 1978, Nosofsky 1986, Krushke 1992)
  – categories represented in terms of remembered objects (exemplars)
  – Similarity is measured between input and all exemplars
  – think non-parametric density estimation

• Vanevar Bush (1945), Memex (MEMory EXtender)
  – Inspired hypertext, WWW, Google…
Bush’s Memex (1945)

- Store publications, correspondence, personal work, on microfilm
- Items retrieved rapidly using index codes
  - Builds on “rapid selector”
- Can annotate text with margin notes, comments
- Can construct a trail through the material and save it
  - Roots of hypertext
- Acts as an external memory
Visual Memex, a proposal
[Malisiewicz & Efros]

Nodes = instances
Edges = associations

types of edges:
• visual similarity
• spatial, temporal co-occurrence
• geometric structure
• language
• geography
• ..

New object
“What is this?”

“What is this like?”

Malisiewicz & Efros, CVPR’08
Visual Associations

• How are objects similar?

- Shape
- Color
Distance “Similarity” Functions

• Positive Linear Combinations of Elementary Distances Computed Over 14 Features

\[ D_e(z) = w_e \cdot d_{ez} \]
Learning Distance Functions

“similar” side

“dissimilar” side

Dcolor

Decision Boundary

Dshape

Focal Exemplar

Don’t Care
Visualizing Distance Functions (Training Set)

Query: car

Top Neighbors with Tex-Hist Dist:
- car
- car
- car
- car

Top Neighbors with Learned Dist:
- car
- car
- car
- car
- car
Visualizing Distance Functions (Training Set)
### Labels Crossing Boundary

<table>
<thead>
<tr>
<th>Label</th>
<th>Label</th>
<th>Confusion Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop sign</td>
<td>sign</td>
<td>7.8%</td>
</tr>
<tr>
<td>pole</td>
<td>streetlight</td>
<td>6.7%</td>
</tr>
<tr>
<td>motorcycle</td>
<td>motorbike</td>
<td>6.2%</td>
</tr>
<tr>
<td>mountains</td>
<td>mountain</td>
<td>6.2%</td>
</tr>
<tr>
<td>ground grass</td>
<td>sidewalk</td>
<td>3.7%</td>
</tr>
<tr>
<td>grass</td>
<td>lawn</td>
<td>3.6%</td>
</tr>
<tr>
<td>road highway</td>
<td>road</td>
<td>3.4%</td>
</tr>
<tr>
<td>painting</td>
<td>picture</td>
<td>3.4%</td>
</tr>
<tr>
<td>sidewalk</td>
<td>road</td>
<td>3.2%</td>
</tr>
<tr>
<td>cloud</td>
<td>sky</td>
<td>3.1%</td>
</tr>
<tr>
<td>grass</td>
<td>ground grass</td>
<td>3.1%</td>
</tr>
<tr>
<td>mountain</td>
<td>mountains</td>
<td>2.7%</td>
</tr>
</tbody>
</table>

Table 2: Top dozen label confusions discovered after distance function learning.
Image Parsing with Context

Figure 1: The Visual Memex graph encodes object similarity (solid black edge) and spatial context (dotted red edge) between pairs of object exemplars. A spatial context feature is stored for each context edge. The Memex graph can be used to interpret a new image (left) by associating image segments with exemplars in the graph (orange edges) and propagating the information.
Torralba’s Context Challenge
Torralba’s Context Challenge

Slide by Antonio Torralba
Torralba’s Context Challenge
Our Challenge Setup

Figure 2: Torralba’s Context Challenge: “How far can you go without running a local object detector?” The task is to reason about the identity of the hidden object (denoted by a “?”) without local information. In our category-free Visual Memex model, object predictions are generated in the form of exemplar associations for the hidden object. In a category-based model, the category of the hidden object is directly estimated.

Malisiewicz & Efros, NIPS’09
3 models

Visual Memex: exemplars, non-parametric object-object relationships
  • Recurse through the graph

Baseline: CoLA: categories, parametric object-object relationships

Reduced Memex: categories, non-parametric relationships
Qual. results
Figure 3: a.) Context Challenge confusion matrices for the 3 methods: Visual Memex, KDE, and CoLA. b.) Recognition Precision versus Recall when thresholding output based on confidence. c) Side by side comparison of the 3 methods’ accuracies for 30 categories.
Next Step: top-down segmentation
Take Home Message

• Categorization is not a goal in itself
  – Rather, it is a means for transferring knowledge onto a new instance

• Skipping explicit categorization might make things easier, not harder
  – The “harder intermediate problem” syndrome

• Keeping around all your data isn’t so bad...
  – you never know when you will need it
Questions?