To Categorize or Not to Categorize

Tamara Berg

CSE 591 Recognizing People, Objects, and Actions

Many slides from Antonio Torralba & Alyosha Efros
Object naming -> Object categorization

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

Slide by Fei Fei, Fergus & Torralba
Object recognition

We do it effortlessly yet no one knows how. The theories are appealing, but don’t explain much.

Definitions
Basic categories (Rosch)
Gestalt (Wertheimer)
Features (Barlow, Treisman, …)
Bayesian (Marr, Poggio, …)
What is recognition?

- Assigning an image of an object to a category. Calling a chair a chair, despite variations in style, viewpoint, rendering, and surrounding clutter.

- This match can serve various purposes:
  - Naming: “moose” “squirrel”.
  - Individual identification: “Bullwinkle” “Rocky”.
  - Recognition memory: I saw that moose before.
  - Matching: this moose and that moose are the same moose.
Invariance of recognition
Novel examples need to be recognized...
Occlusion: recognition when only part of an object is visible...
Categorization

A category exists whenever two or more distinguishable objects or events are treated equivalently. This equivalent treatment may take any number of forms, such as labeling distinct objects or events with the same name, or performing the same action on different objects. Stimulus situations are unique, but organisms do not treat them uniquely; they respond on the basis of past learning and categorization. In this sense, categorization may be considered one of the most basic functions of living creatures.

Mervis and Rosch 1981
Why Categorize?

1. Knowledge Transfer
2. Communication
Why do we care about categories?

Perception of function: We can perceive the 3D shape, texture, material properties, without knowing about objects. But, the concept of category encapsulates also information about what can we do with those objects.

“We therefore include the perception of function as a proper –indeed, crucial- subject for vision science”, from Vision Science, chapter 9, Palmer.
Why do we care about categories?

When we recognize an object we can make predictions about its behavior in the future, beyond what is immediately perceived.
The perception of function

• Direct perception (affordances): Gibson

 Flat surface  
 Horizontal  
 Knee-high  
...

Sittable  
 upon

• Mediated perception (Categorization)

 Flat surface  
 Horizontal  
 Knee-high  
...

Chair  
Chair

Sittable  
 upon

One caveat of this comparison: deciding that something is a chair might require access to more features than the ones needed to decide that we can sit on something... (it is a different level of categorization)
Direct perception

Some aspects of an object function can be perceived directly

• Functional form: Some forms clearly indicate a function (“sittable-upon”, container, cutting device, …)

It does not seem easy to sit-upon this...

Slide credit: A. Torralba
Direct perception

Some aspects of an object function can be perceived directly

• Observer relativity: Function is observer dependent
Limitations of Direct Perception

Objects of similar structure might have very different functions.

Not all functions seem to be available from direct visual information only.

The functions are the same at some level of description: we can put things inside in both and somebody will come later to empty them. However, we are not expected to put inside the same kinds of things...

Slide credit: A. Torralba
Limitations of Direct Perception

Visual appearance might be a very weak cue to function

- Propulsion system
- Strong protective surface
- Something that looks like a door
- Sure, I can travel to space on this object

Slide credit: A. Torralba
Indirect perception of function by categorization

Well... this requires object recognition (for more details, see entire course)
So, what do we use direct or indirect?

“It seems exceedingly unlikely (though logically possible) that we categorize everything in our visual fields”, Palmer.

**Hypothesis:** we categorize the objects that are relevant for a specific task that we have at hand, but we only extract affordances from the others.
How many categories?
“Muchas”
How many object categories are there?

~10,000 to 30,000

Biederman 1987
How many categories?

- Probably this question is not even specific enough to have an answer

Slide credit: A. Torralba
Which level of categorization is the right one?

Car is an object composed of:
  a few doors, four wheels (not all visible at all times), a roof,
  front lights, windshield

If you are thinking in buying a car, you might want to be a bit more specific about your categorization.

Slide credit: A. Torralba
From Wordnet, categories can be organized in hierarchies (tree structures are commonly used).

Categorical hierarchies

Slide credit: A. Torralba
Organizing “things” into categories

(1) Feature-based

- Definition: 
  disassembling a concept into a set of featural components

- Each feature is an essential element of the category: “for a thing to be an X, it must have that feature. Otherwise it is not an X”
Prototype Theory

• According to the prototype view, an object will be classified as an instance of a category if it is sufficiently similar to the prototype.

• Similarity ~ the number of features shared between an object and the prototype (however, some features should be weighted more heavily as being more central to the prototype than are other features).

Figure 7.3. Schematic of the prototype model. Although many exemplars are seen, only the prototype is stored. The prototype is updated continually to incorporate more experience with new exemplars.
Levels of Categorization

The idea of prototypes and typicality led to the study of levels of categorization.

- Rosch et al: “albeit concepts exist at many different levels of a hierarchy, one level is fundamental: basic level.”

- Basic level: the best compromise between grouping together similar objects, and distinguish among objects from the same category.

• Willingham (247)
# Levels of Categorization

**SUPERORDINATE LEVEL CATEGORIES**

**BASIC-LEVEL CATEGORIES**

**SUBORDINATE LEVEL CATEGORIES**

<table>
<thead>
<tr>
<th>Superordinate Level</th>
<th>Basic Level</th>
<th>Subordinate Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musical instrument</td>
<td>Guitar</td>
<td>Folk guitar</td>
</tr>
<tr>
<td></td>
<td>Piano</td>
<td>Grand piano</td>
</tr>
<tr>
<td>Fruit</td>
<td>Peach</td>
<td>Freestone peach</td>
</tr>
<tr>
<td></td>
<td>Grapes</td>
<td>Concord grapes</td>
</tr>
<tr>
<td>Tree</td>
<td>Maple</td>
<td>Silver maple</td>
</tr>
<tr>
<td></td>
<td>Birch</td>
<td>River birch</td>
</tr>
<tr>
<td></td>
<td>Oak</td>
<td>White oak</td>
</tr>
</tbody>
</table>

*Table 7.3. Examples of Nested Category Structures*

*Slide credit: A. Torralba*
Rosch’s Levels of Categorization

Definition of Basic Level:

- **Similar shape**: Basic level categories are the highest-level category for which their members have similar shapes.

- **Similar motor interactions**: ... for which people interact with its members using similar motor sequences.

- **Common attributes**: ... there are a significant number of attributes in common between pairs of members.

Similarity declines only slightly going from subordinate to basic level, and then drops dramatically.
Levels of Categorization

• Rosch et al (1976) found that when people are shown pictures of objects, they identify objects at a basic level more quickly than they identified objects at higher or lower levels.

• Objects appear to be recognized first at their basic level, and only afterwards they are classified in terms of higher or lowers level categories.
Typicality effects

- Typicality: how good or common an item is a member of a given category.

- The typical exemplar is like a representation of the average (or central tendency).

- But, the representation of a category vary with experience, so does the “typical” exemplar.

### Table 7.1. Some Results of Rosch’s (1973) Typicality Ratings

<table>
<thead>
<tr>
<th>Category</th>
<th>Member</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruit</td>
<td>Apple</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>Plum</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Pineapple</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Strawberry</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Fig</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>Olive</td>
<td>6.2</td>
</tr>
<tr>
<td>Sport</td>
<td>Football</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Hockey</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Wrestling</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Archery</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>Gymnastics</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>Weight-lifting</td>
<td>4.7</td>
</tr>
<tr>
<td>Bird</td>
<td>Robin</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Eagle</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Wren</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Chicken</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Ostrich</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Bat</td>
<td>5.8</td>
</tr>
<tr>
<td>Vehicle</td>
<td>Car</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Boat</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>Scooter</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Tricycle</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>Horse</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>Skis</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Slide credit: A. Torralba
Entry-level categories
(Jolicoeur, Gluck, Kosslyn 1984)

• Typical members of a basic-level category are categorized at the expected level

• Atypical members tend to be classified at a subordinate level.

A bird
An ostrich

Slide credit: A. Torralba
We do not need to recognize the exact category

A new class can borrow information from similar categories
Are Categories Necessary?

Leci n'est pas une pipe.

Alexei (Alyosha) Efros
CMU
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
Are categories necessary?
[Hays & Efros, SIGGRAPH’07]
2 Million Flickr Images
Lots of data available
• 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
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]

New object

Nodes = instances
Edges = associations

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

Slide credit: A. Efros
“What is this?”

Input Image


Slide credit: A. Efros
“What is this *like*?”

Malisiewicz & Efros, CVPR’08
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.
Visual Associations

• How are objects similar?

Slide credit: A. Efros
Distance “Similarity” Functions

- Positive Linear Combinations of Elementary Distances Computed Over 14 Features

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

Building e Distance Function

Slide credit: A. Efros
Learning Distance Functions

- "similar" side
- "dissimilar" side
- Decision Boundary
- Focal Exemplar
- Dshape

Slide credit: A. Efros
Visualizing Distance Functions (Training Set)

Query: car

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

Top Neighbors with Learned Dist: car, car, car, car

Distance Function:
- Bot Height
- Top Height
- Abs Mask
- Color-Hist
- Color Std
- Mean Color
- Interior Tex-Hist
- Tex-Hist Bot
- Tex-Hist Left
- Tex-Hist Top
- Tex-Hist Right
- Size
- BB
- Centered Mask

Slide credit: A. Efros
Visualizing Distance Functions (Training Set)
Contextual Associations
Evaluation:
Torralba’s Context Challenge
Torralba’s Context Challenge
Qual. results

Input Image + Hidden Region

Visual Memex Exemplar Predictions

Categorization Results

Visual Memex
- car
- trash
- streetlight

KOE
- car
- wall
- sidewalk
glass

COLA
- person
car
building

Visual Memex
- table
- door
- wall
door

KOE
- table
- door
- wall
door

COLA
- sidewalk
car
road
person

Visual Memex
- mountain
- ground
- streetlight
- road

KOE
- ground
- grass
- mountains
- building

COLA
- building
dry
- tree
- mountains

Slide credit: A. Efros
Quant. 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.
Conclusions?