Are Categories Necessary?

Alexei (Alyosha) Efros
CMU

Before We Begin…

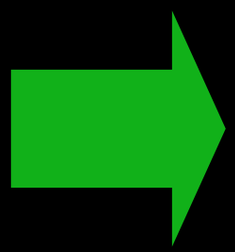
- Not an AI/learning person
- My work is in Computer Vision and Computer Graphics
- But I want to utilize lots of data

A Confession

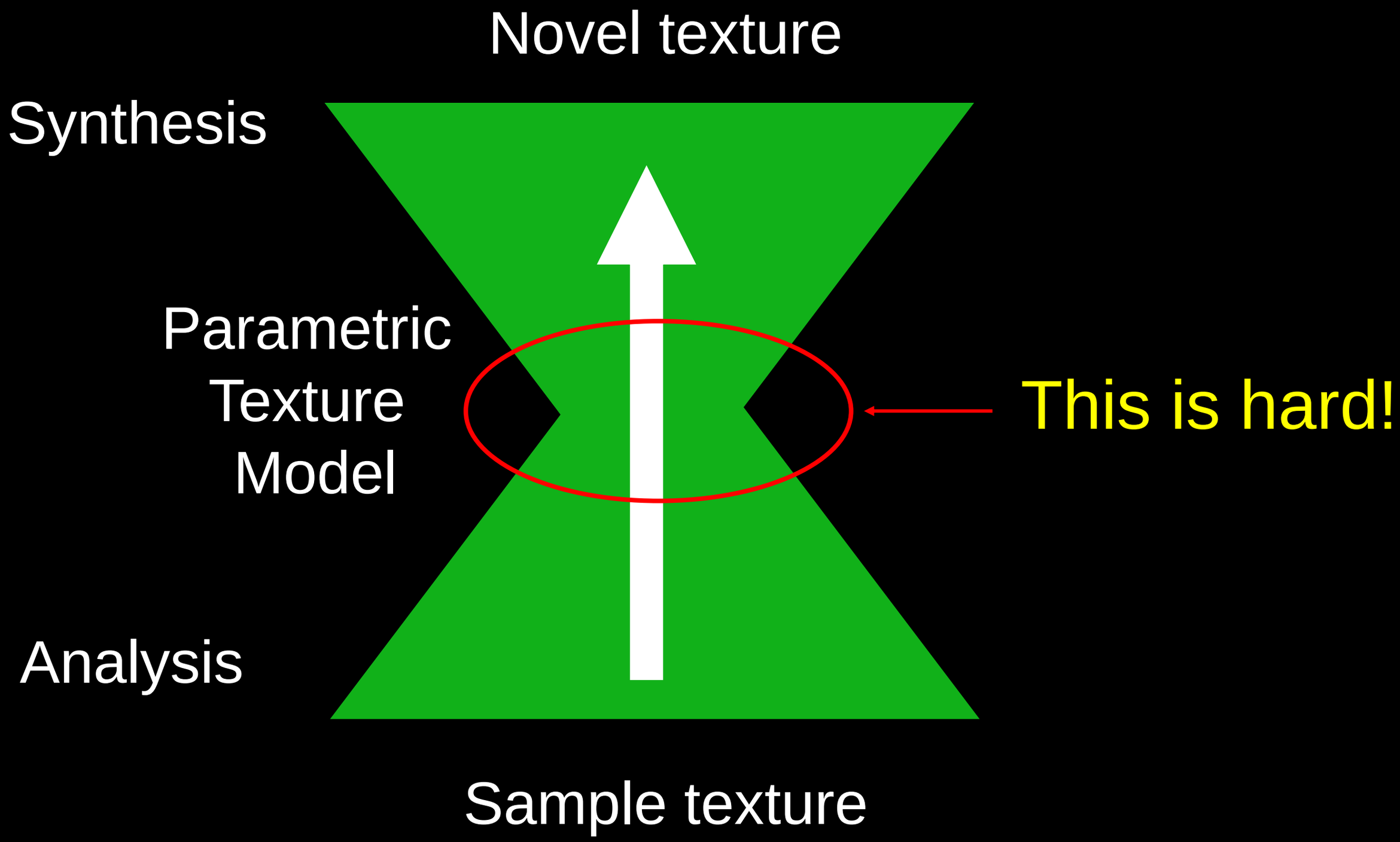
I have a problem…

…I am a nearest-neighbor addict!

Texture Synthesis



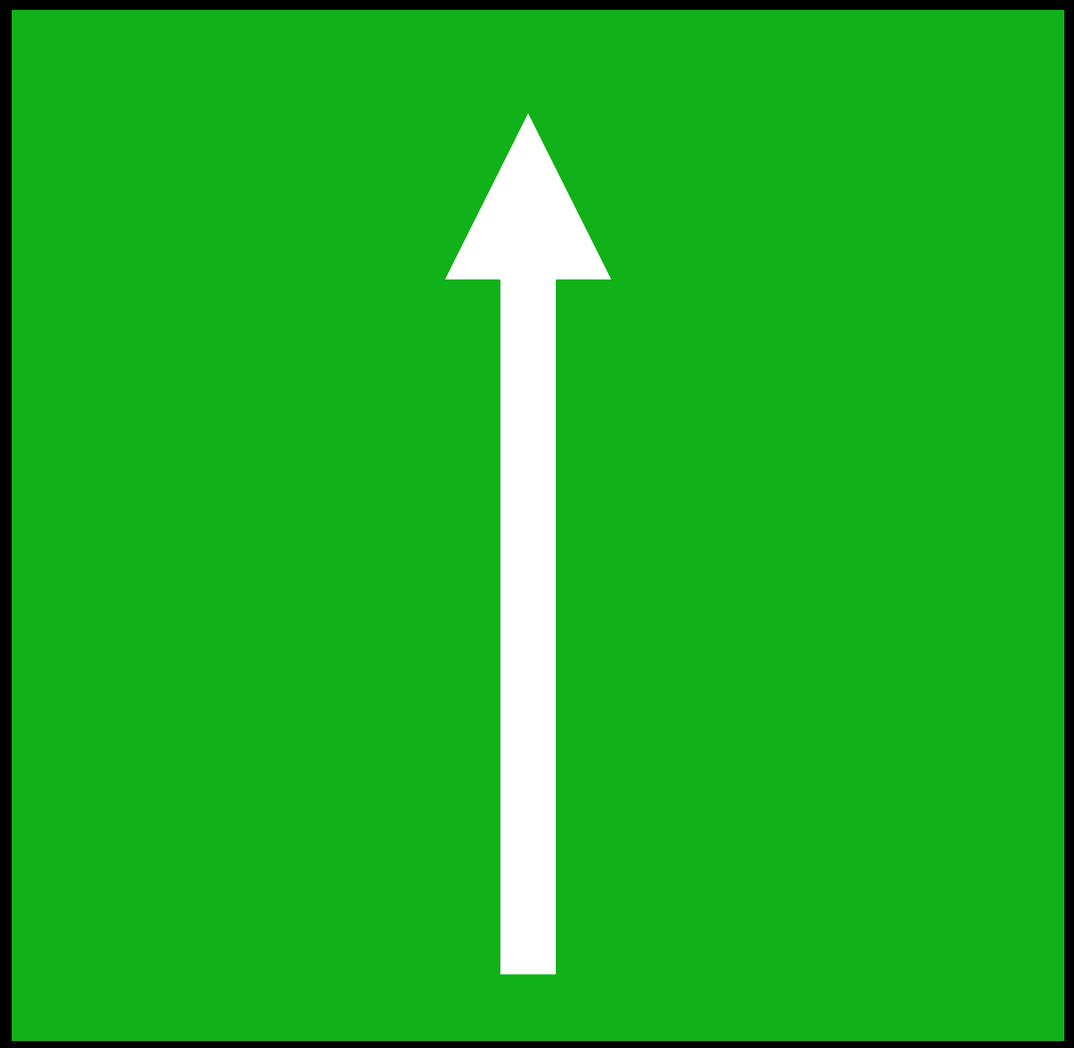
Classical Texture Synthesis



Non-parametric Approach

Novel texture

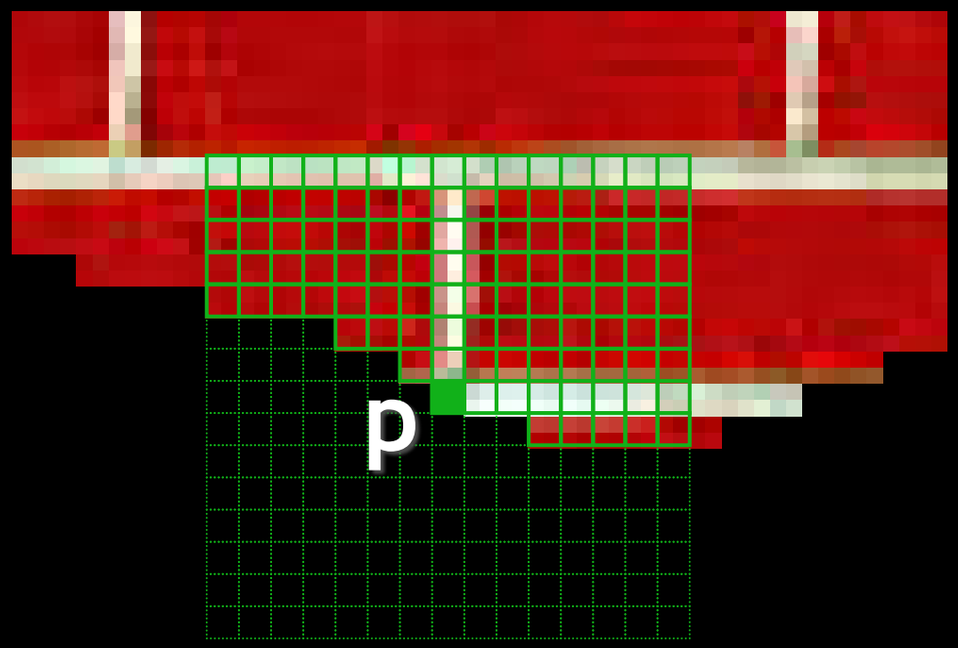
Synthesis



Analysis

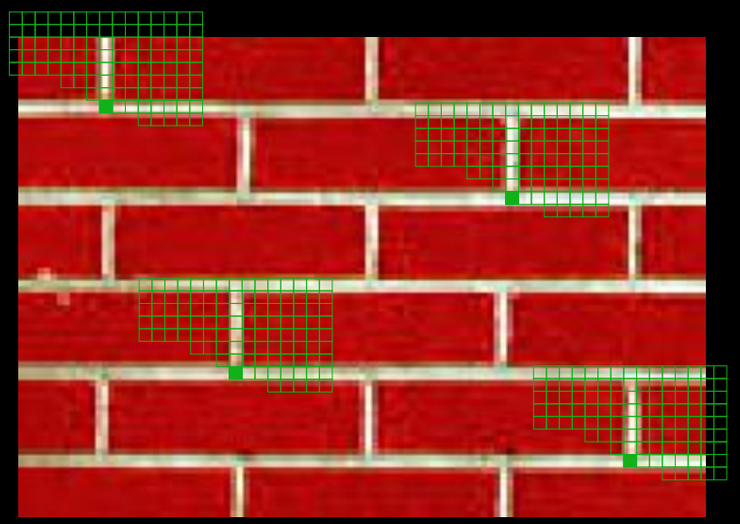
Sample texture

Non-parametric Synthesis [Efros&Leung'99]



Synthesizing a pixel

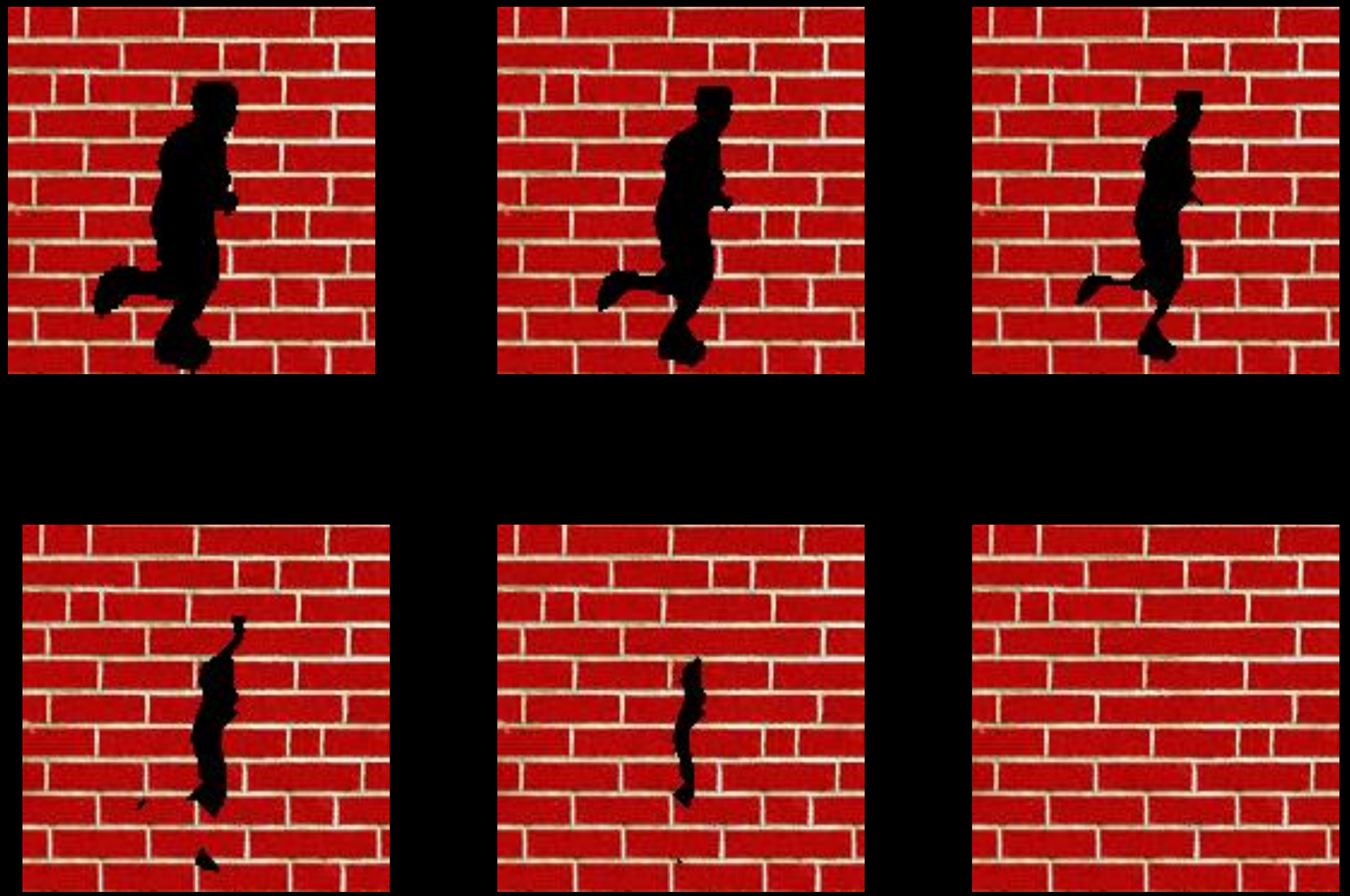
non-parametric sampling

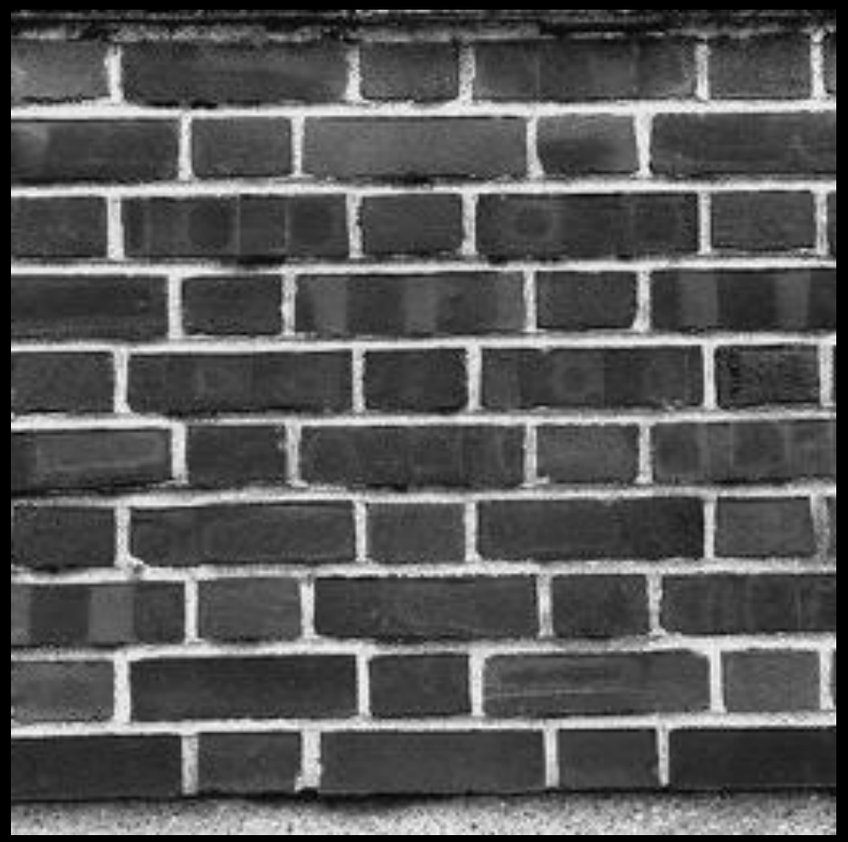


Input image

- 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

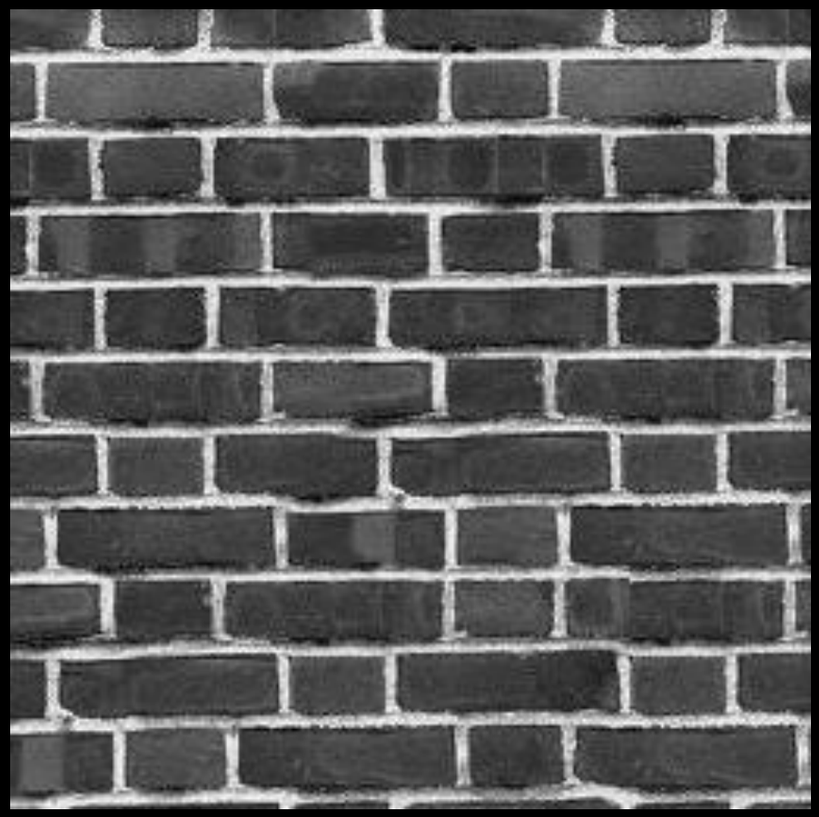
Hole Filling



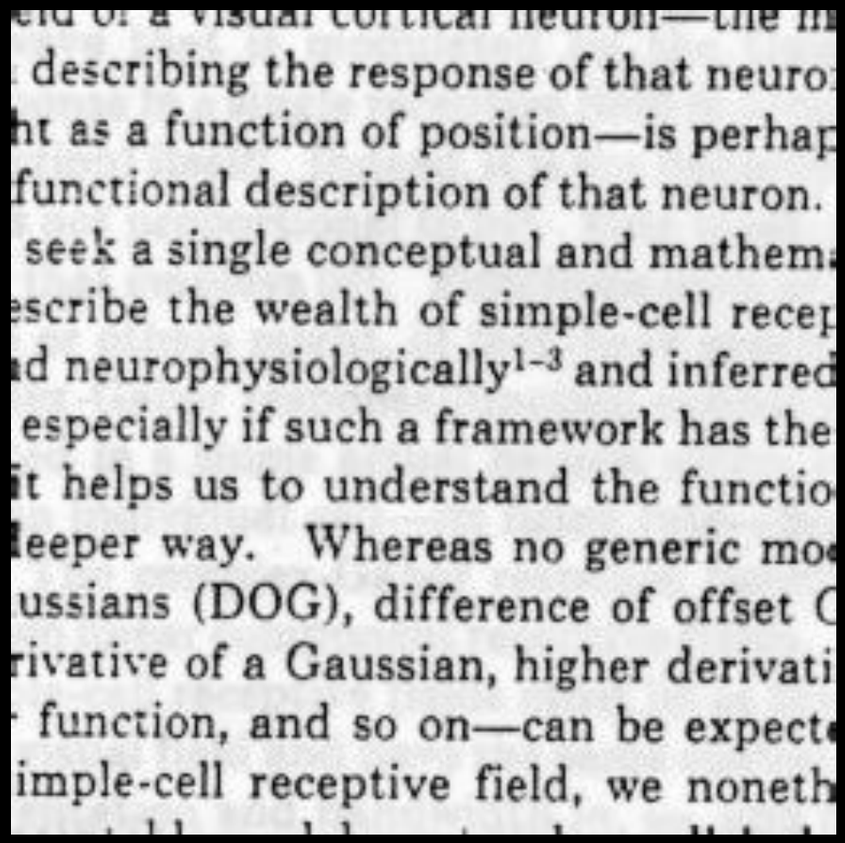


input image

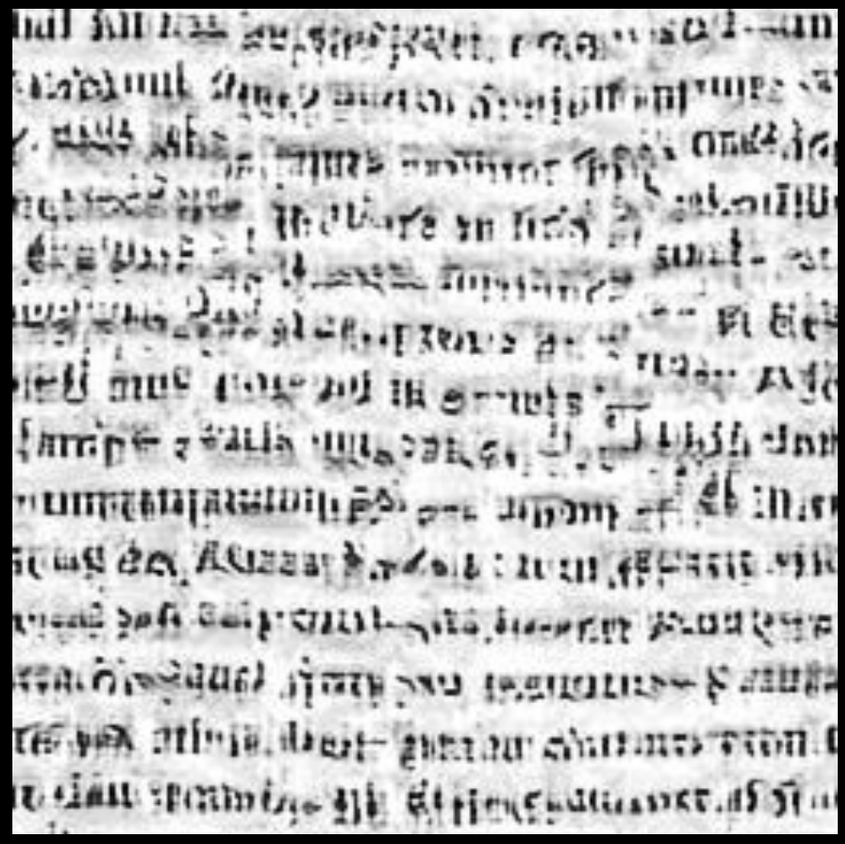
Portilla & Simoncelli



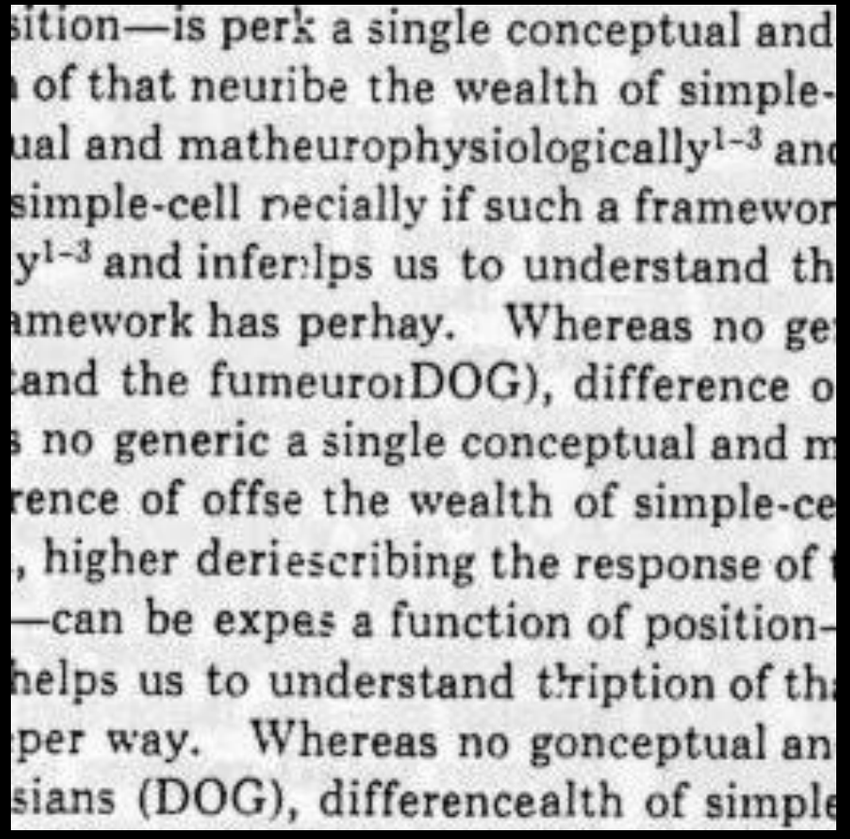
Our algorithm


eld of a visual cortical neuron—the m
describing the response of that neuro
ht as a function of position—is perhap
functional description of that neuron.
seek a single conceptual and mathem
escribe the wealth of simple-cell recep
d neurophysiologically[1–3] and inferred
especially if such a framework has the
it helps us to understand the functio
leeper way. Whereas no generic mo
ussians (DOG), difference of offset G
rivative of a Gaussian, higher derivati
function, and so on—can be expecte
imple-cell receptive field, we noneth

input image

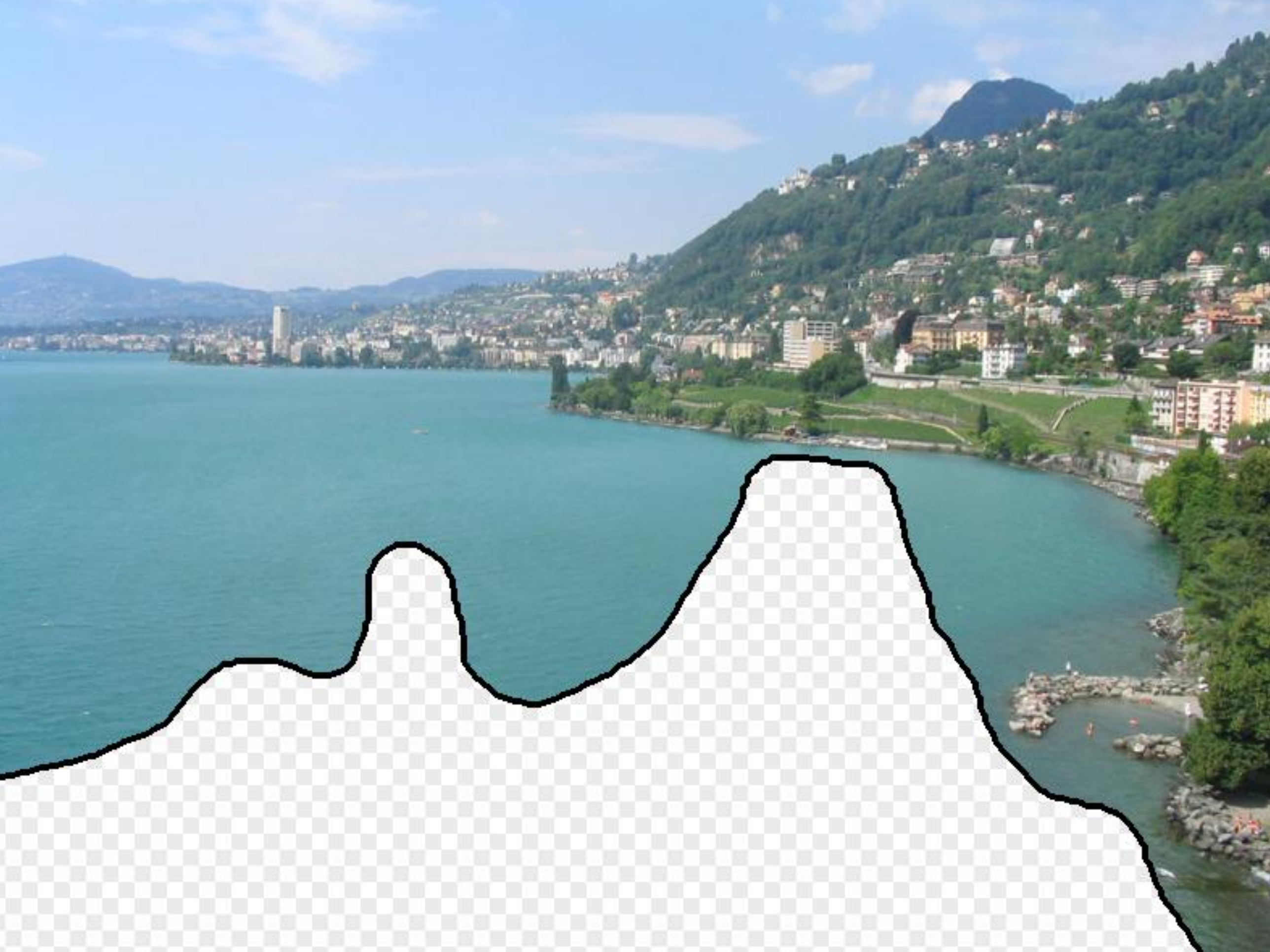


Portilla & Simoncelli

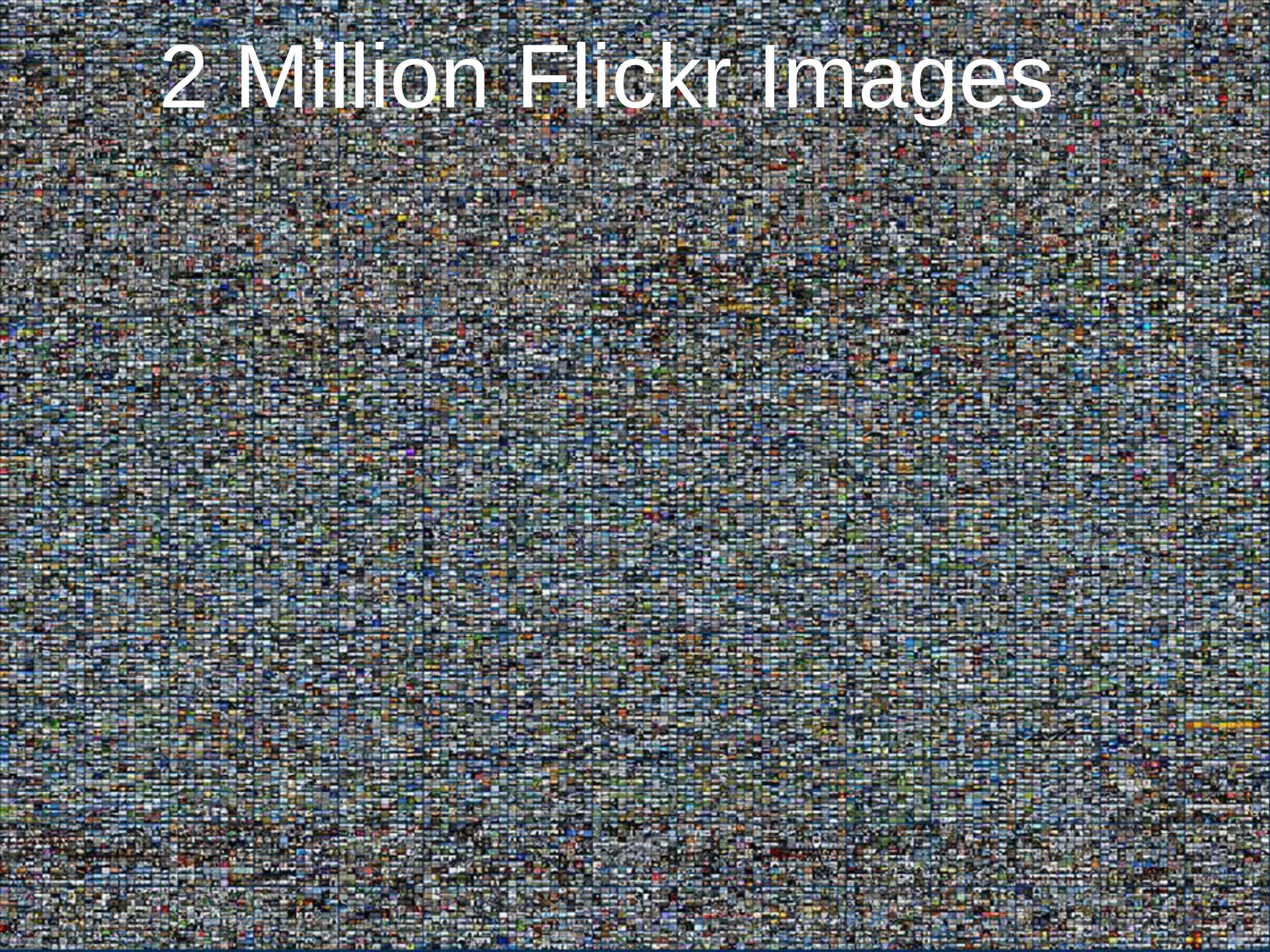

sition—is perk a single conceptual and
of that neuribe the wealth of simple-
ual and matheurophysiologically[1–3] and
simple-cell necially if such a framewor
y[1–3] and infer lps us to understand th
amework has perhay. Whereas no ge
and the fumeuroDOG), difference o
s no generic a single conceptual and m
rence of offse the wealth of simple-ce
, higher deriescribing the response of
—can be expes a function of position-
helps us to understand tription of th
per way. Whereas no gonceptual an
sians (DOG), differencealth of simple

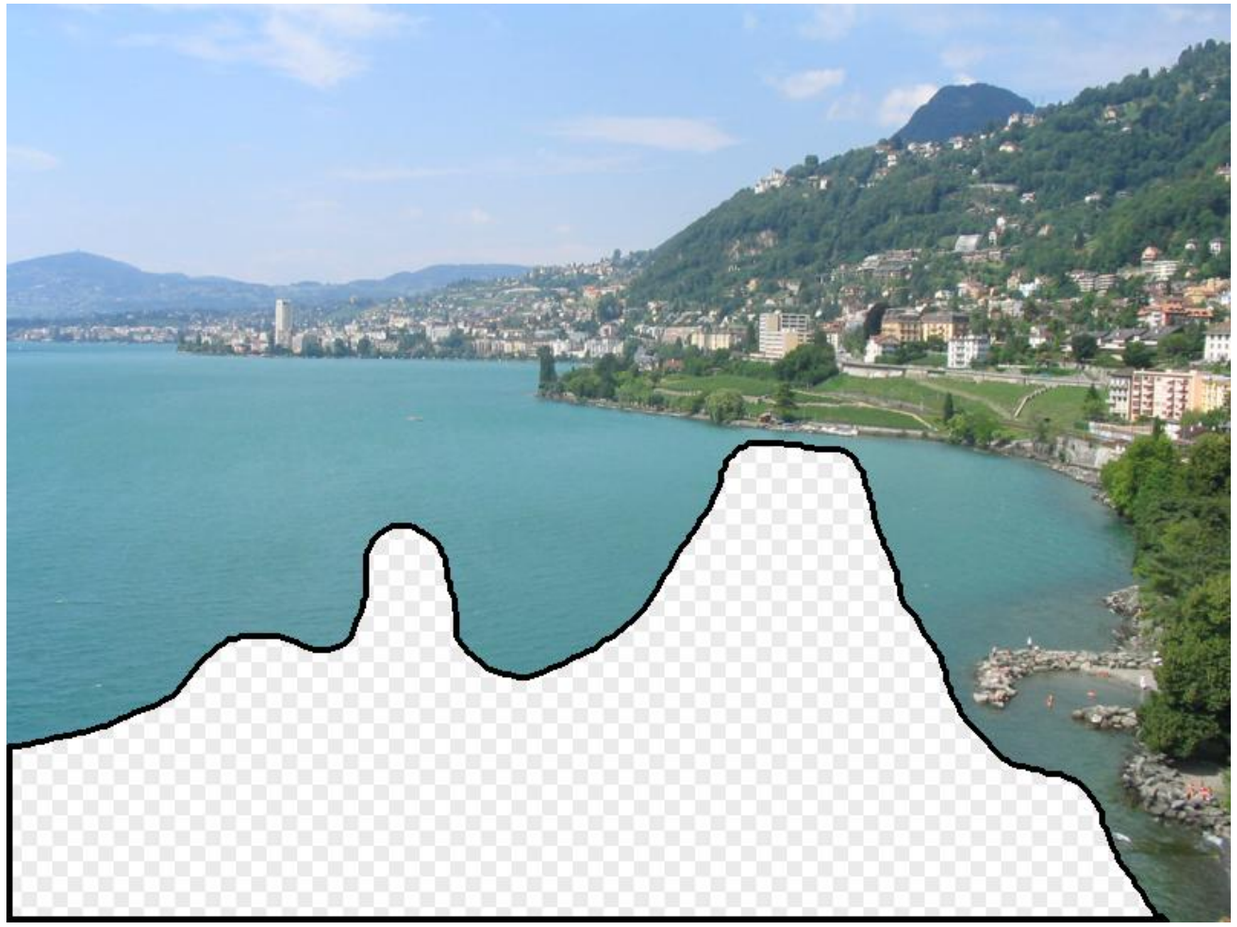
Our algorithm

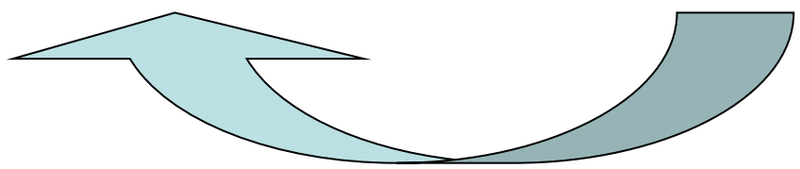
[Hays & Efros, SIGGRAPH'07]



2 Million Flickr Images



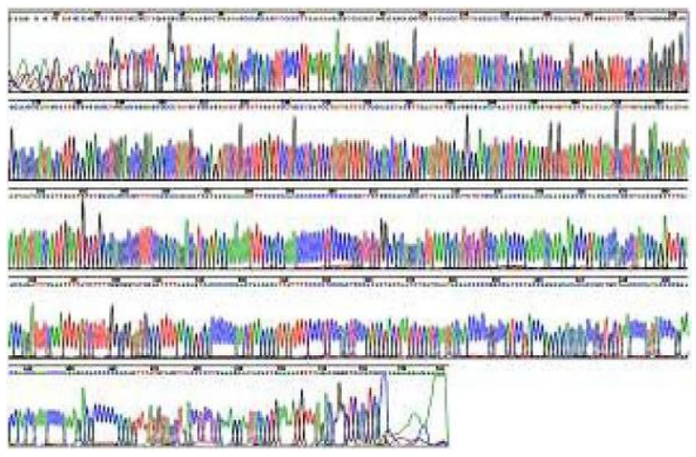




Two Kinds of Things in the World

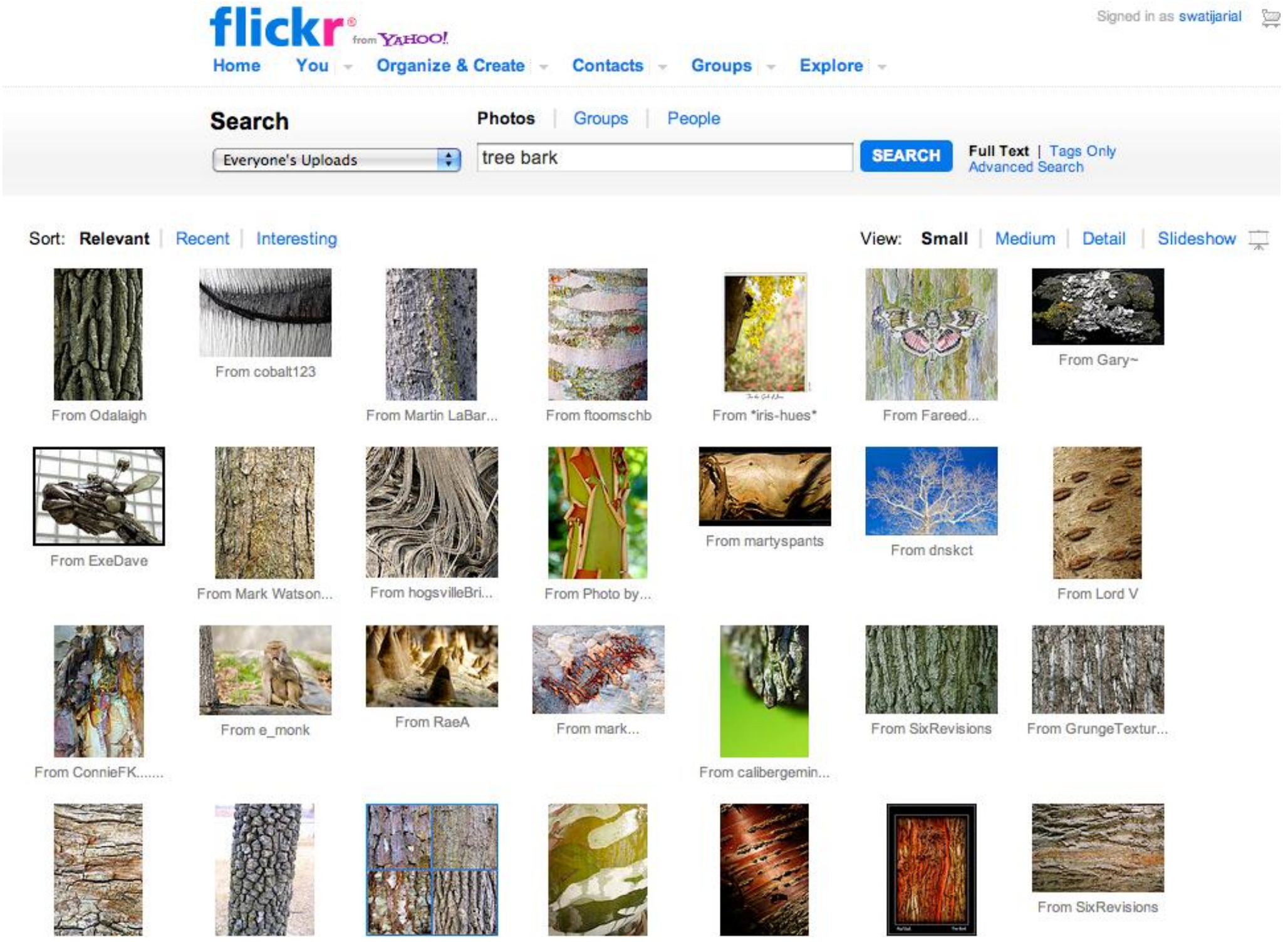
Navier-Stokes Equation

$$\frac{\partial \mathbf{u}}{\partial t} = -\left(\mathbf{u} \cdot \nabla\right)\mathbf{u} + v\nabla^2\mathbf{u} - \frac{1}{d}\nabla p + \mathbf{f}$$



+ 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

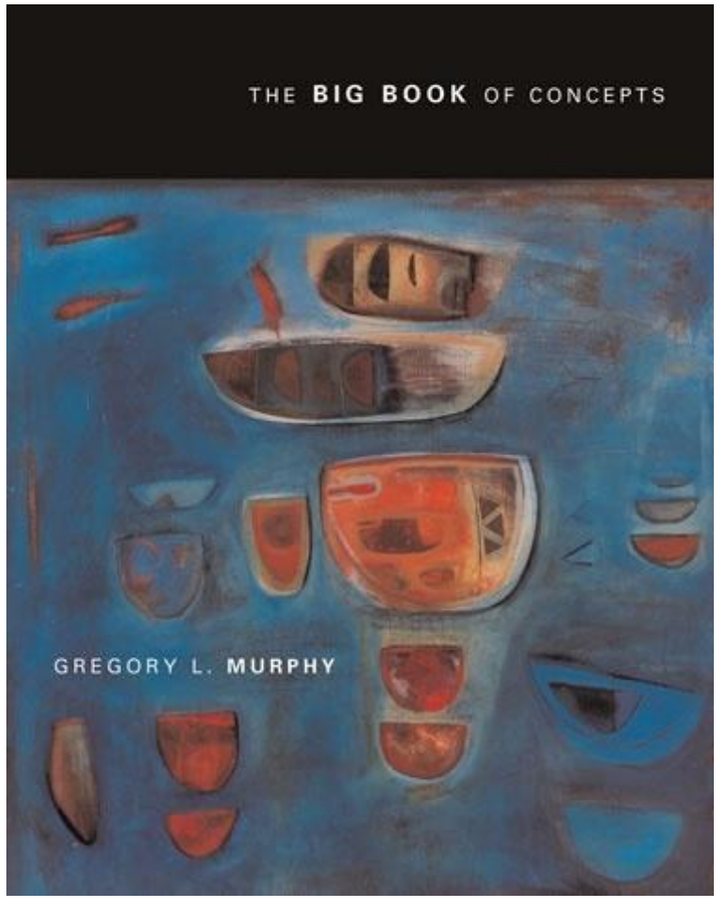
- A.I. for the postmodern world:
  - all questions have already been answered…many times, in many ways
  - Google is dumb, the "intelligence" is <u>in the data</u>

Are Categories Necessary?

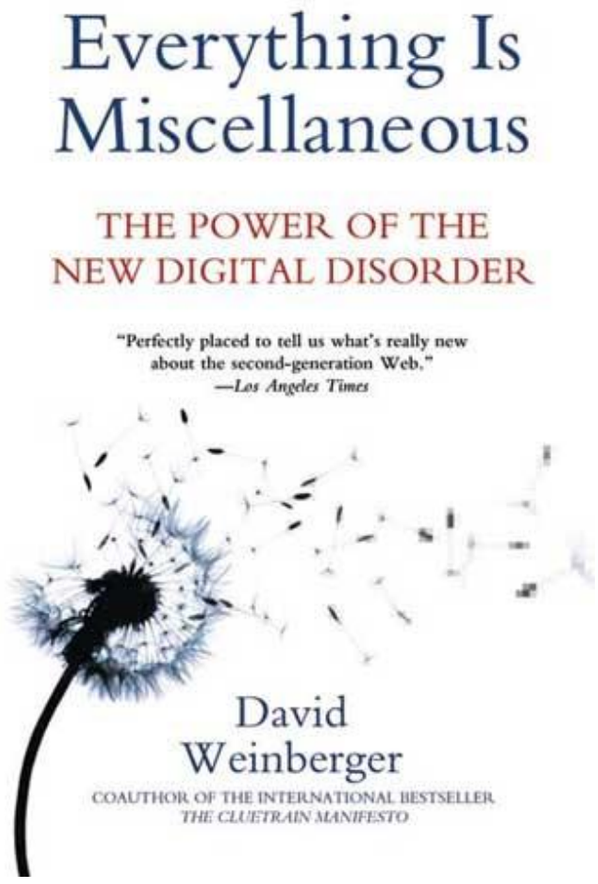
Alexei (Alyosha) Efros
CMU

Joint work with
Tomasz Malisiewicz

Acknowledgements



Murphy
Big Book of Concepts



Weinberger
Everything is Miscellaneous

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

slide by Fei Fei, Fergus & Torralba

Object categorization

sky

building

flag

face

banner

wall

street lamp

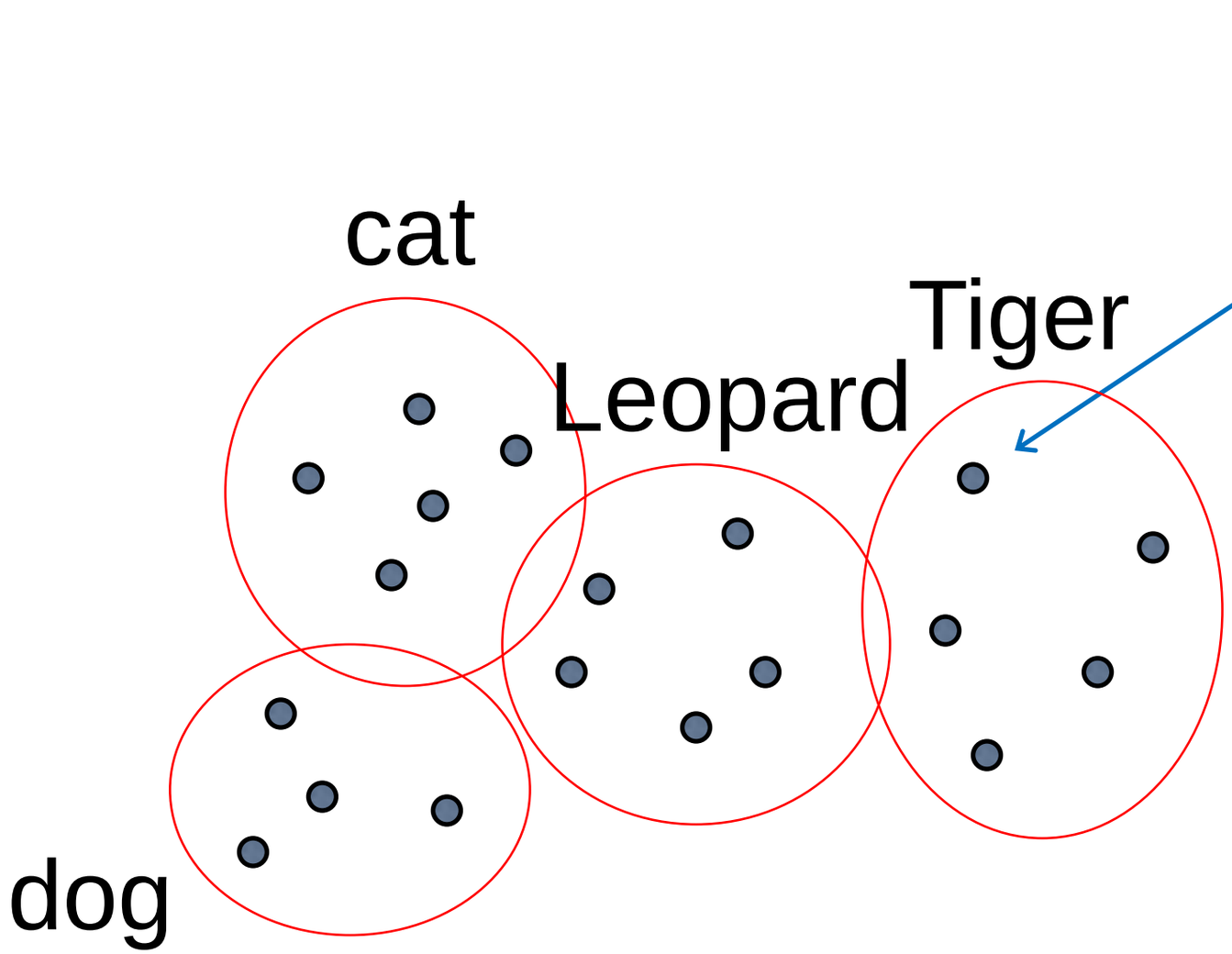
bus

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

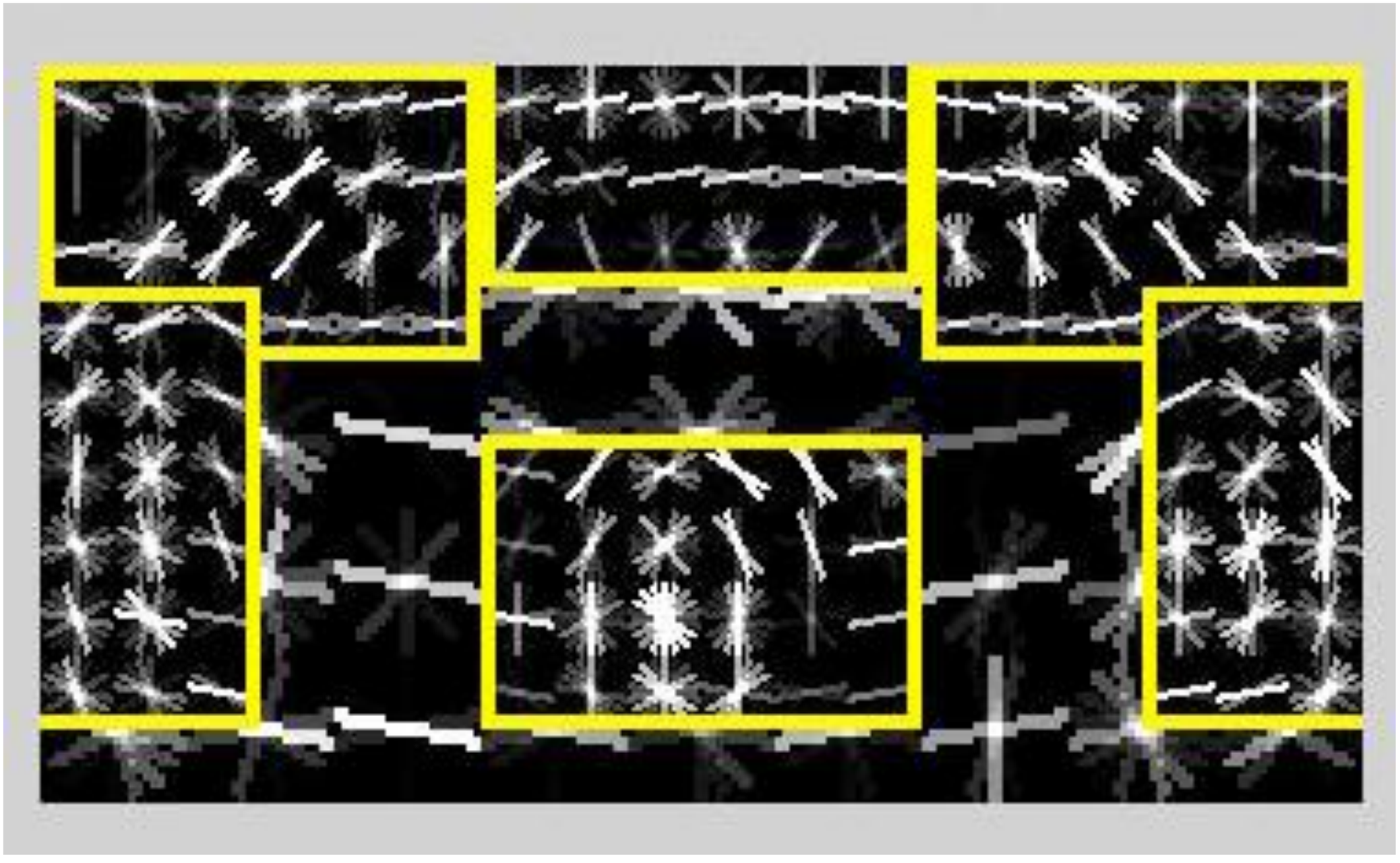
Chair

- World is too varied

- Categories are 3D, but images are 2D

car

Typical HOG car detector



Felzenszwalb et al, PASCAL 2007

Why not?

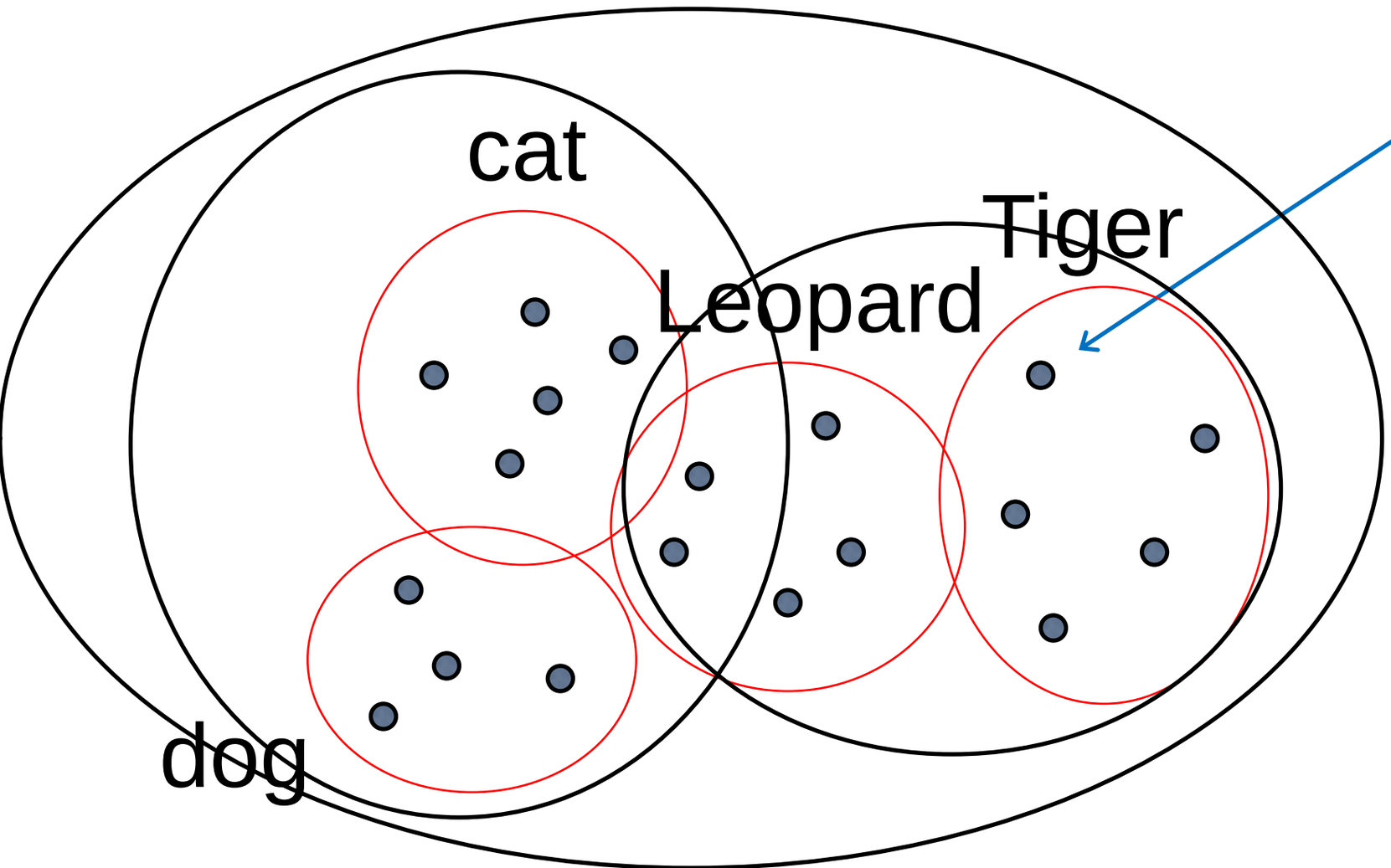
ABSOLUT
Country of Sweden
VODKA
This superb vodka
was distilled from grain grown
in the rich fields of southern Sweden.
It has been produced at the famous
old distilleries near Åhus
in accordance with more than
400 years of Swedish tradition.
Vodka has been sold under the name
Absolut since 1879.
40% ALC./VOL. (80 PROOF) 1 LITER
IMPORTED
PRODUCED AND BOTTLED IN AHUS, SWEDEN
BY THE ABSOLUT COMPANY
A DIVISION OF V&S VIN&SPRIT AB.

+

Solution: hierarchy?

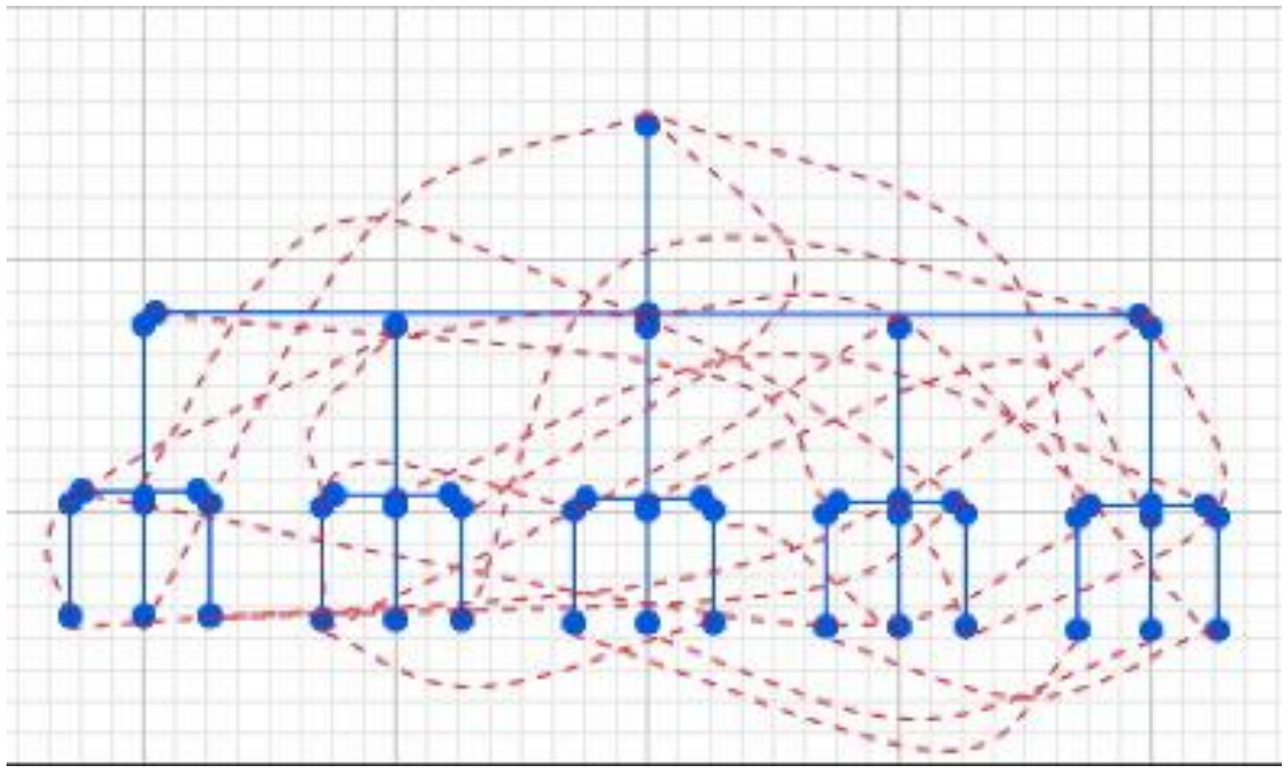
Ontologies, hierarchies, levels of categories (Rosch), etc.

WordNet, ImageNet, etc etc



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!

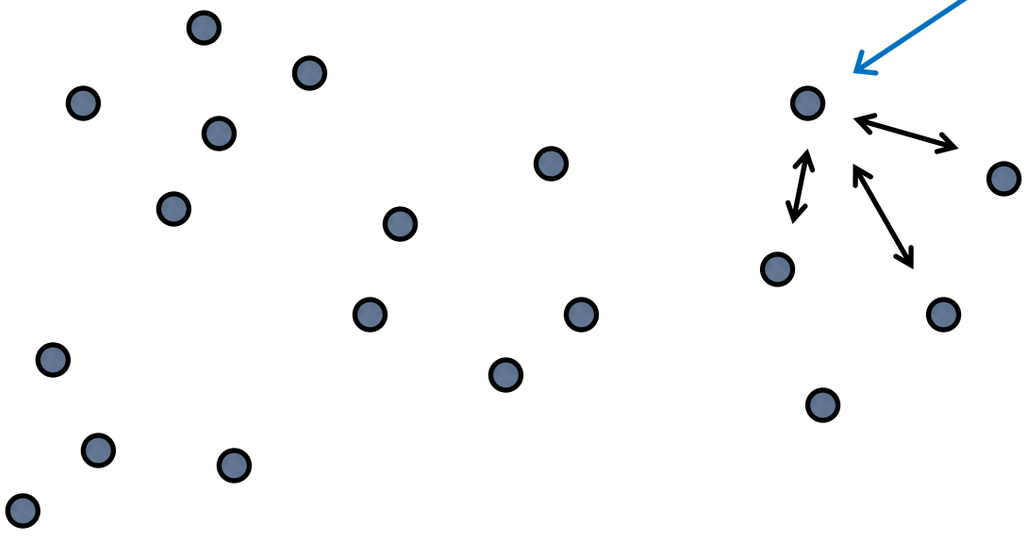
The Dictatorship of Librarians

categories are losing…

vs.

On-the-fly Categorization?

1. Knowledge Transfer
2. ~~Communication~~



Association instead of categorization

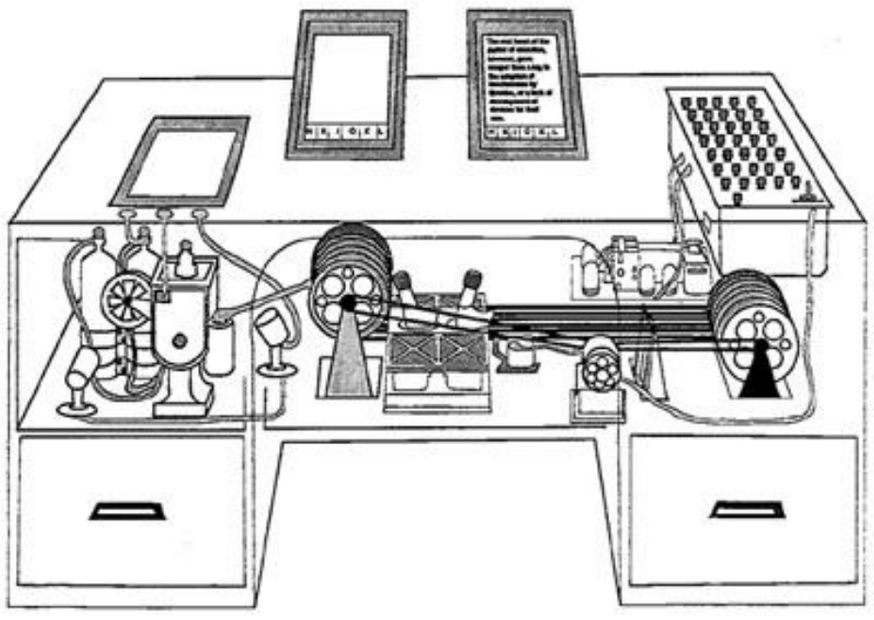
Ask not "what is this?", ask "what is this <u>like</u>"

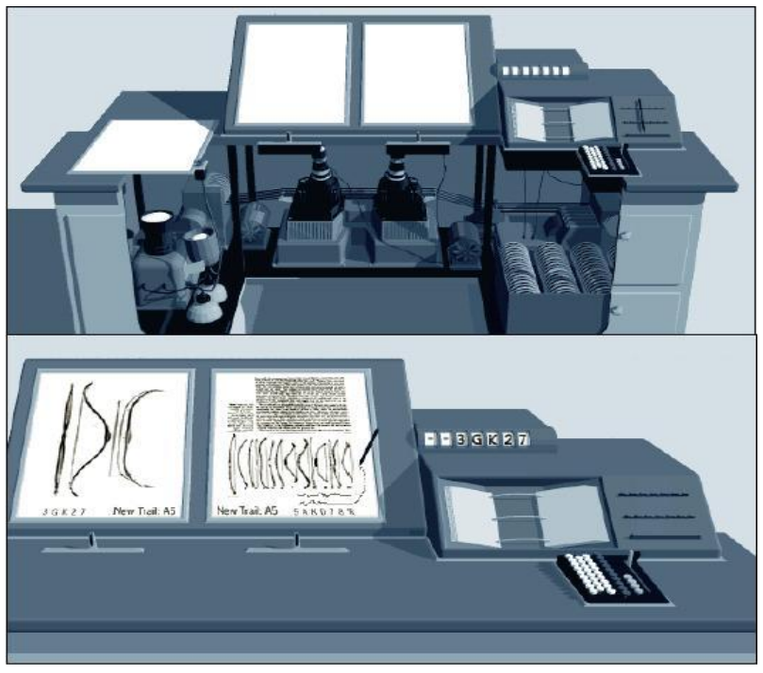
– Moshe Bar

- 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), <u>Memex</u> (MEMory EXtender)
  - –Inspired hypertext, WWW, Google…

Bush's Memex (1945)

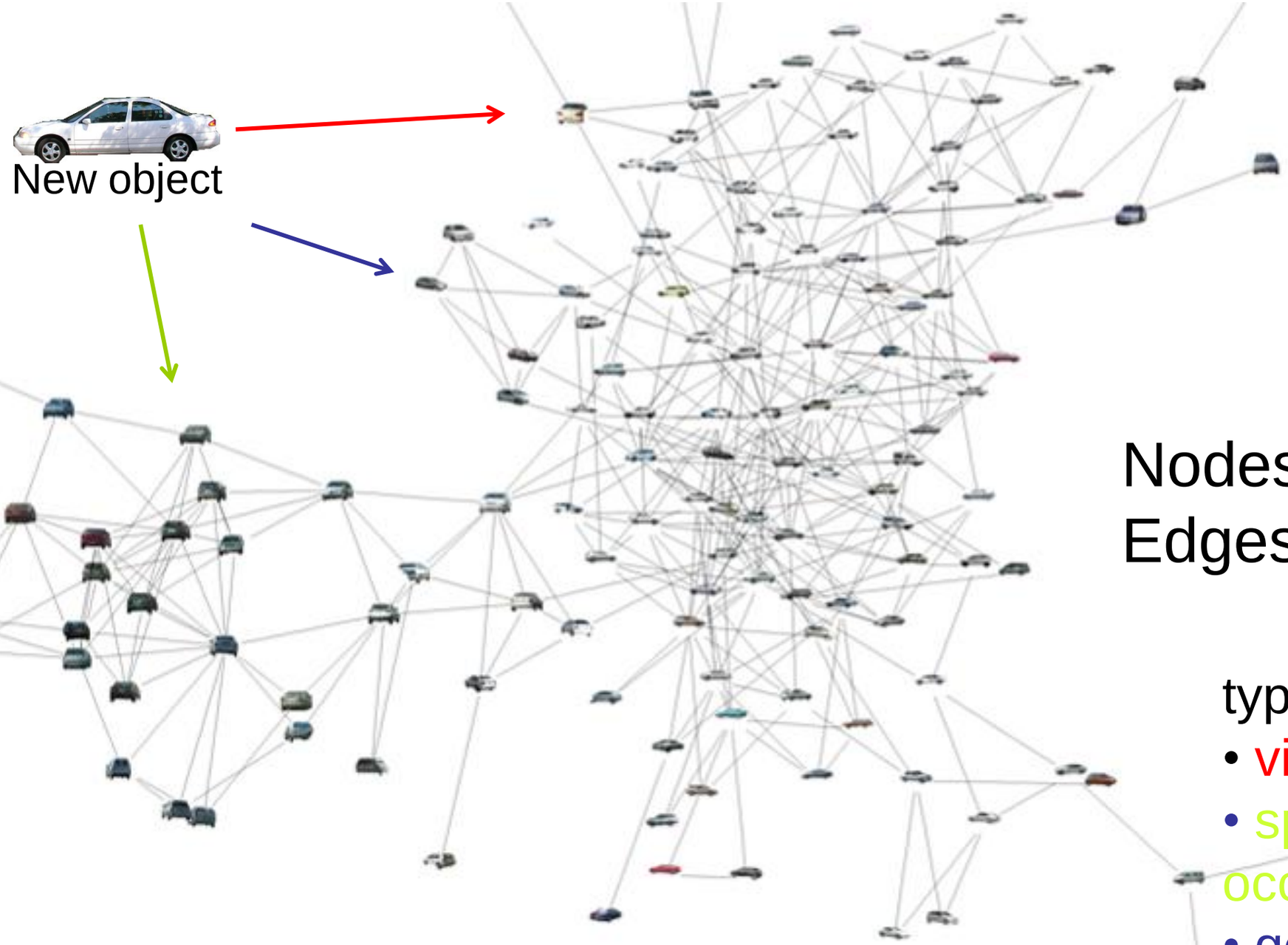
- Store publications, correspondence, personal work, on microfilm
- Items retrieved rapidly using index codes
  - Builds on "rapid selector"
- Can annotate text with margin notes, comments
- Can construct a *trail* through the material and save it
  - Roots of hypertext
- Acts as an external memory





Visual Memex, a proposal

[Malisiewicz & Efros]



Nodes = instances
Edges = associations

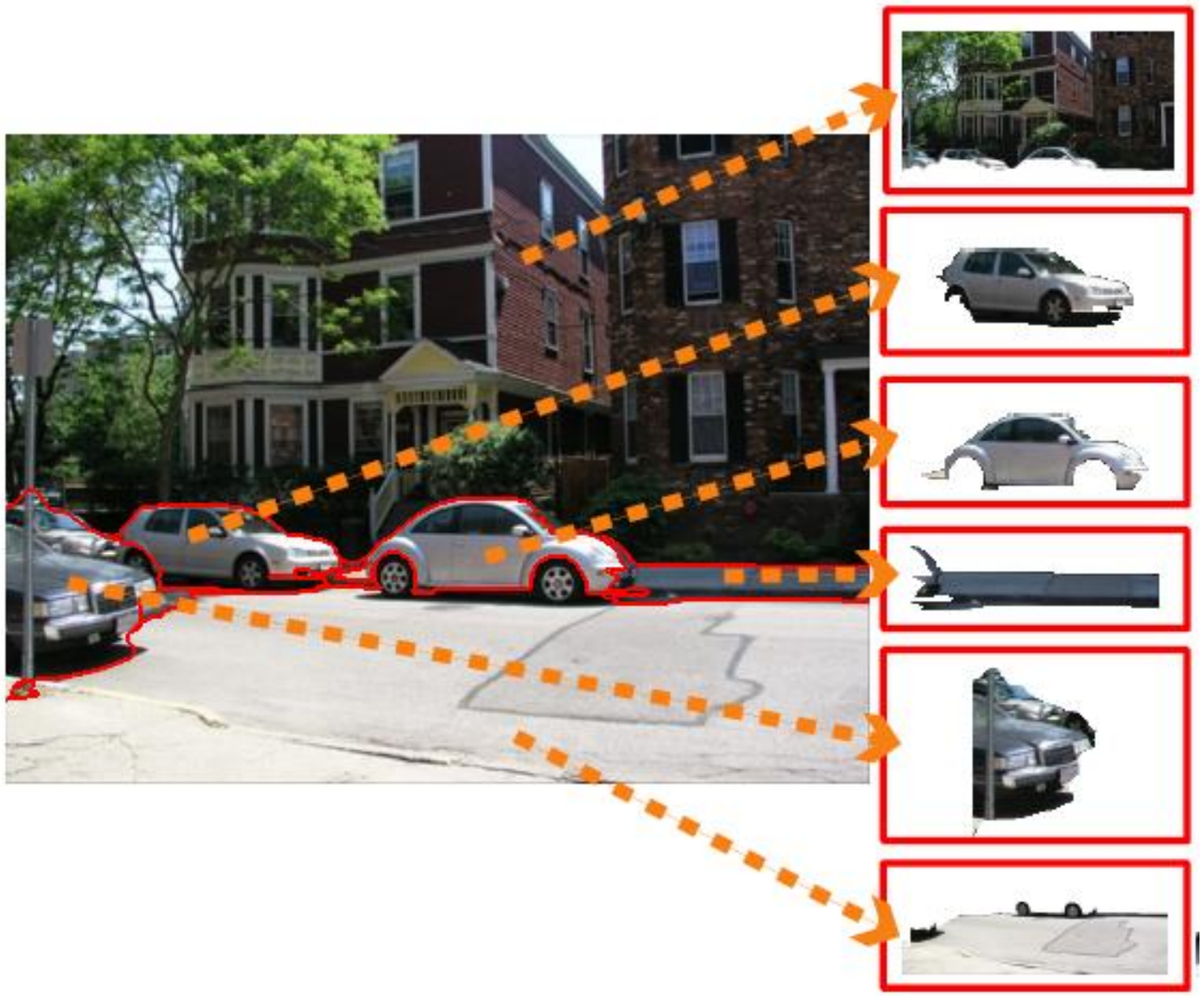
types of edges:
- visual similarity
- spatial, temporal co-occurrence
- geometric structure
- language
- geography
- ..

“What is this?”

Input Image

He 2004, Tu 2004, Shotton 2006, Galleguillos 2008, Fei-Fei 2009, Gould 2009, etc.

“What is this *like*?”



Malisiewicz & Efros, CVPR’08

Visual Associations

- How are objects similar?

Shape

~~Color~~

~~Shape~~

Color

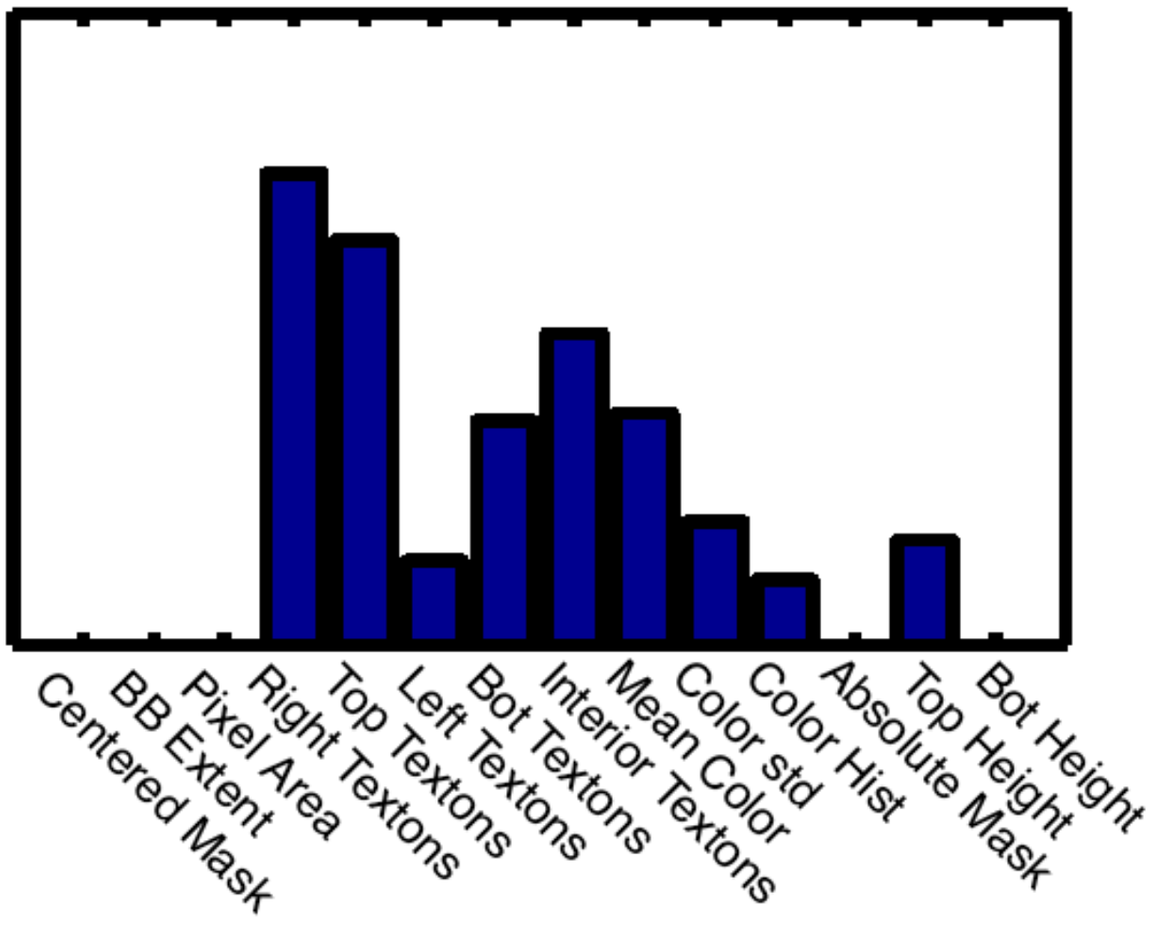
Distance “Similarity” Functions

- Positive Linear Combinations of Elementary Distances Computed Over 14 Features

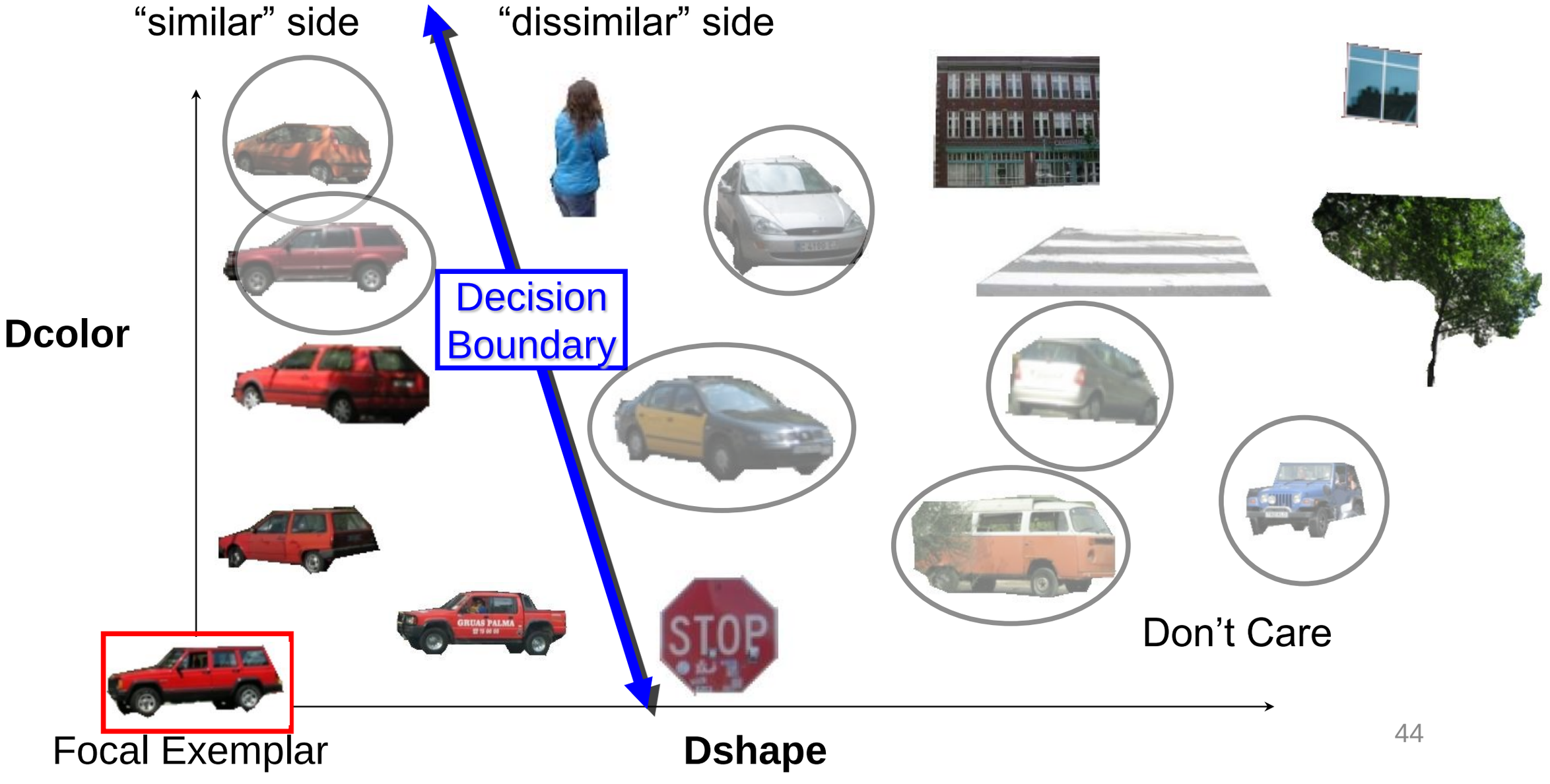
$$D_e(z) = \mathbf{w}_e \cdot \mathbf{d}_{ez}$$

Building e

Building e Distance Function



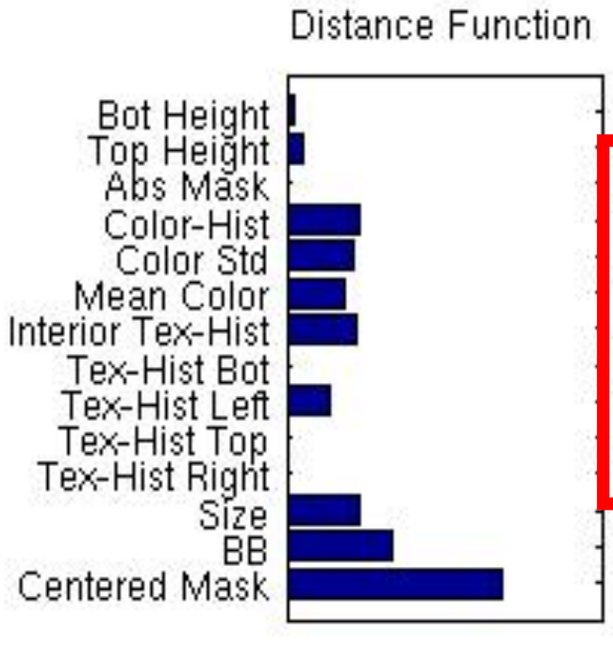
Learning Distance Functions



Visualizing Distance Functions (Training Set)

Query

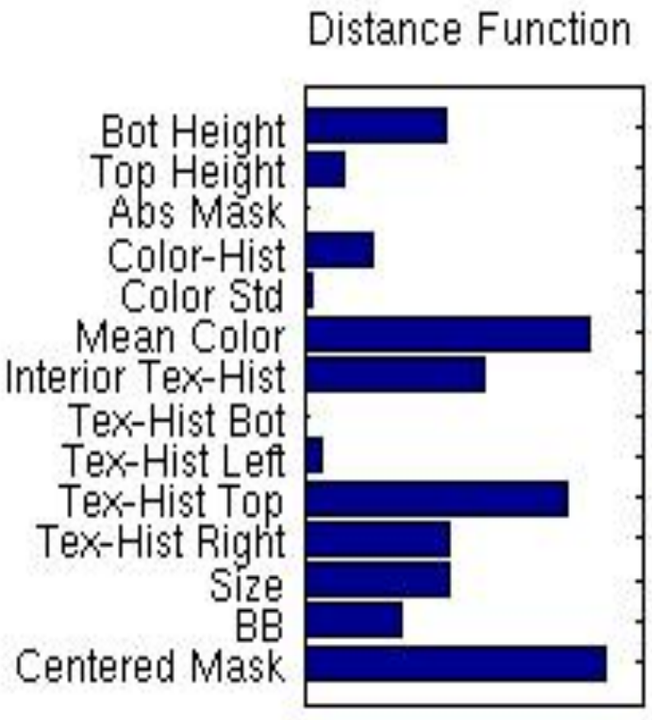
Top Neighbors with Tex-Hist Dist



Query

Top Neighbors with Learned Dist

Visualizing Distance Functions (Training Set)



Labels Crossing Boundary

stop sign	sign	7.8%
pole	streetlight	6.7%
motorcycle	motorbike	6.2%
mountains	mountain	6.2%
ground grass	sidewalk	3.7%
grass	lawn	3.6%
road highway	road	3.4%
painting	picture	3.4%
sidewalk	road	3.2%
cloud	sky	3.1%
grass	ground grass	3.1%
mountain	mountains	2.7%

Table 2: Top dozen label confusions discovered after distance function learning.

Image Parsing with Context

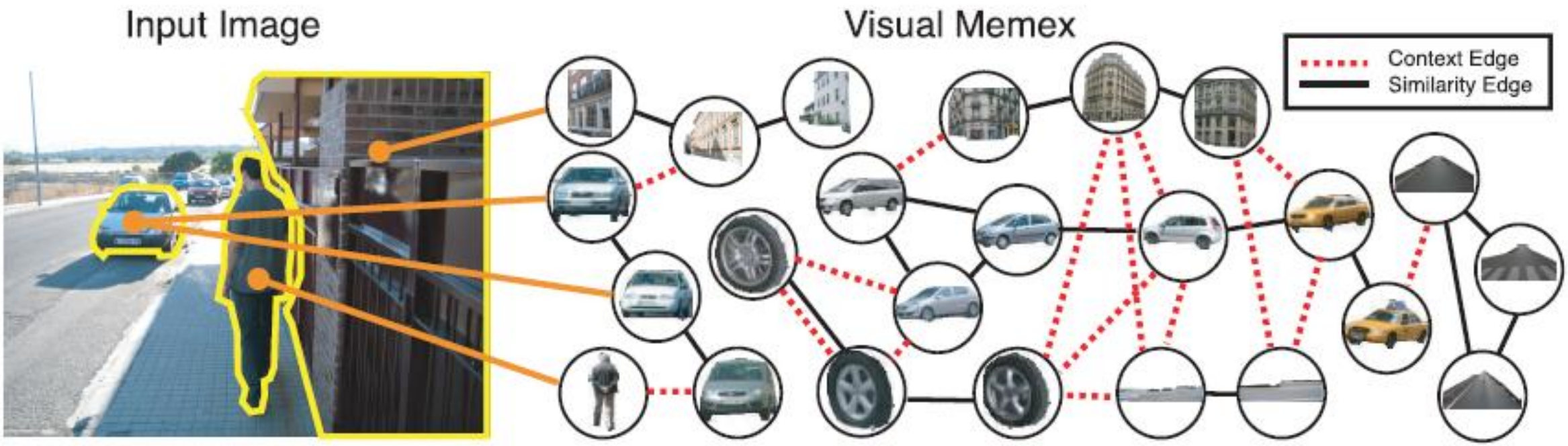


Figure 1: The **Visual Memex** graph encodes object similarity (solid black edge) and spatial context (dotted red edge) between pairs of object exemplars. A spatial context feature is stored for each context edge. The Memex graph can be used to interpret a new image (left) by associating image segments 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

Slide by Antonio Torralba

Our Challenge Setup

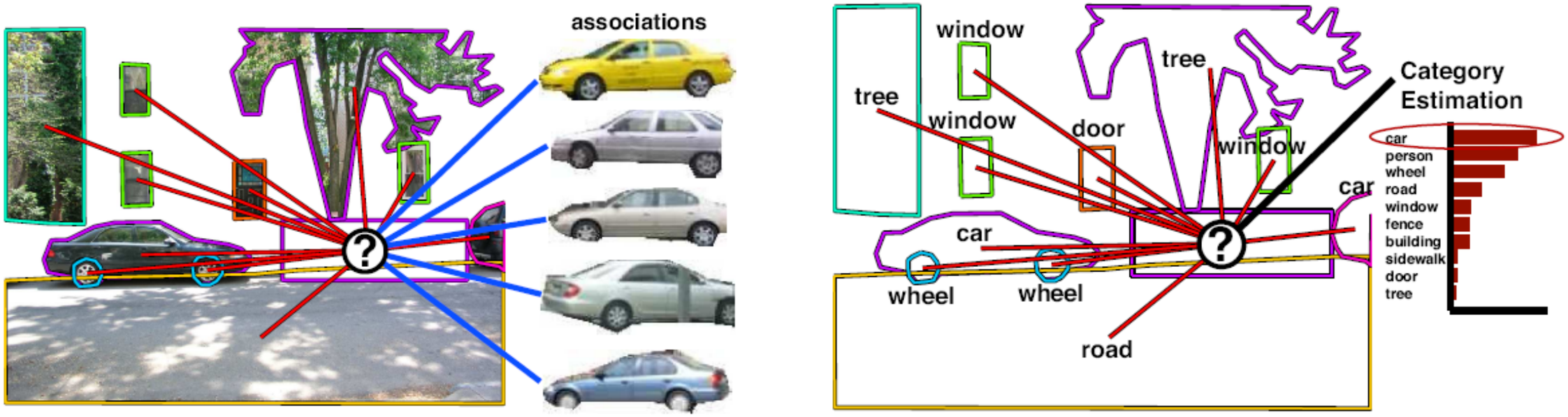


Figure 2: Torralba's Context Challenge: "How far can you go without running a local object detector?" The task is to reason about the identity of the hidden object (denoted by a "?") without local information. In our category-free Visual Memex model, object predictions are generated in the form of exemplar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

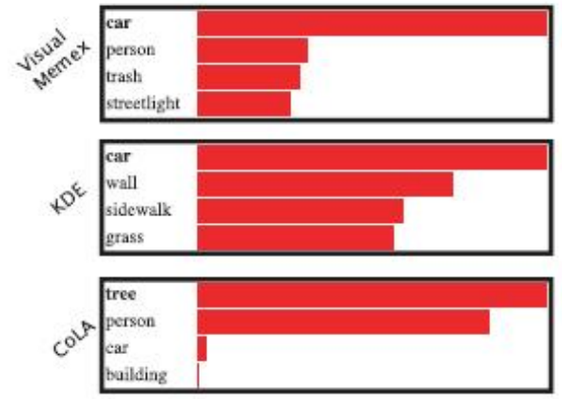
Reduced Memex: categories, non-parametric relationships

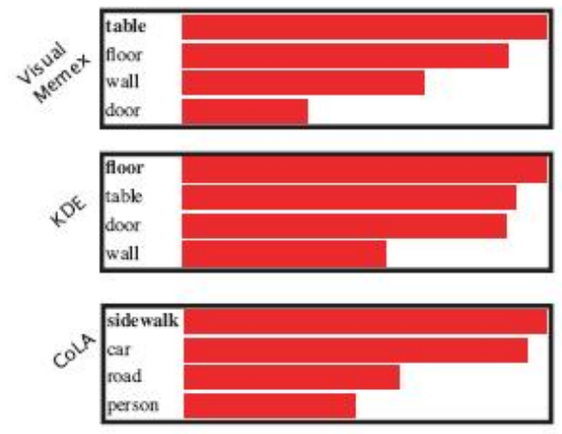
Qual. results

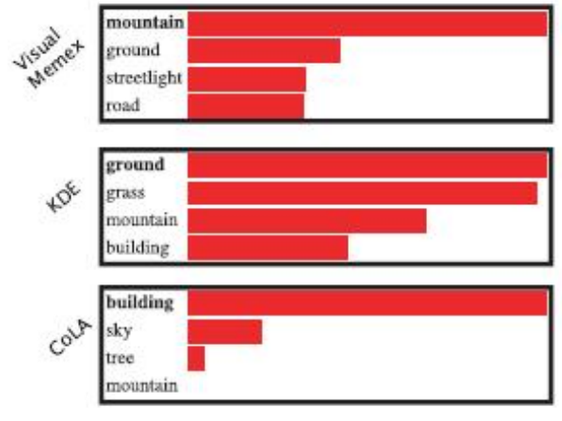
Input Image + Hidden Region

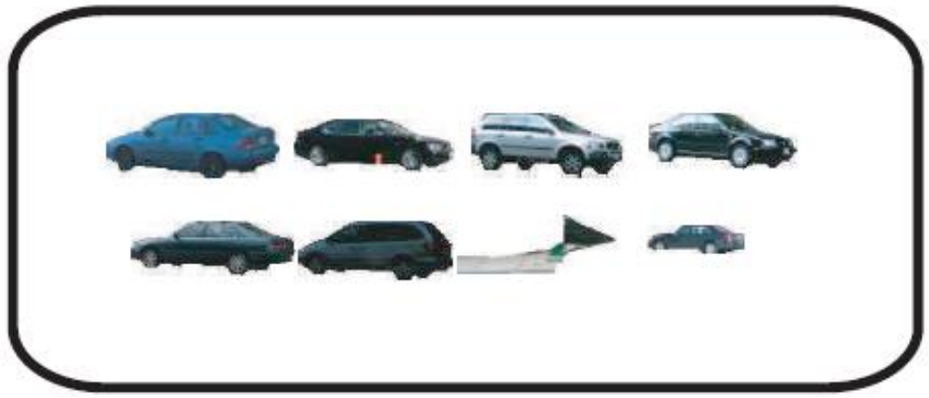
Visual Memex Exemplar Predictions

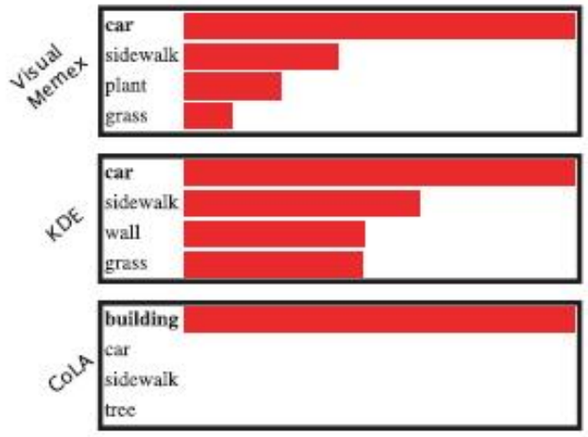
Categorization Results

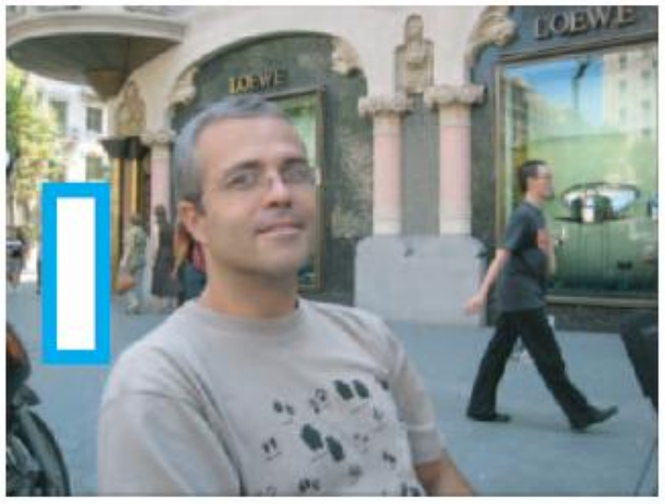


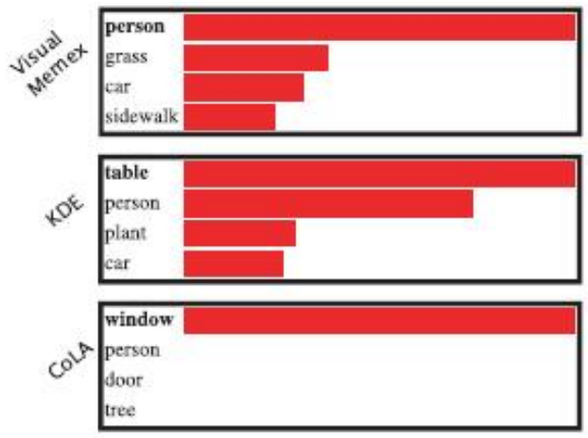


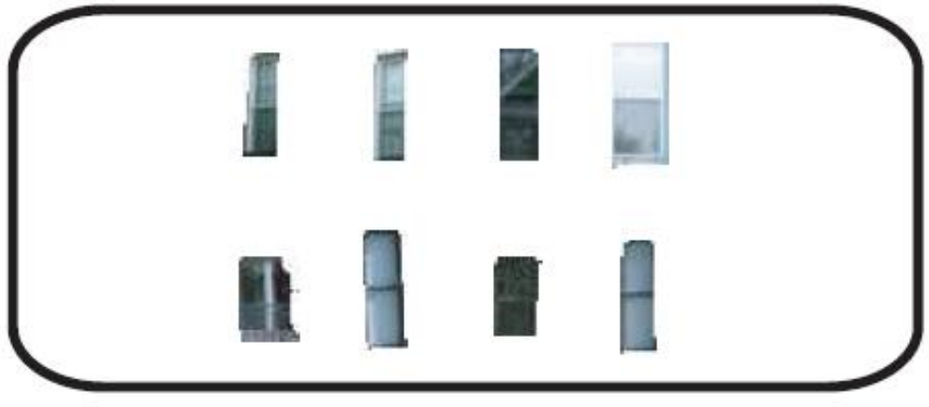


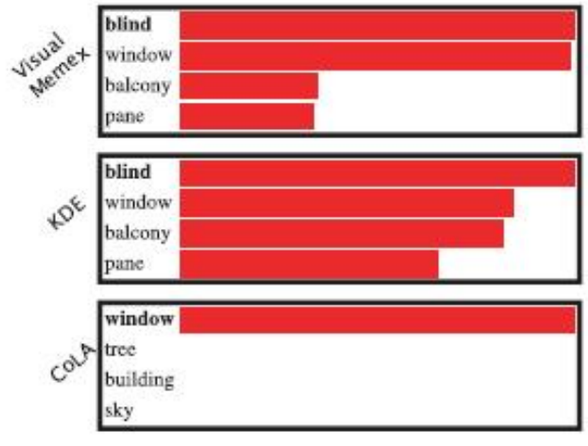



Visual Memex
car
sidewalk
plant
grass
KDE
car
sidewalk
wall
grass
CoLA
building
car
sidewalk
tree




Visual Memex
person
grass
car
sidewalk
KDE
table
person
plant
car
CoLA
window
person
door
tree




Visual Memex
blind
window
balcony
pane
KDE
blind
window
balcony
pane
CoLA
window
tree
building
sky

Quant. results

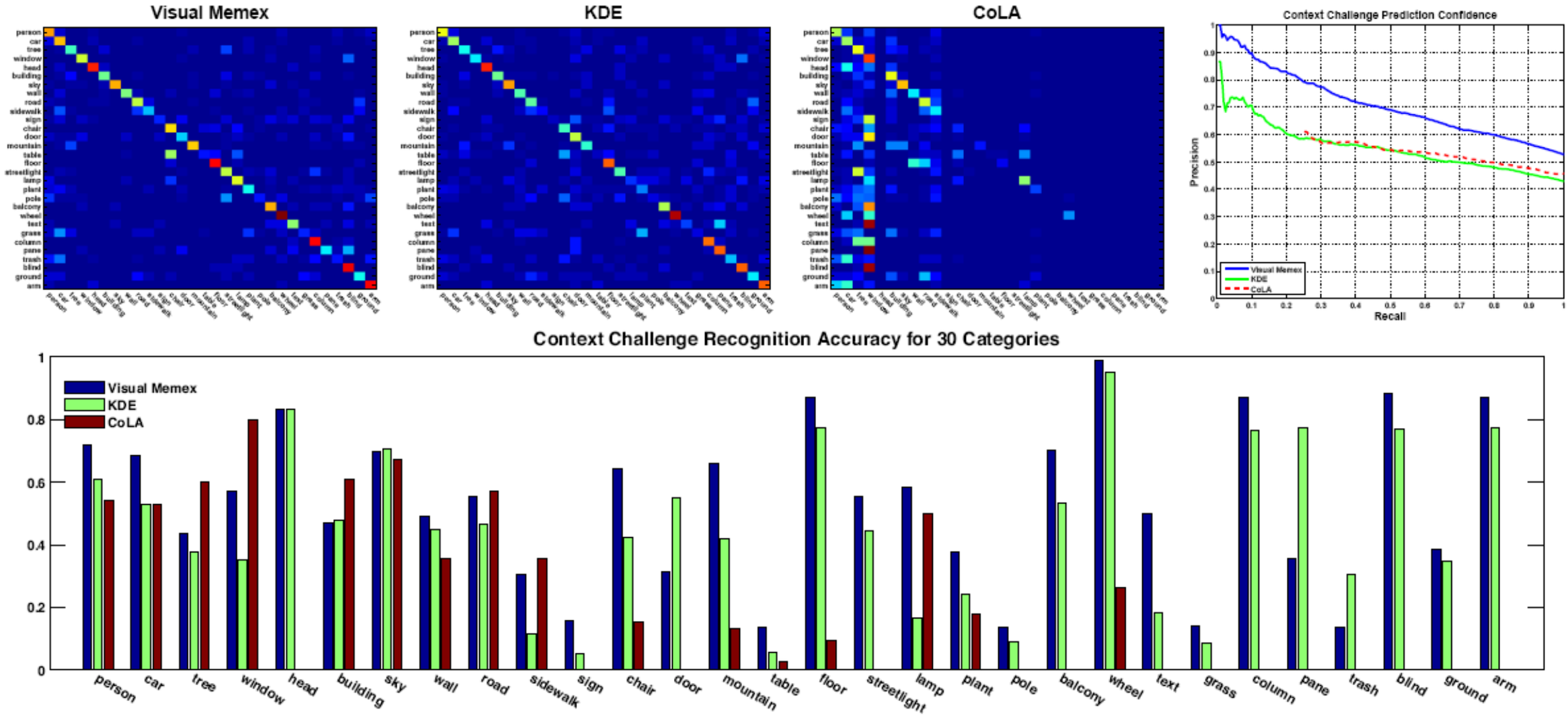
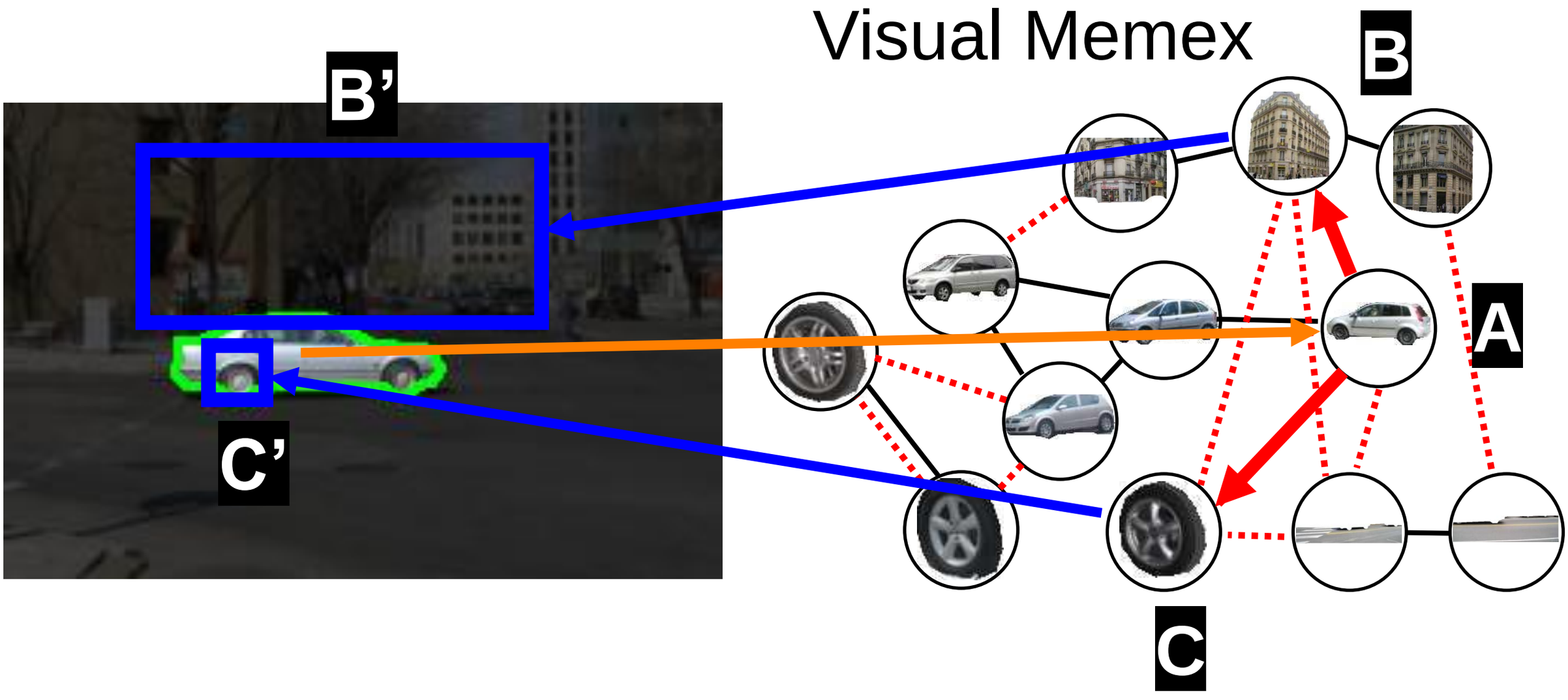


Figure 3: a.) Context Challenge confusion matrices for the 3 methods: Visual Memex, KDE, and CoLA. b.) Recognition Precision versus Recall when thresholding output based on confidence. c) Side by side comparison of the 3 methods' accuracies for 30 categories.

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?

ABSOLUT
Country of Sweden
VODKA
This superb vodka
was distilled from grain grown
in the rich fields of southern Sweden.
It has been produced at the famous
old distilleries near Åhus
in accordance with more than
400 years of Swedish tradition.
Vodka has been sold under the name
Absolut since 1879.
40% ALC./VOL. (80 PROOF) 1 LITER
IMPORTED
PRODUCED AND BOTTLED IN AHUS, SWEDEN
BY THE ABSOLUT COMPANY
A DIVISION OF V&S VIN&SPRIT AB.
+