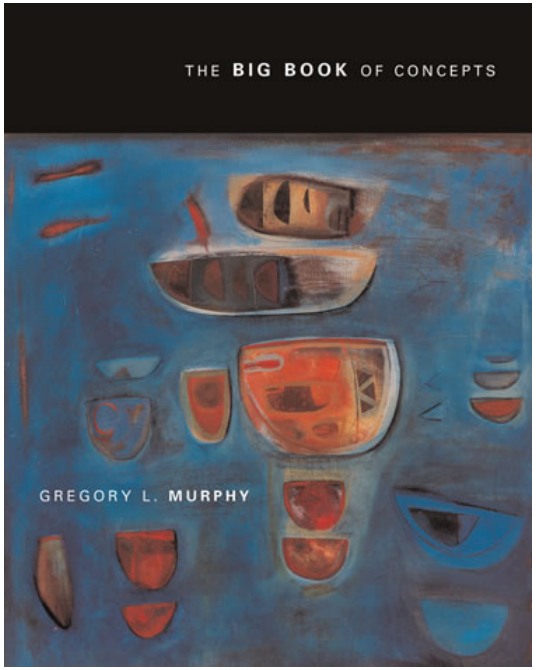
Beyond Scene Categories… Toward a "Visual Memex"

© Quint Buchholz

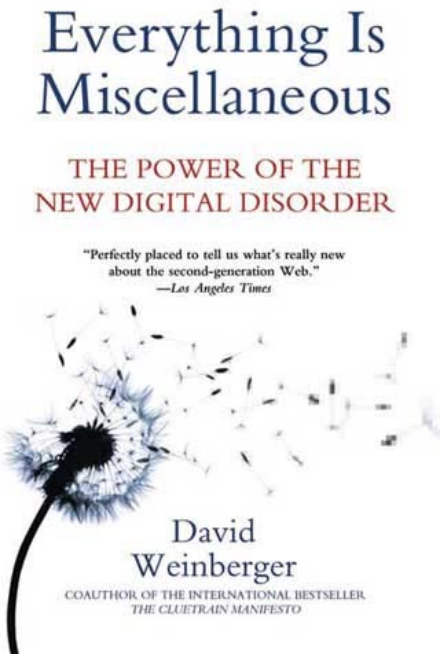
Alexei (Alyosha) Efros
CMU Pittsburgh / ENS Paris

Acknowledgements

Talk by Moshe Bar on the Proactive Brain, ICCV'07



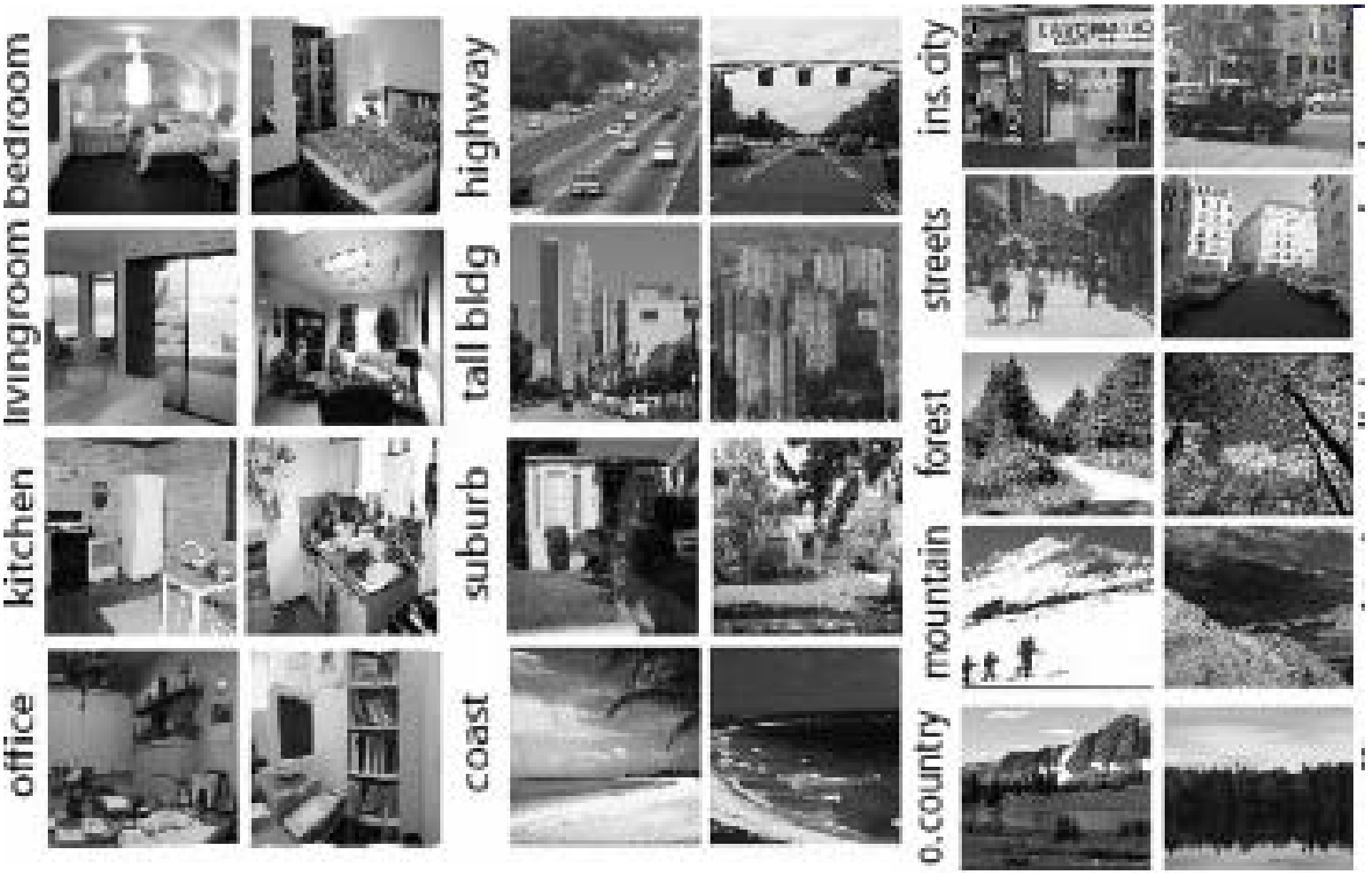
Murphy
Big Book of Concepts



Weinberger
Everything is Miscellaneous

Many great discussions with colleagues, especially Tomasz Malisiewicz, James Hays, and Derek Hoiem

Classic Scene Categorization



Categorization example:

Categorization example:

“a picture is worth a thousand words?”

Theories of Categorization

- 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?
  - Typicality
    - e.g. Chicken -> bird, but bird -> eagle, pigeon, etc.
  - Intransitivity
    - e.g. car seat is chair, chair is furniture, but …

It only gets worse!

- Multiple category membership (it's not a tree, it's a forest!)
  - e.g. Tolstoy's "War and Peace" belongs to:
    - love story
    - Napoleonic wars
    - long Russian novels with lots of French dialog
- Doesn't work even in human-defined domains
  - e.g. Is Pluto a planet?

problems with Visual Scene Categories: *A tale of two **cities***

- There is semantic but not visual similarity

- Hugely unfair to the poor classifier!!!

it’s like “*vodka & potato” classifier!*

+

Why do we need categories anyway?

- To transfer knowledge!
- …can we do it without explicit categorization?

Association instead of categorization

Ask not "what is this?", ask "what is this like"

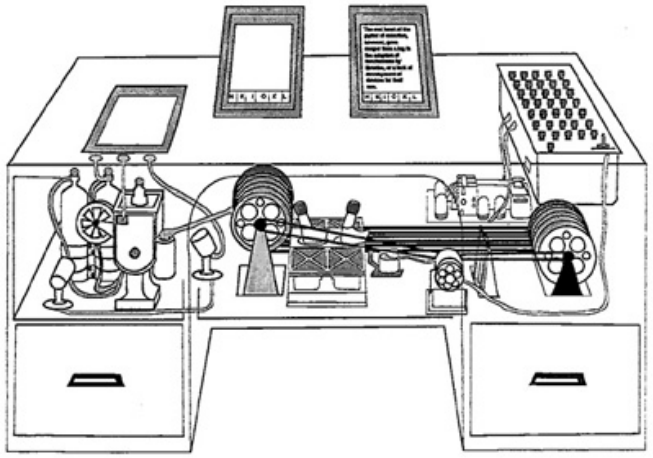
– Moshe Bar

- 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)

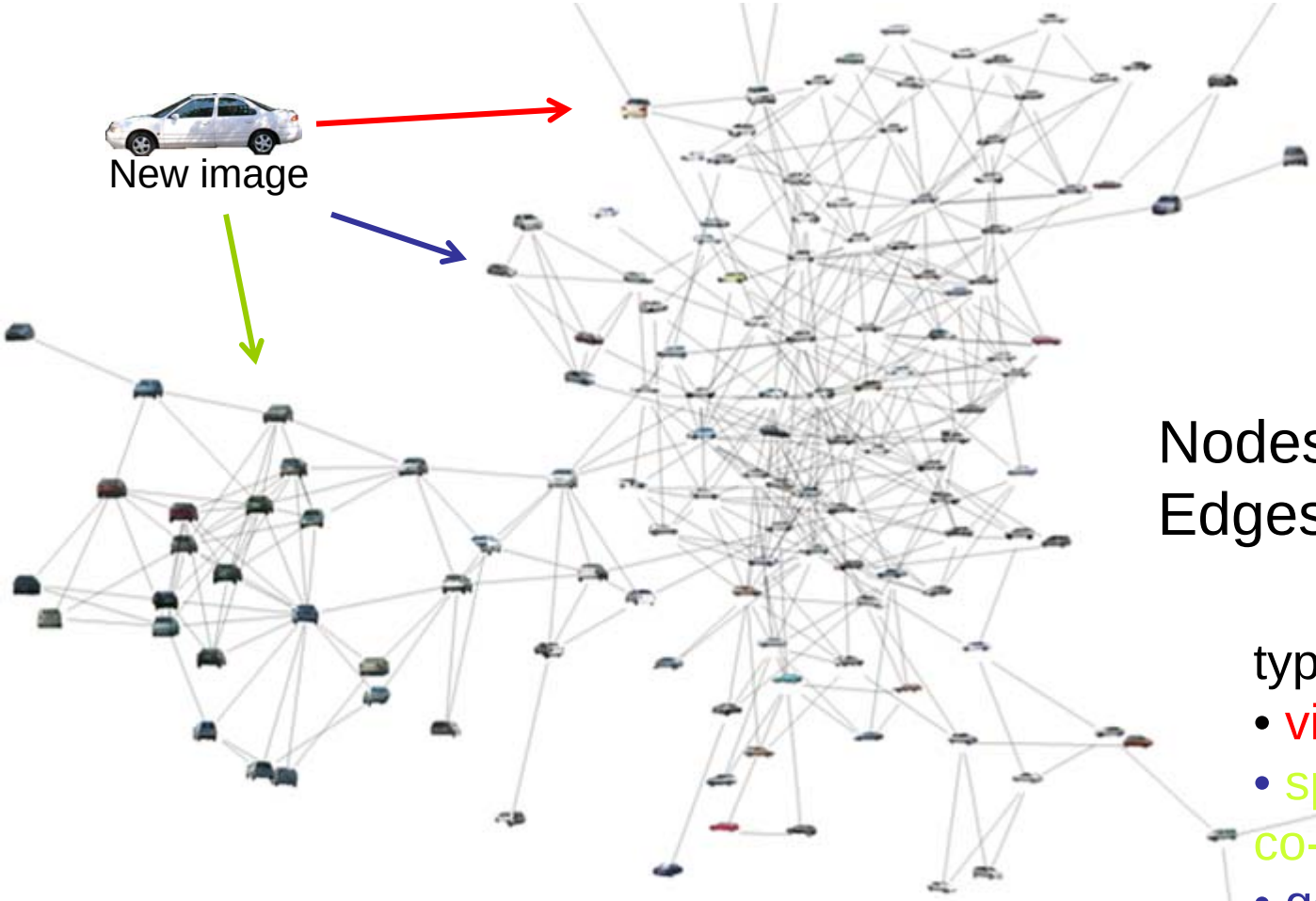
- Store publications, correspondence, personal work, on microfilm
- Items retrieved rapidly using index codes
  - Builds on "rapid selector"
- Can annotate text with margin notes, comments
- Can construct a *trail* through the material and save it
  - Roots of hypertext
- Acts as an external memory



Slide by Eylon Caspi et al

Visual Memex, a proposal

[Malisiewicz & Efros, CVPR'08, ongoing]



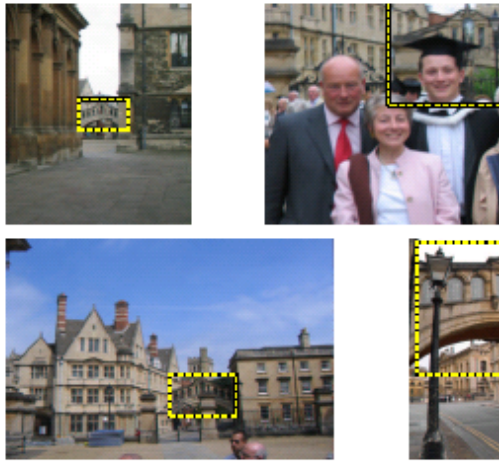
Nodes = instances
Edges = associations

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

1. exact place matching

query

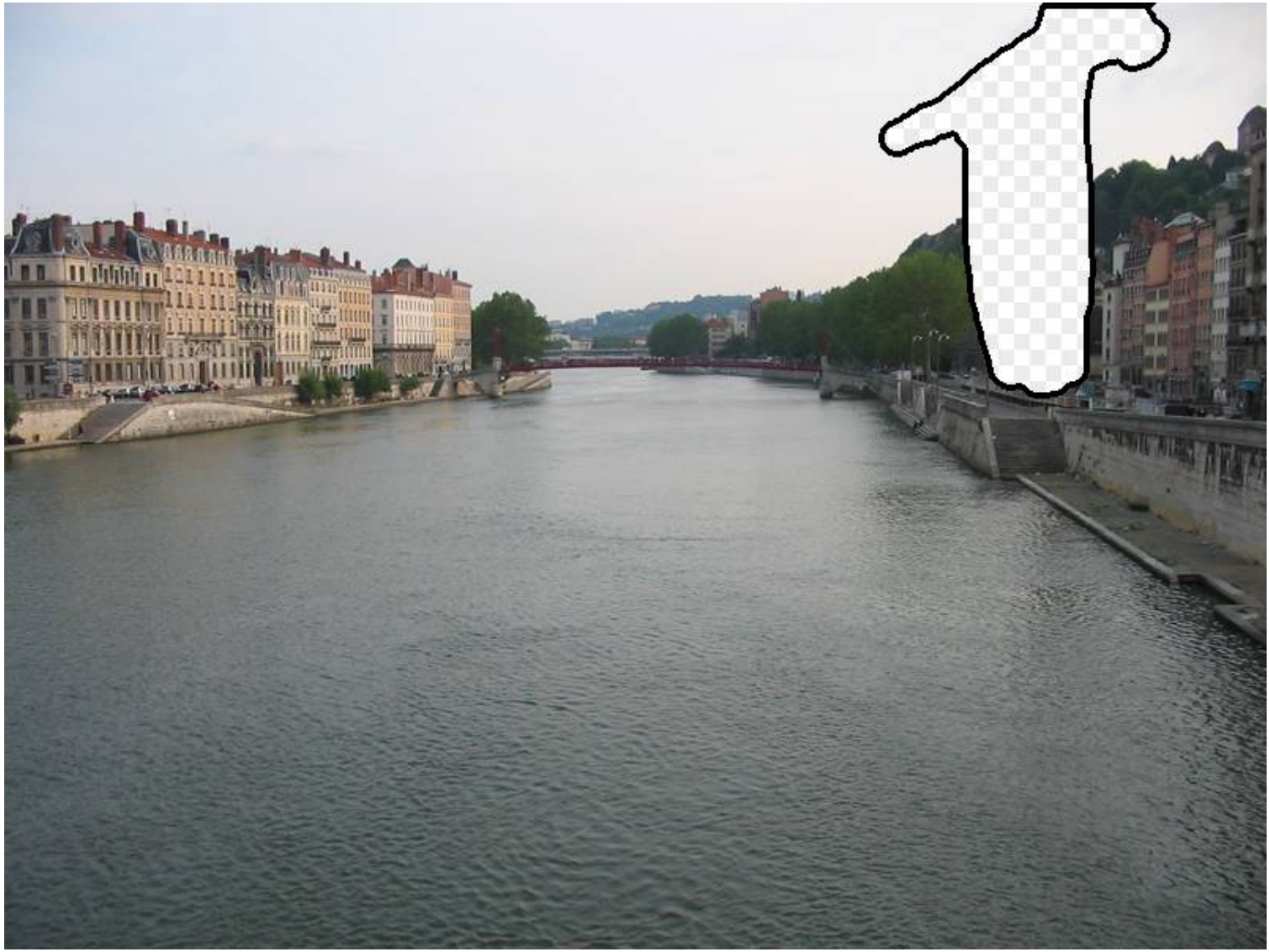
Video Google
[Sivic et al]



Total Recall
[Philbin et al]

2. (extreme) visual similarity

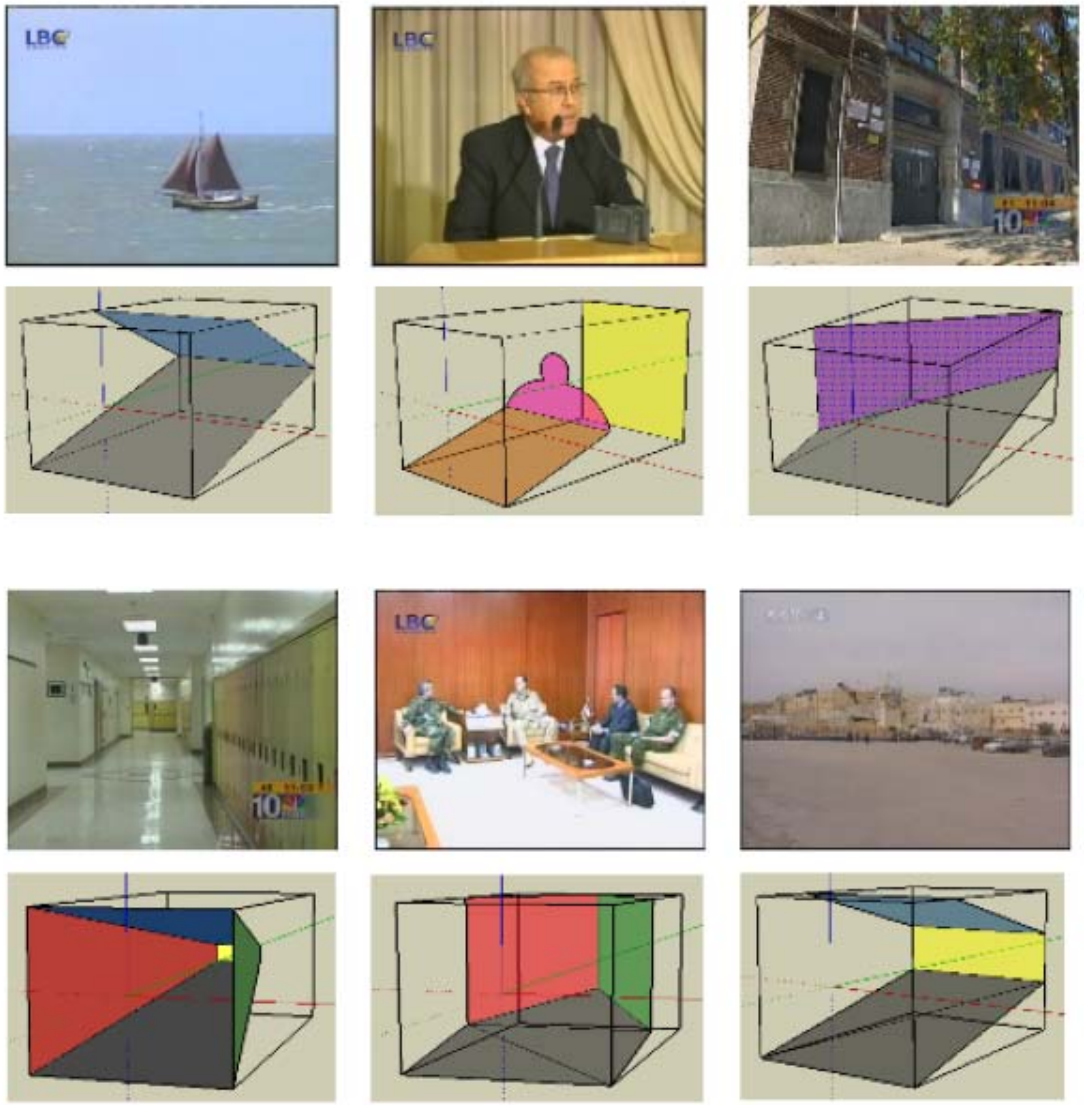
Hays and Efros, SIGGRAPH’07



Hays and Efros, SIGGRAPH’07

Hays and Efros, SIGGRAPH’07

3. Rough Geometrical Structure



Nedovic et al, ICCV'07

Scene Gist

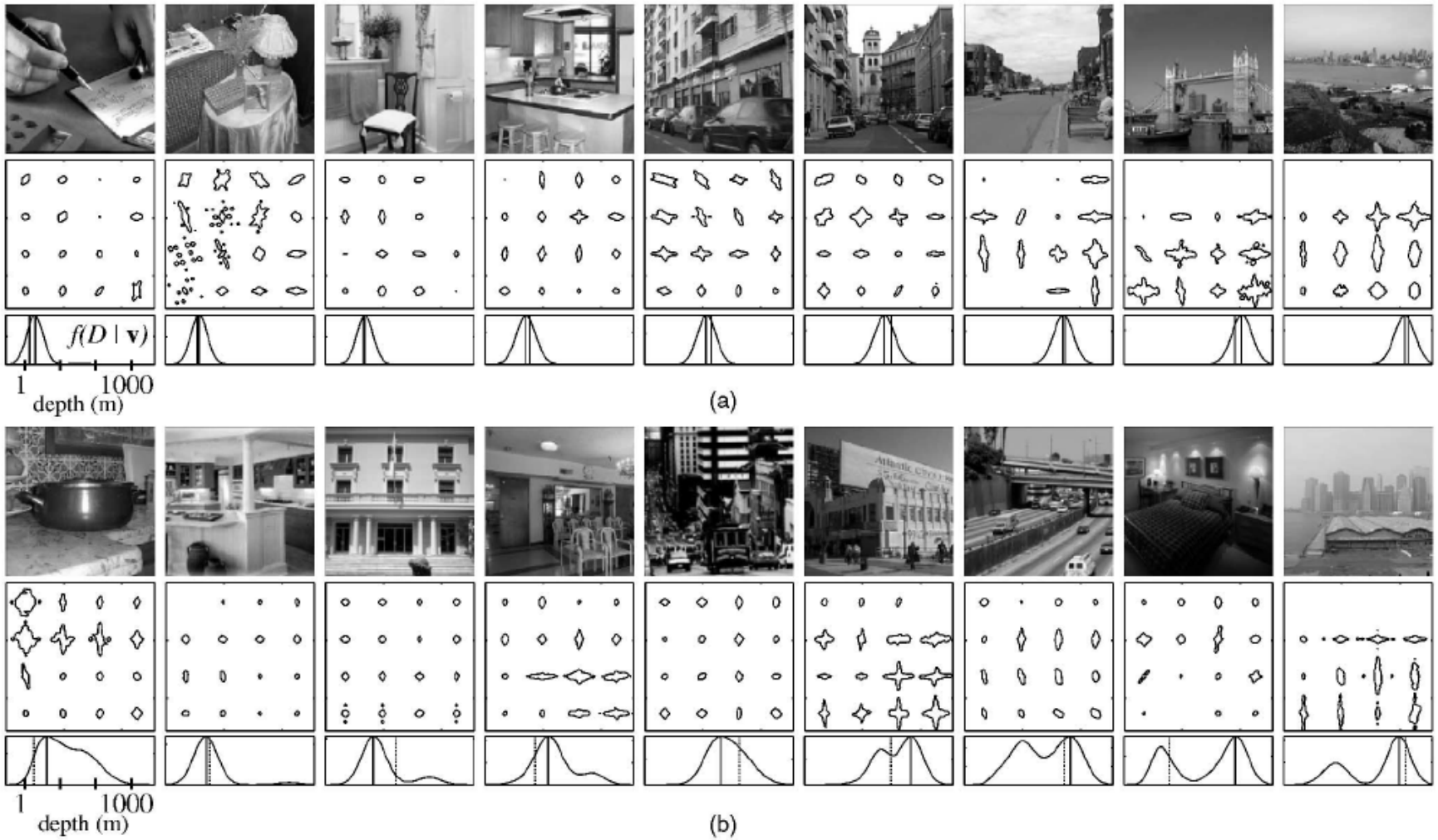
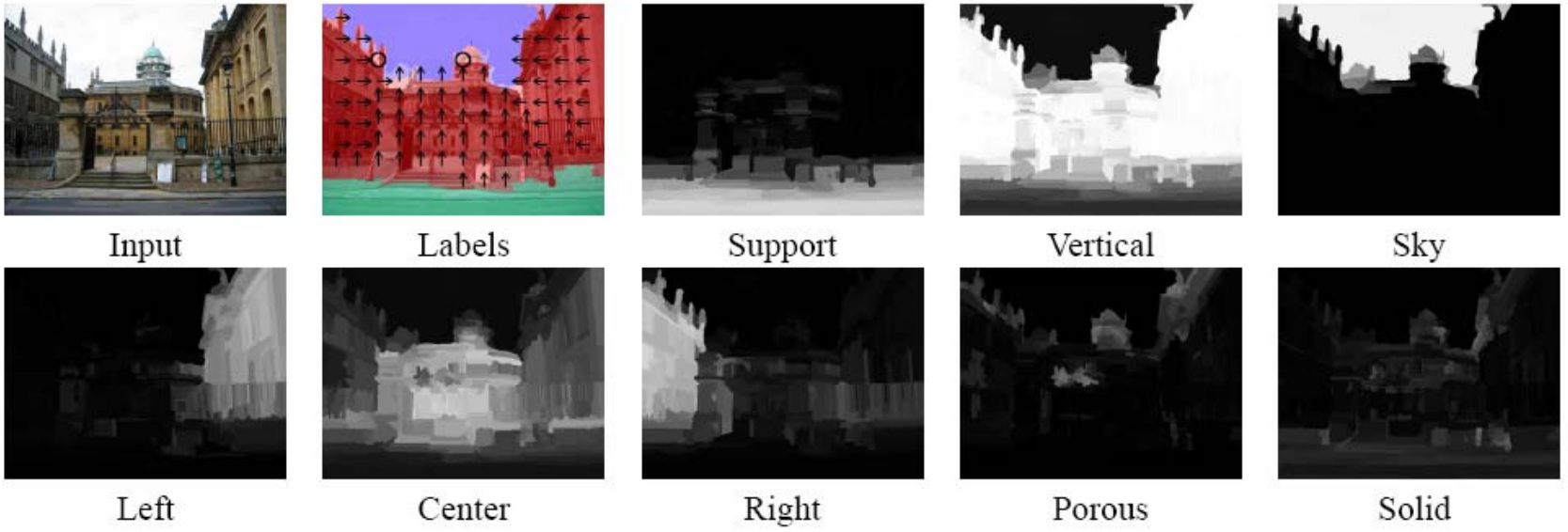
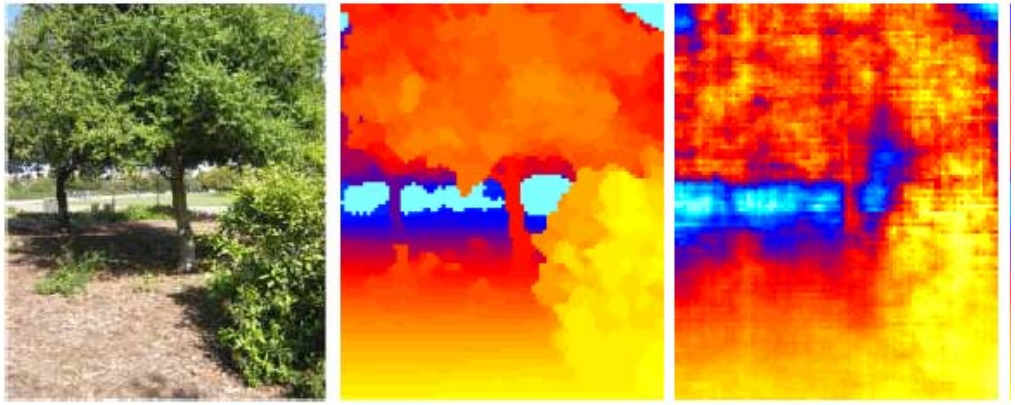


Fig. 8. Examples of man-made scenes sorted according to mean depth. (a) Pictures selected among the 85 percent of pictures with the highest confidence estimations. Middle line shows the 80 percent level section of the local amplitude spectrum and the bottom line shows the conditional PDF of depth obtained from the features $A_M(\mathbf{x}, k)$. (b) Pictures selected among the images with the lowest confidence estimations.

Torralba & Oliva, 2001

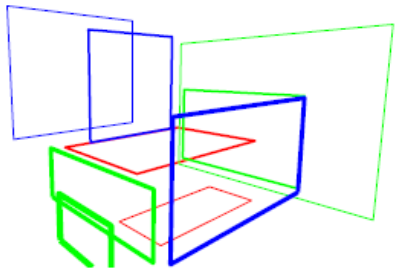


Hoiem, Efros, Hebert, 2005

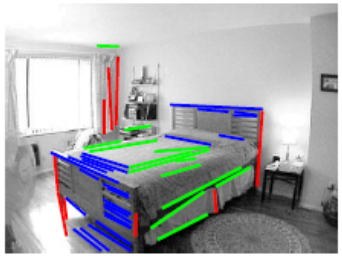


Saxena, Ng, 2005

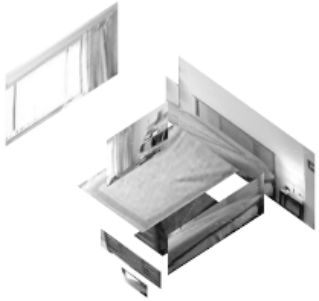
Delage, Lee, Ng, CVPR’06



a. depth-ordered planes

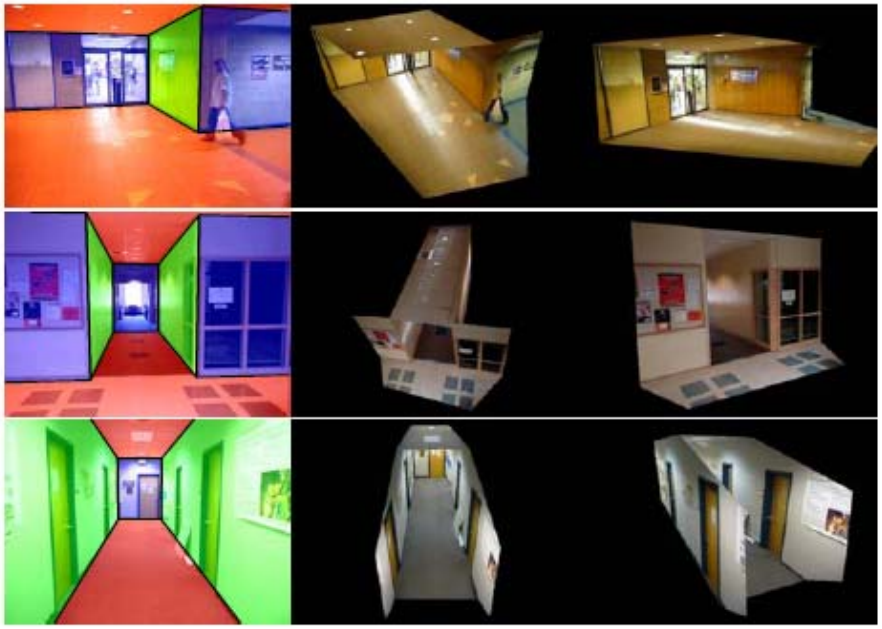


b. occluders

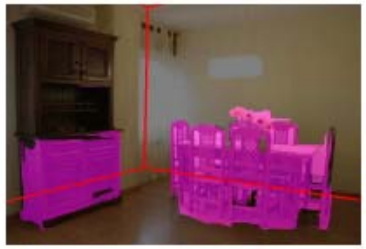


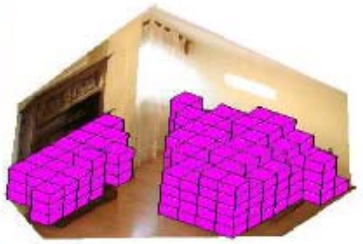
visualization in 3D

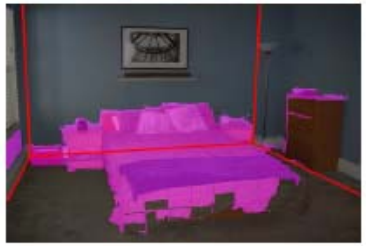
Yu, Zhang, Malik, CVPR’08

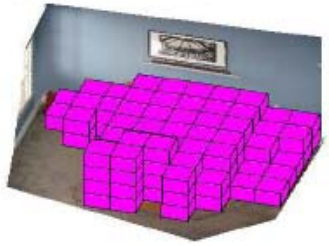


Lee, Hebert, Kanade, CVPR’09





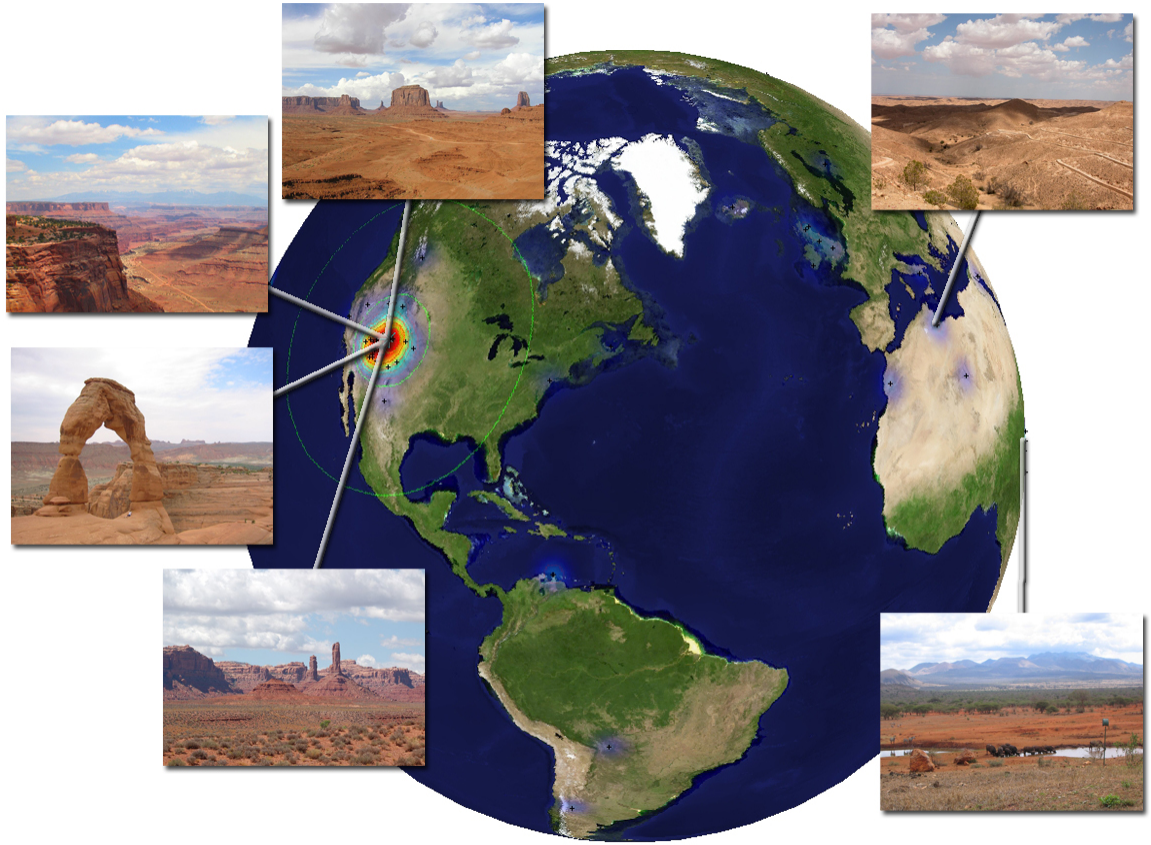




Hedau, Hoiem, Forsyth, ICCV’09

4. Geographic Properties

Query Photograph



Visually Similar Scenes

Im2gps [Hays & Efros, CVPR'08]


Argentina


Andorra


Andorra


Iceland


Idaho


Switzerland


Argentina


Bolivia


Nevada


Hawaii


Hawaii


Egypt


China

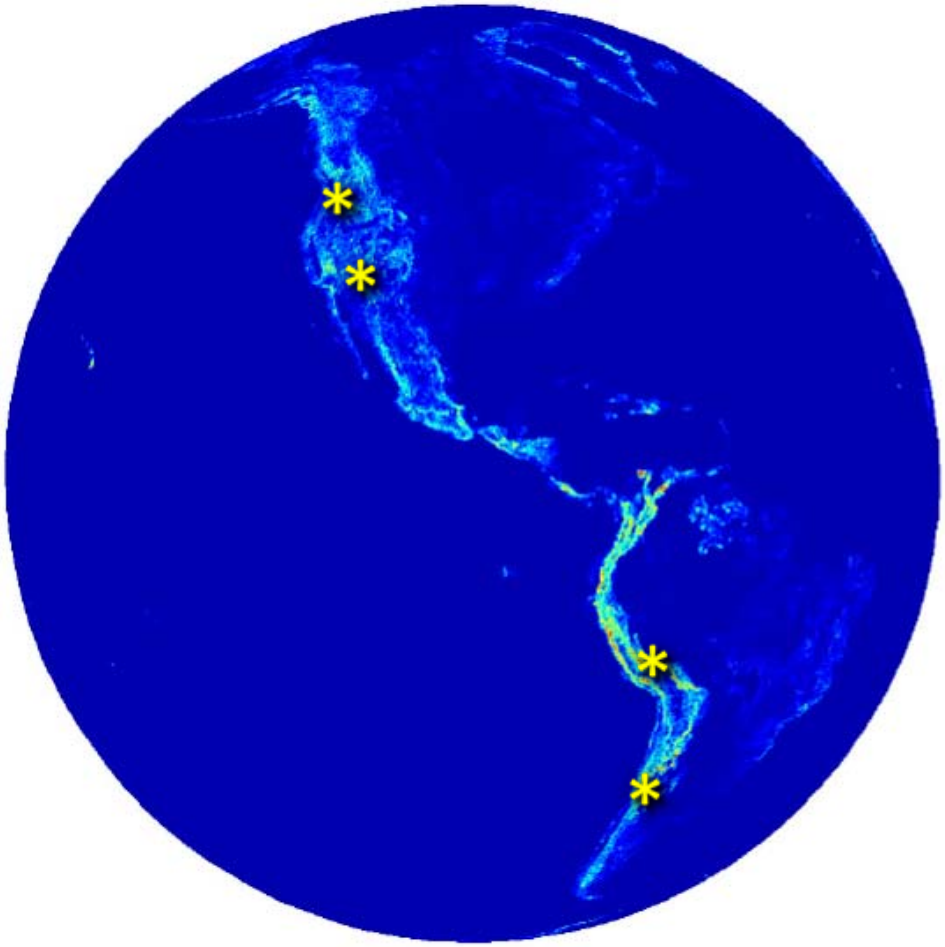

Arizona

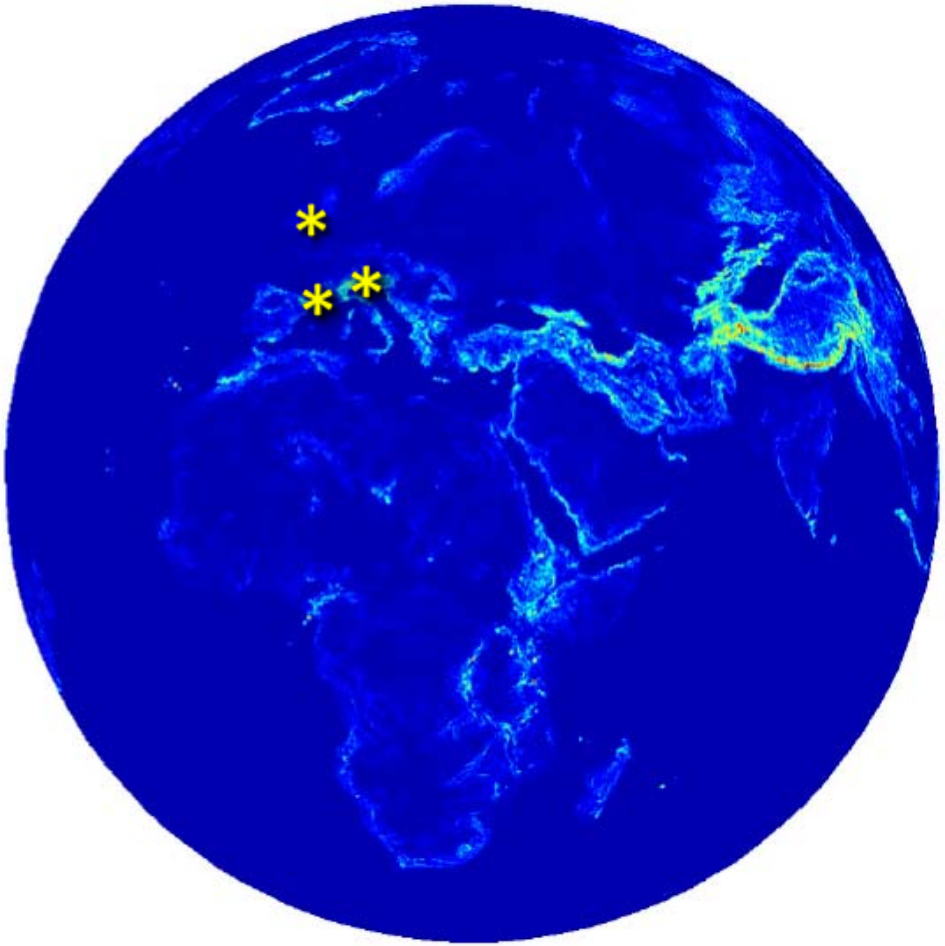

Peru


Oregon

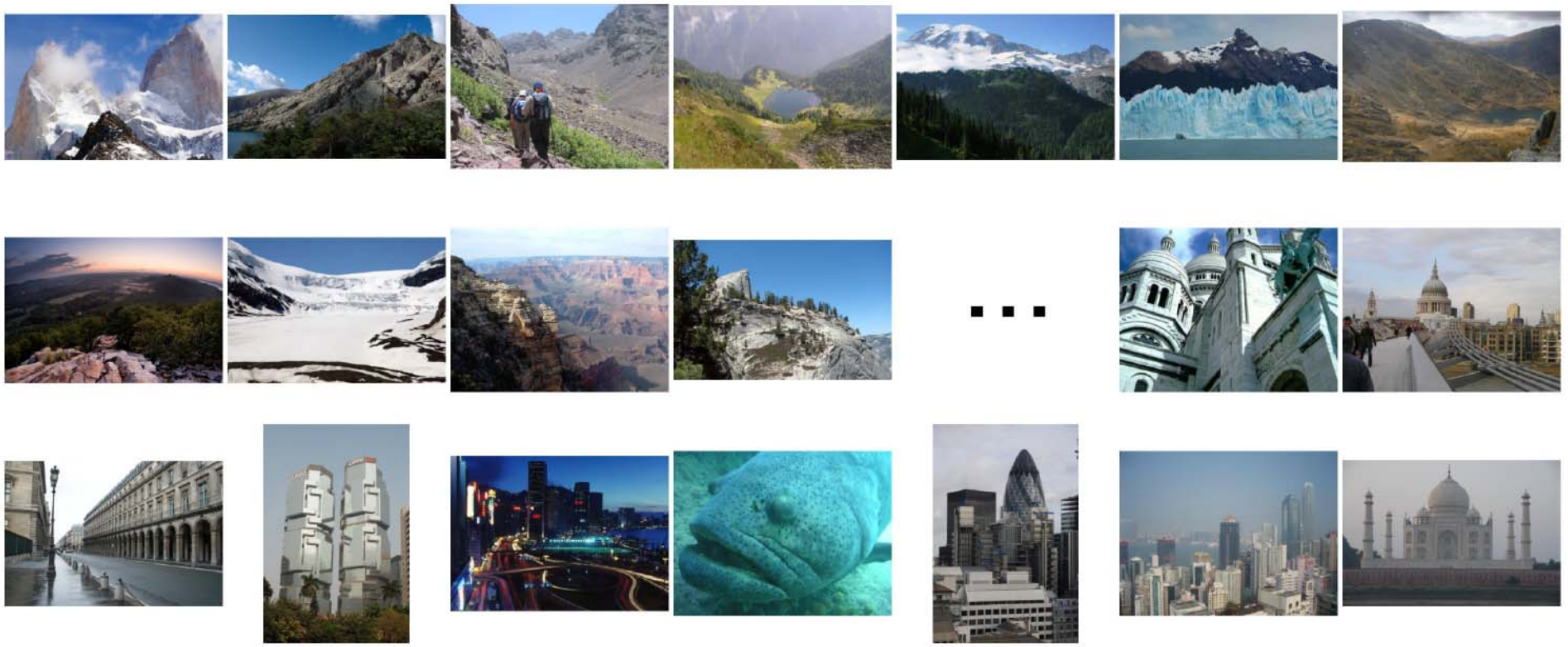

Argentina
Andorra
Andorra
Iceland
Idaho
Switzerland
Argentina
Bolivia
Nevada
Hawaii
Hawaii
Egypt
China
Arizona
Peru
Oregon

Elevation gradient = 112 m / km





Elevation gradient magnitude ranking



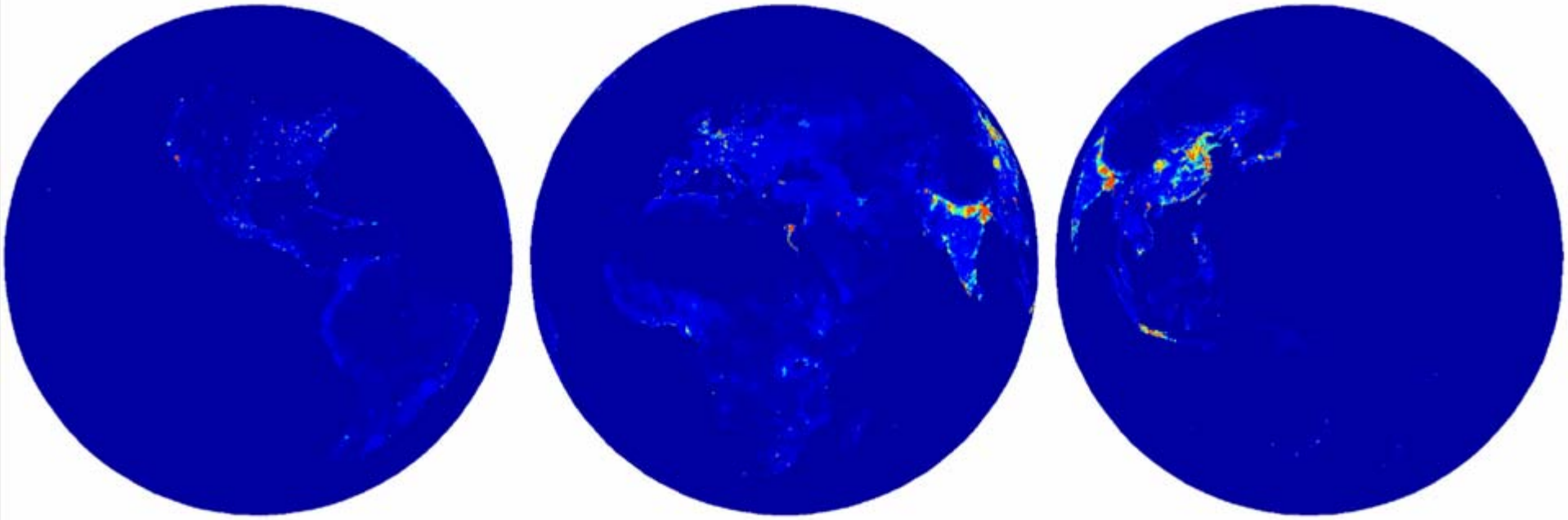
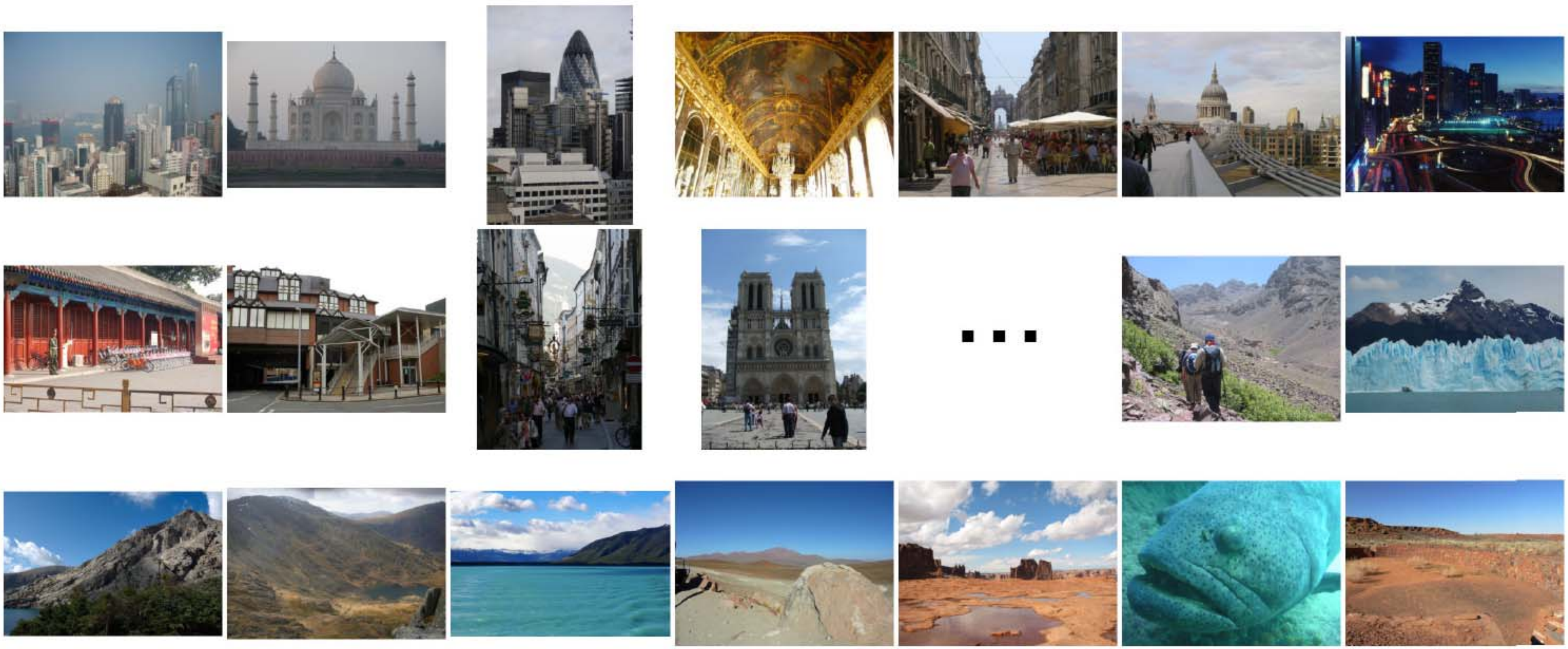


Figure 2. Global population density map.

Population density ranking



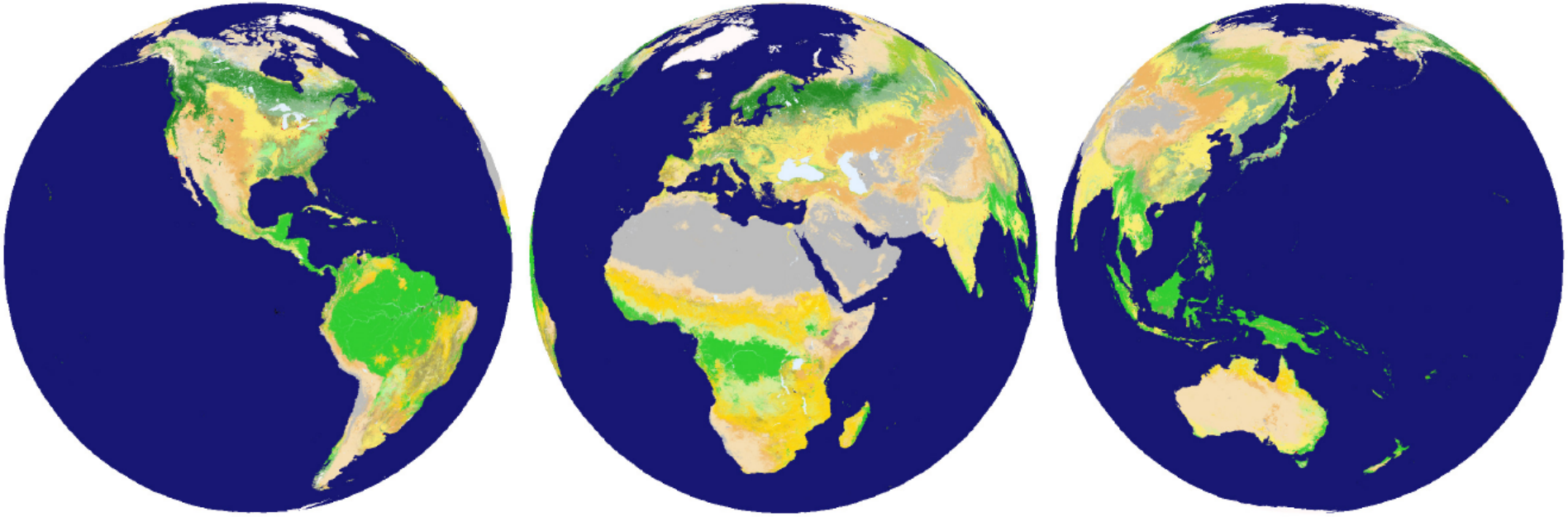
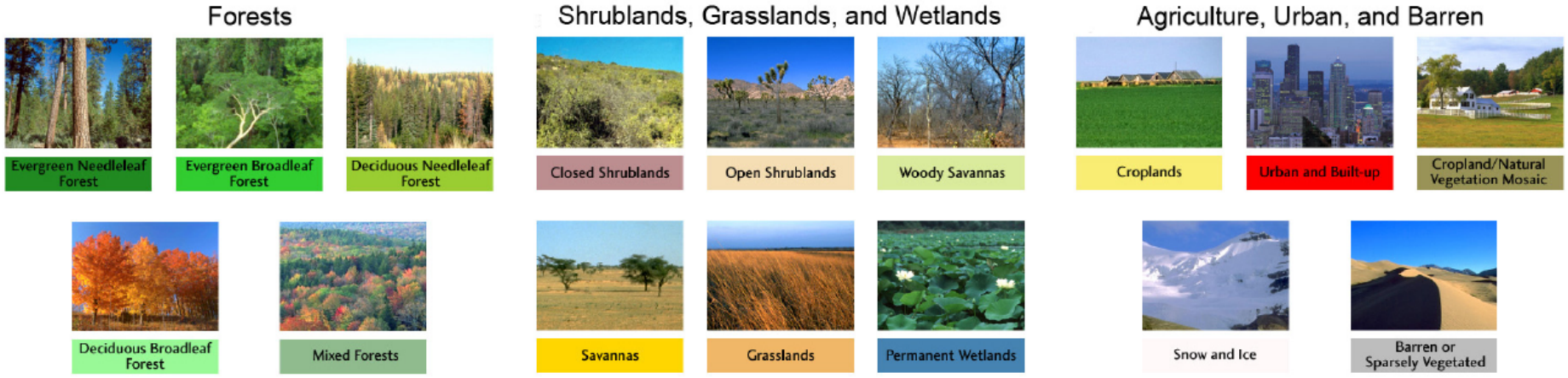
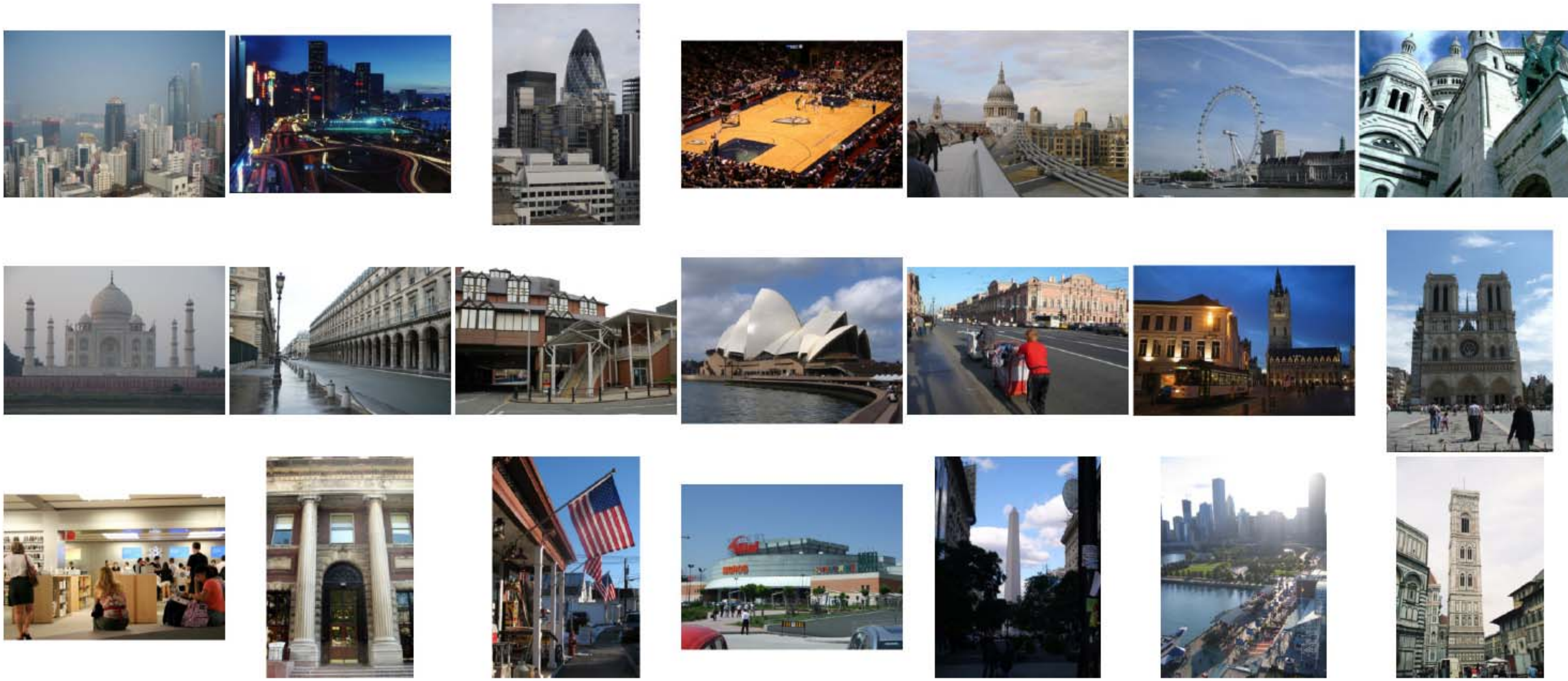


Figure 4. Global land cover classification map.

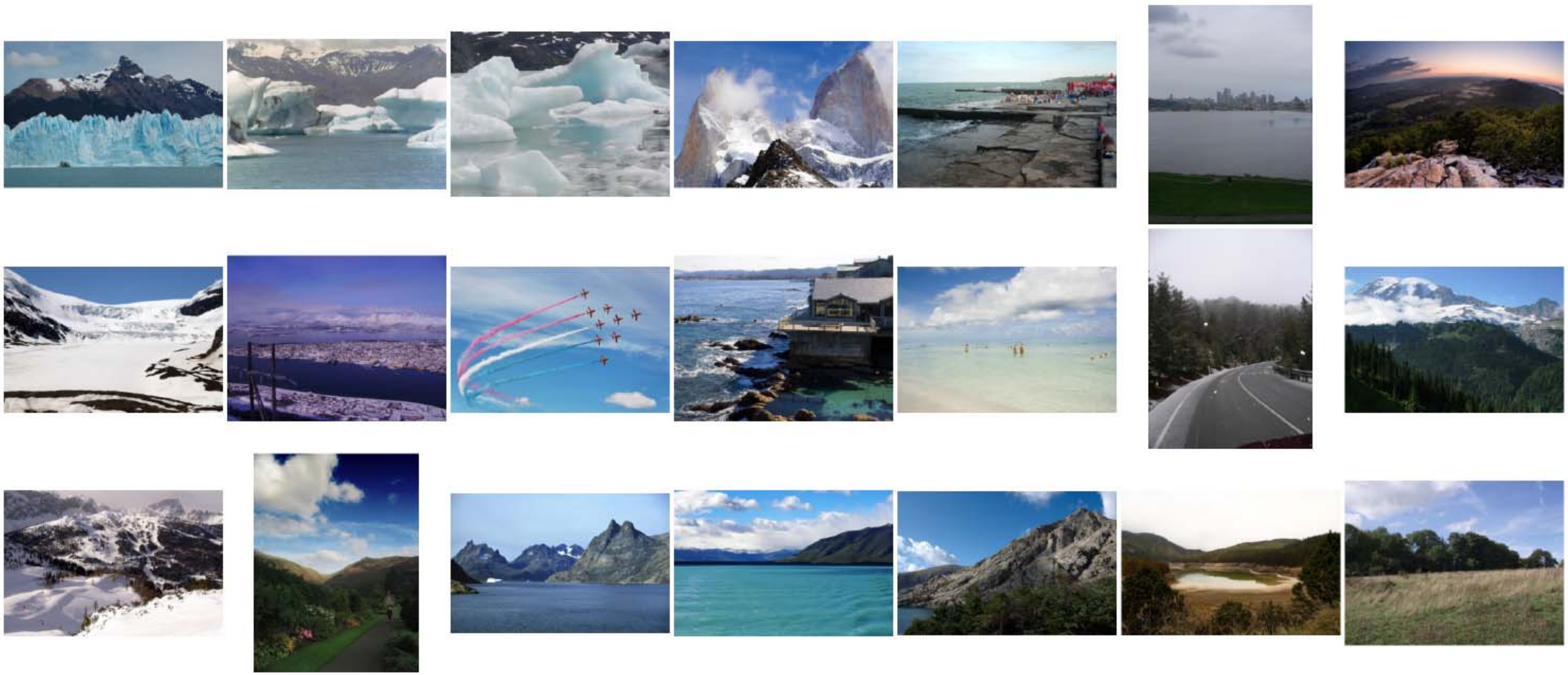


Barren or sparsely populated

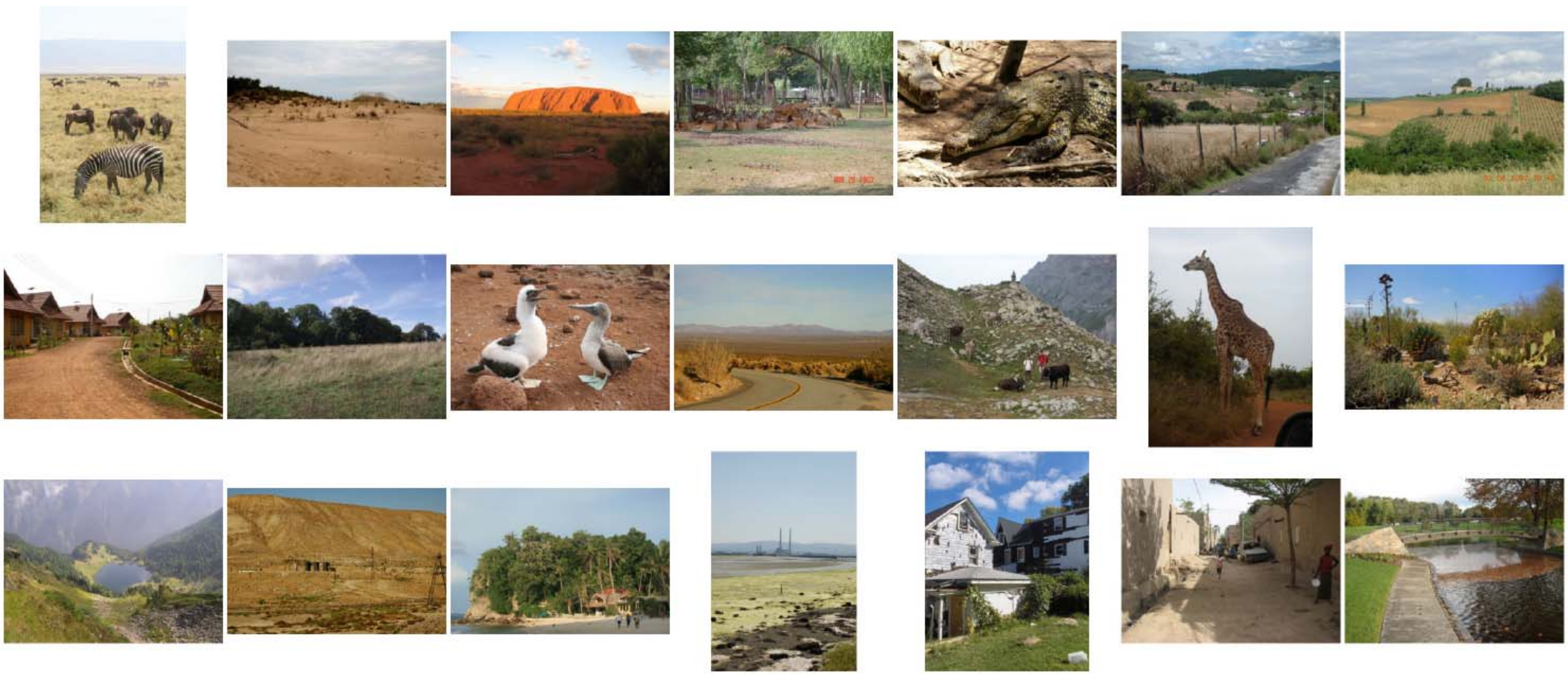
Urban and built up



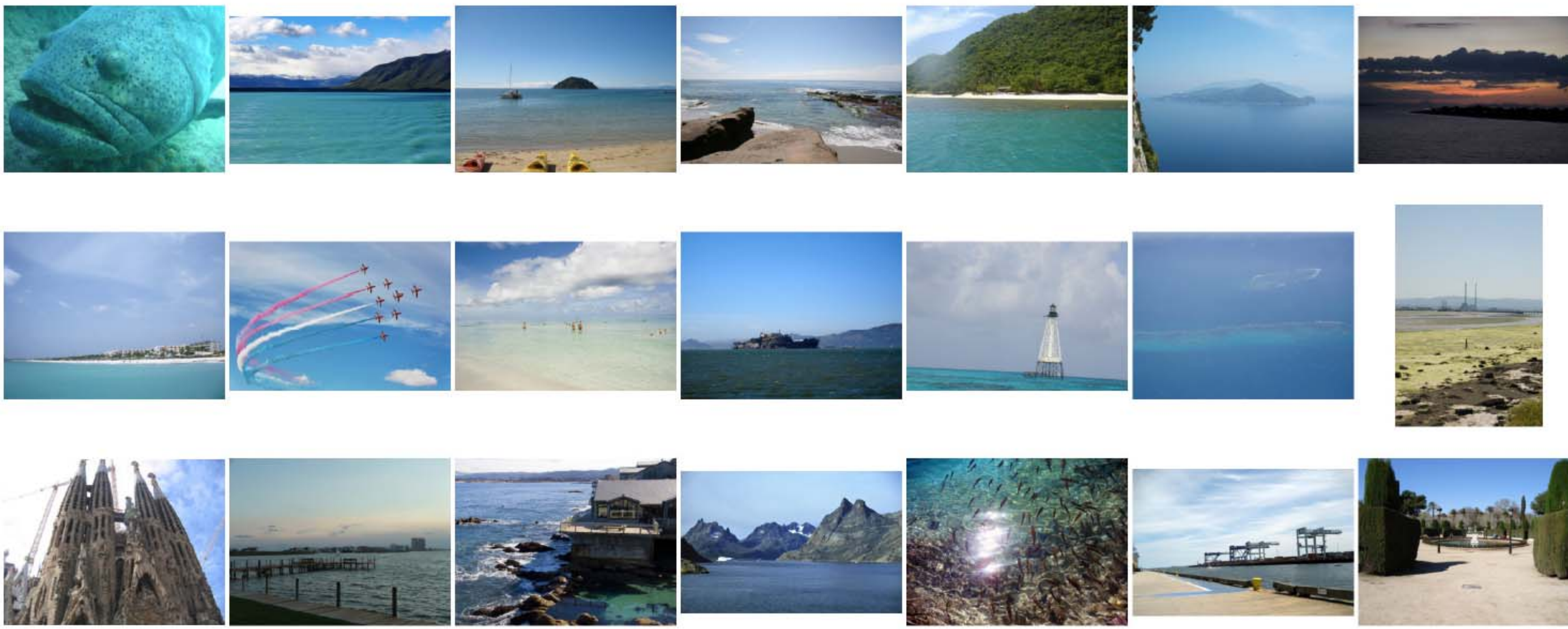
Snow and Ice



Savannah



Water



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

Thank you


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 ÅHUS, SWEDEN
BY THE ABSOLUT COMPANY
A DIVISION OF V&S VIN&SPRIT AB

+