Seeing the \{Closed+Camouflage+Natural\=Forest\} for the Trees: Rapid Scene Categorization can be mediated by conjunctions of global scene properties

Michelle R. Greene & Aude Oliva
Perceptual Science Group, Brain & Cognitive Science Department, MIT

Introduction

Local object database norming

Two observers hand annotated all regions in the 200 image database:

- Field
- Mountain
- Waterfall
- Low mountain
- Lake
- Low lake
- Ocean
- Desert
- River

Low 29%
8%
水面

Local 23% sky
35% water
18% trees
12% mountain
1% hill
1% clouds
23% grass

Global =
41% camouflage
61% moving
47% navigable
32% hot
66% open
64% expansive
74% deep

Use of global properties

These global properties are used to predict human image classification, but are they employed by human observers when doing the task?

Hypothesis: If global property information is used by people in rapid semantic classification, then responding to a target among distractors that share the target category’s global properties should yield more false alarms than distractor sets that do not share global properties with the target category.

Method: 8 category \* 7 global properties - 2 magnitude confusion matrix

Discussion

Global properties are holistic properties related to the shape of scene space, interactions humans can have in that space or properties of the surfaces in the scene.

Global properties have an earlier perceptual availability than the semantic category of images.

Global properties are sufficient to predict human semantic categorization performance and the specific errors made. It seems that people use global properties to do rapid image classification.

Error analysis

Humans and model OK:

- Field
- Lake
- River
- Mountain
- Low mountain
- Ocean
- Desert
- River

Humans and model miss:

- Field
- Lake
- River
- Mountain
- Low mountain
- Ocean
- Desert
- River

Perceptual availability of global properties

What is the perceptual availability of global properties relative to semantic categories?

We compared GLOBAL property classification to SEMANTIC category classification in a staircased yes-no forced choice task.

Threshold is presentation time permitting 70% correct classifications.

The model using an object representation does not predict as well human semantic classification performance (r=0.67, p=0.06) by category. It also predicts fewer of the specific errors that humans make (62%).

Ranking Results

- Desert
- Forest
- Field
- Lake
- Waterfall
- Ocean
- River
- Mountain

Global properties

- Temperature
- Depth
- Navigable
- Forest
- Expansive
- Material
- Openness
- Depth
- Mountain
- River

Semantic categories

- Forest
- River
- Field
- Desert
- Ocean
- Mountain
- Waterfall
- Low mountain
- Low lake
- Lake
- Low lake
- Ocean
- Desert
- River

Mean\(\overline{21.3}\)
Mean\(\overline{21.3}\)

Sufficiency of global properties

To what extent is global property information or object information alone sufficient to predict human categorization performance?

Humans OK, model misses:

- Model OK, humans miss:

Questions:
1. What is global information in a scene?
2. How is this information used by human observers?
3. How does it compare to local (object-centered) information in scenes?

Global scene properties

A set of properties that describe the space that a scene subtends, the possible interactions a human can have in the space, or properties of the surfaces in the scene.

Space
- Degree of openness
- Mean depth
- Navigability

Interaction
- Degree of camouflage
- Degree of movement

Surface
- Temperature
- Degree of expansion

Method:

- 200 images from 8 semantic categories were ranked along these 7 dimensions by 50 observers.

Degree of openness

- Field
- Mountain
- Waterfall
- Low mountain
- Lake
- Low lake
- Ocean
- Desert
- River

Depth

- Field
- Mountain
- Waterfall
- Low mountain
- Lake
- Low lake
- Ocean
- Desert
- River

Navgational

- Field
- Mountain
- Waterfall
- Low mountain
- Lake
- Low lake
- Ocean
- Desert
- River

Discussion

Thanks

Grateful thanks to Antionio Torralba, Molly Potter and Josh Tenenbaum for suggestions and discussion. Object annotations using LabelMe. MRG is supported by NSF-GRF.

mrgreenemit.edu olivamit.edu http://cvcl.mit.edu

MRG is supported by NSF-GRF.