• Project **presentations** (next Monday/Wednesday)
  • Presentations assignments at random
  • Each person (or team) will get 7 (or 10) mins to present
    - Problem statement, preliminary results, data analysis, todo
  • Final report due on Dec. 13 (hard deadline)

• Course evaluations
  • A show of hands who might be missing next class?
Overview of the course so far

Early vision:
- image formation
- light and color perception
- basic image processing
  - edges, corners and blobs

Mid-level vision:
- texture
  - synthesis and representation
- grouping
  - segmentation
  - alignment

High-level vision:
- recognition and learning
- image representation
  - features, etc
- object detection

Misc. topics:
- deep learning
- memorability [Khosla]
- human-centric vision
- optical flow and tracking
Overview

• Motivation
• Levels of categorization
• Visual 20q game [Branson et al., ECCV 2010]
• Similarity comparisons based recognition
  • global similarity [Wah et al., CVPR 2014]
  • localized similarity [Wah et al., WACV 2015]
What type of bird is this?
What type of bird is this?
What type of bird is this?

Computer Vision

?
What type of bird is this?
What type of bird is this?

Computer Vision

Chair?

Bottle?
• Field guides difficult for average users
• Computer vision doesn’t work perfectly (yet)
• Research mostly on basic-level categories
What kind of bird is this?

Parakeet Auklet
Levels of Categorization

Basic-Level Categories

Airplane? Chair? Bottle? ...

[Griffin et al. ‘07, Lazebnik et al. ‘06, Grauman et al. ‘06, Everingham et al. ‘06, Felzenzwalb et al. ‘08, Viola et al. ‘01, ... ]
Levels of Categorization

Subordinate Categories

American Goldfinch?  Indigo Bunting? ...

[Belhumeur et al. ‘08, Nilsback et al. ’08, ...]
Levels of Categorization

Parts and Attributes

Yellow Belly? Blue Belly?...

[Farhadi et al. ‘09, Lampert et al. ’09, Kumar et al. ‘09]
Hard classification problems can be turned into a sequence of easy ones.
Computers: reduce number of required questions
Humans: drive up accuracy of vision algorithms
Research Agenda

2014
- Heavy Reliance on Human Assistance
  - Blue belly? no
  - Cone-shaped beak? yes
  - Striped Wing? yes
  - American Goldfinch? yes

2020
- More Automated
  - Striped Wing? yes
  - American Goldfinch? yes

Computer Vision Improves

2025
- Fully Automatic
  - American Goldfinch? yes
## Field Guides

### Attribute Groups
- **Basics**
- Pattern
- Head
- Body
- Flight
- Illustrations
- Scientific Name
- Extras

### Attributes

#### Basics
- **Location**
  - Common
- **Location**
  - Uncommon
- **Shape**
- **Size**
- **Color**
  - Primary
- **Color**
  - Secondary
- **Habitat**
- **Bill Shape**

[www.whatbird.com](http://www.whatbird.com)
You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

**Example Questions**

What is the color of the underparts of the bird?

Select at least one. If the underparts aren't visible, make your best guess, then select "Guessing". If the color is a mixture of two colors, select both (e.g., for blue-green select blue and green). If the underparts have multiple regions or patterns with multiple colors, select all relevant colors (e.g., for yellow with black stripes, select yellow and black).
You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

What is the color of the underparts of the bird?

Select at least one. If the underparts aren't visible, make your best guess, then select "Guessing". If the color is a mixture of two colors, select both (e.g., for blue-green select blue and green). If the underparts have multiple regions or patterns with multiple colors, select all relevant colors (e.g., for yellow with black stripes, select yellow and black).
You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

**What is the color of the underparts of the bird?**

Select at least one. If the underparts aren't visible, make your best guess, then select "Guessing". If the color is a mixture of two colors, select both (e.g., for blue-green select blue and green). If the underparts have multiple regions or patterns with multiple colors, select all relevant colors (e.g., for yellow with black stripes, select yellow and black).
You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

What is the color of the underparts of the bird?

Select at least one. If the underparts aren't visible, make your best guess, then select "Guessing". If the color is a mixture of two colors, select both (e.g., for blue-green select blue and green). If the underparts have multiple regions or patterns with multiple colors, select all relevant colors (e.g., for yellow with black stripes, select yellow and black).
Example Questions

You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

What is the **pattern of the breast** of the bird?

Select one. If the breast isn't visible, make your best guess, then select "Guessing".

- Solid
- Multi-Colored
- Striped
- Spotted

[Go Back] [Guessing] [Probably] [Definitely]
Example Questions

You will be asked to answer a series of questions based on identifying visual features from the bird image on the left. Closely follow the specific instructions for each question. Holding the mouse over each selectable option for 1 second will provide additional instructions or examples.

What is the shape of the bill/beak?

Select one. If the beak isn't visible, make your best guess, then select "Guessing".

- Cone
- Curved (up or down)
- Dagger
- Hooked
- Hooked Seabird
- Needle
- Spatulate
- Specialized
Basic Algorithm

Input Image ($x$)

**Question 1:** Is the belly black?

A: NO

Max Expected Information Gain

**Question 2:** Is the bill hooked?

A: YES

Max Expected Information Gain

...
Without Computer Vision

Input Image ($x$)

Class Prior

Max Expected Information Gain

Question 1: Is the belly black?

A: NO

Max Expected Information Gain

Question 2: Is the bill hooked?

A: YES

\[
\begin{align*}
&\text{Class Prior} \\
&\text{Max Expected Information Gain} \\
&\text{Question 1: Is the belly black?} \\
&\quad \text{A: NO} \\
&\quad \text{Max Expected Information Gain} \\
&\text{Question 2: Is the bill hooked?} \\
&\quad \text{A: YES} \\
&\ldots
\end{align*}
\]

\[
\begin{align*}
p(c) &\quad p(c | u_1) \\
p(c | u_1) &\quad p(c | u_1, u_2)
\end{align*}
\]
Select the next question that maximizes expected information gain:

- Easy to compute if we can to estimate probabilities of the form:

\[ p(c \mid x, u_1, u_2 \ldots u_t) \]

Object

Image

Sequence of user responses
Basic Algorithm

\[ p(c \mid x, u_1, u_2 \ldots u_t) \]

\[ \approx p(u_1, u_2 \ldots u_t \mid c) p(c \mid x) / Z \]

- Model of user responses
- Computer vision estimate
- Normalization factor
Basic Algorithm

\[ p(c \mid x, u_1, u_2 \ldots u_t) \]

\[ \approx \frac{p(u_1, u_2 \ldots u_t \mid c) p(c \mid x)}{Z} \]

- Model of user responses
- Computer vision estimate
- Normalization factor
Modeling User Responses

• Assume: \( p(u_1, u_2 \ldots u_t | c) \approx \prod_{i=1}^{t} p(u_i | c) \)

• Estimate \( p(u_i | c) \) using Mechanical Turk

What is the color of the belly?

Pine Grosbeak
Incorporating Computer Vision

- Use any recognition algorithm that can estimate: \( p(c | x) \)
- Two simple methods:

\[
p(c | x) \propto \exp \{ \gamma \cdot m(x) \}
\]

1-vs-all SVM

\[
p(c | x) \propto \prod_i p(a_i | c)
\]

Attribute-based classification

[Lampert et al. ’09, Farhadi et al. ‘09]
• Use combination of features to learn linear SVM classifiers

- Geometric Blur
- Self Similarity
- Color SIFT, SIFT
- Color Histograms
- Color Layout
- Bag of Words
- Spatial Pyramid

Multiple Kernels

[Vedaldi et al. ’08, Vedaldi et al. ’09]
Birds 200 Dataset

- 200 classes, 6000+ images, 288 binary attributes
- Why birds?

Black-footed Albatross
Groove-Billed Ani
Parakeet Auklet
Field Sparrow
Vesper Sparrow
Arctic Tern
Forster’s Tern
Common Tern
Baird’s Sparrow
Henslow’s Sparrow
Birds 200 Dataset

- 200 classes, 6000+ images, 288 binary attributes
- Why birds?

Black-footed Albatross
Groove-Billed Ani
Parakeet Auklet
Field Sparrow
Vesper Sparrow

Arctic Tern
Forster’s Tern
Common Tern
Baird’s Sparrow
Henslow’s Sparrow
200 classes, 6000+ images, 288 binary attributes

Why birds?
Comparing Different User Models
Perfect Users: 100% accuracy in $8 \approx \log_2(200)$ questions if users answers agree with field guides...
Results: Without Computer Vision

Real users answer questions

MTurkers don’t always agree with field guides...
Results: Without Computer Vision

Real users answer questions

MTurkers don’t always agree with field guides...

Q: Is the belly red?  yes (Def)
Q: Is the breast black? yes (Def.)
Q: Is the primary color red? yes (Def.)
Results: Without Computer Vision

Probabilistic User Model: tolerate imperfect user responses

Q: Is the belly red?  yes (Def)
Q: Is the breast black? yes (Def.)
Q: Is the primary color red? yes (Def.)
Results: With Computer Vision
Results: With Computer Vision

Users drive performance: 19% → 68%

Just Computer Vision

19%
Results: With Computer Vision

Computer Vision Reduces Manual Labor: 11.1 → 6.5 questions
Without computer vision:
   Q #1: Is the shape perching-like? no (Def.)
With computer vision:
   Q #1: Is the throat white? yes (Def.)
User Input Helps Correct Computer Vision

**Magnolia Warbler**

Is the breast pattern solid?

- no (definitely)

- **Magnolia Warbler**

**Common Yellowthroat**

- **computer vision** → Common Yellowthroat
Recognition is Not Always Successful

- Acadian Flycatcher
- Unlimited questions
- Least Flycatcher

- Parakeet Auklet
- Least Auklet
- Is the belly multi-colored? yes (Def.)
Recognition of fine-grained categories

Users drive up performance

More reliable than field guides

Computer vision reduces manual labor

19% increase in performance

11.1 → 6.5 questions
Summary

11.1 → 6.5 questions

More reliable than field guides

Users drive up performance

Recognition of fine-grained categories

Computer vision reduces manual labor

19%
Summary

Recognition of fine-grained categories

Users drive up performance

More reliable than field guides

Computer vision reduces manual labor

11.1 → 6.5 questions

19%
Summary

Recognition of fine-grained categories

Users drive up performance

More reliable than field guides

Computer vision reduces manual labor

11.1 → 6.5 questions

19%
Drawbacks

• Relies on part and attribute questions

• This may be:
  • Hard to define for some categories, e.g. handbags, sofas, etc
  • Hard to annotate due to lack of domain-specific expertise or language barriers
Similarity comparison based framework

Click on the bird to the right that is the most similar species to the bird above.

Instructions
Interactive categorization

Optimal display found by maximizing information gain

Wah et al, CVPR 14
Learned embeddings

Each image is visualized in as a point in 2d space
Interactive categorization

Deterministic users
No computer vision

Deterministic users
With computer vision

Simulated noisy users
With computer vision

Wah et al, CVPR 14
Localized similarity

Which of the two images is more similar to this image?

Nonlocalized similarity comparison

Localized similarity comparison

Wah et al, WACV 15
Localized similarity comparisons

(a) Localized comparison

(b) Nonlocalized comparison
Localized vs. Non-localized

Parts were found by clustering HOG features

11.53 → 9.85 questions
Annotations are also faster
Related Work

20 Questions Game [20q.net]
onMoby [IQEngines.com]
Field Guides [whabird.com]

Botanist’s Electronic Field Guide [Belhumeur et al. ‘08]
Oxford Flowers [Nilsback et al. ‘08]

Attributes [Lampert et al. ‘09] [Farhadi et al. ‘09] [Kumar et al. ‘09]

Many Others: Crowdsourcing, Information Theory, Relevance Feedback, Active Learning, Expert Systems, ...
Further thoughts and readings ...

- Papers discussed today:
  - Visual recognition with humans in the loop, Branson et al., ECCV 2010
  - Similarity comparisons for interactive fine-grained categorization, Wah et al., CVPR 2014
  - Learning localized perceptual similarity metrics for interactive categorization, Wah et al., WACV 2015

- Minimize annotation effort:
  - Active learning, better user interfaces

- Learning perceptual similarity:
  - Stochastic Triplet Embedding, L van der Maaten, K Weinberger

- Computer vision and human computation workshop, CVPR 2014