EE/CNS 148

Lecture 1 - Introduction
Spring 2004 - Caltech
Applications of visual recognition

- Biometrics: Iris, Fingerprint, Face recognition
- Toys, home robots: Recognition of people, rooms and household objects
- Manufacturing: quality control, assembly
- Surveillance, security: Detect/recognize people, objects
- Military: tanks, people, targets…
- Web, consumer: index into image databases
- Manufacturing: assembly, detect bad parts
Toys and home robots
Fingerprints

http://www.digitalpersona.com/
Iris
Face detection and recognition

http://www.ius.cs.cmu.edu/IUS/ursrp0/har/FaceDemo/gallery-inline.html
Example: Groundhog Day

73 keyframes retrieved
53 correct, first incorrect ranked 27

Rank: 12  35  50  69
Example: Groundhog Day
Retrieve shots from Lola and Groundhog Day
Retrieved shots in groundhog day for search on Sony logo
Fundamental problems

- Individual objects
- Object categories
- Scenes: specific and category
- Discrimination vs. detection vs. recognition

- How many objects and categories?
- Taxonomical organization.
Recognition, Detection, Discrimination

Tell me what you see.
Can you find faces?
Which face is Hope’s?
Objects, categories, ...

individual objects

`visual' object classes

`functional' classes
From shape to function
How many object categories can you recognize?
How many object categories can you recognize?

None, it’s a delusion
100
1000
10,000
100,000
1,000,000
More than that!

Current best estimate: 30,000
(3000 entry-level)
Biederman ‘84
What is this?
What is this?
Taxonomies and entry-level categories

Bird

Penguin

sparrow
Visual taxonomy

- OBJECTS
  - ANIMALS
    - VERTEBRATE
      - MAMMALS
        - TAPIR
      - BOAR
    - BIRDS
  - PLANTS
  - INANIMATE
    - NATURAL
    - MAN-MADE
      - CAMERA

- ...
Challenges in recognition

- Variation in size, position, rotation, lighting, color
- Occlusion / presence-absence of features
- Deformations
- Clutter
- Brand-new objects
Lighting invariance

http://cvc.yale.edu/projects/yalefaces/yalefaces.html
Viewpoint invariance
Feature invariance
Presence/absence of features

www.corbis.com

occlusion
Presence/absence of features
Correspondence and caricatures
Variability within a category

Intrinsic

Deformation
Deformation
Clutter
More deformation
Learning

• Need for learning
• It happens all the time
• Learning amongst clutter
• How many examples do we need? One-shot learning.
• Category formation
Meet a student in vision
Learn a new class X
Spot the X
Meet the Boletus Edulis
Supervised learning
(Almost) unsupervised

Positive examples

Negative examples
Representation

• Models vs images

• Matching models to images
Part similarity
Importance of `mutual position'
Model: constellation of Parts

Fischler & Elschlager, 1973

• Yuille, ‘91
• Brunelli & Poggio, ‘93
• Lades, v.d. Malsburg et al. ‘93
• Cootes, Lanitis, Taylor et al. ‘95
• Amit & Geman, ‘95, ‘99
• Perona et al. ‘95, ‘96, ‘98, ‘00

Tanaka et al., 1993

Perrett & Oram, 1993
### Generative probabilistic model

#### Model (Parameters)

<table>
<thead>
<tr>
<th>Object shape pdf</th>
<th>Appearance pdf and prob. of detection</th>
<th>Clutter pdf</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x) = G(x</td>
<td>\mu, \Sigma)$</td>
<td>$p(x) = A^{-1}$ (uniform)</td>
</tr>
</tbody>
</table>

#### Example

1. **Object Part Positions**
2. **Part Absence**
3a. **N false detect**
3b. **Position f. detect**

- $p(\text{N detect.}) = \text{Poisson}(N_1|\lambda_1)$
- $p(\text{N detect.}) = \text{Poisson}(N_2|\lambda_2)$
- $p(\text{N detect.}) = \text{Poisson}(N_3|\lambda_3)$