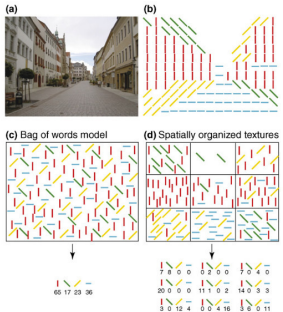
Biological Vision and Applications
Module 07-03: Part-based recognition

Hiranmay Ghosh

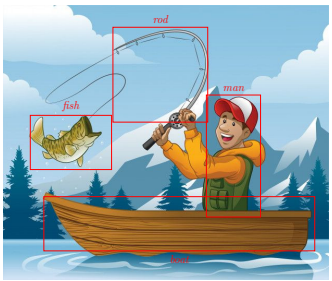
Visual recognition

Holistic vs. Part-based

Holistic representation



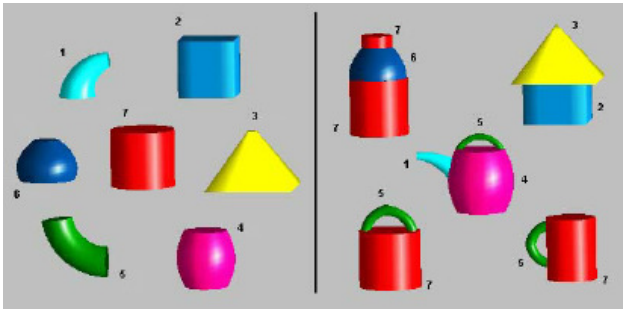
Structured (part-based) representation



EdPuzzle: Dual Process Theory

Part-based object recognition

An object is composed of some elementary 3D parts



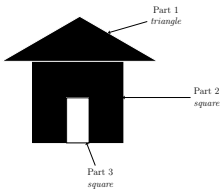
geons

- Composability as in natural language
  - Parts = visual words (elementary units for visual recognition)

Irani & Ware. The Effect of a Perceptual Syntax on the Learnability of Novel Concepts

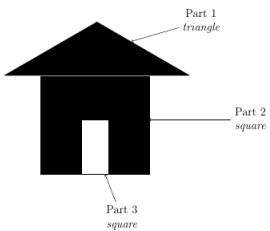
How to define a part?

- Operational definition: Parts are what can be detected by part detectors
- Definition based on principle of simplicity: Parts are polynomial surfaces approximating closed, non-overlapping image regions that optimally partition the image (MDL)



The elementary parts (geons) are characterized by their shapes, size, colors and textures.

Perceptual organization of objects



Each part (geon) is characterized by shape and relative size

Structural relations:

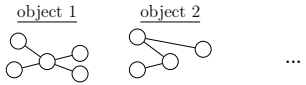
Part-1 *above* Part-2
Part-3 *contained-in* Part-2

- Geon-diagrams with similar structural composition represent same kind of objects
  - Appearance model for a geon: (shape, rel size)
  - Relations between geons (structure)
- Geon-diagrams with identical geons, but with different structural composition may represent different kind of objects

- Semantics:
  - A *above* B $\rightarrow$ B *supports* A
  - A *contained-in* B $\rightarrow$ A *part-of* B

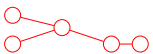
Object recognition with perceptual model

An object can be represented like a graph (geon diagram)



Each node represents a geon
- *Characterized by shape and rel size*

Each edge represents a relation between two geons
- *Above, contained-in, etc.*



Observed object Graph

Which object graph explains the observation best ?

Graph matching: ... (short paper) *

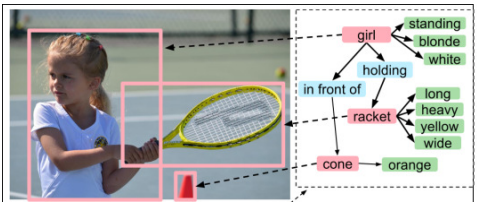
Probabilistic model and learning

- Let G be a finite vocabulary of geons, R a finite set of relations
- Each object is characterized by a probabilistic geon-graph: (V, E)
  - $v \in G$: a geon, characterized by a probabilistic appearance model $(type, size)$
  - $e \in R$: Probabilistic relations between a pair of vertices
  - Parameters can be learned over a large number of observations
- An observation is also a geon-graph
- Inference: Which object explains the observed graph the best ?

- Probabilistic model & learning: Crandall & Huttenlocher. Weakly Supervised Learning of Part-Based Spatial Models ...
- Application to neural network: Krause, et al. Learning Features and Parts for Fine-Grained Recognition ...

Activity recognition and Scene Graph

- Girl playing tennis
- objects + locations + interactions

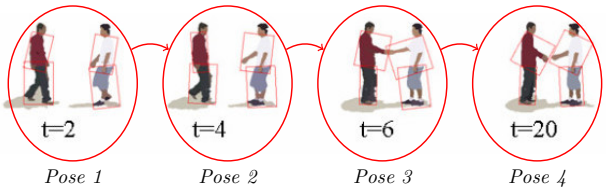


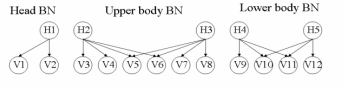
Chang, et al. Survey of scene graphs

Johnson, et al. Image Retrieval using scene graph

Event recognition

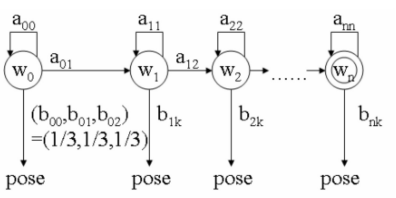
Extending Scene Graph in time





H1: Head ={0:front, 1:left, 2:right, 3:rear}
H2: ArmV={0:high, 1:mid-high, 2:mid-low, 3:low}
H3: ArmH={0:withdrawn, 1:intermediate, 2:streethed}
H4: LegV={0:high, 1:middle, 2:low}
H5: LegH={0:withdrawn, 1:intermediate, 2:streethed}

V1: angle of vector from head to face
V2: ratio of face to head
V3: y-position of hand blob
V4: y-pos of far most point of upper body
V5: upper body ellipse ratio
V6: upper body ellipse rotation
V7: x-position of hand blob
V8: x-pos of far most point of upper body
V9: y-pos of far most point of lower body
V10: lower body ellipse ratio
V11: lower body ellipse rotation
V12: x-pos of far most point of lower body



Hidden Markov Model (HMM) representation of an event

Rayoo & Aggarwal. Recognition of Composite Human Activities ...

Quiz

Quiz 07-03

End of Module 07-03