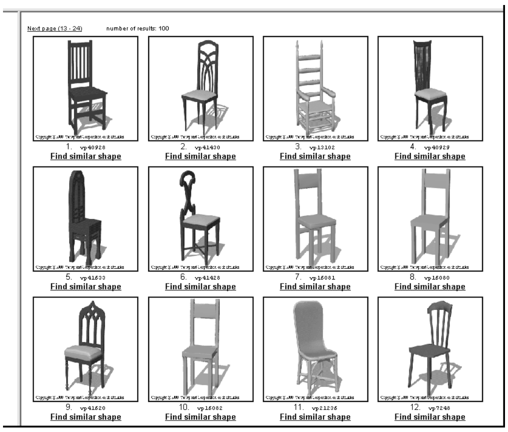
Approaches to Representing and Recognizing Objects

Visual Classification

CMSC 828J – David Jacobs

What the course is about

- Visual Classification
  - Recognizing nouns and verbs from images.
- This is one of the key problems of vision/cognition.
  - ~ half of cerebral cortex is vision.
  - Vision divides roughly into *what* and *where*.
    - This class is about *what*.



This is very hard

- What is a class?
  - Ill-defined.
  - Tremendous variability.
- And how do we relate images to objects or actions?
- Current solutions grossly inadequate.

So how do we have a class about this?

1. Learn some fundamental things relevant to visual classification
2. See how researchers have tried to apply these to visual classification.

Fundamental things

- Lectures (but also discussion of some important papers).
- Much of it mathematical and computational
  - Geometry of projection and invariance; PCA; shape spaces and shape matching; stochastic models of classes; learning theory.
- But we draw from other fields too:
  - Philosophy: what is a class?
  - Biology: how does shape vary in nature?
  - Psychology: How do people do classification?

Application of these ideas to visual classification

- Read papers and discuss.
- Shows how fundamental ideas can be used.
- How math and computation interact.
- Don’t solve big problem, but often useful for smaller problems.

Class Goal: Prepare us to solve problems of visual classification

- Learn fundamental concepts important for vision.
- Get us to think about what classification is.
- Understand state-of-the-art attempts to solve it.

Approaches to Visual Classification

- Definitional: a class is defined by the presence or absence of properties (a point in feature space).
- A class is a subspace of images.
- Class is determined by similarity of images.
  – To prototypes or exemplars
- Class represented by a generative model.
- Class described by object parts and their relations.
- Classes and generic learning.

A tour of the syllabus

How this might change

- Probably way too much material.
- Lectures may be longer than indicated.
- I am open to suggestions about other papers or topics

Requirements (1)

- Read papers before they are presented.
- Paper Questions
  - On classes with paper discussions (all except first, last and midterm) you must provide an answer to a posted question.
  - Your answer should demonstrate knowledge of the papers and critical thought.
  - You may skip when you are giving a presentation and one other time.
  - 15% of grade.

Requirements (2)

- Presentations
  - For each presentation class, two teams of students will debate
  - Advocate one of two competing research directions. Examples:
    - 3D knowledge vs. 2D
    - Better descriptors vs. better ways to use them
    - What is the best direction for building a robot that can recognize household objects?
    - How can one represent complex classes of objects, such as chairs? Machine learning, manifolds, graphical models.
  - 15% of grade.

Requirements (3)

- Midterm and Final
  - Will cover materials in lectures.
  - 50% of grade.

Requirement (4)

- Project: Implement two recognition algorithms.
  - Choose a data set. Examples:
    - Parkinson's disease, PIE, face recognition with weight change, leaves, CalTech, PASCAL, …
  - Pick two algorithms. At least one non-trivial.
  - Implement and evaluate them

Your Background

- Calculus, linear algebra, probability is essential.
- Math that makes you really learn these topics is important.
- Other math very helpful: geometry, stochastic processes, optimization.
- Knowledge of vision may help a little.

First Homework

- Readings for Wednesday.
- Review *is* due.

What is a category?

Related Papers

- Women, Fire and Dangerous Things by Lakoff, Chapters 1 and 2.
- S. Laurence and E. Margolis, ``Concepts and Cognitive Science'', in *Concepts* edited by E. Margolis and S. Laurence, MIT Press.
- L. Wittgenstein, *Philosophical Investigations,* sections 65-78.

Categorization is a critical issue in many disciplines

- Linguistics – what do words mean
- Psychology – All these issues are central, as in computer vision.
- Philosophy – What do statements mean, what does it take for them to be true.

Key questions a theory of categorization should answer

- Acquisition – how is a category learnable.
- Application – how do we decide on a category
- Composition of categories
  - Big question in language and philosophy, since we combine categories so flexibly, and can understand a category we've never heard of before (blind, green fish).
  - Interesting in vision: big fish, red chair.
    - Is a category the sum of its parts?
    - Wooden spoon. Not just a spoon that is wooden, as most wooden spoons have different shapes. How do we combine information about spoons, wood, and wooden spoons.
- What categories can occur?
  - In language, any category seems understandable. But some may be more natural. In vision, what categories are visually understood.
- Others less relevant to us.

So, what is a category? What is a chair? A bird?

Classical Theory

- Categories have definitions – bachelor = unmarried adult male
- Necessary and sufficient conditions
- May be defined in terms of other concepts, but these ultimately bottom out in statements verified by our senses.
  - Ie, *red* is visually observable by reactions of cones to light.
- Seems to offer answers to all core questions.
- Part of a larger philosophical program.
  - A statement has meaning because it can be translated into logic, with verifiable atoms.
- Held with little question for ~2,000 years.
- Initial focus of AI, cognitive science: eg., Schank, Hayes, expert systems, anthropology.

Problems with the classical theory: Plato's problem

- Reading from Wittgenstein
- Fodor's example of *painting.*
  - X covers Y with paint – exploding paint factory
  - X is an agent and X covers Y with paint – kicking over a bucket of paint
  - X is an agent and X intentionally covers the surface of Y with paint – when you paint a painting, you are not painting the canvas
  - X is an agent and X intentionally covers the surface of Y with paint and X's primary intention is to cover Y with paint – dipping painting brush into paint isn't painting it.
- Definitions are, at the least, quite difficult

Problems with classical theory: Categories have structure

- Prototype effects.
  - We'll discuss this more next week.
  - some elements seem more typical than others.
    - Sparrow is a more typical bird than ostrich
    - This is psychologically real. That is, we can answer questions like "is this a bird" faster for more typical elements.
  - Conceptual fuzziness. Membership in category seems ill-defined
    - Example: is carpet furniture?
- Are these really problems for classical theory? At least they suggest we don't use definitions to perform classification.

Other problems

- Stability of concepts: less relevant to us but interesting
- Measles: thought to be due to evil spirits, now thought to be due to virus. These both seem like core components of the concept, part of the definition. But we want to say both are talking about the same concept, not two different concepts.

Prototype theory

- One version, a category has features, (maybe weighted) and an object must have enough features to belong.
  - Explains typicality, fuzziness of categories.
  - Acquisition and application are pretty straightforward.
  - Not clear how composition works. Why is a goldfish a prototypical pet fish, but not a prototypical pet or fish
    - This can perhaps be solved with a more sophisticated view of categories as distributions that are combined.

How does this fit in with how a computer scientist sees concepts

- Concept = some subspace of a feature space; or a distribution on it.
- Definition implies concept is an axial rectanguloid in feature space.
- More general notion possible spaces might fit prototype theory. Easy to think of prototype theory as a Gaussian distribution in feature space.

Problems in CS view

- What subsets of feature space can be a concept?
  - Must be complex enough to represent real categories.
  - To learn and generalize a concept can't be too complex.
    - Notion of VC dimension is one way to formalize this.
  - Possible candidates: Gaussian, mixture of Gaussian, manifold, convex subspace.
  - Convexity is especially nice. Interpolation possible.
- Maybe mix of necessary conditions and feature subspaces. (Neoclassical theory)
  - You must touch something to pick it up.
- Where do the features come from?