CSSE463: Image Recognition

Matt Boutell
Myers240C x8534
boutell@rose-hulman.edu

What is image recognition?

In the 1960's, Marvin Minsky assigned a couple of undergrads to spend the summer programming a computer to use a camera to identify objects in a scene. He figured they'd have the problem solved by the end of the summer. Half a century later, we're still working on it.

http://xkcd.com/1425/

Agenda: Introductions to…

- The players
- The topic
- The course structure
- The course material

Introductions

- Roll call:
  - Your name
    - Pronunciations and nicknames
    - Help me learn your names quickly
  - Your major
  - Your hometown
  - Where you live in Terre Haute

Note to do a quiz question during this slide→ Q1-2

About me

Matt Boutell

U. Rochester
PhD 2005

Kodak Research intern 4 years

11th year here. CSSE120 (& Robotics), 220, 221, 230, 325; 479; 483, ME430, ROBO4x0, 4 senior theses, many ind studies

Kodak

Personal Info

Agenda

- The players
- The topic
- The course structure
- The course material

What is image recognition?

- Image understanding (IU) is “Making decisions based on images and explicitly constructing the scene descriptions needed to do so” *(Shapiro, Computer Vision, p. 15)*
- Computer vision, machine vision, image understanding, image recognition all used interchangeably
  - But we won’t focus on 3D reconstruction of scenes, that’s CSSE461 with J.P. Mellor’s specialty.
- IU is not image processing (IP; transforming images into images), that’s ECE480/PH437.
  - But it uses it
- IU isn’t pattern classification: that’s ECE597
  - But it uses it

IU vs IP

- Knowledge from images
  - What's in this scene?

- Enhancing images
  - Sharpen the scene!

- It's a sunset
- It has a boat, people, water, sky, clouds

Why IU?

- A short list:
  - Photo organization and retrieval
  - Control robots
  - Video surveillance
  - Security (face and fingerprint recognition)
  - Intelligent IP
- Think now about other apps
  - And your ears open for apps in the news and keep me posted; I love to stay current!

Agenda

- The players
- The topic
- The course structure
- The course material

What will we do?

- Learn theory (lecture, written problems) and “play” with it (Friday labs)
- See applications (papers)
- Create applications (2 programming assignments with formal reports, course project)
- Learn MATLAB. (Install it asap if not installed)
  - Instructions here: \\rose-hulman.edu\dfs\Software\Course Software\MATLAB_R2015a

Course Resources

- Moodle is just a gateway to website (plus dropboxes for labs and assignments)
- Bookmark if you haven't

http://www.rose-hulman.edu/class/csse/csse463/201620/

- Schedule:
  - See HW due tomorrow and Wednesday
- Syllabus:
  - Text optional
  - Grading, attendance, academic integrity

Agenda

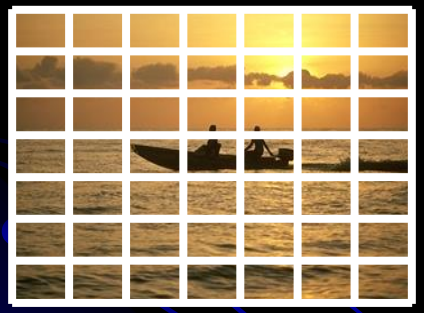
- The players
- The topic
- The course structure
- The course material

Sunset detector

- A system that will automatically distinguish between sunsets and non-sunset scenes
- I use this as a running example of image recognition
- It's also the second major programming assignment, due at midterm
  - **Read the paper tonight** (focus: section 2.1, skim rest, come with questions tomorrow; I'll ask you about it on the quiz)
  - We'll discuss features in weeks 1-3
  - We'll discuss classifiers in weeks 4-5
- A "warm-up" for your term project
- A chance to apply what you've learned to a known problem

Pixels to Predicates

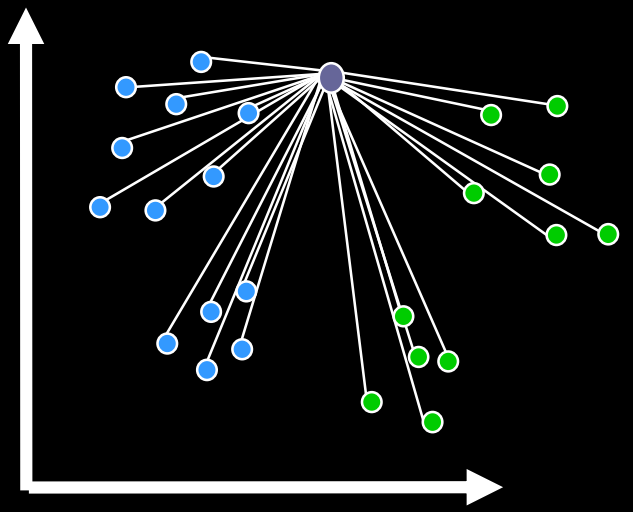
1. Extract features from images



$$x = \begin{pmatrix} 0.4561 \\ 0.1928 \\ \dots \\ 0.2756 \end{pmatrix}$$

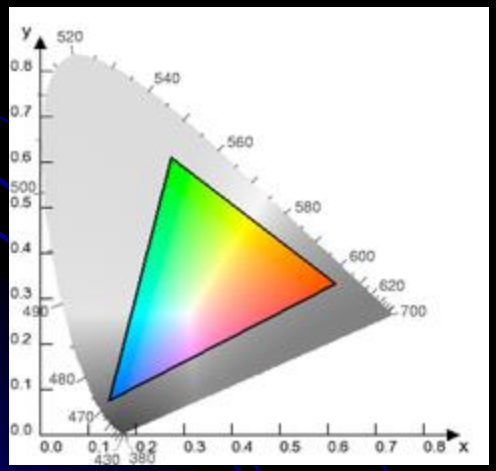
Color
Texture
Shape
Edges
Motion

2. Use machine learning to cluster and classify



Principal components
Neural networks
Support vector machines
Gaussian models

Basics of Color Images



Source: Wikipedia

- A color image is made of red, green, and blue *bands* or *channels*.
  - Additive color
    - Colors formed by adding primaries to black
  - RGB mimics retinal cones in eye.
  - RGB used in sensors and displays
  - Comments from graphics?

What is an image?

- Grayscale image
  - 2D array of pixels
    - (row,col), not (x,y)! Starts at top!
  - **Matlab demo** (preview of Friday lab):
    - Notice row-column indexing, 1-based, starting at top left
- Color image
  - 3D array of pixels. Takes 3 values to describe color (e.g., RGB, HSV)
- Video:
  - 4th dimension is time. “Stack of images”
- Interesting thought:
  - View grayscale image as 3D where 3rd D is pixel value