Admistrivia

If you don’t have a CS account, get one!

Once you have a CS account, run “apply”, ASAP.

Access to lab in BSE 328 will require these two steps

Course page is now up: http://www.cs.arizona.edu/classes/cs477/spring04
(Linked from instructor’s home page (http://kobus.ca))

Lectures and assignments will require either connecting from a UA machine, OR a login id (“me”) and password (“vision4fun”)
Course Intro Continued
Part III: Early Vision in Multiple Images

• The geometry of multiple views
  – Where could it appear in camera 2 (3, etc.) given it was here in 1 (1 and 2, etc.)?

• Stereopsis
  – What we know about the world from having 2 eyes

• Structure from motion
  – What we know about the world from having many eyes
  • or, more commonly, our eyes moving.
3D Reconstruction from multiple views

• Multiple views arise from
  – stereo
  – motion

• Strategy
  – “triangulate” from distinct measurements of the same thing

• Issues
  – Correspondence: which points in the images are projections of the same 3D point?
  – The representation: what do we report?
  – Noise: how do we get stable, accurate reports
Part IV: Mid-Level Vision

- Finding coherent structure so as to break the image or movie into big units
  - Segmentation:
    - Breaking images and videos into useful pieces
      - finding video sequences that correspond to one shot
      - finding image components that are coherent in internal appearance
  - Tracking:
    - Keeping track of a moving object through a long sequence of views
Segmentation

- Which image components “belong together”?
- Belong together = lie on the same object
- Cues
  - similar colour
  - similar texture
  - not separated by contour
  - form a suggestive shape when assembled
Tracking

- Use a model to predict next position and refine using next image
- Model:
  - simple dynamic models (second order dynamics)
  - kinematic models
  - etc.
- Face tracking and eye tracking now work rather well
Part V: High Level Vision (Geometry)

- The relations between object geometry and image geometry
  - Model based vision
    - find the position and orientation of known objects
  - Smooth surfaces and outlines
    - how the outline of a curved object is formed, and what it looks like
  - Aspect graphs
    - how the outline of a curved object moves around as you view it from different directions
  - Range data
Part VI: High Level Vision
(Probabilistic)

• Using classifiers and probability to recognize objects
  – Templates and classifiers
    • how to find objects that look the same from view to view with a classifier
  – Relations
    • break up objects into big, simple parts, find the parts with a classifier, and then reason about the relationships between the parts to find the object.
  – Geometric templates from spatial relations
    • extend this trick so that templates are formed from relations between much smaller parts
Some applications of recognition

- Digital libraries
  - Find me the picture of JFK and Marilyn Monroe embracing
- Surveillance
  - Warn me if there is a mugging in the grove
- HCI
  - Do what I show you
- Military
  - Shoot this, not that
What are the problems in recognition?

• Which bits of image should be recognized together?
  – Segmentation.

• How can objects be recognized without focusing on detail?
  – Abstraction.

• How can objects with many free parameters be recognized?
  – No popular name, but it’s a crucial problem anyhow.

• How do we structure very large model bases?
  – Again, no popular name; abstraction and learning come into this
Matching templates

- Some objects are 2D patterns
  - e.g. faces
- Build an explicit pattern matcher
  - discount changes in illumination by using a parametric model
  - changes in background are hard
  - changes in pose are hard
Relations between templates

- e.g. find faces by
  - finding eyes, nose, mouth
  - finding assembly of the three that has the “right” relations
Some typical results in cluttered images are shown in the following figures.
Horse grouper
Returned data set
Part VII: Some Applications in Detail

- Finding images in large collections
  - searching for pictures
  - browsing collections of pictures
- Image based rendering
  - often very difficult to produce models that look like real objects
    - surface weathering, etc., create details that are hard to model
    - Solution: make new pictures from old
Querying from 25000 images (2000 returned by the filter).
Official Start of Course
Image Formation

Images = spatial sampling (typically of light on a plane)
Typically encoded as an array of values

Images = interaction of light with camera

They carry information about the world when light form a source has interacted with the world

We will briefly study in turn
Light
Camera (geometry)
Camera (response)
The interaction of light and the world

§1 (focus on §1.1.1, §1.4.2), highlights of §4
Geometrically approximate light by rays (vectors)

Typical scene has light going in a multitude of directions

The light from a point in the world imaged on a camera plane (or retina) got there by being the bit that went the right way (obvious, so the part on projection should be trivial!)

A bit of light has additional characteristics (energy/wavelength, polarization)

The light in a certain direction is a mix, so we get a spectrum over wavelengths

Spectrum records how much power is at each wavelength

Visible portion is about 400 to 700 nm (Certain applications may require modeling some IR and/or UV also).
Two disparate source spectra

**Fig. 4.1.** Wavelength composition of light from a tungsten-filament lamp [typified by CIE ILL A (Sect. 4.6)]. Relative spectral power distribution curve. Color temperature: 2856 K

**Fig. 4.2.** Wavelength composition of light from a daylight fluorescent lamp. Typical relative spectral power distribution curve. Correlated color temperature: 6000 K. (Based on data of Jerome reported in [Ref. 3.14, p. 37])
Sensors (including those in your eyes) have a varied sensitivity over wavelength. Different variations lead to different kinds of sensor responses ("colors" in a naïve sense).

The sensor sensitivity spectra for a particular color camera (they vary a lot).
Image Formation (Spectral)

\[(R, G, B) = (R, G, B) \ast d\]