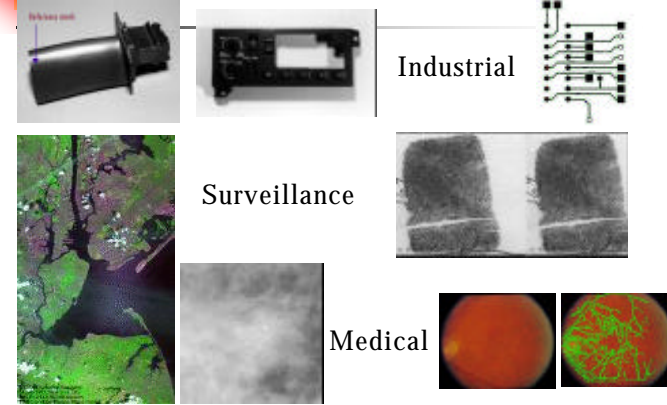


Overview of Computer Vision



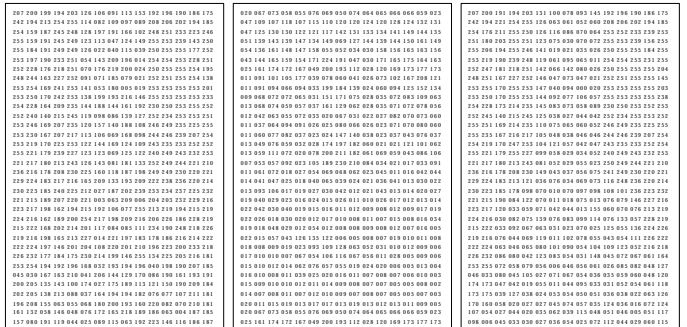
C306, fall 2001
Martin Jagersand
Zach Dodds, Greg Hager

Uses of Computer Vision



The "Vision Problem"

Input → Vision Algorithm → Output



The "Vision Problem"

Input → Vision Algorithm → Output



Allies and Inspiration



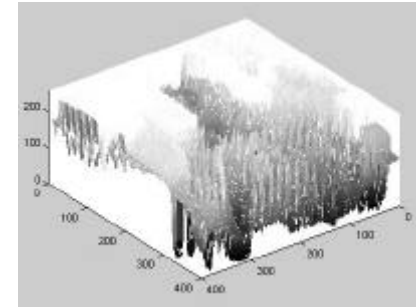
- Image Processing
- Cognitive Science
- Physics, Optics
- Engineering
- Biology, Psychology
- Computer Science
- Graphics, AI, NNets, ...

Recognizing objects

What is an image? What is an object?

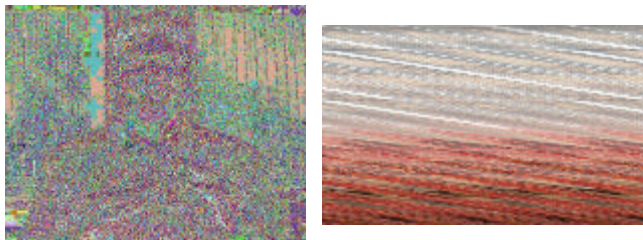
```

287 208 184 194 253 138 100 895 187 122 102 186 186 175
242 194 222 224 255 124 874 882 872 876 208 208 202 194 185
254 178 204 255 248 122 122 125 111 881 252 252 252 252 258
252 172 261 252 248 122 895 848 884 188 252 252 252 126 254
251 187 182 255 248 123 864 877 876 832 258 252 251 181 255
252 218 188 258 258 888 882 148 128 812 254 254 252 258 255
252 258 188 258 251 188 141 181 188 884 258 252 251 251 288
248 248 188 257 252 111 888 128 881 881 252 251 252 251 142
252 252 171 254 252 142 857 122 888 817 252 252 252 254 281
252 258 178 252 252 128 124 127 128 878 252 252 252 254 237
254 258 188 252 252 144 122 888 888 128 258 252 254 254
252 244 148 215 242 123 852 842 881 877 252 254 252 252 252
254 258 188 212 222 117 188 882 128 878 248 248 252 252 252
254 254 187 252 257 128 878 888 888 878 248 248 258 257 254
252 218 178 258 252 112 128 188 882 872 242 252 252 252 252
252 222 178 248 257 111 882 842 881 881 248 248 242 252 252
252 217 188 212 242 188 878 848 188 842 248 248 242 252 252
258 218 178 258 252 124 877 882 128 888 248 248 258 258 251
252 254 182 112 122 852 882 124 881 122 248 252 252 254
258 252 182 182 112 878 888 124 128 128 128 252 252 252 252
252 212 128 188 124 871 888 821 887 252 252 252 252 252
252 217 122 848 872 876 858 848 812 182 872 876 882 252 252
254 218 841 182 888 182 878 112 118 184 888 178 888 252 252
252 222 848 111 877 872 888 848 888 888 128 888 128 254 258
252 218 882 842 874 142 812 171 128 872 882 882 142 217 222
252 224 878 841 852 122 882 128 112 188 128 882 252 258
258 232 118 188 841 182 128 128 887 172 878 888 887 842 174
252 254 878 872 128 888 828 888 182 874 882 118 188 882 881
842 834 888 842 142 827 128 188 888 888 888 112 887 882 137
182 182 872 882 812 882 818 882 118 832 888 888 888 871 134
182 182 812 122 882 828 878 881 882 818 888 888 888 878 138
184 181 888 812 812 838 848 882 872 838 142 848 832 887 138
128 874 838 882 838 874 812 142 142 888 848 888 882 142
122 818 837 872 812 822 848 882 838 884 128 842 832 112 131
  
```



Recognizing objects

What is an image? What is an object?



Computer Vision vs Image processing

- Image Processing
 - Mostly concerned with *image-to-image* transformations
 - Filtering
 - Enhancement
 - Compression
- Computer Vision
 - Concerned with how images *reflect the 3D world*
 - Filtering *for feature extraction*
 - Enhancement *for recognition/detection*
 - Compression *that preserves geometric information in images*

Tasks in Vision

Perception:

- **What:** Label what is in a scene: What people, animals, objects...

Action:

- **Where:** Determine the coordinates of where something is. (Mishkin, Ungerleider)
- **How:** Use vision to perform some physical manipulation task. (Goodale etc.)

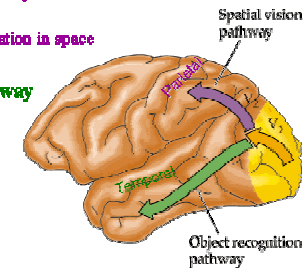
Compare in Human vision: Dorsal and Ventral Pathways

Dorsal (magno) Pathway

- to parietal lobe
- spatial vision – localization in space
- “WHERE”

Ventral (parvo) Pathway

- to temporal lobe
- object recognition
- “WHAT”



What: Recognizing people and objects

How do we determine that these are the same objects?

Two approaches:

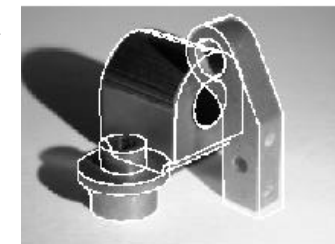
1. Shape
2. Texture



Model based recognition: Find and match the shape

Example:

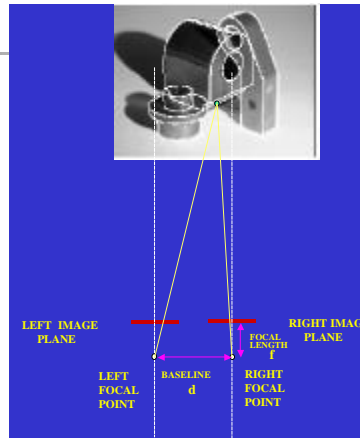
- Find the outlines of objects.
- Try to generate 3D or some pseudo 3D geometric description
- Check database for a similar geometric object



Model based recognition: Techniques.

Stereo:

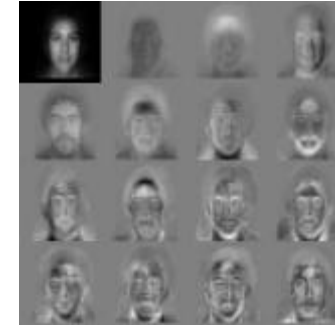
- From two (or more) images, determine the geometry of the scene by *matching* corresponding areas of the images



Apperance based recognition Match the 2D visual "texture"

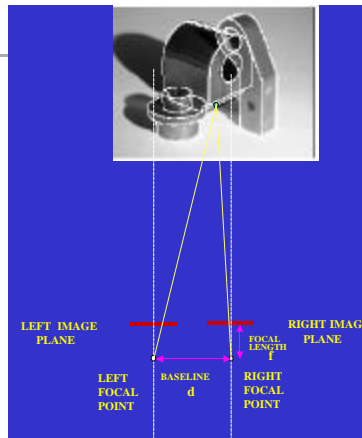
Methods:

- Spatial:** Determine "likeness" by correlation.
- "Frequency":** Transform and measure in fourier or KL space.



Where

- Use stereo to determine 3D location with respect to camera.
- If we know where the camera is we can determine 3D world position



How (to act in the world)?



Motion estimation and tracking

Detecting motion:



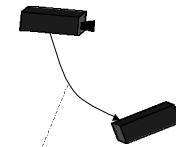
> 50

Candidate areas for motion

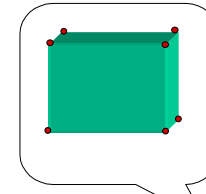


Moving the camera

- Like stereo!



The change in spatial location between the two cameras (the "motion")



Locations of points on the object (the "structure")

Tracking

- Goal: Stabilizing motion.
- Method: Find same region in consecutive images



Vision and Sensory based motion control



Hand-Eye coordination in Man and Machine

- New exciting course next semester (by yours truly) covering:
 1. **Visual motion**: how to track and analyze it.
 2. **Motor movement**: how to make robots do useful tasks.
 3. **How does the brain do it?** Human visual-motor function, anatomy and physiology.
 4. Putting things together: **Man-Machine interfaces** using vision and physical interaction.

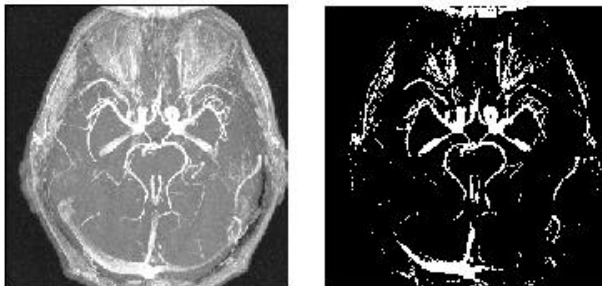
Main focus of c610:

- Most other courses are about static vision.
What is in the image? Im → Processing → Symbol Interp
- This course: Vision for registering, understanding and generating motion. **How to act based on visual information.**

```

      Images
        ↓
      Agent doing physical motion
        ↓
      Real world
    
```

Applications of Computer Vision: Medical Imaging



Applications of Computer Vision: Image Databases

(Courtesy D. Forsyth & J. Ponce)



From a search for horse pix in 100 horse images and 1086 non-horse images

Applications of Computer Vision: Data Acquisition



Welcome to Visual Computing

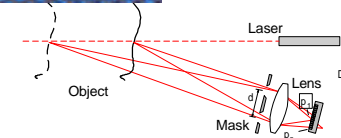
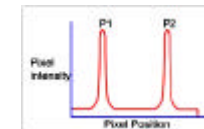
Seminars:
Wed Lunch



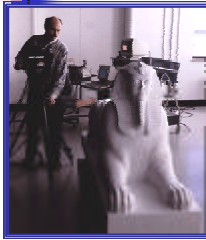
Rendering

- Basic techniques from the forward modeling perspective (traditional rendering)
 1. Input: 3D description of 3D scene & camera
 2. Solve light transport through environment
 3. Project to camera's viewpoint
 4. Perform ray-tracing
- Basic technique from the inverse modeling perspective (image-based rendering)
 1. Collect one or more images of a real scene
 2. Warp, morph, or interpolate between these images to obtain new views

3D sensors: Laser scanner



3D: Direct Replica of Large Museum Objects



3-D scan of the sculpture



STL model



Replicas produced
by a 3-D Laser Sintering
Machine (Kalplast Inc.)

Computer Vision Summary:

The engineering solution: (Does not work well)

- 3D model based recognition and location determination solves what & where.

The biological solution? (Not proven yet)

- What: 2D recognition by pattern matching
- How: 2x 2D image based techniques

Your solution?