Image Basics

Bob Fisher
School of Informatics
University of Edinburgh

Many slides in this lecture are due to other authors; they are credited on the bottom right
The dream of a machine that can see

First encounter

What do you see?

credit V. Ferrari
Second encounter

And now?
First encounter

Rapid visual recognition is important!
The goal of computer vision

To extract “meaning” from pixels

Humans are remarkably good at this…

Source: “80 million tiny images” by Torralba et al.
What kind of information can be extracted from an image?

- 2D/3D measurements
- Semantics / meaning
- Counts
- Quantities used for decisions (e.g. medical)
Vision as measurement device

Despite the reliance on vision as a measurement device, real-time stereo and structure from motion (SfM) remain challenging tasks. However, recent studies have shown promising results.

1. NASA Mars Rover
   - Relating images
   - Structure & Motion recovery
   - Dense Matching
   - 3D Model Building

2. Internet photo collections
   - Input sequence
   - Feature matches
   - 3D features and cameras
   - Dense depth maps
   - 3D surface model

Pollefeys et al.

Goesele et al.

Credit: V. Ferrari
Vision as a source of semantic information
Object categorization

- sky
- building
- flag
- banner
- face
- street lamp
- wall
- bus
- cars

Slide credit: Fei-Fei, Fergus & Torralba

Credit V. Ferrari
Scene and context categorization

- outdoor
- city
- traffic
- ...

slide credit: Fei-Fei, Fergus & Torralba
credit V. Ferrari
Qualitative spatial information

slanted

non-rigid moving object

vertical

rigid moving object

horizontal

slide credit: Fei-Fei, Fergus & Torralba

credit V. Ferrari
Why is computer vision difficult?
Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba
Challenges: illumination

image credit: J. Koenderink
Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba
Challenges: non-rigid deformation

Xu, Beihong 1943
Challenges: occlusion

Magritte, 1957
Challenges: background clutter

Emperor shrimp and commensal crab on a sea cucumber in Fiji

Photograph by Yim Lamal

© 2007 National Geographic Society. All rights reserved.

credit V. Ferrari
Challenges: object intra-class variation
Inherent ambiguity of the problem

• Many different 3D scenes could have given rise to a particular 2D image

• Possible solutions
  – Bring in more constraints (more images)
  – Use prior knowledge about the structure of the world

• Need a combination of geometric and statistical methods
Connections to other disciplines

- Artificial Intelligence
- Robotics
- Machine Learning
- Computer Graphics
- Cognitive science
- Neuroscience
- Image Processing

slide credit: Lazebnik
What have we learned?

• Images are numbers – hard to interpret
• Many kinds of results extractable from images
• Many issues make image analysis difficult
• Computer vision links to (and requires) many topics