Lecture: Introduction to Computer Vision
Juan Carlos Niebles and Ranjay Krishna
Stanford Vision and Learning Lab
Welcome to CS131
CS131 is the introductory course for computer vision

- **CS131 (fall, 2017):**
  - Enthusiastic undergrads
  - Want to get to know this exciting technology
  - Pre-req to more advanced vision classes

- **CS231a (winter, 2018, Prof. Silvio Savarese):**
  - Advanced Computer Vision
  - Seniors, masters, and PhDs

- **CS231n (spring, 2018):** Deep Learning and Convolutional Neural Networks
Today’s agenda

• Introduction to computer vision
• Course overview
Quiz?
What about this?
What is (computer) vision?

Image (or video) → Sensing device → Interpreting device → Interpretations

garden, spring, bridge, water, trees, flower, green, etc.
The goal of computer vision

• To bridge the gap between pixels and “meaning”
What is (computer) vision?

Image (or video) → Sensing device → Interpreting device → Interpretations

- garden, spring, bridge, water, trees, flower, green, etc.
1981: Nobel Prize in medicine

Hubel & Wiesel
Human vision is superbly efficient

Potter, Biederman, etc. 1970s
150 ms!!

Change blindness

Rensink, O’regan, Simon, etc.
Change blindness
segmentation
Perception
What is (computer) vision?

Image (or video) → Sensing device → Interpreting device → Interpretations

garden, spring, bridge, water, trees, flower, green, etc.
The goal of computer vision

• To bridge the gap between pixels and “meaning”
Origins of computer vision: an MIT undergraduate summer project

THE SUMMER VISION PROJECT
Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".
What kind of information can we extract from an image?

• Metric 3D information
• Semantic information
Vision as measurement device

Pollefeys et al.

Goesele et al.
Vision as a source of semantic information

- **Objects**: The Wicked Twister, Cedar Point, Lake Erie, tree, tree, bench, deck, Ferris wheel, ride, umbrella, people sitting on ride, maxair, pedestrians, text/writing, faces, gestures, motions, emotions...

- **Activities**: Deck, people waiting in line, carousel, people sitting on ride, maxair, umbrellas, pedestrians.

- **Scenes**: Amusement park, sky, water.
Why study computer vision?

- Vision is useful: Images and video are everywhere!

Surveillance and security

Medical and scientific images
Special effects: shape and motion capture
3D urban modeling

Bing maps, Google Streetview

Source: S. Seitz
3D urban modeling: Microsoft Photosynth

http://photosynth.net

Source: S. Seitz
Face detection

- Many digital cameras now detect faces
  - Canon, Sony, Fuji, ...

Source: S. Seitz
The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

Sony Cyber-shot® T70 Digital Still Camera

Source: S. Seitz
Face recognition: Apple iPhoto software

http://www.apple.com/ilife/iphoto/
Biometrics

How the Afghan Girl was Identified by Her Iris Patterns

Source: S. Seitz
Biometrics

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
iphone X just introduced face recognition
Optical character recognition (OCR)

Technology to convert scanned docs to text
- If you have a scanner, it probably came with OCR software

Digit recognition, AT&T labs

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Source: S. Seitz
Toys and Robots
Mobile visual search: Google Goggles

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

- Landmark
- Book
- Contact Info
- Artwork
- Places
- Wine
- Logo
Mobile visual search: iPhone Apps

Query Images

- Perspective
- Zoom
- Rotation
- Coverage
- Lighting
- Logos
- Occlusion
- Blur
- Zoom

Matched Image
• **Mobileye**: Vision systems in high-end BMW, GM, Volvo models
  
  – “In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians.”

Source: A. Shashua, S. Seitz
Vision in supermarkets

**LaneHawk by EvolutionRobotics**

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it… “

*Source: S. Seitz*
Amazon Go
Vision-based interaction (and games)

- Microsoft’s Kinect
- Sony EyeToy
- Assistive technologies

Source: S. Seitz
Augmented Reality
Virtual Reality
Vision for robotics, space exploration

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.