Lecture 1: Introduction to “Computer Vision”

Dr. Juan Carlos Niebles
Stanford AI Lab

Prof. Fei-Fei Li
Stanford Vision Lab
Welcome to CS131
CS131 is the introductory course for computer vision

- CS131 (fall, 2015):
  - Enthusiastic undergrads
  - Want to get to know this exciting technology
  - Pre-req to more advanced vision classes

- CS231a (spring, 2016, Prof. Silvio Savarese)
  - Similar to existing CS231a
  - Seniors, masters, and PhDs

- CS231b (spring, 2016): Cutting Edge Computer Vision
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.
What is it related to?

Computer Vision

- Neuroscience
- Psychology
- Biology
- Computing Science
- Engineering
- Mathematics
- Physics
- Robotic Systems
- Speech
- Image Processing
- Machine Learning
- Cognitive Sciences
- Information Retrieval
- Graphics, Algorithms, System Theory, …
The goal of computer vision

- To bridge the gap between pixels and "meaning"

What we see | What a computer sees
---|---

La Gare Montparnasse, 1895

Source: S. Narasimhan

Fei-Fei Li

Lecture 1 - 9

23-Sep-15
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
Change blindness

Rensink, O’regan, Simon, etc.
Change blindness

Rensink, O’regan, Simon, etc.
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”

What we see

What a computer sees

Source: S. Narasimhan

Fei-Fei Li

Lecture 1 - 22 23-Sep-15
Origins of computer vision: an MIT undergraduate summer project

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.

Fei-Fei Li

Lecture 1 - 25 23-Sep-15
Vision as a source of semantic information

Objects
Activities
Scenes
Locations
Text / writing
Faces
Gestures
Motions
Emotions...

The Wicked Twister
Cedar Point
Lake Erie
Ferris wheel
12 E
sky
amusement park
water
tree
ride
people waiting in line
people sitting on ride
umbrellas
carousel
pedestrians
maxair
desk
bench
Why study computer vision?

• Vision is useful: Images and video are everywhere!
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
Smile detection

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/
How the Afghan Girl was Identified by Her Iris Patterns
Biometrics

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Source: S. Seitz

Fei-Fei Li  Lecture 1 - 35  23-Sep-15
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
Automotive safety

- **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
Vision-based interaction (and games)

Microsoft’s Kinect

Sony EyeToy

Assistive technologies

Source: S. Seitz
Vision for robotics, space exploration

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

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.
Today’s agenda

• Introduction to computer vision
• Course overview
Contacting instructor and TAs

• Instructors:
  – Dr. Juan Carlos Niebles
  – Prof. Fei-Fei Li

• Teaching Assistants
  – Ranjay Krishna, Masters, CS
  – Zelun (Alan) Luo, Masters, CS
  – TBD
Contacting instructor and TAs

• All announcements, Q&A in Piazza
  – http://piazza.com/stanford/fall2015/cs131

• ALL EMAIL CORRESPONDENCES TO ANYONE OF US:
  – cs131-fall1516-staff@lists.stanford.edu
Overall philosophy

• Breadth
  – Computer vision is a huge field
  – It can impact every aspect of life and society
  – It will drive the next information and AI revolution
  – Pixels are everywhere in our lives and cyber space
  – CS131 is meant as an introductory course, we will not cover all topics of CV
  – Lectures are mixture of details techniques and high level ideas
  – Speak our “language”

• Depth
  – Computer vision is a highly technical field, i.e. know your math!
  – Master bread-and-butter techniques: face recognition, corners, lines, features, optical flows, clustering and segmentation
  – Programming assignments: be a good coder AND a good writer
  – Theoretical problem sets: know your math!
  – Final Exam: your chance to shine!
Syllabus

• Go to website...
  
  http://vision.stanford.edu/teaching/cs131_fall1516
Grading policy

- Program Set 0: 8%
  - Normalizing background knowledge
- 3 theoretical Problem Sets: 12% x 3 = 36%
- 3 programming assignments: 12% x 3 = 36%
- Final Exam: 20%

Late policy
- 5 free late days – use them in your ways
- Afterwards, 25% off per day late
- Not accepted after 3 late days per PS

Collaboration policy
- Read the student code book, understand what is ‘collaboration’ and what is ‘academic infraction’
Homeworks

- All homework must be submitted on Gradescope ([http://gradescope.com](http://gradescope.com)) as PDF.

- All code must be submitted via email to cs131.submissions@gmail.com

- PS0 is live, you can start working on it immediately.