CS 143: Introduction to Computer Vision

Instructor: James Hays
TAs: Hari Narayanan (HTA), Libin “Geoffrey” Sun, Greg Yauney, Bryce Aebi, Charles Yeh, Kurt Spindler
Today’s Class

• Introductions
• What is Computer Vision?
• Computer Vision at Brown
• Specifics of this course
• Questions
Thesis: Large Scale Scene Matching for Graphics and Vision

Thesis
hays_thesis.pdf, 107MB

Committee
- Alexei A. Efros (chair)
- Martial Hebert
- Jessica K. Hodgins
- Takeo Kanade
- Richard Szeliski, Microsoft Research
Scene Completion

[Hays and Efros. Scene Completion Using Millions of Photographs. SIGGRAPH 2007 and CACM October 2008.]
Nearest neighbor scenes from database of 2.3 million photos
Graph cut + Poisson blending
My Research

IM2GPS: estimating geographic information from a single image

An Empirical Study of Context in Object Detection
Categories of the SUN database
CS 143 TAs

Hari Narayananan (HTA)
Libin “Geoffrey” Sun
Greg Yauney
Bryce Aebi
Charles Yeh
Kurt Spindler
What is Computer Vision?

• What are examples of computer vision being used in the world?
Computer Vision

Make computers understand images and video.

What kind of scene?
Where are the cars?
How far is the building?
...

Vision is really hard

• Vision is an amazing feat of natural intelligence
  – Visual cortex occupies about 50% of Macaque brain
  – More human brain devoted to vision than anything else
Why computer vision matters

Safety

Health

Security

Comfort

Fun

Access
Ridiculously brief history of computer vision

• 1966: Minsky assigns computer vision as an undergrad summer project
• 1960’s: interpretation of synthetic worlds
• 1970’s: some progress on interpreting selected images
• 1980’s: ANNs come and go; shift toward geometry and increased mathematical rigor
• 1990’s: face recognition; statistical analysis in vogue
• 2000’s: broader recognition; large annotated datasets available; video processing starts
• 2030’s: robot uprising?
How vision is used now

- Examples of state-of-the-art
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
http://www.research.att.com/~yann/

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

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, ...
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
3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009
Object recognition (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… “
Vision-based biometrics

“How the Afghan Girl was Identified by Her Iris Patterns”  Read the [story](https://wikipedia)
Login without a password...

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/
Object recognition (in mobile phones)

Point & Find, Nokia

Google Goggles
Special effects: shape capture
Special effects: motion capture

*Pirates of the Caribbean*, Industrial Light and Magic
Sports

*Sportvision* first down line

Nice [explanation](http://www.howstuffworks.com) on [www.howstuffworks.com](http://www.howstuffworks.com)

Smart cars

- **Mobileye**
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers
Google cars

June 24, 2011. "Nevada state law paves the way for driverless cars". Financial Post. Christine Dobby
Aug 9, 2011, "Human error blamed after Google's driverless car sparks five-vehicle crash". The Star (Toronto)
Interactive Games: Kinect

- Object Recognition: [http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o](http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o)
- Mario: [http://www.youtube.com/watch?v=8CTJL5lUjHg](http://www.youtube.com/watch?v=8CTJL5lUjHg)
- 3D: [http://www.youtube.com/watch?v=7QrnwoO1-8A](http://www.youtube.com/watch?v=7QrnwoO1-8A)
- Robot: [http://www.youtube.com/watch?v=w8BmgtMKFbY](http://www.youtube.com/watch?v=w8BmgtMKFbY)
Vision in space

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.
Industrial robots

Vision-guided robots position nut runners on wheels
Mobile robots

NASA’s Mars Spirit Rover

http://www.robocup.org/

Saxena et al. 2008
STAIR at Stanford
Medical imaging

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT
Computer Vision and Nearby Fields

• Computer Graphics: Models to Images
• Comp. Photography: Images to Images
• Computer Vision: Images to Models
Computer Vision at Brown

Pedro Felzenszwalb  James Hays  Erik Sudderth

Thomas Serre  Stu Geman  David Mumford  Gabriel Taubin

David Cooper  Ben Kimia  Joe Mundy

See also: Brown Center for Vision Research (CVR)
Course Syllabus (tentative)

• http://www.cs.brown.edu/courses/csci1430/
Grading

• 80% programming projects (5 total)
• 20% quizzes (2 total)
Scope of CS 143

Computer Vision

Robotics

Human Computer Interaction

Medical Imaging

Neuroscience

Image Processing

Feature Matching

Recognition

Machine Learning

Graphics

Computational Photography

Optics
Course Topics

• Interpreting Intensities
  – What determines the brightness and color of a pixel?
  – How can we use image filters to extract meaningful information from the image?

• Correspondence and Alignment
  – How can we find corresponding points in objects or scenes?
  – How can we estimate the transformation between them?

• Grouping and Segmentation
  – How can we group pixels into meaningful regions?

• Categorization and Object Recognition
  – How can we represent images and categorize them?
  – How can we recognize categories of objects?

• Advanced Topics
  – Action recognition, 3D scenes and context, human-in-the-loop vision...
Textbook

Computer Vision: Algorithms and Applications

© 2010 Richard Szeliski, Microsoft Research

http://szeliski.org/Book/
Prerequisites

• **Linear algebra**, basic calculus, and probability
• Experience with image processing or Matlab will help but is not necessary
Projects

• Image Filtering and Hybrid Images
• Local Feature Matching
• Scene Recognition with Bag of Words
• Object Detection with a Sliding Window
• Boundary Detection with Sketch Tokens
Proj1: Image Filtering and Hybrid Images

• Implement image filtering to separate high and low frequencies
• Combine high frequencies and low frequencies from different images to create an image with scale-dependent interpretation
Proj2: Local Feature Matching

• Implement interest point detector, SIFT-like local feature descriptor, and simple matching algorithm.
• Feed feature matches to a structure-from-motion system
Proj3: Scene Recognition with Bag of Words

- Quantize local features into a “vocabulary”, describe images as histograms of “visual words”, train classifiers to recognize scenes based on these histograms.
Proj4: Object Detection with a Sliding Window

- Train a face detector based on positive examples and "mined" hard negatives, detect faces at multiple scales and suppress duplicate detections.
Proj5: Boundary Detection with Sketch Tokens

• Quantize human-annotated boundaries into “sketch tokens”, train a multi-way classifier to recognize such tokens.