Today

- Introduction to CSE 527
- Administrations
- What is Computer Vision?
- Few concepts in
  - Linear Algebra
  - Calculus
  - Probability
- OpenCV
CSE 527: Intro to Computer Vision

Grad version of CSE 327

Switch to *hacker mode*:

- This is an intensive hands-on class with weekly theoretical and practical programming assignments, which will give students the confidence to tackle computer vision problems in the wild.

Goals of the class:

- Gain knowledge of theory and practice in Computer Vision,
- Implement solutions for real-world problems in image and video analysis
- Hands-on experience of classical vision problems
- Evaluate implementations on standard datasets, and compare to state-of-the-art
Administrations | Time, Location

Time: TuTh 10:00-11:20 AM (we begin 10:05)
Location: OCS 2120

Bring: Laptop, pen & notebook, questions (your brain)

Don’t bring: (unmuted) phone, snacks, pets, attitude
Grading, Projects, Passing

Grade Components:

- HW 40%
- Exams
  - Midterm 15%
  - Final / Project 30%
- Attendance 15%

Projects vs. Finals

- Project is optional. Also for 522
- Final will likely take less time

Cut-offs, thresholds for a passing grade:

- Attendance skip > 20% = F
- HWs missing > 3 = F
- Exams missing != F

HW late fees: 15%/day

Collaboration:

- OK to collaborate on assignments, submission is individual
- OK to collaborate on projects, submission is communal (groups of <= 3 students)
Prerequisites

Mathematics / Computation:

- Linear algebra
- Calculus
- Probability theory
- Algorithms
- Machine learning

Don’t need to be an expert, but have the basics down.

Programming.

Options:

- C++
- Python

Non-options:

- C++: OpenCV, Boost, CMake
- Python: OpenCV, NumPy, Scikit
- TensorFlow

OS (Windows, Linux, Mac) is up to you.
Textbooks That I Use In This Class

Szeliski
2010

Prince
2012

Marr
2010
(1979)

Géron
2017
Textbooks That Are Good To Know About

- Forsyth & Ponce
  2011

- Nixon
  2012

- Hartley & Zisserman
  2004
Course Timeline

We learn:
Human/Computer Vision
Cameras, Optics
Image Processing
Features, Object Detection
Tracking, Segmentation

We learn:
Stereo / Multi-view
Structured light
3D Reconstruction
SLAM

We learn:
Deep Learning
GPU

Final Exam - 12/7
Project Presentations

Today  Oct. 10th  Oct. 31st  Nov. 30th  Dec. 7th
2D  3D  Beyond  Final

11 6 8  Classes
5 2 3  HWs
Who’s here to help

My office hours:
Thursdays 1-4pm, Room 145
Schedule via [Google Calendar](#)

TAs: TBD
Their office hours: TBD
(Will know within a week or so. Will post in Piazza)

My lab students:
Room 140

Do not disturb, unless you bring food.
Course and Contact Information

Course website: [http://hi.cs.stonybrook.edu/cse-527](http://hi.cs.stonybrook.edu/cse-527)

Piazza: [piazza.com/stonybrook/fall2017/cs527/home](piazza.com/stonybrook/fall2017/cs527/home)

Email: roys@cs.stonybrook.edu

Phone: (631) 632-2753

Office: Room 145

Lab: Room 140
What is Computer Vision?
NO ONE CAN BE TOLD WHAT COMPUTER VISION IS
YOU MUST SEE IT FOR YOURSELF
“... vision is an inverse problem, in which we seek to recover some unknown given insufficient information to fully specify the solution. [...] In computer vision we are trying to [...] describe the world in one or more images and to reconstruct its properties, such as shape, illumination, and color distributions.”

-- Rick Szeliski

“We observe an image and from this we extract measurements. [...] The vision problem or goal is to use the measurements to infer the world state. [...] To accomplish the goal we build a model [...] a relationship between the measurements and the real world.”

-- Simon J.D Prince
Origins | 1960s, 1970s

Early cognitive & neural science 1920s-1950s.

1960s: The MIT AI crowd: Minsky, Winston, Marr, Poggio, Ullman
1980s

Pyramids.

More stereo, 3D reconstruction.

Active contours.

MRFs.
1990s

- SfM, 3D reconstruction, multi-view.
- Optical flow, dense stereo.
- Tracking: level sets, particle filters
- Segmentation: graph cuts, mean shift
- Eigenfaces
2000s

Image-based modeling / rendering

Texture synthesis

Computational photography, HDR

Feature-based recognition & detection

Non-rigid tracking, Mean-shift
2010s
(Sampler)

And then there was *Deep learning*!

CNN, RNN, LSTM, ResNets, ...

A lot of [human] pose estimation

Crowdsourcing vision

Non-Rigid SfM, reconstruction

Deconvolution, Dehazing
Cross-Discipline

- Neuroscience
- Behavioral Studies
- Psychophysics
- Software Engineering
- Algorithms
- Physics
- Optics
- Biology
- Electrical Engineering
- Robotics
- Mechanical Engineering
- Signal Processing
- Communication Engineering
- Computer Graphics
- Computational Geometry
- Machine Learning
- Probability, Statistics
<table>
<thead>
<tr>
<th>Related Field:</th>
<th>Description:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Computer Vision:</strong></td>
<td>From Images to Models</td>
</tr>
<tr>
<td>Computer Graphics:</td>
<td>From Models to Images</td>
</tr>
<tr>
<td>Machine Learning:</td>
<td>Data-driven Models</td>
</tr>
<tr>
<td>Artificial Intelligence:</td>
<td>Neuro-inspired Models</td>
</tr>
<tr>
<td>Computational Geometry:</td>
<td>Models for geometry</td>
</tr>
<tr>
<td>Biomedical Engineering:</td>
<td>Models for human, non-rigid geometry</td>
</tr>
<tr>
<td>Visualization:</td>
<td>From Models to Chart Images</td>
</tr>
</tbody>
</table>
Business

- **Mobileye**: Sold to Intel for >$15Bn
- **Magic Leap**: ~$4.5Bn valuation
- **NVIDIA**: $7Bn Yearly revenue
- **tobii**: $80M/y revenue
- **Aptina**: Sold for $400M
- **C emotions**: Raised $50M
Applications

Automotive, Industrial, Film & Media, Surveillance, Biomedical, Biometric, Retail, Entertainment, Photography, ...............
How Generative Models are changing the computer vision world…

$$P(x|w)$$

$$P(w|x)$$
Deep Learning | On the Rise

Deep Learning in CVPR, AAAI and NIPS proceedings
Academia | Lots of Major Venues
## Top Computer Science Conferences

**Ranking is based on Conference H-index >= 12 provided by Google Scholar Metrics**

<table>
<thead>
<tr>
<th>#</th>
<th>H-index</th>
<th>Publisher</th>
<th>Conference Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>158</td>
<td>IEEE</td>
<td><strong>CVPR : IEEE Conference on Computer Vision and Pattern Recognition, CVPR</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Jun 18, 2018 - Jun 16, 2018 - Salt Lake City, United States</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://cvpr2018.thecvf.com/submission/timeline">http://cvpr2018.thecvf.com/submission/timeline</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Deadline : Wed 08 Nov 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dec 4, 2017 - Dec 4, 2017 - Long Beach Convention Center, United States</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="https://nips.cc/Conferences/2017/CallForPapers">https://nips.cc/Conferences/2017/CallForPapers</a></td>
</tr>
<tr>
<td>3</td>
<td>98</td>
<td>Springer</td>
<td><strong>ECCV : European Conference on Computer Vision</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oct 8, 2016 - Oct 8, 2016 - Amsterdam, Netherlands</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://www.eccv2016.org/">http://www.eccv2016.org/</a></td>
</tr>
<tr>
<td>4</td>
<td>91</td>
<td>IMIS</td>
<td><strong>ICML : International Conference on Machine Learning (ICML)</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aug 6, 2017 - Aug 11, 2017 - Sydney, Australia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://icml.cc/2017">http://icml.cc/2017</a></td>
</tr>
<tr>
<td>5</td>
<td>89</td>
<td>IEEE</td>
<td><strong>ICCV : IEEE International Conference on Computer Vision</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oct 22, 2017 - Oct 29, 2017 - Venice, Italy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://iccv2017.thecvf.com/">http://iccv2017.thecvf.com/</a></td>
</tr>
<tr>
<td>6</td>
<td>85</td>
<td>Association for Computing Machinery</td>
<td><strong>CHI : Computer Human Interaction (CHI)</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Apr 21, 2018 - Apr 21, 2018 - Montréal, Canada</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="https://chi2018.acm.org/">https://chi2018.acm.org/</a></td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------------------</td>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>Molecular Systems Biology</td>
<td>journal</td>
<td>8.366</td>
</tr>
<tr>
<td>2</td>
<td>International Journal of Computer Vision</td>
<td>journal</td>
<td>8.269</td>
</tr>
<tr>
<td>3</td>
<td>Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition</td>
<td>conference and proceeding</td>
<td>7.342</td>
</tr>
<tr>
<td>4</td>
<td>Proceedings of the IEEE International Conference on Computer Vision</td>
<td>conference and proceeding</td>
<td>7.336</td>
</tr>
<tr>
<td>5</td>
<td>MIS Quarterly: Management Information Systems</td>
<td>journal</td>
<td>6.687</td>
</tr>
<tr>
<td>6</td>
<td>IEEE Transactions on Pattern Analysis and Machine Intelligence</td>
<td>journal</td>
<td>6.298</td>
</tr>
<tr>
<td>7</td>
<td>BMVC 2014 - Proceedings of the British Machine Vision Conference 2014</td>
<td>conference and proceeding</td>
<td>5.177</td>
</tr>
<tr>
<td>8</td>
<td>Journal of Statistical Software</td>
<td>journal</td>
<td>5.139</td>
</tr>
</tbody>
</table>
Fastest Ever Math Review
I Assume

Things You Know

Linear Algebra
- Vectors, Matrices
- Basic operations and operators: dot, cross
- Norms and normalization

Calculus
- Integration & Derivation
- Multivariate operators

Probability
- Random Variables
- Distributions: Uniform, Normal, ...
- Bayes Rule
- Combinatorics

Algorithmics
- Data structures
- Some graph theory
- Elementary algorithms, like sorting
Examples
OpenCV
Live Demo
Write an OpenCV program to do the following things:

- Read an image from a file and display it to the screen
- Add to, subtract from, multiply or divide each pixel with a scalar, display the result.
- Resize the image uniformly by \( \frac{1}{2} \)

Due: Thursday 9/7 9am, before class

Submission mechanics: TBD (see Piazza)