CS 4476A/6476B: Intro to Computer Vision

Instructor: Frank Dellaert
Tas: Heyley Gatewood(head TA), and Kritika, Nikith, Junyan, Sarath, Shashank, Tongshu, Vince.
And almost 265 of you! 165u+100g
Today’s Class

Who are we?

Specifics of this course

What is Computer Vision?
Teaching Assistants

Heyley Gatewood (head TA)
A bit about me

Originally from Belgium
1989 EE in Leuven
1993 M.Sc. ECE at CWRU
2001 Ph.D. CS, Carnegie Mellon

Georgia Tech since August 2001
Teaching Computer Vision etc.:

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Name</th>
<th>Years</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>7641</td>
<td>Machine Learning</td>
<td>02, 03, 04</td>
<td>45</td>
</tr>
<tr>
<td>8803</td>
<td>3D Reconstruction and Mapping</td>
<td>02, 09, 10, 12</td>
<td>70</td>
</tr>
<tr>
<td>4495</td>
<td>Computer Vision, undergraduate</td>
<td>04, 05, 06, 07, 19, 21</td>
<td>441</td>
</tr>
<tr>
<td>8803</td>
<td>Intro to Perception and Robotics</td>
<td>05, 06, 07, 13, 14, 20</td>
<td>485</td>
</tr>
<tr>
<td>6476</td>
<td>Computer Vision, graduate</td>
<td>06, 07, 11, 12, 13, 21</td>
<td>453</td>
</tr>
<tr>
<td>4475</td>
<td>Computational Photography</td>
<td>07, 09, 18</td>
<td>155</td>
</tr>
<tr>
<td>4480</td>
<td>Digital Video Special effects</td>
<td>08, 09, 10, 11</td>
<td>146</td>
</tr>
</tbody>
</table>

https://dellaert.github.io/
Monte Carlo Localization, at Carnegie Mellon!

Dellaert, Fox, Burgard & Thrun, ICRA 1999
Fox, Dellaert, Burgard & Thrun, AAAI 1999
Spatiotemporal Reconstruction

Historical Image Collection

-supported by NSF CAREER, Microsoft

Recent revival: NSF NRI award on 4D crops for precision agriculture…

Grant Schindler
Factor Graphs -> GTSAM!
Silicon Valley intermission at Skydio
The Scene Understanding and Modeling Challenge

RGB-D 360 degree image

Object-based representation of a room

Organizers
Daniel Huber (Facebook)
Lyne Tchapmi (Stanford University)
Frank Dellaert (FB / Georgia Tech)

Vision Problems Addressed
Object segmentation
Object completion
Instance labeling
6-DOF pose estimation
Appearance modeling
Layout estimation

Silicon Valley intermission at Facebook
Our research

Researchers across Google are innovating across many domains. We challenge conventions and reimagine technology so that everyone can benefit.

Publications

Google publishes hundreds of research papers each year. Publishing our work enables us to collaborate and share ideas with, as well as learn from, the broader scientific community.

Research Areas

From conducting fundamental research to influencing product development, our research teams have the opportunity to impact technology used by billions of people every day.

Tools & datasets

We make tools and datasets available to the broader research community with the goal of building a more collaborative ecosystem.
What’s next? Robot Art! Dynamics!

https://dellaert.github.io/NeRF/
Course Website/Syllabus

Fall ’19 Computer Vision

https://dellaert.github.io/21F-x476/
Project 0: Linear Algebra with Pytorch

Out: 6.30 today
Due: Friday September 3, midnight
Late policy: 10% per day

Learning Objectives:
• Set up a conda environment
• Understand how to start a jupyter notebook
• Linear algebra in python using pytorch
• Test your code using unit tests
• Use gradescope to submit your work
Project 1: 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
Project 2: CNN Image Segmentation

Learning Objectives:
1. Understanding convolutional neural networks for inference
2. Construct a basic CNN for image segmentation
3. Understand the use of some basic layers used in CNNs
4. Set up the inference workflow in Pytorch.
Learning Objectives:
1. Understanding the rationale behind data pre-processing
2. Construct a basic CNN for multi-class classification
3. Understand some more basic layers used in CNNs
4. Set up the **training** workflow in Pytorch.
Project 4: Local Feature Matching

Implement interest point detector, SIFT-like local feature descriptor, and simple matching algorithm.
Project 5: Projection Matrix, F+Ransac

Understand geometry of pose estimation, and use a random sampling algorithm to do a detailed matching based on the fundamental matrix constraint.
Project 6: Object and Pose Detection

Application of open source libraries (Google Mediapipe) to a practical problem, using your own imagery.

https://google.github.io/mediapipe/solutions/objectron.html