Cornell Bowers C·IS
College of Computing
and Information Science

Undergraduate Introduction To Computer Vision

Spring 2025

CS, 4670/5670

Course website: https://canvas.cornell.edu/courses/74091

Enrollment questions: courses@cis.cornell.edu

Faculty Name(s): Bharath Hariharan and Wei-Chiu Ma
Faculty Email: bh497@cornell.edu and wm347@cornell.edu
Faculty Office Hours: TBA (check the course website for most up-to-date information)
Course Staff and Course Staff Office Hours:
This course will have ~20 teaching assistants. Times and venues for office hours will be posted on the course website the first week of classes.

Prerequisites/Corequisites: *knowledge of linear algebra (recommended), programming and probability/statistics*
Time and Location: Mondays/Wednesdays/Fridays 1:25 - 2:15 pm in Baker Laboratory 200.

Course Description

This course will introduce the core problems of computer vision and discuss classical approaches based on the geometry and physics of image formation, as well as introduce modern techniques using deep learning. Topics include stereopsis and 3D reconstruction, image segmentation, object recognition, feature representations of images and patches, and convolutional networks.

Course Objectives/Student Learning Outcomes

After taking this course, students will be able to:

- *[Basics]*
    1. *Describe intuitively and mathematically the geometry and physics of image formation.*
    2. *Explain intuitively what information gets lost in image formation*
- *[Image processing]*
    1. *Implement convolution and understand what kind of filtering operations can be implemented as a convolution.*

2. *Fourier transforms Relate physical properties of the image to the ``frequency'' in a Fourier transform and explain its impact on various image processing operations and its interaction with convolution*
3. *Use the language of Fourier transforms to explain what happens during subsampling or upsampling, and when subsampling can lead to aliasing.*

- *[Grouping]*
  1. *Implement basic edge detection using convolution and explain why image noise can be a problem.*
  2. *Explain when edges can be hard to detect*
  3. *Explain how edges can be sharpened using non-max suppression and implement the same.*
- *[Reconstruction]*
  1. *Enumerate what it means to recover 3D structure of the scene: the position and orientation of the camera and the 3D locations of all points*
  2. *Explain why recovering 3D structure from an image is an ill-posed problem.*
  3. *Explain the additional information that is needed, and thus describe the problem of camera calibration, pose estimation, stereo and structure-from-motion.*
  4. *Identify the reasons why estimating correspondences between images is hard.*
  5. *Contrast different ways of measuring patch/pixel similarity in terms of their invariance and discriminability.*
  6. *Explain why feature detection is useful, and what kind of features are useful to detect.*
  7. *Mathematically derive and implement in code the Harris corner detector and derive the connection between the second moment matrix and the nature of the feature and its orientation.*
  8. *Implement the MOPS feature descriptor and explain how SIFT improves upon MOPS.*
  9. *Derive how camera parameters and 3D structures can be mathematically obtained from a set of correspondences (camera calibration, stereo, structure from motion).*
  10. *Derive the special case of two-view stereo when the two views are related by a simple translation along X; explain the relationship between disparity and depth.*
  11. *Explain and implement the plane-sweep stereo algorithm.*
  12. *Understand the need for removing outliers and derive the RANSAC algorithm.*

13. *Explain what radiance means, how it relates to pixel values, and explain mathematically its relationship to surface normal, surface albedo and lighting.*
14. *Derive and implement photometric stereo to get normal and depth from a set of images from the same camera but different lights.*

- *[Recognition]*
  1. *Write down the ERM principle and how it relates to generalization, overfitting and underfitting.*
  2. *Derive the gradient descent update and SGD update for a general loss function and machine learning model.*
  3. *Explain the rationale behind the bag-of-words feature representation.*
  4. *Explain the rationale behind convolutions and subsamplings as critical layers in a neural network.*
  5. *Design and implement a neural network for classification.*
  6. *Explain the semantic segmentation task.*
  7. *Design model architectures and loss functions for semantic segmentation.*
  8. *Explain the object detection task.*
  9. *Describe the R-CNN, Fast R-CNN and Faster-RCNN object detection approaches.*

Course Materials

Course materials in the form of lecture notes and zoom recordings will be available through the Canvas webpage

Method of Assessing Student Achievement

- **Deliverables:**
  a. Group projects: This course will have **4 projects** on the topics of image processing, estimating correspondences, 3D reconstruction and vision language respectively.
     i. The assignments must be done in groups of 2. Students wanting to do the projects alone must let the instructor know.
     ii. Students taking 4670 must do at least 3 out of the 4. If they do all 4 then their grade will be based on best 3 of 4.
     iii. Students taking 5670 must do all 4 projects.

b. Individual homeworks: This course will have 2 homeworks to be done individually. The homework will involve both written and programming components.
c. Exams: This course will have one prelim and one final exam. **Both exams will be in person.**

- **Grading policies**

1. *Late work:* You will have 10 slip days for the entire course. Once these are exhausted, you will lose 5% of the assignment grade for every day of delay.
2. *Missed work:* If you miss assignments, homeworks or exams due to emergencies or unforeseen health problems (e.g., COVID), contact the instructor of the course with the reason. If the instructor finds the reason justified, you will be given the option of simply rescaling whatever work you have done to calculate your grade. However, this option will only be available if you have submitted at least one assignment and one exam. If you are unable to do this minimum amount of work, then you will be encouraged to take an INC and finish the work later, or switch to S/U.
3. *Grade distribution:*

Assignment, Assessment or activity	**Percentage of grade or points**
Programming Assignments	35%
Written Homeworks	25%
Prelim	20%
Final	20%

- **Grading scale**

Note that the grading table below only goes to C-. To pass the course, you need to do one programming assignment and one exam in good faith.

95-100%	A+	50-65 %	C+	
90-95%	A	30-50 %	C	

85-90%	A-	<30%	C-	
80-85%	B+			
75-80%	B			
65-75%	B-			

Course Management

ACADEMIC INTEGRITY:

Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student in this course for academic credit will be the student's own work.

In particular, since modern language model-based tools (like ChatGPT) can copy without citation any text in their training data, the use of such tools will be considered as plagiarism and is therefore strictly prohibited.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES:

Students with Disabilities: Your access in this course is important. Please give me your Student Disability Services (SDS) accommodation letter **and email me a note** early in the semester so that we have adequate time to arrange your approved academic accommodations. If you need an immediate accommodation for equal access, please speak with me after class or send an email message to me and/or SDS at sds_cu@cornell.edu. If the need arises for additional accommodations during the semester, please contact SDS. Student Disability Services is located at Cornell Health Level 5, 110 Ho Plaza, 607-254-4545, sds.cornell.edu.

INCLUSIVITY:

Computer vision is a technology fraught with many ethical issues in its current practice. As new entrants into this field, you have the power to change this for the better. We can start by keeping our course an inclusive environment that supports everyone's learning, maintains a civil discourse, and respects what every one of us brings to the table.

MENTAL HEALTH AND STRESS MANAGEMENT RESOURCES

If you are feeling overwhelmed, or worried about a friend, please reach out to one of your instructors or your academic advisor.
Please look at this guide that collects all the resources that you can avail of.

Note that Cornell has trained peer mentors available to listen and help: Empathy, Assistance, and Referral Service , Also trained counselors: Cornell Health's Counseling and Psychological Services (CAPS, 607-255-5155), and Let's Talk.