CS2831: Advanced Computer Vision

Fall 2024 Edition

Synchronous Meeting Time: Tuesday and Thursday 12:45-2:00pm
Synchronous Meeting Location: SEC LL2.223

Instructor: Prof. Todd Zickler (zickler@seas.harvard.edu)

Teaching Fellow: Dean Hazineh (dhazineh@g.harvard.edu)

Office Hours: TBD

Course Description. Computer vision is about making systems that "see" by turning measurements of light into useful information. This course provides a comprehensive foundation for understanding and creating such systems. Topics include: camera geometry; radiometry and light transport; elements of biological vision; and classical and neural-network methods for extracting information about 3D shape, materials, dynamics and semantics. The course balances breadth and depth, and it blends theory and practice.

Required Preparation. CS 2831 assumes students have had significant programming experience with Python and NumPy, ideally with some exposure to PyTorch or JAX. Experience with Matlab or Julia can be a viable substitute for students who are self-motivated to pick-up Python on the fly. CS 2831 also assumes students have taken an undergraduate course in computer vision (e.g., ES 1431), or that they have taken a course in each of: probability (e.g. ES150 or Stat 110), linear algebra (e.g., AM22a or Math 21b) and signal processing (e.g., ES 156 or AM 104).

Evaluation. A small portion of the evaluation is based on participation, both in asynchronous discussions and during synchronous meeting times. (Note that attendance during synchronous meeting times is expected.) The remaining portion is based on weekly assignments, an in-class exam, and a final project with written report. These are weighted as follows:

10%	Participation
30%	Assignments
30%	In-class exam (mid-November)
30%	Final project: written report and code

Assignments. The assignments are combinations of (i) mathematical analysis and (ii) coding with Python in Google Colab. Assignments are completed independently and submitted electronically. Examples of coding tasks include: single-view 3D reconstruction, automatic panoramic stitching, structure from motion, shadow removal, texture synthesis, depth from stereo, and designing and training neural networks for classification, segmentation and object detection.

Textbooks. There are two required textbooks for this class, and both of them are available electronically:

- R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd Ed., 2022. (PDF publicly available here.)
- R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd Ed., Cambridge University Press, 2003. (Access PDF through Harvard Library. See the *Library Reserves* tab on this page.)

We'll use the Szeliski book most. It provides a broad and modern overview of the field, as well as a good learning tool for certain topics. The Hartley & Zisserman book is an essential reference and learning tool for the geometry of Computer Vision.

In addition, there are four textbooks that can serve as useful supplements for learning and reference purposes. Most of them are available electronically for short-term check-out in the *Library Reserves* tab:

- A. Torralba, P. Isola and W. T. Freeman, *Foundations of Computer Vision*, MIT Press, 2024.
- D. Forsyth, and J. Ponce, *Computer Vision: A Modern Approach*, 2nd Ed., Prentice, 2012.
- S.J. Gortler, *Foundations of 3D Computer Graphics*. MIT Press, 2012.
- B.K.P. Horn, *Robot Vision*. MIT Press, 1986.

The books by Torralba et al. and by Forsyth & Ponce provide broad overviews like Szeliski. In terms of learning tools, Torralba et al. is particularly good for image processing and for various neural network models, while Forsyth & Ponce is particularly good for color and radiometry. The book by Gortler contains chapters on color and light (radiometry) that are great learning tools for those topics. The classic book by Horn is a good learning tool for topics on radiometry, shape from shading, and optical flow.

Ed Discussion. CS 2831 uses *Ed Discussion* for communication. You can access it from the left menu on this page. Ed Discussion is the way to ask any sort of question about the course. The teaching staff will also use its chat function to direct-message students.

You will get faster answers from staff and peers on Ed Discussion than you will through email. Here are some tips when using it:

- Search before you post
- "Heart" questions and answers you find useful
- Answer questions you feel confident answering
- Share interesting course-related content with staff and peers
- Set discussion threads to "private" when needed (see Getting Help below)

For more information, you can refer to the [Ed Discussion Quick Start Guide].

Getting Help. There are generally two ways to get help: by attending office hours and by posting to Ed Discussion. *Office hours* are weekly time slots (listed at the top of this page) when you can talk to a member of the teaching staff. You do not need to have a "good question" or reason to attend office hoursâ€”you can just pop in to say hello!

In most cases, the best way to get an answer outside of office hours is to use *Ed Discussion* as described above. This gets the quickest response because answers can come from other students in addition to the teaching staff. Posting your question on Ed Discussion also helps other students who have the same question as you.

You can also use Ed Discussion to create private threads with the teaching staff. We will respond to private threads as promptly as possible during 09:00 AM-10:00 PM ET. Within these times, you can generally expect a response within four hours of your post. In cases where we receive a private question about class logistics or material that does not concern a private matter and whose answer will benefit other students, we may choose to broadcast the question and answer to the rest of the class.

Note that *email* to the instructor or TFs should be used rarely and only for exceptionally private matters. An example is when you experience an extraordinary personal or family challenge that causes a large disruption which cannot be accommodated by the standard late policy (see below). Like private Ed Discussion threads, we will respond to email as promptly as possible during 09:00 AM-10:00 PM ET. Within these times, you can generally expect a response within 24 hours. If you are comfortable sending your email to the TFs in addition to the instructor, doing so will often reduce the response time.

Late Policy and Extensions. There may be times during the semester when unexpected challenges disrupt your ability to complete an assignment on time. The *late policy* covers these situations by providing some flexibility to each student. Each student has a budget of three "late days" that can be used during the semester, with each one providing a 24-hour extension beyond the stated deadline. Up to one "late day" can be used for any one assignment. "Late days" cannot be applied to the final project report.

There may be exceptional circumstances when unexpected challenges cause disruptions that are more significant than can be accommodated by the late policy. In these cases, you should contact the instructor by email to request an extension (see Getting Help above). Situations where an extension is likely to be granted include personal or family events that are unexpected and extraordinary. Situations when an extension is unlikely to be granted include predictable inconveniences, such as an athletic tournament or a co-occurrence of multiple academic deadlines. Extensions are also unlikely to be granted in situations that are self-inflicted by lack of care, such as losing a file that wasn't backed up, or by waiting until the very last moment to submit before a deadline and then suffering from a "computer crash".

Inclusion and belonging. CS 2831 believes that diversity of thoughts, backgrounds, perspectives, and experiences speed the invention of creative solutions and improve our community as a whole. The CS 2831 teaching staff recognizes our studentsâ€™ identities, including but not limited to race, gender, class, sexuality, socioeconomic status, religion, and ability. We strive to create a learning environment where every student belongs and feels welcomed and valued. We need your help to accomplish this goal. If you see an opportunity for the course to be made more inclusive, please share your ideas with the teaching staff. If something is said in a class or meeting, by anyone, that makes you feel uncomfortable, or if there is course material that feels insensitive, please talk to the instructor or TFs about it, even if anonymously.

Accommodations. CS 2831 values inclusive excellence and providing equal educational opportunities for all students. Our goal is to remove barriers for disabled students related to inaccessible elements of instruction or design in this course. If reasonable accommodations are necessary to provide access, please contact the Disability Access Office (DAO). Accommodations do not alter fundamental requirements of the course and are not retroactive. Students should request accommodations as early as possible, since they may take time to implement. Students should notify DAO at any time during the semester if adjustments to their communicated accommodation plan are needed.

Collaboration and Academic Integrity. Assignments are to be completed independently. Some of them are challenging and will cause you to get stuck, perhaps beyond what can be resolved through individual perseverance. In these situations, it is acceptable (even encouraged) to collaborate with other students in planning and designing solutions to the assignment.

However, no collaboration is allowed in writing up solutions or writing code. You are allowed to work with others in the form of discussing, brainstorming, and walking through strategies for solving assignment problems. But when you have finished interacting, you must write your solutions and code independently, and you may not compare your written solutions or code against each other or against notes that were taken during the collaboration. Specifically:

- When you are stuck on a written problem, you may show your work to a collaborator in order to get help in identifying the error, but you may not look at a collaborator's own written answers.
- When you are stuck on a coding problem, you may show your code to a collaborator in order to get help in identifying the bug, but you may not look at a collaborator's own code.
- When you are finished a problem, you may check your answer verbally with a collaborator, but there may be no direct comparison of code or written solutions between collaborators.

Before consulting others (peers, TFs, instructors) make sure you have made a genuine effort to solve the problems by yourself. Getting "unstuck" by yourself is a powerful way of learning, and individual perseverance is important for identifying your personal roadblocks so you can focus on them.

One final important point: Some of the assignment problems will be taken from textbooks, other courses, and previous offerings of this course. It is not acceptable to simply find pre-existing solutions to these problems, from any source whatsoever, and treat them as "collaborators".

Use of Generative Artificial Intelligence Tools (this will be updated periodically; last update 08/14/2024). Generative AI tools such as ChatGPT, Copilot and Gemini are evolving rapidly, and so is our access to them through Harvard (see HUIT's AI landing page). These tools can provide useful answers to questions about course material and topics, as you might expect from office hours. They can explain code and error messages in plain language, they can summarize textbook content, and they can generate long segments of code based on textual descriptions of desired functionality. They can also correct grammar and modify writing style, and they can generate long segments of text in response to a prompt or outline.

CS 2831 recognizes that, when used appropriately, they can be useful tools for learning and productivity. CS 2831 also recognizes that they can *impede* learning, by providing inaccurate information and by doing too much work for students and thereby robbing them of opportunities to assimilate new knowledge and practice new skills.

The AI policy in this course reflects its learning goals. The deliverables are designed to help students learn new concepts in computer vision by creatively applying them to solve problems and by creating code that exhibits prescribed visual processing behaviors. The deliverables are also meant to help students improve their ability to effectively communicate their technical ideas. (The learning goals do *not* include the development of low-level coding skills or of written grammar skills.)

Thus, for most assignments and for the final project in this course, students may use generative AI to:

- gain an understanding of key topics and concepts
- modify code they have created, for example to correct syntax and style
- obtain suggestions for code that implements their intentions

Students may not:

- submit AI-generated code that they cannot precisely explain, or whose validity and correctness they have not verified
- submit AI-generated text or mathematical equations that they cannot precisely explain, or whose validity and correctness they have not verified

Additionally, students may not use generative AI tools during the in-class exam.

Students must recognize that AI-generated material is prone to errors and bias, and students are responsible for verifying the accuracy of any submitted material that has been generated by AI.

We emphasize that the AI policy for this course reflects its specific goals. We fully understand that different courses will have different policies, and we expect that AI tools will continue to change during the semester. Please reach out to the teaching staff if you are unclear about what our policy is or why, and when it is not clear whether you are allowed to use generative AI for a particular purpose.

Schedule (subject to changes)

No.	**Date**	**Title**
1	Tu 9/3	Projective geometry and model-fitting I
2	Th 9/5	Projective geometry and model-fitting II
3	Tu 9/10	Projective geometry and model-fitting III
4	Th 9/12	Single-view measurement
5	Tu 9/17	Cameras I
6	Th 9/19	Cameras II
7	Tu 9/24	Epipolar geometry
8	Th 9/26	Structure from motion
9	Tu 10/1	Filtering and Fourier Transform
10	Th 10/3	Edges, blobs, corners
11	Tu 10/8	Interest points
12	Th 10/10	Multi-resolution representations
13	Tu 10/15	Radiometry and reflection
14	Th 10/17	Lighting
15	Tu 10/22	Color
16	Th 10/24	Texture
17	Tu 10/29	Segmentation
18	Th 10/31	Stereo
19	Tu 11/5	Optical flow [proposals for final projects are due]
20	Th 11/7	Neural networks and back-propagation
21	Tu 11/12	NNs for image classification
	Th 11/14	In-class exam
22	Tu 11/19	NNs for pixel-wise prediction
23	Th 11/21	NNs for detection
24	Tu 11/26	Advanced topics I
	Th 11/28	Holiday, class does not meet
25	Tu 12/3	Advanced topics II
	12/??	Final project reports due on date to be set by registrar