Computer Vision

CS 482 | Spring 2023 | Professor Gregory J. Stein

Course Information

3 Credits
Lectures: Merten Hall 1200
    Tue (T) 10:30AM–11:45AM
    Thu (θ) 10:30AM–11:45AM
Office Hours:
    Prof. Stein, Fri 10AM–11:30AM, ENGR 4439
    TA: TBD

Instructor Information

Prof. Gregory J. Stein
He, Him, His
gjstein@gmu.edu

Course Description

The aim of computer vision is to compute properties of the three-dimensional world from images so that it may be understood by machines. Topics in this class include how to create panoramic images, to build a 3D reconstruction of an environment from videos, and to recognize familiar people and objects, all through analysis of images and video clips.

This course is at the senior undergraduate level and aims to introduce students to the field of computer vision through hands-on programming projects. By the end of the course, students should be familiar with the standard tools used by the computer vision community, and be able to approach research papers in the field.

See the detailed class schedule below for more details on what will be covered during the course.

Course Structure

Especially because this course will be taught online, ensuring that the students are engaging with the material and with one another can be difficult. As such, many of the lectures—particularly in the first half of the course—will be broken into two parts: a lecture module roughly 50 minutes in length followed by a (short) breakout session during which students can collaborate to begin with an *un-graded* short example problem in small groups. Breakout sessions are meant to give the students some hands-on experience during lecture before reconvening in the main room to ask questions about the material. The breakout sessions will often be shorter and simpler versions of the programming assignments students will be graded on. Remaining class time will consist of a shorter lecture covering a related topic or research area.

Prerequisites

- **CS 310: Data Structures** As many (if not all) of the computer vision techniques we discuss will be algorithmic in nature, understanding how to implement simple algorithms and understand object-oriented code will be an important skill in this course. Big-O notation will also come up in the course, though students will not be expected to dive deep into computational complexity.
- **MATH 203: Linear Algebra** This is the most important prerequisite, as many of the tools developed and derived in this course involve matrix manipulation. Understanding eigenvalues and eigenvectors will also be key skills. Other matrix factorization techniques will also be used in this course, though prior knowledge (e.g., of singular value decomposition) is not required.
- **STAT 344: Probability and Statistics for Engineers and Scientists I** Probability will come up on occasion throughout the course, so a working knowledge of fundamentals (e.g., properties of Gaussian distributions, Bayes Rule) will be useful. Specific concepts associated with hypothesis testing (and the Chi-squared test) will be useful, but are also covered in the course.

In addition, students are expected to have some familiarity with the Python programming language (including numpy) in which all programming projects and assignments will be done.

Assignments & Grading

The grading in this course will come from three main components:

- 5 programming assignments (50%)
- 5 take home "quizzes" (20%; 5% each, since lowest grade dropped)
- 1 final project (30%)

All assignments will be turned in via Blackboard.

Programming Assignments

In each programming assignment, students will be expected to implement some of the algorithms we will discuss during class. Assignments will be given in the form of Jupyter Notebooks and will often include some partially written code for students to complete.

Students will be expected to write up a report for each programming assignment, preferably in LaTeX. As this is senior level course, some creativity on the solution will occasionally be required; some of the problems will have open-ended prompts

and students should expect to explore the parameter space of the algorithms they implement and report on their findings. Students may collaborate on the programming assignments in small groups (no more than 3 or 4 students) but solutions must be written up independently.

Quizzes

To supplement the programming assignments, which test mostly practical knowledge, the quizzes will be shorter and ask two or three conceptual questions related to the course material. We prefer that you type your solutions (e.g., via LaTeX) but neatly handwritten and scanned submissions will also be accepted. Students may collaborate on the quizzes in small groups (no more than 3 or 4 students) but solutions must be written up independently.

There will be a total of 5 take home quizzes assigned during the semester. At the end of the term the lowest quiz score (including a missed quiz) will be dropped.

Final Project

While the details for the final project are still TBD (and will be provided in the middle of the term), the final project will be an open-ended project, the aim of which will be to explore some topic or application that goes beyond what we have done in the programming assignments.

Participation & Lectures

Lectures will regularly include *breakout sessions* during which students will be expected to work collaboratively to tackle some small problem related to the lecture. However, I have chosen *not* to include participation or attendance as a part of the overall grade. You are still strongly encouraged to come to lecture, as the breakout sessions are an integral part of the course experience.

Lateness Policy

Every assignment (except the Final Project) can be turned in 3 days late without penalty. I ask that you make a best effort to turn in the assignments on time, and use the additional three days if you get stuck or if “life happens” (which it seems to do a lot these days). Any additional late days will result in a 10%/day penalty for all assignments, up to a maximum of 1 week after the original deadline, after which the assignment will not be accepted. If you feel you need an extension in addition to the 3 free late days, please try to ask me *before the original deadline*, as I would like to avoid emails asking for extensions with only hours before late penalties start

to accumulate.[1] **If you turn in the assignments on time, you will get a 2% bonus on that assignment's grade (up to the maximum).**

Course Resources

Optional Textbook & Readings

I have made my best effort to make the course as self-contained as possible. As such, there are no *required* textbooks for this course. If interested, you might consider purchasing *Concise Computer Vision: An Introduction into Theory and Algorithms* by Reinhard Klette, which I will occasionally refer to throughout the course for further reading. The book has nice illustrations of many of the concepts and has relatively easy-to-understand prose. I will also refer to *Computer Vision: Algorithms and Applications*, by Richard Szeliski (available as a PDF online for free), though that text is rather more advanced and can be difficult to follow if you do not already know the concepts. Additional supplemental readings, usually in the form of papers will be included either as links in the lecture slides or as PDFs uploaded to Blackboard.

Lecture Slides and Breakout Session Materials

I will be providing lecture slides as PDF documents after each lecture via Blackboard. In addition to the lecture slides, many of the lectures in the first two-thirds of the course will have breakout sessions for you to code or follow along with during lecture; these will also be uploaded to the Blackboard in advance. Lecture materials for the week will be uploaded on Mondays.

Discussion

This term we will be using Piazza for class discussion. The system is highly catered to getting you help fast and efficiently from classmates and myself. Rather than emailing questions to the teaching staff, I encourage you to post your questions on Piazza. Find our class signup link at: https://piazza.com/gmu/spring2023/cs482 . The access code is available for enrolled students on Blackboard.

[1] I have tried my best to come up with a late policy that is as fair and flexible as possible. Deadlines are largely there for your protection: in the past, I granted some assignments extensions early on in the class to students that were then behind for much of the rest of the term. Please try to start on the assignments early and ask for extensions only if you think you really need them, so that this does not happen to you.

Detailed Course Schedule

The following schedule is mostly complete, but some dates and lecture content may change.

The course is largely broken into three Units:

1. **Images** (Lectures 1–10) In which we will discuss the fundamental mathematical tools used to process images and identify common *features* between multiple images
2. **Structure** (Lectures 11–20) In which we will study how we might understand and reconstruct the 3D world.
3. **Modern Applications** (Lectures 21–28) Devoted to discussing state-of-the art research tools and applications, including Place Detection, Simultaneous Localization and Mapping, and Convolutional Neural Networks (and applications).

The following schedule is *tentative* and subject to change.

LEC	DATE	TOPICS	PROJECTS & QUIZZES
1	1/24 T	Introduction	P1 Out
2	1/26 θ	Fundamentals & Image Filtering	
3	1/31 T	Fourier Transforms	Q1 Out
4	2/02 θ	Resampling & Image Pyramids	
5	2/07 T	Feature & Corner Detection	P1 Due, P2 Out
6	2/09 θ	Feature Invariance	
7	2/14 T	Image Transformations	Q1 Due, Q2 Out
8	2/16 θ	Feature Descriptors	
9	2/21 T	Image Alignment	P2 Due, P3 Out
10	2/23 θ	RANSAC	
11	2/28 T	Camera Models	Q2 Due, Q3 Out
12	3/02 θ	Blender Tutorial & Camera Models cont.	
13	3/07 T	Panoramas	P3 Due, P4 Out
14	3/09 θ	Single-view Modeling	Final Project Description Out
—	3/14 T	No Class; Spring Break	
—	3/16 θ	No Class; Spring Break	
15	3/21 T	Two-view Geometry	Q3 Due, Q4 Out
16	3/23 θ	Stereo Vision	Final Project Proposals Due
17	3/28 T	Light	P4 Due, P5 Out
18	3/30 θ	Photometric Stereo	
19	4/04 T	Multi-view Stereo Applications	Q4 Due, Q5 Out
20	4/06 θ	Structure from Motion	
21	4/11 T	Neural Networks	P5 Due
22	4/13 θ	Convolutional Neural Networks	
23	4/18 T	Image Classification	
24	4/20 θ	Bag of Words Methods	
25	4/25 T	Simultaneous Localization & Mapping (SLAM)	Q5 Due

26	4/27 θ	Modern SLAM
27	5/02 T	Ethical Considerations
28	5/04 θ	Generative Methods (and GANs) Course Summary

Inclusion & Integrity

I stand by Mason's commitment to diversity and inclusion and hope to foster an inclusive environment in which all feel welcome in my class.

> *True diversity is defined not only as differences in individual backgrounds, personal identities, intellectual approaches, and demographics; it is also the removal of barriers and the creation of space that allow individuals to fully engage in the life of the university.*

Every student in this class is exactly where they belong and it is my honor to welcome each of you to join us in learning throughout this semester. Every student in this class, regardless of background, sex, gender, race, ethnicity, class, political affiliation, physical or mental ability, veteran status, nationality, or any other identity category, is an equal member of our class.

You have the right to be called by whatever name you wish, to be referred to by whatever pronoun you identify, and to adjust these at any point. If you feel uncomfortable in any aspect of our instruction that results in any barrier to your inclusion in this course, please contact me (your professor) directly.

Honor Code Statement

The GMU Honor Code is in effect at all times. In addition, the CS Department has further honor code policies regarding programming projects, which are detailed here. Any deviation from the GMU or the CS department Honor Code is considered an Honor Code violation.

Disability Accommodation

If you have a documented learning disability or other condition which may affect academic performance, make sure this documentation is on file with the Office of Disability Services and then discuss with the professor about accommodations. Submitting the paperwork at the deadline for a project or quiz is far too late! Even if you don't know if you plan on utilizing the accommodations ahead of time, it's in your best interest to prepare them ahead of time.

Mental Wellness

Graduate School can be a stressful environment and the realities of remote work can amplify these stresses. My “door” is always open; if you are struggling with the course work or would like someone to talk to, feel free to reach out to me. GMU also provides many mental health resources that I encourage you to look at.

Sexual Harassment and Interpersonal Violence

As a faculty member and designated "Responsible Employee," I am required to report all disclosures of sexual assault, interpersonal violence, and stalking to Mason's Title IX Coordinator per university policy 1412. If you wish to speak with someone confidentially, please contact the Student Support and Advocacy Center (703-380-1434), Counseling and Psychological Services (703-993-2380), Student Health Services, or Mason's Title IX Coordinator (703-993-8730; cde@gmu.edu).

Privacy and Email

Students must use their Masonlive email account to receive important University information, including communications related to this class. I cannot respond to messages sent from or send messages to a non-Mason email address.

To protect your privacy, I also cannot list your Masonlive email address on any public forum or provide it to any other students. You may, of course, give your email address to any other students.

Video recordings of class meetings that are shared only with the instructors and students officially enrolled in a class do not violate FERPA or any other privacy expectation.

All course materials posted to Blackboard or other course site are private; by federal law, any materials that identify specific students (via their name, voice, or image) must not be shared with anyone not enrolled in this class.