CMSC 426-0101: Computer Vision

Spring '24

Instructor:

Jia-Bin Huang, jbhuang@umd.edu. Section 0101.

Office: IRB 4234

Course Overview: This course introduces computer vision and computational photography. The course will cover basic principles of image processing, image recognition using classical methods and deep learning, and multi-view geometry for visual navigation. It will explore the topics of image formation, image features, image stitching, image segmentation, motion estimation, tracking, and object and scene recognition.

The course is organized around several projects. Through these projects you will learn the theory and practical skills required to obtain a computer vision engineering job.

Text: All concepts will be covered during the in-class lecture. However, we also recommend the following books as good references:

- Richard Szeliski, *Computer Vision: Algorithms and Applications, 2nd ed.*, 2022 online version.
- D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice-Hall, 2003 (available online).
- Rafael Gonzalez, and Richard Woods, *Digital Image Processing*, Prentice Hall, 2008.
- Richard Hartley, and Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2008.

Prerequisites: Each student is expected to know the basic concepts of Python programming (in particular numpy), linear algebra, and multivariable calculus.

Course Work and Exams: Coursework will consist of individual homework assignments and a final project. For the homework assignments, you may discuss problems and general solution strategies with classmates, but you *must* write up the solutions yourself. We will be using the Python programming language for this course.

Assignments will be turned in on ELMS, https://www.elms.umd.edu.

Do not post any of your project/homework code publicly online (e.g., GitHub public repository). Doing so is a violation of the academic integrity policy.

Late Policy: The first 7 days of late homework submissions are *free* throughout the semester. After that, assignments turned in late will incur a 10% penalty per day. An assignment submitted 1 second after the deadline will be one day late. We will be ruthless in enforcing this policy. Assignments turned in late because of excused absences (e.g., COVID, family emergency) will be accepted penalty-free for up to 7 days. Non-excusable examples include internship interview, other coursework, and traveling.

Final Project: The final project will be due on Tuesday, May 14 at 1:30 PM (ET).

Piazza: We will be using Piazza (www.piazza.com), a question-and-answer system designed to streamline discussion outside of the classroom. It supports LaTeX, code formatting, embedding of images, and attaching of files. It will be moderated by the instructor and TAs, but students are encouraged to answer questions and will receive (limited) extra credit for doing so. All posts will be public so that everyone can benefit from question answers and so that TAs aren't answering the same question many times. Do not post code or answers to Piazza.

ELMS: We will be using ELMS to view lecture slides, hand in assignments, and to see grades.

Grading: Final grades will be based on individual homework assignments (60%), midterm exam (15%), participation (5%), and the individual final project (20%). We aim to have a standard 90/80/70/… breakdown for letter grades, but reserve the right to curve up.

Disability Support Services: Any student eligible for and requesting reasonable academic accommodations due to a disability is requested to provide, to the instructor via email, a letter of accommodation from the Office of Disability Support Services (DSS) within the first two weeks of the semester.

Religious Holidays: If you observe any religious holidays that conflict with the course schedule, please let the instructor know during the first two weeks of the semester so that accommodations can be made.

Course Evaluations: The Department of Computer Science takes student course evaluations very seriously. Evaluations will usually be open during the last few weeks of the course. Students can go to www.courseevalum.umd.edu to complete their evaluations.

Copyright: Students are permitted to use course materials for their own personal use only. Course materials may not be distributed publicly or provided to others (except other students in the course), in any way or format.

Class Concerns: If you have any class concerns, feel free to contact the instructor. If an issue arises with the instructor, report it using the form available at https://www.cs.umd.edu/classconcern

Syllabus: This is the current version of the syllabus. The instructors reserve the right to change it at any time.

Office Hours:

Jia-Bin Huang: TBD

Topics: The following is a *tentative* list of topics and readings in *approximate* order.

1. Introduction to Computer Vision, Linear Algebra and Python
2. Image Processing / Correlation / Convolution / Edge Detection
3. Feature Detection / Corner detection (Harris) / SIFT
4. Classification with Classical Methods / Bag of Features / Nearest Neighbor / SVM
5. Neural Networks
6. Convolutional Neural Networks
7. Segmentation
8. 2D Transformations / RANSAC
9. Optical Flow
10. Tracking
11. Multi-view Geometry / Stereo / Structured Light / Structure from Motion
12. Computational Photography
13. Recent Trends in Computer Vision