ECE 5554/4554: Computer Vision – Fall 2016

Electrical and Computer Engineering, Virginia Tech

Class Time and Place

Where: Room 230, New Classroom Building (NCB)
When: Tuesday and Thursday 3:30 – 4:45 PM

People

Instructors	**Email**	**Office**	**Office Hours**
Jia-Bin Huang	jbhuang	Whittemore 440	Friday 10 – 11 AM
Akrit Mohapatra (TA)	akrit	Whittemore 236	Wednesday 10:30 – 11:30 AM Monday 10:30 – 11:30 AM (on HW due day)

Important links

Webpage:	http://bit.ly/vt-computer-vision-fall-2016
HW submission:	https://canvas.vt.edu/courses/29158/ for ECE 5554 https://canvas.vt.edu/courses/31947/ for ECE 4554
Discussion board:	http://piazza.com/vt/fall2016/ece5554ece4554/home
Google calendar:	http://bit.ly/ece5554-4554-fall16
Feedback:	https://goo.gl/forms/JQMQCzFoMvBxQKZ12

Overview

Computer vision aims to develop algorithms to enable machines to understand and analyze visual data, i.e., teaching machines to see". Applications include 3D reconstruction, autonomous vehicle navigation, medical image analysis, multimedia search, face detection/recognition, entertainment, and security.

In this introductory course, we will cover many of the core concepts and algorithms of computer vision: image formation, linear filters, interest points, correspondence and alignment, single-view and multi-view geometry, grouping, and recognition. Through programming assignments and short answer questions, you will put these core concepts into practice. The course aims to provide broad coverage of the topics and will not go into great depth on any specific area of research. However, having successfully completed this course, you should be prepared for any advanced vision-related applications.

General Information

Textbook Lectures are not based on any particular textbook. Our primary reference for this course is:

- Computer Vision: Algorithms and Applications by Richard Szeliski, 2010. The electronic copy of the book in PDF is freely available on the web page: http:szeliski.org/Book

Other useful references includes

- Computer Vision: A Modern Approach (2nd edition) by David Forsyth and Jean Ponces
- Multiple View Geometry in Computer Vision by Hartley and Zisserman
- Concise Computer Vision: An Introduction into Theory and Algorithms by Reinhard Klette
- Computer Vision, Shapiro and Stockman (a nice introduction to computer vision)
- Linear Algebra and its Applications, Gilbert Strang (excellent book on linear algebra)
- Vision Science: Photons to Phenomenology, Stephen Palmer (great book on human perception)
- Digital Image Processing, 2nd edition, Gonzalez and Woods (a good general image processing text)

Prerequisites. Good knowledge of linear algebra and calculus. Previous experience with MATLAB will be helpful as all homeworks involve programming in MATLAB.

Attendance. Regular attendance is expected. I will post lecture slides on the course website. However, the slides will be difficult to interpret without attending lectures.

Disability-related academic adjustments. To obtain disability-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor and the Services for Students with Disabilities (SSD) as soon as possible. To contact SSD you may visit Suite 310 at Lavery Hall, or contact SSD via email ssd@vt.edu or here.

Learning objectives Having successfully completed this course, the students will be able to

- Become familiar with both the theoretical and practical aspects of image processing and analysis techniques.
- Describe the foundation of image formation and image analysis.
- Understand basics of measurements and robust detection of local features in images.
- Describe various methods used for registration, alignment, and matching across images.
- Understand the basics of 2D and 3D Computer Vision.
- Program software of core computer vision techniques such as edge detection, shape registration, multi-view reconstruction, tracking, and image categorization.
- Get an exposure to advanced concepts leading to object and scene categorization.
- Develop practical skills that are necessary for building computer vision applications in other domains.
- Understand basic ideas of deep convolutional neural networks and their applications in computer vision.

Assignments and Grading

- **Homeworks (60% of final grade)**: There are in total five homework assignments.
- **Final project (25% of final grade)**: Do a final project of your choice in groups of 2-4 people.

- **Final exam (15% of final grade)**: The exam will cover a variety of conceptual and paper-and-pencil problems related to the topics of lectures.

Graduate credits Graduate students enrolled in ECE 5554 will be expected to do additional work for each homework assignment. Each homework assignment is worth up to 100 points, so you can earn 500 points through the standard assignments. In each assignment, we will also list several extra credit opportunities available. ECE 4554 students are graded out of 525 points. ECE 5554 students are graded out of 600 points.

Academic honesty : Feel free to discuss homeworks with your classmates, but please refrain from showing or sharing any code. Any existing code from the Internet cannot be used in your project assignments unless it is specifically approved by the course instructor. Be sure to acknowledge any help that you do get from other students or outside works, even if its just a small suggestion. Note that violations of academic integrity will go on record at the university, and zero points for the entire project assignment. For details, please review the honor code manual and policy online.[1]

Due dates : All problem sets/reports are to be submitted through https://canvas.vt.edu/ by the due date noted on the assignment. Deadlines are firm.

Late policy : You are expected to do assignments on time. Late assignments will be assigned a penalty of 20% per day. Throughout the term you have an allowance of **four** free late days for your submissions, meaning you can accrue up to four days in late submissions with no penalty.

Final project: The final project is a chance to further explore a topic of interest. Groups of up to four are highly encouraged. More is expected of larger groups. Projects will include a project report webpage and a poster presentation. Various types of projects are possible. You could implement a paper that you find interesting, something discussed in class, a significant extension of one of the course projects, or something entirely of your own design. The work does not have to be of publishable originality. However, you are encouraged to submit the revised versions of projects to top computer vision conferences.

- *Research project*: Perform a project in a topic of your choice. Formulate a goal, devise an approach, and evaluate. When proposing, indicate what dataset you will use for evaluation. For example, you could base your project on an existing paper and try to improve the accuracy or speed with some modification. You could also apply existing algorithms to your own field (e.g., robotics).

- *Review and implement a paper*: Choose a paper or set of papers and write a scholarly review. Then, implement and evaluate the algorithm. If done in a group, more than one paper should be implemented and compared. Reviews should be written independently for each person, but the group can collaborate on implementation and evaluation.

[1] http://www.honorsystem.vt.edu/HSConstitution/honor_code_manual_and_policy.html

Class Organization

The class is organized into six sections, with one homework for each of the first five sections. Homeworks will include some derivation, concept questions, and some implementation. Assignments may, for example, include building an edge detector; detecting and tracking interest points; using tracked interest points to do 3D reconstruction; parameter learning with EM; object instance recognition; graph cut segmentation; and image categorization.

- **Weeks 1-3: Interpreting Intensity**
  - Light, shading, and color
  - Image filters, template matching, and image pyramids
  - Edge detection
- **Weeks 4-5: Correspondence and Alignment**
  - Interest point detection and feature tracking
  - Fitting and registration of objects and images
- **Weeks 6-7: Perspective and 3D Geometry**
  - Camera models
  - Single-view and multi-view geometry
  - Stereo and structure-from-motion
- **Weeks 9-10: Grouping and Segmentation**
  - Clustering, EM
  - Segmentation
- **Weeks 11-13: Recognition**
  - Face recognition
  - Image features and categorization
  - Statistical templates and part-based recognition
- **Weeks 14-15: Special Topics**
  - Action recognition
  - Convolutional neural networks
  - 3D scene and contexts