Course Syllabus

Instructor: Matthew Dailey; Office Computer Science building room 103; Phone 02 524 5712;
Email `mdailey@ait.asia`; Office hours 2–3 PM Mondays.

Lectures: Tue Thur 13:00–16:00, Room 106 Computer Science Building.

Course Web Page: `http://www.cs.ait.ac.th/~mdailey/class/vision` — all course information, handouts, and assignments will be posted there.

Course Goals: Machine vision is concerned with the image processing, geometry, and statistical inference tools necessary for extracting useful information about the world from two-dimensional images.

We have arrived at a curious stage of technological development, in which it is now possible for a high school student to use modern neural network development tools to build a model capable of recognizing objects in images at a level of accuracy rivaling humans. The current pace of improvement in visual processing capabilities of computer systems was unimaginable only 10 years ago.

However, many problems, such as understanding how a modern neural network model works, or embedding a computer vision model in a real-world software system, or developing new models for new problems, or extracting useful 3D information about the world from 2D video streams, require advanced knowledge of and tools for statistical modeling, optimization, projective geometry, signal processing, and software engineering.

This course is an advanced survey of the state of the art in machine vision, focused primarily on robotics applications and human-computer interfaces. The course is a mixture of lectures on fundamentals, student presentations of research from the primary academic literature, and group projects involving application of machine vision technology to real-world problems. The course prepares students to do thesis research in the field.

This year we will focus on topics related to 3D reconstruction of objects and scenes from video or still images, camera motion estimation from video, object detection and recognition, and tracking.

This is *not* a basic image processing course. We will briefly cover necessary background for the methods we study in class, and Octave and OpenCV will be introduced by example along the way, but for a complete treatment of image processing you should take existing courses in ISE or the UG program.

Readings: The main required textbooks are Hartley and Zisserman's classic beautiful guide to 3D computer vision and Goodfellow, Bengio, and Courville's guide to deep learning:

- R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge University Press, 2004.
- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016. Electronic copy available from Amazon.com; HTML version freely available at `http://www.deeplearningbook.org`.

Additional material will be drawn from several textbooks and readings from the academic literature:

- R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010. Available on Web.
- C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- C. Tomasi, *Mathematical Modelling of Continuous Systems.* Available on Web.
- A. Kaehler and G. Bradski, *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*, O'Reilly, 2016.

- Primary academic journals and conferences:
  - International Journal of Computer Vision (IJCV)
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)
  - Image and Vision Computing
  - Pattern Recognition
  - Pattern Recognition Letters
  - Computer Vision and Image Understanding
  - IEEE Computer Vision and Pattern Recognition Conference (CVPR)
  - International Conference on Computer Vision (ICCV)
  - European Conference on Computer Vision (ECCV)
  - Asian Conference on Computer Vision (ACCV)
  - IEEE International Conference on Robotics and Automation (ICRA)
  - IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)

Prerequisites: You need to be able to program in some high-level language like C or Java, and be comfortable with engineering mathematics.

Grading:

- (20%) Homework
- (30%) Midterm
- (20%) In-class presentations
- (30%) Project

Presentations: In most meetings of the class, we will have one or two student presentations on an academic research paper describing applications of machine vision technology, possibly in the student's anticipated thesis research area. Presentations should be planned for 20 minutes, allowing 10 minutes for discussion. Each student taking the course for credit must give two presentations. You should select your papers from the list of suggested papers on the course Web site.

Well in advance of your presentation, you should prepare a presentation of about 15 slides. I recommend LaTeX with the "Beamer" document class, but you can also use OpenOffice, Powerpoint, etc. The slides must be your own original work. DO NOT PLAGIARIZE PRESENTATIONS YOU FIND ON THE NET. Send me your presentation in advance in PDF format so I can post it to the course Web site. In class, as you give your talk, your peers will do a written evaluation of the presentation, offering suggestions for improvement. Finally, within a week after your talk, after incorporating feedback from the peer evaluations, you should give the final presentation in PDF format to me so I can post it to the course web site.

The level of the talk should be very technical and geared towards getting your colleagues to understand the approached being used, as deeply as possible given time limitations. The talk should present the problem the paper addresses, the technique used to solve the problem, an overview of any experimental results, a critique of the method's strengths and weaknesses, and a discussion of how this work is related to your future or ongoing research. Your presentation grade will be based on the organization and content of your slides, their effectiveness in presenting the main points of the paper you studied, and the effectiveness of your presentation.

Papers for presentation: This semester, the readings for student presentations will be drawn from *International Journal of Computer Vision* (IJCV) and *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* issues from 2017 and 2018. These journals are the top journals specific to computer vision, and they both have some of the highest journal impact factors of all journals in computer science.

This means the papers will be very high in quality but also quite advanced. To deeply understand the paper you have selected for presentation, you may need to spend considerable time researching the background literature. You can follow the references in your paper or come to me for help.

AIT subscribes to IJCV and IEEE PAMI online. You should be able to freely download papers as long as you have an AIT IP address. If you have any trouble, please contact the library.

Discussion and presentation feedback: All students are expected to read the assigned papers and come prepared for discussion. You should try to jump into the discussion in every session of the class, if nothing else by asking a question about something you read carefully but could not understand. As you listen to the presentation, try to understand the material and think about how the techniques being discussed might be useful in your own research or other domains entirely.

Homework: I will assign 4 or 5 short homework assignments to make sure you understand the important concepts being taught in the class. You will submit a brief writeup of your work in class on the due date. The usual honesty policy applies (see below). In general, you can use code available online, but you must give credit to the original authors when you do.

Midterm: In the midterm, you will be assigned some practical problems to solve within the fixed exam period and demonstrate your solutions to me at the end of the exam.

Project: If you are taking the course for credit, you will do an individual implementation project in machine vision. Discuss what you want to do in your project with me.

The project deliverables are

- A project plan presentation explaining what you will do, to be presented in week 2;
- Weekly project status reports explaining your progress on the project;
- A 20-minute presentation of your project and its results, in the last week of class;
- A 6-page final report on your project in IEEE conference format, due in the last week of class.

You must use LaTeX for the final report. You will probably hate it at first (unless you are a math and programming geek like me) but after the initial learning curve you will be very happy when you are breezily preparing your final thesis document using the AIT LaTeX document class and the rest of your colleagues are struggling to get consistent formatting and pretty equations in other word processors.

Auditing: If you audit the course, I expect you to keep up with the reading, contribute to the discussions, give presenters feedback through evaluations, and give one presentation.

Honesty policy: The two cardinal crimes of academia are fabricating results and plagiarism. Either could get you fired from your job, depending on what you do. Don't be tempted to do either.

Fabricating results includes things like modifying your numerical results to make your work look better than it really is, or leaving out measurements that are inconvenient, and so on. This is the worst form of dishonesty in academia. In this class, you must report your results in a complete and honest way.

Plagiarism, or taking someone else's work and representing it as your own, or not giving them full credit, constitutes lying, cheating, and stealing. In your presentation slides, it should be clear what material is yours and what material came from the paper or the work of others. For example, it is fine to capture a figure or equation from a paper and put it in your presentation as long as it is clear that the figure is from the paper (you should add a footnote saying which figure it is). It is NOT fine to take the text of someone else's slides on the same paper and put it in your presentation UNLESS you very clearly specify that you are directly quoting someone else's work and provide a citation. Similarly, in your project, it is fine to take public source code and use it as part of your system, as long as in your final project presentation and report you give proper credit to the source. If you have any questions about what is acceptable vs. unacceptable use of someone else's work, just ask me.

Violations of the honesty policy will, at the very least, result in no credit for the work in question and a letter to the Dean. Depending on the seriousness of the infraction, penalties might include an F for the course and more severe punishment.

Course schedule: We will mainly focus on 3D reconstruction, machine learning (especially deep learning models) in vision, and sequential state estimation (tracking). Matlab/Octave and the OpenCV library will be used throughout for implementation examples. Student presentations will cover additional techniques and applications. Time allowing, we will have one or two lectures on academic writing. See the course web site for an up-to-date lecture schedule.