ECE 497: An Invitation to 3-D Vision: From Images to Models

Yi Ma
Department of Electrical & Computer Engineering

Fall 2002

Course description
This is a new advanced graduate-level course on the geometry and dynamics of computer vision. It will cover both geometric and algorithmic aspects of recovering three-dimensional motion and structure from multiple (or a sequence of) two-dimensional images. The first part of the course consists of a mathematical introduction to multi-view geometry relying on a brand new rank techniques from linear algebra, with a special emphasis on a global approach to multiple view analysis (as opposed to most extant methods). New geometry and algebra associated to multiple views of multiple moving objects will be reached in the end, with applications to motion segmentation and reconstruction (which no vision course or book so far has ever covered).

This is an advanced course in computer vision and it can be considered as a follow-up course of ECE449 or CS443, since the topics covered will be narrower but much deeper. However, the course is self-contained, necessary background knowledge in rigid body motion, camera models, feature extraction and correspondence, two and three view geometry will be briefly covered in the beginning. Since they are not the focus, softwares and codes will directly be given to students to get hands-on experience. The goal is however for them to learn how to deal with multiple images and image sequences to achieve much more demanding reconstruction tasks.

The essential prerequisite is linear algebra (Math318, Math381 or ECE415). Some familiarity with geometry (Math423 or Math424), control theory (ECE415), estimation theory(ECE461) and rigid body kinematics (ECE389/GE389) would increase your appreciation but not crucial. The course targets at the following students:

1. Graduate students in ECE/CS in the areas of computer vision and computer graphics interested in a more unified and simple treatment of reconstruction from multiple images or video sequences.
2. Graduate students in ECE or ME in the areas of control, robotics and image processing interested in vision based robotic control.
3. Graduate students in Mathematics interested in applications of modern geometry and multi-linear algebra to computer vision.

General information
Primary text:

1. Invitation to 3-D Vision: from Images to Models (to be provided by the instructor), Yi Ma, S. Soatto, J. Košcká and S. Sastry, to appear, Springer, 2002.

Recommended readings:


**Grading policy:** Weekly homework (60%) and Final Project (40%). The final project can be done in a group of 2 or 3 students. The project can be theoretical, experimental or a mix of both. It consists of a midterm proposal, a final presentation (in class) and a web based report.

**Course syllabus and schedule**

1. **Representation of a three-dimensional moving scene (3 hours):** Rigid motion, canonical exponential coordinates, Rodrigues formula, Euclidean, affine and projective transformations.

2. **Image formation (1.5 hours):** Mathematical model for ideal perspective projection and the pinhole camera, other geometric projection models.

3. **Image primitive and correspondence (1.5 hours):** Photometric features and geometric features, image correspondences and optical flows, feature selection, matching and tracking.

4. **Two view geometry (3 hours):** Epipolar geometry (discrete and continuous cases), geometric characterization of the essential matrix, and the eight-point, seven-point and six-point algorithms, optimal estimation and reconstruction from two views.

5. **Camera calibration and self-calibration (6 hours):** Camera calibration from a rig, uncalibrated epipolar geometry, the fundamental matrix, camera self-calibration from Kruppa’s equation and from projective, affine and Euclidean stratification.

6. **Introduction to multiple view geometry (3 hours):** Two and three view geometry review, relation to bi-focal and tri-focal tensors, introduction to the rank techniques.

7. **Midterm project proposal (7th week):** A 3 page project proposal from each team.

8. **Multiple view geometry of point and line features (4.5 hours):** Multiple view constraints in terms of matrix rank, reconstruction based on pure point and line features.

9. **Batch multiple view reconstruction (3 hours):** A universal rank condition for multiple images of point, line and plane, a multiple view factorization algorithm based on the rank condition.

10. **Multiple view geometry in high dimensional spaces (3 hours):** Generalized rank conditions for multiple images with applications to dynamical scenes, motion segmentation, group and formation detection, occlusion.

11. **Recursive reconstruction from image sequences (4.5 hours):** Motion and structure as a filtering problem, observability, realization, stability, implicit extended Kalman filter (IEKF).

12. **Selected topics in graphics and robotic control (6 hours):** Three-dimensional map/model building from images, vision based navigation (of mobile robot), vision based landing (of helicopter), tracking and estimation of human motion.

13. **Final project presentation and report (1.5 hours):** 20 minutes in class presentation for each project.