Syllabus (Spring 2018, 5980)

Course title

Multiview 3D Geometry in Computer Vision

Description

Multiple cameras are continually capturing our daily events involving social and physical interactions in a form of first person camera (e.g., google glass), cellphone camera, and surveillance camera. Multiview geometry is a core branch in computer vision that studies the 3D spatial relationship between cameras and scenes. This technology is used to localize and plan robots, reconstruct a city, e.g. Rome, from internet photos, and understand human behaviors using body-worn cameras. In this course, we will focus on 1) fundamentals of projective camera geometry; 2) implementation of 3D reconstruction algorithm; and 3) research paper review. The desired outcome of the course is for each student to have his/her own 3D reconstruction algorithm called "structure from motion". This will cover core mathematics of camera multiview geometry including perspective projection, epipolar geometry, point triangulation, camera resectioning, and bundle adjustment. This geometric concept will be then, in parallel, implemented to directly apply to domain specific research such as robot localization. This course includes a final term project that uses the multicamera system to build a creative system such as Robotics/AR/VR/Vision/Graphics.

Instructor
Hyun Soo Park

Course Website
Slides/Schedule/HW/Papers/Code will be released through the website:
`http://www-users.cs.umn.edu/~hspark/CSci5980/csci5980_3dvision.html`

Code
`http://www-users.cs.umn.edu/~hspark/CSci5980/code/`

Lecture
Mon/Wed 4:00pm-5:15pm (Ford Hall B15)

Office hour
Tue/Thr 5:00pm-6:00pm (Keller Hall 5-225E)

Prerequisites
CSci 2033 (Linear Algebra); MATLAB

Reference Book
Hartely and Zisseman, “Multiple View Geometry in Computer Vision”, 2nd Edition

Topic

1. Computer Vision Introduction
2. Single View Geometry
    - Camera Model
    - Camera Projection Matrix
    - Projective Line
    - Where am I? (Vanishing Point)
    - Single View Metrology
    - Image Transform
    - Linear Least Sqaures / Homography
    - Where am I? (Homography)
    - Camera Calibration
    - Rotation Representation
    - Where am I? (Perspective-n-Point)
3. Multiview Geometry
    - Epipolar Geometry
    - Fundamental Matrix
    - Where am I? (Essential Matrix)
    - Triangulation
    - Stereo
    - Feature Descriptor/Matching/Tracking
    - Robust Estimation (RANSAC)
    - Nonlinear Refinement (Gradient Decent Optimization)
    - Bundle Adjustment I (Geometric Error)
    - Bundle Adjustment II (Sparse Optimization and Analytical Jacobian)
    - Bundle Adjustment III (Practical Issues/Ceres Solver/SBA Solver)
    - Nonrigid Structure from Motion
4. Future of Computer Vision

Grading Policy

- 50%: 5 homework assignments
- 20%: exam (April 4)
- 5%: project proposal presentation and document
- 10%: project related work presentation and document
- 15%: project final presentation and document
- Late submission: 20% off from each extra late day

Homework Assignment

5 programming tasks will be assigned in the course. The homework will be announced two weeks before the deadline, and students will submit their codes with a brief report. **No team work or collaboration is allowed.**

- Building a camera obscura and dolly zoom: students will design their own camera obscura and simulate dolly zoom effect with their cellphone camera.
- Creating a 360 panoramic image: students will write a code that creates a panoramic image using multiple images taken by their own cellphone camera.
- Virtual tour into your photo: students will write a code that navigates their image in 3D using homography.
- Camera Pose Estimation: students will use multiple images of a 3D scene to compute the relative transformation between two images.
- Bundle Adjustment: students will build a full 3D reconstruction code optimized by bundle adjustment with analytic Jacobian.

Exam

One exam will be taken place. The exam will be closed-book, and no calculators or other electronic devices will be allowed. Students will be permitted to use one double-sided, 8.5"x11" note sheet for the exam.

Course Project

Two students can form a team for their course project. The project is designed to write a computer vision paper (6-8 pages) using the multi-camera system in Shepherd Laboratory (`https://www.youtube.com/watch?v=3tDhRTUfxBQ`). Your team will come up with a novel idea of 3D reconstruction using multiple view images.

The course project will be evaluated in terms of three essential parts of a technical paper:

- project proposal (10 min presentation): the team will present the idea and associated data, and write introduction of the paper.
- project related work (10 min presentation): the team will present at least two seminal papers relevant to the project and write related work of the paper.
- project final (20 min presentation): the team will present full paper including method, result, and conclusion.

The CVPR paper format can be found in: `http://cvpr2018.thecvf.com/submission/main_conference/author_guidelines`

Scholastic Misconduct

Scholastic misconduct is broadly defined as "any act that violates the right of another student in academic work or that involves misrepresentation of your own work. Scholastic dishonesty includes, (but is not necessarily limited to): cheating on assignments or examinations; plagiarizing, which means misrepresenting as your own work any part of work done by another; submitting the same paper, or substantially similar papers, to meet the requirements of more than one course without the approval and consent of all instructors concerned; depriving another student of necessary course materials; or interfering with another student's work."