Division of Engineering Science
UNIVERSITY OF TORONTO

ROB501: Computer Vision for Robotics

Course Information and Syllabus

Fall 2019

Course Description and Learning Objectives

This course provides an introduction to aspects of computer vision specifically relevant to robotics applications (i.e., *robotic* vision). Topics of study will include the geometry of image formation, basic image processing operations, camera models and calibration methods, image feature detection and matching, stereo vision, structure from motion, and 3D reconstruction from a moving platform. More recent developments in the field, such as deep learning for robotic vision, will also be explored later in the course.

The objective of the course is to enable learners to understand and implement the fundamental algorithms necessary for robots to perform a variety of common tasks (e.g., safely navigating through the hallways of a building) using visual perception. By the end of the course, you will be able to:

- explain the image formation process and interpret the models used to describe common monocular and stereo camera sensors;
- select and apply various image filtering and detection operations to extract the relevant information from image sequences needed for common robotic tasks;
- compare and contrast several 'classical' computer vision methods with learning-based techniques, including deep learning, in terms of effectiveness and complexity;
- implement, modify, and evaluate the performance of algorithms for visual motion estimation, navigation and mapping, and object recognition; and
- integrate visual processing as part of the standard perceive-plan-act loop running on board a real (mobile) robot platform.

As an advanced undergraduate class, the lectures will involve a seminar-style component—we will discuss seminal and novel research in the field (which is moving very quickly). The goal is to provide solid preparation for further study and research at, for example, the graduate level or in industrial laboratories.

Instructor

- Prof. Jonathan Kelly, ✉ jkelly@utias.utoronto.ca, ☎ (416) 667-7708 (UTIAS office)

Teaching Assistants

- Mr. Oliver Limoyo (Primary), ✉ oliver.limoyo@robotics.utias.utoronto.ca
- Ms. Chih-Chun (Katherine) Chen (Grading), ✉ chihchun.chen@mail.utoronto.ca

Schedule

- **Lectures:** Tuesday 12:00–14:00, MY380, and Thursday 9:00–11:00, MY380.
- **Office Hours:** By email appointment; additional office hours will be scheduled before both of the exams.

Course Website

We will use the University of Toronto Quercus platform for course administration:

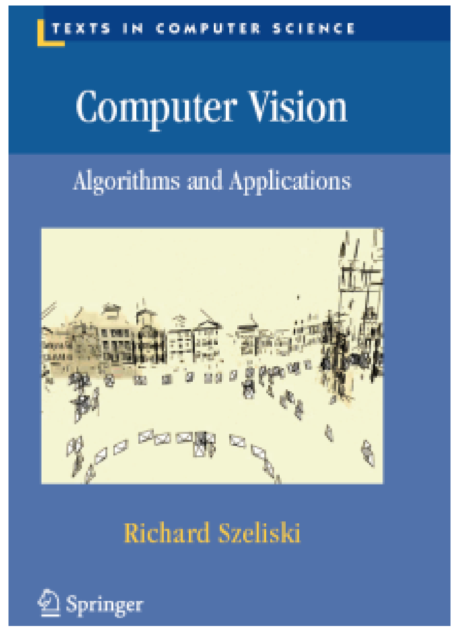
http://q.utoronto.ca/

Relevant slides, notes, papers, and other materials will be posted on Quercus. Please ensure that you regularly check the *Discussions* section for information related to projects, grading, upcoming review sessions, etc.

Course Textbook and Reference Materials

Our primary reference text for the course will be the popular book by Richard Szeliski:

> R. Szeliski, *Computer Vision: Algorithms and Applications*, 1st ed., ser. Texts in Computer Science. London, United Kingdom: Springer-Verlag London, Nov. 2010, ISBN: 978-1-84882-934-3. DOI: 10.1007/978-1-84882-935-0.



This text is widely used in many introductory computer vision courses, at both the undergraduate and graduate levels. It's also available for free in PDF format online at http://szeliski.org/Book/, along with a listing of errata. Although the book is available online, the print copy is beautifully formatted and a pleasure to read.

To complement the Szeliski book, and to provide materials related specifically to robotic vision, we will augment the main text with excerpts from other references (notably the 2012 book by Prince), as well as readings from important papers in the field. A listing of all of the readings will be available as a separate document on Quercus at the start of the semester.

Prior to most lectures, PDF versions of the slides will also be posted on Quercus. However, on occasion, we may work directly on the chalkboard—summary notes will then be posted after the lecture session.

Outline and Syllabus

A *rough* outline of the topics that will be covered in the lectures is provided below. Each week, the tutorial portion of the class will be dedicated to a) providing examples of course concepts implemented in code and/or in operating robotic systems, and b) discussing seminal papers in robotic vision.

#	Date	Topic	Readings
1	Sep. 5	**Overview and Introduction to the Course** Brief history of robotic vision, material covered in the course.	S § 1
2	Sep. 10	**Mathematical Fundamentals** Coordinate frames, geometric representations, transforms.	S § 2.1 Lecture Notes
3	Sep. 12	**Probability Theory and Estimation Techniques** Probabilities, multivariate Gaussians, least squares, RANSAC.	P § 2 Lectures 2 & 3: [1]
4	Sep. 17	**Image Formation and Optics** Projection from 3D to 2D, camera models, optical effects, HDR.	S § 2.1–2.3, 10.1–10.2 Lecture 4: [1]

5	Sep. 19	**Image Processing and Transformations** Point operations, linear/nonlinear transforms, regularization.	S § 3.1–3.3, 3.7.1
6	Sep. 24	**Image Features: Detection and Description** Detection of salient points, variant and invariant descriptors.	S § 4.1, 4.2.1 Lecture 6: [1–3]
7	Sep. 26	**Image Features: Matching and Tracking** Matching criteria, windowed tracking (e.g., KLT).	S § 4.1 Lecture 7: [1]
8	Oct. 1	**Camera Pose Estimation** Linear and iterative methods for the PnP problem.	S § 6.1–6.2
9	Oct. 3	**Stereo Vision** Epipolar geometry, fronto-parallel projection, stereo matching.	S § 11.1–11.5 Lecture 9: [1–5]
10	Oct. 8	**Monocular and Stereo Camera Calibration** Intrinsic and extrinsic calibration techniques.	S § 6.3 Lecture 10: [1–7]
11	Oct. 10	**Structure from Motion: Incremental Methods** Frame-to-frame motion estimation and visual odometry.	S § 7.1–7.2 Lecture 11: [1–2]
12	Oct. 15	**Case Study: Computer Vision on Mars** ♂ Vision for EDL, navigation, and science discovery.	Lecture 12: [1–5] Lecture Notes
13	Oct. 17	**Active Learning Session** Solving the calibration problem without an initial guess.	Handout
14	Oct. 22	**Open Review Session** Summary of course content (so far), papers to understand.	*Course + Exam Q & A*
15	Oct. 24	***Midterm Examination*** *Covers material up to and including Lecture 13.*	*Review Session*
16	Oct. 29	**Structure from Motion: Bundle Adjustment** History, batch bundle adjustment framework, characteristics.	S § 7.4–7.5 Lecture 16: [1]
17	Oct. 31	**Optical Flow and Motion Tracking** The flow problem, tracking motion using flow, examples.	S § 8.4 Lecture 17: [1]
18	Nov. 5	**Dense and Photometric Mapping** Building fully dense 3D representations from multiple views.	S § 12.1, 12.3–12.6 Lecture Notes
19	Nov. 7	**Image Segmentation** Active contours, split & merge, k-means.	S § 5.1–5.3
20	Nov. 12	**Recognition for Robotics** Object detection, instance and category recognition.	S § 14.1–14.4 Lecture 20: [1–2]
21	Nov. 14	**Visual Servoing** Image-based cost functions, integration with control loops.	C § 15 Lecture Notes
22	Nov. 19	**Deep Learning: An Introduction for Robotic Vision** History, networks structures, learning methods.	G § 6, 9 Lecture Notes
23	Nov. 21	**Deep Learning: SfM, Localization, and Mapping** Structure from motion as a learning problem, localization.	Lecture 23: [1–2] Lecture Notes
24	Nov. 26	**Deep Learning: Semantic Segmentation and Recognition** Semantic segmentation of places and objects, recognition tasks.	Lecture 24: [1–2] Lecture Notes

25	Nov. 28	**Case Study: Self-Driving Cars and the Road Ahead** 🚘 Use of vision for autonomous driving, state-of-the-art systems.	Lecture Notes
26	Dec. 3	**Open Review Session** Summary of course content, papers to understand.	*Course + Exam Q & A*
	TBA	***Final Examination*** ✎ *Covers* ***all*** *course material, including lectures and tutorials.*	*Review Session*

In the 'Readings' column, the abbreviation 'S' refers to the text by Szeliski, 'P' to the text by Prince, 'C' to the text by Corke, and 'G' to the text by Goodfellow et al. Complete bibliographic information for all of these books is available in the course reading guide.

Programming Projects

There will be five programming 'mini'-projects assigned during the course. The projects will allow students to gain experience working with basic robotic vision algorithms. In general, projects will involve development 'from scratch,' (e.g., usually, no use of computer vision libraries will be allowed). The release dates and due dates for the projects will be posted on Quercus early in the term. Note that *each student must design, implement, and submit her or his own solution for every project,* although discussion amongst your classmates is encouraged.

For project submissions, each student in the course will be given **two** complementary late days total, to be used as she or he sees fit (i.e., just tell the TA when submitting your project that you're 'cashing in' a late day). *After these two days are used, no further late project submissions will be accepted, and a mark of zero will be recorded.*

Our primary programming tool will be Python (version 3.7), which is a highly versatile language for general purpose and scientific computing. We will also use several Python libraries, including NumPy, Matplotlib, and OpenCV (with Python bindings). Students will submit their projects for testing and grading through an online system; more details will be provided at the start of the semester.

Active Learning Exercise

This semester, students will also participate in an active learning exercise (worth 5% of your total grade). Active learning is *participatory* and involves engagement beyond passive listening; you can find out more about active learning on the relevant Wikipedia page.

You will spend one two-hour class session working together with your classmates in groups (five to six students per group) to solve an interesting robotic vision problem. During the session, the instructor and the TAs will be available to answer questions and to provide guidance. This is an opportunity for 'hands-on' robotic vision and should be lots of fun! More details will be provided prior to the session (scheduled for October 17).

Relationship to Other Robotics Courses

The ROB501 course is designed to provide further insight into one of the most important sensing modalities, vision, in the context of robotics. You will use concepts from ROB301 (*Introduction to Robotics*) in your work, and this course will show, in turn, how vision can be used to complete several of the tasks discussed in ROB301 (navigation, mapping, etc.). We will draw on mathematical tools (Bayesian statistics and optimization) from ROB310 (*Mathematics for Robotics*) in our work. In AER521 (*Mobile Robotics and Perception*), you will learn how vision can be incorporated into more advanced mobile robotic systems.

Grading

Grades will be assigned according to the following weighting scheme:

- Programming Projects – 30% (6% each $\times$ 5)
- Active Learning Exercise – 5%
- Midterm Exam – 25%
- Final Exam – 40%

The allowed aid for the midterm exam will be a single sheet of paper (letter-sized). Students may write on both sides of the aid sheet, without restriction. Nothing may be affixed or appended to the sheet. The allowed aid for the final exam will also be a single sheet of paper (letter-sized), again without anything affixed or appended. No calculators will be allowed (nor required).

"The highest activity a human being can attain is learning for understanding, because to understand is to be free."

– Baruch Spinoza