Engineering Tripos Part IIB, 4F12: Computer Vision, 2024-25

Module Leader

Prof R Cipolla [1]

Lecturers

Prof R Cipolla and Dr Matthew Johnson [2]

Timing and Structure

Michaelmas term. 16 lectures (including 3 examples classes). Assessment: 100% exam

Aims

The aims of the course are to:

- introduce the principles, models and applications of computer vision.
- cover image structure, projection, stereo vision, structure from motion and object detection and recognition.
- give case studies of industrial (robotic) applications of computer vision, including visual navigation for autonomous robots, robot hand-eye coordination and novel man-machine interfaces.

Objectives

As specific objectives, by the end of the course students should be able to:

- design feature detectors to detect, localise and track image features.
- model perspective image formation and calibrate single and multiple camera systems.
- recover 3D position and shape information from arbitrary viewpoints;
- appreciate the problems in finding corresponding features in different viewpoints.
- analyse visual motion to recover scene structure and viewer motion, and understand how this information can be used in navigation;
- understand how simple object recognition systems can be designed so that they are independent of lighting and camera viewpoint.
- appreciate the commerical and industrial potential of computer vision but understand its limitations.

Content

- ***Introduction*** (1L)
  Computer vision: what is it, why study it and how ? The eye and the camera, vision as an information processing task. 3D interpretation of 2D images. Geometrical, statistical and learning frameworks for vision. Applications.

- ***Image structure*** (4L)
  Image intensities and structure: edges, corners and blobs. Edge detection, the aperture problem and corner detection. Image pyramids, blob detection with band-pass filtering. The SIFT feature descriptor for matching. Characterising textures.

- ***Projection*** (4L)
  Orthographic projection. Planar perspective projection. Vanishing points and lines. Projection matrix, homogeneous coordinates. Camera calibration, recovery of world position. Weak perspective and the affine camera. Projective invariants.

- ***Stereo vision and Structure from Motion*** (2L)
  Epipolar geometry and the essential matrix. Recovery of depth by triangulation. Uncalibrated cameras and the fundamental matrix. The correspondence problem. Structure from motion. 3D shape examples from multiple view stereo.

- ***Deep Learning for Computer Vision*** *(5L)*
  Basic architectures for deep learning in computer vision. Object detection, classification and semantic segmentation. Object recognition, feature embedding and metric learning. Transformer architectures and self-supervised learning.

- ***Example classes***
  Discussion of examples papers and past examination papers will be integrated with lectures.

Booklists

Please refer to the Booklist for Part IIB Courses for references to this module, this can be found on the associated Moodle course.

Examination Guidelines

Please refer to Form & conduct of the examinations [3].

Last modified: 04/02/2025 12:03

Links
[1] mailto:rc10001@cam.ac.uk
[2] mailto:rc10001@cam.ac.uk and mj293@cam.ac.uk
[3] https://teaching24-25.eng.cam.ac.uk/content/form-conduct-examinations