3D Models in Computer Vision, 5 ECTS **UJM semester 2**

Course instructors: Prof. Alain Trémeau **Language of instruction: English**

Overview

Computer Vision is an important field related to Artificial Intelligence concerned with questions such as "how to extract information from 2D or 3D image(s) or videos, or how to reconstruct 3D scenes from 2D views". Recent explosive growth of digital imaging technology, advanced computing, and deep learning makes the problems of automated image interpretation even more exciting and much more relevant than ever.

This course introduces students to fundamental problems in computer vision, as well as state-of-the-art solutions. Topics covered in detail include: image formation, camera geometry, feature detection, geometric frameworks for vision, single view and two views geometry; 3D visual reconstruction, camera calibration; stereo vision, image classification and object recognition, deep learning and neural networks for computer vision etc.

The course features extensive practical components including computer labs and practical projects that provide students with the opportunity to practice and refine their skills in computer vision.

Learning outcomes

On successful completion of this course, students should have the skills and knowledge to:

- Understand and master basic knowledge, theories and methods in 2D and 3D computer vision;
- Identify, formulate and solve practical problems in computer vision;
- Critically review and assess scientific literature in the field and apply theoretical knowledge to identify the novelty and practicality of proposed methods.
- Design and develop practical and innovative computer vision applications or systems.
- Conduct themselves professionally and responsibly in the areas of computer vision image processing and deep learning.

Content

- Introduction to computer vision: What is computer vision? Examples and applications.
- Notations and definitions: 3D Euclidean space, Cartesian coordinate frames and homogeneous coordinates.
- Image formation: Projective geometry, Camera models, Pinhole camera model
- Recovering 3D from images: Visual cues, perception of objects and scenes. Shape from X.
- Fundamentals of objects perception and recognition. Categorization.
- The Marr paradigm and scene reconstruction, Model-based vision. Gestalt cues. Other paradigms for image analysis: bottom-up, top-down, neural network, feedback.
- Pixels, lines, boundaries, regions, and object representations. "Low-level", "intermediate-level", and "high-level" vision.
- Object recognition model-based methods. Appearance-based methods. Invariant features.
- From scenes to objects, emergent features, scene categorization.The importance of the context.
- Recovering 3D from stereovision & Multiview: Introduction to Mutli-view Geometry, Stereovision,
- Two view geometry: Epipolar geometry, 3D reconstruction ambiguities.

- Computation of the Essential Matrix and Fundamental Matrix (linear methods, iterative methods, robust methods), Structure computation, Rectification methods.
- Camera Geometry and Single View Geometry, Calibration and auto-calibration in Stereovision.
- Depth from Triangulation, Two-View Geometry, N-View Geometry, Depth estimation and 3D reconstruction.
- Primitive description from lines, edges, corners, interest points. Correlation methods, energy minimization methods.
- Recovering camera and geometry up to ambiguity (affine approximation, Algebraic methods, Factorization methods).
- Shape from stereovision & N-views, Shape from Motion. Multi-view geometry: computational models, auto-calibration.
- Introduction of Motion Field, Optical Flow. Motion Analysis. Motion detection.
- Deep learning and neural networks -based methods for computer vision.

Teaching methods

- Lectures: 24 hours
- Practical work: 18 hours
- Project work: 18 hours

Study materials

- An invitation to 3D vision, From Images to Geometric Models, Ma Y., Soatto S., Kosecka J. and Sastry S.S. Springer, 2004.
- Computer Vision: Algorithms and Applications, Richard Szeliski, Springer, 2010.
- Multiple view geometry in computer vision, R. Hartley et A. Zisserman, vol. 2., Cambridge Univ Press, 2000
- Computer Vision: Models, Learning, and Inference, Simon J.D. Prince, Cambridge University Press, 2012.
- Programming Computer Vision with Python, Jan Erik Solem, 2012

Assumed Knowledge

- Basic calculus, linear algebra and basic probability theory.
- Entry-level computer programming experience in either Matlab, Python, or C/C++.
- Previous knowledge of digital image processing will be helpful, but is not essential.

Evaluation criteria

- Written exam 40%
- Written assignments / Labs 30%
- Project work 20%
- Seminar presentations 10%