

2015/2016

## Visual Recognition

Code: 43090

Credits: 6

Type: OB/OT/TFM Course: 1 Semester: 2

### Contact

Name: Maria Vanrell Martorell

Email: maria.vanrell@uab.cat

### Lecturers

Gloria Haro (Module coordinator and Project coordinator)

Coloma Ballester

Javier Ruiz

Josep Ram3n Casas

### Use of languages

Principal working language: English

### Prerequisites

Degree in Engineering, Maths, Physics or similar

### Objectives and contextualisation

The goal of this module is to learn the principles of the 3D reconstruction of an object or a scene from multiple images or stereoscopic videos. For that, the basic concepts of the projective geometry and the 3-space are firstly introduced. The rest of the theoretical aspects and applications are built upon these basic tools. The mapping from the 3D world to the image plane will be studied, for that we will introduce different camera models, their parameters and how to estimate them (camera calibration and auto-calibration). The geometry that relates two pair of views will be analyzed. All these concepts will be applied to obtain a 3D reconstruction in the possible settings: calibrated or uncalibrated cameras. In particular, we will learn how to: estimate the depth of image points, extract the underlying 3D points given a set of point correspondences in the images, generate novel views, estimate the 3D object given a set of calibrated color images or binary images, and estimate a sparse set of 3D points given a set of uncalibrated images. The 3D representation in voxels and meshes will be studied. Finally, we will explain the reconstruction and modeling from Kinect data, as a particular model of sensors that provide and image of the scene together with its depths. The concepts and techniques learnt in this module are used in real applications ranging from augmented reality, object scanning, motion capture, new view synthesis, bullet-time effect, robotics, etc.

### Skills and learning outcomes

**E01** - Identify concepts and apply the most appropriate fundamental techniques for solving basic problems in computer vision.

01 - Identify the basic problems to be solved in the recovery of 3D information from scenes, along with the specific algorithms.

**E02** - Conceptualise alternatives to complex solutions for vision problems and create prototypes to show the validity of the system proposed.

02 - Identify the best representations that can be defined for solving problems of 3D information recovery.

**E03** - Choose the most suitable software tools and training sets for developing solutions to problems in computer vision.

03 - Choose the learnt techniques and train them to resolve a particular 3D recovery project

**E04** - Plan, develop, evaluate and manage solutions for projects in the different areas of computer vision.

04 - Plan, develop, evaluate and manage a solution to a particular 3D recovery problem.

**B06** - Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.

05 - Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.

**B07** - Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.

06 - Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.

**B10** - Continue the learning process, to a large extent autonomously

07 - Continue the learning process, to a large extent autonomously.

**T02** - Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.

08 - Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.

**T03** - Accept responsibilities for information and knowledge management.

09 - Accept responsibilities for information and knowledge management.

**T04** - Work in multidisciplinary teams.

10 - Work in multidisciplinary teams.

## Content

1. Introduction and applications. 2D projective geometry
2. 2D projective geometry. Planar transformations.
3. Homography estimation. Affine and metric rectification
4. 3D projective geometry and transformations. Camera models
5. Camera calibration and the image of the absolute conic. Radial distortion
6. Epipolar geometry. Fundamental matrix. Essential matrix. Extraction of camera matrices.
7. Computation of the fundamental matrix. Image rectification
8. Triangulation methods. Depth computation. New View Synthesis
9. Voxel-Based Methods\*. Projective reconstruction with factorization method. Stratified methods.
10. Auto-calibration. Bundle adjustment.
11. 3D sensors (kinect) and 3D model generation (meshes) (I)
12. 3D sensors (kinect) and 3D model generation (meshes) (II)

## Methodology

### Supervised sessions:

- **Lecture Sessions**, where the lecturers will explain general contents about the topics. Some of them will be used to solve the problems.

### Directed sessions:

- **Project Sessions**, where the problems and goals of the projects will be presented and discussed, students will interact with the project coordinator about problems and ideas on solving the project (approx. 1 hour/week)
- **Presentation Session**, where the students give an oral presentation about how they have solved the project and a demo of the results.
- **Exam Session**, where the students are evaluated individually. Knowledge achievements and problem-solving skills

### Autonomous work:

- Student will autonomously study and work with the materials derived from the lectures.
- Student will work in **groups** to solve the problems of the projects with deliverables:
  - Code
  - Reports
  - Oral presentations

## Activities

TYPE	ACTIVITY	HOURS	LEARNING OUTCOMES
Supervised	Project, Presentation and Exam Sessions	10	1, 2, 3, 5, 6, 7, 8, 9, 10
Directed			

Lecture Sessions	20	1, 2
------------------	----	------

#### Autonomous

Homework	120	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
----------	-----	-------------------------------

### Evaluation

The **final marks** for this module will be computed with the **following formula**:

$$\text{Final Mark} = 0.4 \times \text{Exam} + 0.55 \times \text{Project} + 0.05 \times \text{Attendance}$$

where,

**Exam:** is the mark obtained in the Module Exam (must be  $\geq 3$ )

**Attendance:** is the mark derived from the control of attendance at lectures (minimum 70%)

**Projects:** is the mark provided by the project coordinator based on the weekly follow-up of the project and deliverables. All accordingly with specific criteria such as:

- Participation in discussion sessions and in team work (inter-member evaluations)
- Delivery of mandatory and optional exercises.
- Code development (style, comments, etc.)
- Report (justification of the decisions in your project development)
- Presentation (Talk and demonstrations on your project)

### Evaluation activities

TITLE	HOURS	WEIGHTING	LEARNING OUTCOMES
Exam	3	0,4	1, 2, 5, 6, 7, 8, 9
Project	7	0,5	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

### Bibliography

Journal articles:

1. O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*, MIT Press, cop. 1993.
2. O. Faugeras, Q.T. Loung, *The geometry of multiple images*, MIT Press, 2001.
3. R. I. Hartley, A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2000.

Books:

1. D. A. Forsyth, J. Ponce, *Computer vision: a modern approach*, Prentice Hall, 2003.
2. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.