

Cognitive Analysis of Scenes:

**from Computer Vision
to High-Level Descriptions
for Reasoning and Generating
Natural Language Descriptions**

SoSe 2013

Universität Bremen

Program of the Seminar

Dr. -Ing. Zoe Falomir Llansola

Bremen, 1st March 2013

Details

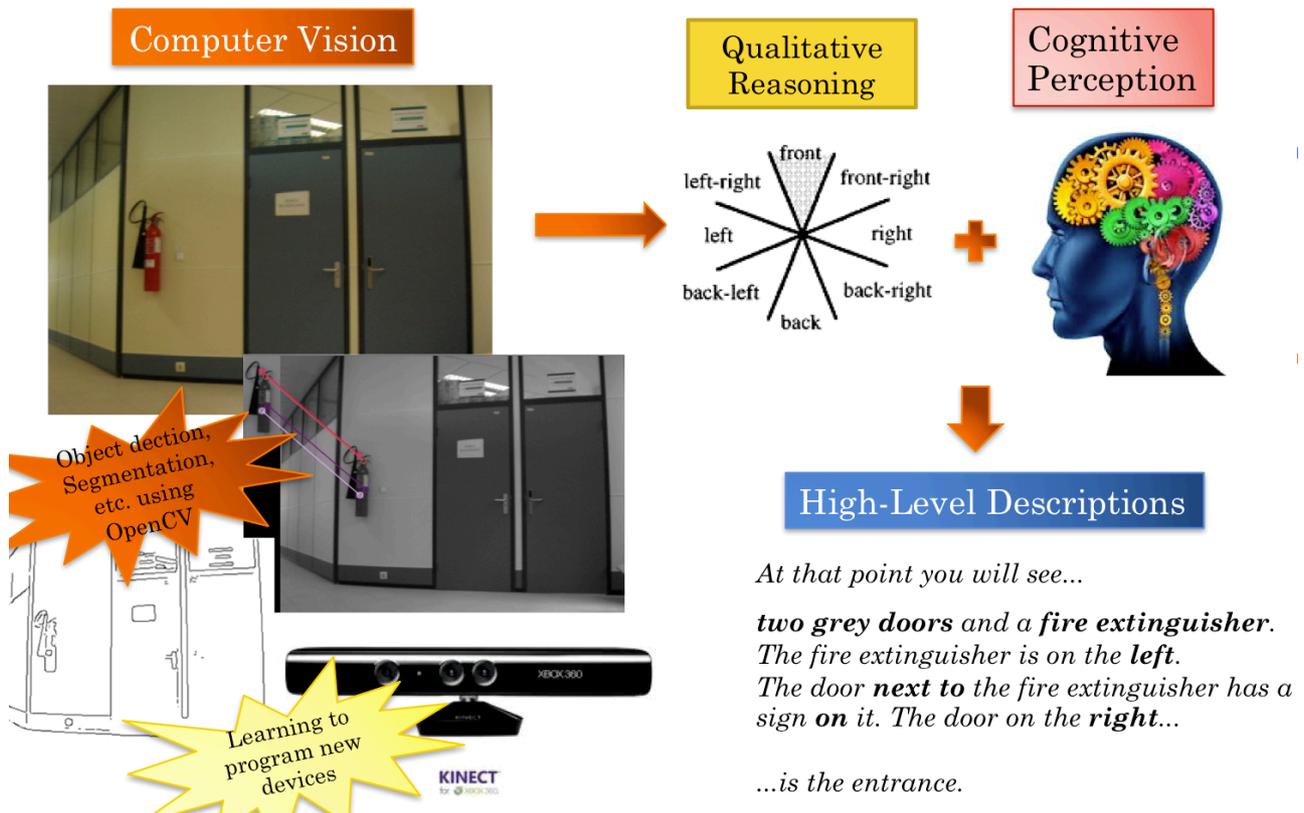
Title	Cognitive Analysis of Scenes: from Computer Vision to High-Level Descriptions for Reasoning and Generating Natural Language Descriptions
Kind of subject	4 ECTS Seminar
Academic Year	2012/2013, Summer Semester SoSe (02.04.13 - 06.07.13)
Degree	Master Informatik at Universität Bremen
Language	English
First Meeting	Thursday 4 th April at 9.00h in 0.41 Room, Cartesium, Groundfloor
Initial Timing	Thursdays 8.30h – 10.00h (Open to negotiation according to students' and teacher's availability)
Seminaroom	0.41 Groundfloor at Cartesium Building
Research group	Cognitive Systems (CoSy)
Teacher	Dr.-Ing. Zoe Falomir Llansola Cartesium Building, Office 3.54 Email: zfalomir@informatik.uni-bremen.de
Office attending hours:	Wednesday 15.00h – 17.00 h

1. Introduction

Digital images are fully integrated within modern daily life. By using digital cameras we can take photographs of trips and holidays, mobile phone cameras allow us to capture any daily scene, and we can use webcams in laptops to show where we are instantly across the network. The digital images and/or videos generated can be easily copied, deleted, edited, sent by email or multimedia messages, included in web pages, etc. and computer systems applications have been developed to provide all these possibilities. However, there is still no system capable of describing a digital image cognitively, that is, in a similar way to how human beings do it.

Psychological studies carried out on how people describe images [1–4] explain that people find the most relevant content in the images and use words to describe it. Usually nouns are used to refer to objects, adjectives to express properties of these objects and prepositions to express relationships between them. And these nouns, adjectives and prepositions are qualitative labels that extract knowledge from images and that communicate and explain image content.

As digital images represent visual data numerically, most image processing has been carried out applying mathematical techniques to obtain and describe image content. Therefore, using a computer vision system to extract information from space and interpreting it in a meaningful way as human beings can do, it is still a challenge.



In this seminar, this challenge is introduced from the perspectives of computer vision, qualitative modelling and cognitive science. And, after given the theoretical and practical fundamentals, students will have the opportunity to carry out a small project in which they will apply their knowledge and creativity. The projects will be developed in teams of two people and there will be class sessions for the students to explain their project to their classmates, for the students to help/be helped by their classmates, and for the students to evaluate/be evaluated by their classmates. This innovative teaching methodology and assessment was previously applied at Universitat Jaume I, Castellón, Spain, on master and bachelor students on Computer Science Engineering and successful results were obtained [5], such as conference papers, master thesis final projects, internships, etc.

The general skills that students will practice and improve in this seminar are:

1. Team working, problem solving, autonomous learning, and critical thinking, and communicating effectively orally and in writing;
2. Learning how to evaluate, how to give/receive critical feedback to/from their equals and how to use it to improve their knowledge and work.

IMPORTANT:

The most outstanding, collaborative and creative student(s) will be given the opportunity to apply for a rewarded training stay at CoSy research group in July 2013 for extending their project and collaborating in the project “Providing human-understandable qualitative and semantic reasoning” funded by the Zentrale Forschungsförderung der Universität Bremen.

Students registered in this seminar would be also given the opportunity to attend the SFB/TR 8 Colloquia on Spatial Cognition (<http://www.sfbtr8.uni-bremen.de/colloquium.htm>) organized by Cognitive Systems research group. And if they write a good summary of the talk relating it to the contents of the seminar, they can ask for a certificate of profitable attendance for their professional CVs.

2. General Objectives

1. Knowing the fundamentals on cognitive vision perception.
2. Understanding computer vision techniques and how and where to apply them.
3. Learning how to use computer vision libraries i.e. OpenCV, CImg, OpenSIFT, OpenSURF, etc.
4. Practising computer vision methods and techniques by means of computer vision libraries.
5. Understanding what is a Qualitative Representation, a Qualitative Model, and what Qualitative Spatial Reasoning involves.
6. Knowing what is a research paper and where and how to find them.
7. Learning the structure of a research paper: how to read it and how to write it.
8. Learning how to make an effective oral presentation.

3. Contents

This seminar is divided into two thematic units. In the first unit of the seminar, the fundamental contents will be explained and, in the second unit, the students will have the opportunity to carry out a small project related to one or more of the previous explained contents.

Unit I: Learning the Fundamentals

1. Motivation (1h)
 - 1.1. The Challenge
 - 1.2. The Contribution to the Research Community
2. Cognitive Vision Perception in Humans (1h)
 - 2.1. Fundamentals
 - 2.2. Mental Imagery
3. Computer Vision Methods (2h)
 - 3.1. Definition and Applications
 - 3.2. Image Formation
 - 3.3. Segmentation Methods
 - 3.4. Object Descriptors and Detectors
 - 3.5. People Pose Descriptors and Detectors
4. Qualitative Representation and Reasoning (QR) (2h)
 - 4.1. Qualitative Representations: Symbolic or Interval-based
 - 4.2. Qualitative Modelling
 - 4.2.1. Qualitative Models of Distance
 - 4.2.2. Qualitative Models of Shape
 - 4.2.3. Qualitative Models of Colour
 - 4.2.4. Qualitative Models of Topology
 - 4.2.5. Qualitative Models of Orientation
 - 4.3. Qualitative Reasoning
5. Applying QR to Analyse Scenes (2h)
 - 5.1. Interesting Approaches in the Literature
 - 5.2. Advantages and Disadvantages
 - 5.3. Generation of Semantic Information
 - 5.3.1. For Users: Natural Language
 - 5.3.2. For Users and Agents: Ontology
 - 5.4. Spatial Reasoning in Scenes

Unit II: Introduction to Research

1. Project Selection and Exposition to Classmates (2h x 2 sessions)

Example of Theoretical Project: State-of-the-art and critical review of the literature on an issue related to the seminar contents, i.e. *Semantic perception with RGB-D cameras/Android devices/etc.*

Example of Practical Project: Using computer vision libraries to detect events in digital images (i.e. objects, people faces, people poses, movement, etc.) and to extract some semantic information from them (i.e. locations, kind of environment, activity, etc.)

2. Project Development (2h x 5 attending sessions)

- 2.1. Group-Working Meeting 1
- 2.2. Group-Working Meeting 2
- 2.3. Group-Working Meeting 3
- 2.4. Group-Working Meeting 4
- 2.5. Group-Working Meeting 5

3. Project Results Presentation and Evaluation (2h x 2 sessions): Final Report

4. Methodology

The methodology applied to the first part of the seminar is mainly **expositive** but also **interactive/communicative** in order to discuss with students what they know about the contents but also to correct their misunderstandings and to orientate them.

The methodology applied to the second part of the seminar is based on **learning by projects** because it is a practical technique that helps students developing important skills, such as independent learning and problem solving, which are very important for a successful professional career.

A class session is dedicated for the students to explain to the rest of their classmates in an oral presentation: (a) their project and (b) the initial approach and means they plan to use to develop it. All the students are invited to make critical and constructive comments on the exposed work and then they are asked to fill in and submit an evaluation of the oral presentation made by each of their classmates. This evaluation is done using questionnaires that teachers provide to students which deal with the interest and difficulty of the problem, the quality of the oral presentation made, the viability of the initial solution proposed, etc.

After defining and presenting their projects, students have time to work in their projects and to discuss their progress in the group-working meetings. These group-working sessions are compulsory. Students can use them to share with the teacher and their classmates the challenges of their problems and to collaborate to find solutions. By means of these group-meetings, the teacher can follow students' work and orientate them towards successful results and can also observe the level of cooperation that each group has.

Finally, students will present and explain their work in a joint session of oral presentations. Both teacher and classmates can ask questions to evaluate the work presented following a predefined chart. Students must also hand out the final report on the project which should be written in English and has a research paper structure: Abstract, Introduction, Methods, Results, Discussion, Conclusion and Future Work. It should not be larger than 10 pages.

5. Assessment and Final Evaluation

Introducing a formative and qualifying assessment by applying a **co-evaluation or a method of evaluation between classmates** [5-6] is intended for students to develop general skills of critical and constructive thinking about their own work, very important also in professional practice.

Moreover, students' work and progress will be followed by means of the group-working meetings. The teacher will guide the students whenever needed (**formative and regulatory assessment**). Students can also use their colleagues and teacher feedback on their presentation and report in order to improve their oral and written skills. Students also can use the teacher's attending hours to solve doubts and make more progress on their projects.

The percentages estimated for the final evaluation of the seminar are:

- Up to **10%** of the final grade is obtained for attending the sessions;
- Up to **10%** of the final grade is obtained for helping their classmates by providing constructive criticism to their work;
- Up to **20%** of the final grade is obtained for handing out a good *Final Report* on the project;
- Up to **20%** of the final grade is obtained for giving a good talk explaining the initial objectives of the project and the final results of the project;
- Up to **40%** of the final grade is obtained taking into account the progress of the project, which will be evaluated by (1) the classmates, and (2) by the teacher. The teacher can ask the team members for a separate interview or *Fachgespräch* in order to evaluate the individual contribution of each member to the project.

Project extensions may add 10-20% to the final grade obtained, depending of the quality and quantity of the extension made.

7. Schedule

The schedule below is for guidance only and it may be adapted to the needs of the final group of students.

Session	Week		Contents
1	4 April	P	Seminar Presentation
2	8-12 April	A	1. Motivation
		R	2. Cognitive Vision Perception in Humans
3	15-19 April	T	3. Computer Vision Methods
4	22-26 April		4. Qualitative Representation and Reasoning (QR)
5	29-3 May	I	5. Applying QR to Analyse Scenes
6	6-10 May		Week for Deciding and Negotiating the Project
7	13-17 May	P	Project Presentation to Classmates
8	21-24 May	A	1st Group-Working Meeting
9	27-31 May	R	2nd Group-Working Meeting
10	3-7 June	T	3rd Group-Working Meeting
11	10-14 June		4th Group-Working Meeting
12	17-21 June	II	5th Group-Working Meeting
13	24-28 June		Week for Preparing Presentations
14	1-5 June		Final Project Presentations and Final Report

References

- [1] C. Jörgensen, Attributes of images in describing tasks, *Information Processing Management: An International Journal* 34 (2–3) (1998) 161–174.
- [2] M. Laine-Hernandez, S. Westman, Image semantics in the description and categorization of journalistic photographs. in: A. Grove, J. Stefl-Mabry (Eds.), *Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology*, vol. 43, 2006, pp. 1–25.
- [3] H. Greisdorf, B. O'Connor, Modelling what users see when they look at images: a cognitive viewpoint, *Journal of Documentation* 58 (1) (2002) 6–29.
- [4] X. Wang, P. Matsakis, L. Trick, B. Nonnecke, M. Veltman, A study on how humans describe relative positions of image objects, in: *Lecture Notes in Geoinformation and Cartography, Headway in Spatial Data Handling*, Springer, Berlin, Heidelberg, 2008, ISBN 978-3-540-68565-4, pp. 1–18. ISSN:1863-2246.
- [5] Falomir Z., Museros L., Escrig M. T., Mixing a teaching methodology based on 'learning by projects' with a 'co-evaluation' assessment for enhancing competences of students in Artificial Intelligence. In *Proc. of International Conference on Education, Research and Innovation (ICERI), organized by the International Association of Technology, Education and Development (IATED)*, ISBN: 978-84-615-3324-4, pp. 006187 – 006192, Madrid, 14-16th November 2011.
- [6] M. Valero-García, L. Díaz de Cerio Ripalda, Autoevaluación y co-evaluación: Estrategias para fomentar la evaluación continuada, *Actas del Congreso Español de Informática (CEDI)*, 2005. <http://epsc.upc.edu/projectes/usuarios/miguel.valero/> (Web access on March 2013).