Project in Computer Vision, Study period 4, 2016

This document contains information about the project course in Computer Vision and suggestions for projects.

- Register to the course by sending e-mails both to the supervisors of the project (see list below) and to the coordinator Magnus Oskarsson

- Produce a short, but informative plan for your project work, about 1 page to all supervisors and to the coordinator.

- Project work according to plan. Contact your supervisors and decide on dates for supervision.

- Produce a written report (3-10 pages). Submit the report as a pdf file to all supervisors and the coordinator

- Oral presentation and review of other project.

Conditions. You may work individually or in teams of two, depending on your preferences. To pass the course you have to present the project work in a written report, about 3-10 pages. Moreover, the results should be presented at a small seminar where all project groups participate. Also, each group will be required to act as opponent of the report an presentation of another groups project. All these elements are considered obligatory. The presentations below are fairly concise. More information and background material can be obtained from the respective project supervisors.

Supervisors:

| CO   | Carl Olsson               | calle@maths.lth.se |
| JF   | Johan Fredriksson         | johanf@maths.lth.se |
| KÅ   | Kalle Åström              | kalle@maths.lth.se  |
| SB   | Simon Burgess             | simon@maths.lth.se  |
| MO   | Magnus Oskarsson          | magnuso@maths.lth.se|
| EB   | Erik Bylow                | erik.bylow@math.lth.se |
| CS   | Cristian Sminchisescu     | cristian.sminchisescu@math.lth.se |
| ML   | Matilda Landgren          | matilda@maths.lth.se |
| VL   | Viktor Larsson            | viktorl@maths.lth.se |
| AR   | Anders Robertsson         | anders.robertsson@control.lth.se |
| MK   | Martin Karlsson           | martin.karlsson@control.lth.se |
Project:

1. **Create Dense 3D models using Depth Sensors.**
   New sensors like the Microsoft Kinect has opened for new possibilities to create dense 3D models. In Structure from Motion you typically get a set of sparse 3D points and it is hard to obtain a surface from these points. In this project you will see how these depth sensors can be used to create a 3D model with surface. In particular you will learn how a surface can be represented through a Signed Distance Function and how to integrate several views into one coherent model. You will implement your own algorithm in either C++ or Matlab.
   
   *Contact*: EB

2. **Mosaics and panoramic views.**
   If we have several images with the same camera center (i.e. taken with a stationary camera) then we know that corresponding points in the any pair of images are related by a homography (that is, a plane projective transformation). Through estimation of these homographies it is possible to glue together several images into a panoramic view.
   
   *Contact*: JF

3. **Computer Hearing.**
   By measuring the Time Difference of Arrival (TDOA) from a sound source to microphones with known position, it is possible to reconstruct the position of the sound. But what happens when both sound sources and microphones have unknown position?

   The aim of the project is to look at the calibration problem, i.e. finding the positions of microphones and/or sound sources from TDOA measurements.

   The project could for instance be used as a part in calibrating a rig of microphones, wifi-array sensor network or locating a sound source without prior calibration of the microphone array.

   Alt. The project could for instance be used for locating an elusive maracas-musician hiding in your office, using just some microphones connected to your sound card.

   *Contact*: SB, KÅ

4. **Solving systems of polynomial equations**
   Many problems in computer vision results in systems of polynomial equations. Examples of such problems are, (i) estimation of the essential matrix using five point correspondences, (i) pose estimation from four points. Recently there has been new results on how to solve such problems based on algebra, algebraic geometry and numerical linear algebra. The purpose of this project is to study such methods and use these techniques to implement a solver for one such problem.

   *Contact*: KÅ.

5. **Augmented reality.**
   An interesting application of computer vision is the so-called augmented reality. From a sequence of images, certain features are extracted and the camera motion is computed. The images of one or more virtual objects are then generated from the positions corresponding to the motion of the camera. These objects are subsequently projected into the original image sequence. In this way the (real) scene is augmented with an number of virtual objects in a believable way. The aim of the project is to perform this procedure for a few simple objects in a relatively short image sequence.

   *Contact*: JF.
6. **Relative positioning with a smartphone**
   Smartphones of today usually have at least two cameras and several inertial sensors. The aim of the project is to implement a prototype system for calculating relative positioning as the camera is faced downwards and moving along. The project could for instance be used to measure distances within a building or from the front of a car to the nearest crossing, to ensure that the car is parked at least 10m from the crossing.
   
   *Contact: KÅ.*

7. **Visual-Based Docking of a Mobile Robot**
   Automatic docking for mobile robots is getting increasingly important. For example, it might be necessary for it to regularly dock to recharge its batteries, or to attach to a tool or another robot, without the involvement of an operator. The challenge with this kind of maneuvers is that high accuracy is required, both in position and orientation. In this project, you will work with a mobile robot with four omnidirectional wheels. It is equipped with laser sensors for simultaneous localization and mapping, with an accuracy of approximately 3 cm. In order to improve the accuracy further, one or several cameras should be mounted on the robot for visual positioning. This could, for example, be combined with visual markers on the docking station.
   
   *Contact: MK,AR.*

8. **Object detection**
   The goal of the project is to explore methods to detect objects in images, for instance a person, a motorbike or an animal like a dog or a cat. Detection would mean, in this case, placing a bounding-box around the object of interest. Existing methods operate by obtaining a classifier using a large training set of positive and negative examples (bounding boxes of objects of the desired type, as well any other objects or structures, not from that class). At test time, the classifier is run exhaustively at different locations and scales in the image, and a bounding box is reported at the location and scale where this fires. The technology works well for some objects (e.g. faces) but fails for many others, like people in general poses or animals. The goal of the project is to explore different, accurate methods for detection that would be applicable to a more diverse set of object categories. The project is best suited to highly motivated students with a strong mathematics and computer vision background and excellent programming skills.
   
   *Contact: CS.*

9. **Prediction of Human Eye Movements and Recognition in Video**
   Recognizing interesting objects and events in images and video (people, objects, animals, as well as actions and interactions) is a central computer vision task. This problem is challenging and is often approached by annotating large volumes of images and video and, learning a detector for each visual category of interest, then scanning the image and video systematically, at multiple locations and scales, in order to find the target. This is computationally demanding and requires annotations which sometimes we do not have at large scale. In this project we will focus on understanding how humans visually process video and replicate that process. Specifically we will use datasets of human eye movement recordings in order to learn statistical models of how to look for objects and actions optimally. The project is best suited to highly motivated students with a strong mathematics and computer vision background, excellent programming skills and interest in doing inter-disciplinary research.
   
   *Contact: CS.*
10. **Human Pose Estimation in Images and Video**
   Analyzing video data involving humans in unconstrained environments is an open and currently active research problem, facing outstanding scientific and computational challenges. The proportions of the human body vary largely across individuals, due to gender, age, weight or race. Aside from this variability, any single human body has many degrees of freedom due to articulation and the individual limbs are deformable due to moving muscle and clothing. Finally, real-world events involve multiple interacting humans occluded by each other or by other objects, and the scene conditions may also vary due to camera motion or lighting changes. All these factors make appropriate models of human structure, motion and action difficult to construct and difficult to estimate from images. In this project we will focus on using large scale datasets of human motion capture as well as images in order to learn statistical models that would allow us to detect a person, recognize their body parts, and predict their 2d and 3d pose from images. The project is best suited to highly motivated students with a strong mathematics and computer vision background and excellent programming skills.
   
   **Contact:** CS.

11. **3D Modeling using Structured Light**
   3D modeling of objects having a uniform color can be difficult because it is hard to match points between different views. This can be remedied by projecting structured light on the object. In fact, such techniques are used to in the film industry to create 3D models of e.g. actors. The aim of this project is a method to create 3D models using a standard projector and two cameras.
   
   **Contact:** CO.

12. **Dense stereo matching using graph cuts**
   Dense stereo matching is the problem of finding corresponding pixels, and their depth, in stereo image pairs. The image of each feature is affected by disparity, i.e., the location of a feature in the left and right image is shifted, compared to one another, depending on its distance from the cameras. If we guess the depth of a feature, we can calculate the disparity of its image. This means that the stereo matching problem can be formulated as trying to assign a depth to each features, such that the observed disparities are explained as well as possible. Estimate solutions to this problem can be computed using graph cuts. The goal of the proposed project is to get to know about stereo matching and graph cuts, and to implement some stereo matching method.
   
   **Contact:** CO

13. **Photometric Stereo.**
   The process of inferring surface shape and reflectance properties by varying illumination is known as Photometric stereo. Using multiple images of an object taken from the same viewpoint under varying lighting conditions it is possible to create a dense depth map of the surface. The goal of this project is to create such a depth map of a simple object.
   
   **Contact:** CO

14. **Your own project proposal**
   You are welcome to submit a project proposal of your own. Before you start, your idea has to be approved, of course. We will then try to find the most suitable supervisor for your project.
   
   **Contact:** MO.