Project in Computer Vision, Study period 4

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 Carl Olsson (`carl.olsson@math.lth.se`).
- 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
KÅ	Kalle Åström	kalle@maths.lth.se
MO	Magnus Oskarsson	magnuso@maths.lth.se
MN	Mikael Nilsson	micken@maths.lth.se
MV	Marcus Valtonen Örnhag	marcus.valtonen_ornhag@math.lth.se
AH	Anders Heyden	heyden@maths.lth.se
NCO	Niels-Christian Overgaard	nco@maths.lth.se
KJB	Kenneth John Batstone	kenneth_john.batstone@math.lth.se
ML	Martin Larsson	martin.larsson@math.lth.se
GF	Gabrielle Flood	gabrielle.flood@math.lth.se
PP	Patrik Persson	patrik.persson@math.lth.se
DG	David Gillsjö	david.gillsjo@math.lth.se
IA	Ida Arvidsson	ida.arvidsson@math.lth.se
MA	Martin Ahrnbom	martin.ahrnbom@math.lth.se
AR	Anders Robertsson	anders.robertsson@control.lth.se

Project:

1. **Structure from Sound**.
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 locationg 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: KJB

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: MO

3. **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: ML

4. **Discrete total variation with application to stereo** Suppose we have observed a noisy grayscale image $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ denotes the image domain. The famous denoising model of Rudin, Osher and Fatemi proposes to recover the "clean" image u_* as the solution of a variational problem:
$$u_*(x,y) = \arg\min_u \lambda \iint_\Omega |\nabla u(x,y)|\, dxdy + \frac{1}{2}\iint_\Omega (f(x,y) - u(x,y))^2\, dxdy\,,$$
where $\lambda > 0$ is a tunable parameter. The term $\iint_\Omega |\nabla u(x,y)|\, dxdy$ is called the *total variation* of the function $u = u(x,y)$, and it is this object that is studied in this project. Since its introduction in 1992, total variation has played an increasingly important role as regularizing term in the ill-posed inverse problems associated with various image restoration problems such as denoising, deblurring and impainting. In the continuous setting total variation is isotropic, but in order to do actual computations one has to discretize the functional, usually on a square grid, and then the desired isotropy is easily lost in practice. The aim of the following project is to understand the recent discretization of total variation proposed by L. Condat (in 2017) that preserves isotropy to a high degree and apply the method to problems in stereo reconstruction.
Contact: NCO

5. **Room estimation using WIFI signal strenght and Received Time Travel**
Using Received Timed Travel between mobile phones and wifis it is posible to map the wifis and the mobile phone positions to a precision of approximately one meter. The signal strength of the wifi signal can then be used to estimate room geometry, i e where are the walls, using the fact that wifi signals are atenuated when passing through a wall.
Contact: ML

6. **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: MO.

7. **Navigation from planar images.**
   In some cases additional information can make 3D reconstruction and camera estimation easier. One such case is when the camera is viewing points that all lie in a plane. In this case we can estimate the camera geometry using homographies. This gives simple means for navigation.
   Contact: MO, MV.

8. **High dynamic range imaging.**
   Most cameras and screens can only capture and display images with low dynamic range. In this project we will focus on how to use several low dynamic range images with different exposures to construct and display a high dynamic range image.
   Contact: MO.

9. **Deep learning for segmentation and object recognition**
   Machine learning and particularly Convolutional Neural Networks has been used to solve many interesting computer vision problems. In this project we will design a Convolutional Neural Network that can distinguish between books and bookshelves in an image.
   Contact: GF

10. **Space Carving**
    Space carving is a 3D reconstruction method where the 3D shape is iteratively found by carving away 3D voxels form a large 3D block until the final shape corresponds to the images observing it. Each voxel (small 3D block) is projected into the cameras observing it. If the projections in the cameras do not match, then the voxel is removed form the large 3D block. This process continues until no more voxels can be removed, which hopefully results in the true shape of the 3D object. You will be given the images and the camera poses and the aim of the project is to find a dense 3D reconstruction of the observed scene.
    Contact: PP

11. **Vision-based pick-and-place with RGBD-camera**
    In an ongoing robotics project an industrial robot arm is to pick up an object in a gripper and very accurately place it at a certain position and orientation. The problem can be decomposed to first perform an object detection and localization in a certain pick-up place (no fixture), and e.g., from a determined bounding box move the robot arm to pick up the object. Even after a successful grasp the object may slide a bit when the gripper closes, so a second stage of pose estimation is needed to accurately determine how the object is held and how to move the robot to place it in the correct position. The project is to be tested on in a lab set up at RobotLab, LTH.
    Contact: AR

12. **Vision-based tracking of masonry joints in old brick walls**
    Sooner or later mortar in brick walls and chimneys start to crumble and has to be cut out and replaced, which is a tedious work to do manually. The project considers calibration between camera/robot/milling tool and visual tracking and planning of (masonry) joints for robotized milling. The project could be extended to milling along a more general surface.
    Contact: AR

13. **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.

14. **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

15. **Deep learning and SfM for depth estimation**.
    Today there are many Convolutional Neural Networks (CNNs) capable of predicting depth from a single RGB image. While impressive it lacks the mathematical precision of a Structure from Motion (SfM) solution. This project will study how such a CNN can be helped by SfM to improve the depth estimate.
    Contact: DG

16. **Room Layout estimation using Deep learning and SfM**.
    Indoor environments has a lot of planar surfaces with little or no texture. This is a challenge for Structure from Motion (SfM) systems as they primarily use image features. Convolutional Neural Networks (CNNs) on the other hand are good at semantic segmentation, which can be used to detect wall, floor and ceiling in images. The aim of this project is to couple semantic information from a Convolutional Neural Network with SfM to create a semantic room layout consisting of planes.
    Contact: DG

17. **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

18. **Deep learning for Prostate Cancer Grading.**
    This project concerns the automatic grading of prostate cancer based on microscopy images of stained biopsies. Such images are usually graded manually using four scales (benign, Gleason scores 3, 4 and 5). In the project we will study machine learning methods, whose parameters are estimated using images of previously graded images.
    Contact: IA

19. **Bat localization and verbalisation.**
    This project is in co-operation with the Max Planck Institute for Ornitology. The goal is to investigate and develop automatic algorithms that can calculate the relative positions of multiple microphones and using such data estimate the trajectories and the vocalization pattern of bat calls.
    Contact: KÅ.

20. **Rank Regularization and Optimization** Many ill-posed problems such as dynamic reconstruction can be regularized by requiring that system matrix is of low rank. In structure from motion this corresponds to the assumption that the observed shape can be written as a linear combination of low number of basis elements. However, optimizing objective functions containing rank constraints/penalties are often difficult since these are discontinuous and non-convex. In this project we will explore ways of relaxing the rank function that allows for easy optimization.
    Contact: MV, CO.

21. **Triangulation with limited baselines** Triangulation, the process of computing a 3D point given corresponding 2D points and camera matrices, is a fundamental problem in computer vision. When the cameras are relatively close to each other, compared to the distance to the points, this problem naturally becomes difficult as small errors in the 2D positions results in a large variation in 3D space. This project will experiment with methods for improving triangulation in a real scene captured by three cameras on a line, given camera matrices, images, and approximate 2D positions. Accurate measurements of some 3D points (that weren't used in calibration) exist, allowing accuracy evaluations. Experiments can include different ways of using all three cameras for triangulation, subpixel strategies to reduce 2D errors and using assumptions on the 3D world to limit the search space.
    Contact: MA, MN.

22. **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: CO.