Student projects

Vejledt af Søren Olsen

Block 3, 2009

Småprojekter 7.5 ECTS:

1. Visuel Adgangskontrol
2. Rigid Structure from Motion
3. Stereo Tracking

Bachelorprojekter 15 ECTS:

1. Genkendelse af Nummerplader

Scecialeprojekter 30 ECTS:

1. None this block

Alle projekter indledes ved blokstart og forudsætter at den/de studerende arbejder mindst halvtids (fuldtid for specialer). Projekterne afleveres ved afslutning af den akutelle blok (for 7.5 ECTS-projekter), eller den næste blok (for større projekter). Det er ikke acceptabelt at den/de studerede samtidigt har erhvervsarbejde.

Vejledning gives gennem billedgruppens fælles projektvejledning, som den/de studerende forventes at tilmelde sig og følge.

Automatic License Plate Recognition - 15 ECTS

Automatic Licence Plate Recognition (LPR) is essential in many existing and future applications such as car park admittance systems, automatic speed control, border control, control and fining of illegally parked cars etc. Several commercial systems are available. This student project aims at developing a simple LPR-system, making experiments to validate the accuracy and robustness of the system, and suggesting possible future improvements.

In a large class of approaches to LPR images are processes one by one. For each image the possible candidate positions for license plates are found. Depending on the application assumptions of image resolution, camera pose, controlled illumination etc may by used. For each candidate LP the text/number area is segmented out. In some application this step is preceded by an geometrical rectification to allow acute camera angles. In the segmenting process prior knowledge on the LP coloring, text composition etc. is used. Next the image areas of the individual letters and digits are found. Often these sub-images are then binarized. In this and the preceding steps many methods apply a set of heuristics to eliminate false detections and to avoid rejection of true candidates. If the student(s) have no background in Image Processing the difficulty of the tasks above may be lessened by assuming well illuminated high resolution images showing only one LP (of only one composition type) at an approximately known position.

Having extracted single digit/letter binary images the next step is classification. A huge amount of classification methods exist. If the student(s) have no background in Pattern Recognition a simple method will be suggested. If focus is on the classification part several methods may be implemented and compared. When all letters/digits are classified a post-processing step often is applied. In this illegal combinations of letters may be detected.

To make experiments with the developed system a large number of appropriate images are required. At DIKU a database of images taken by previous students have been build up. Comming students are expected to contribute to the base.

The aim of the project is to develop, describe, and evaluate an automatic LPR-system. The chosen focus of the project (eg. on the image processing part or the pattern recognition part) should be made clear. A thorough discussion of all important decisions are expected. An evaluation showing the capabilities and limitations of the system should be made. Based on these, the possible improvements of the system should be discussed.

It is a prerequisite that the student(s) have passed the course *Introduktion til digital billedbehandling* or the course *Mønstergenkendelse*. Knowledge within both courses will be a significant advantage.

Projektspecifikke læringsmål:

1. To develop, describe, apply, and evaluate methods for Digital Image processing, Pattern Recognition and Classification, through a practical case-study of Licens Plate Recognition.

Rigid Structure-from-Motion

Probably, one of the most successful methods for 3D-Computer Vision is the factorization approach of Tomasi and Kanade [1] to Rigid Structure from Motion (RSfM). Among the most celebrated applications are fusion of real and animated videos. In this project the original approach [1] as well an upgrade to metric representation, as described by Quan [2], should be implemented and tested.

The original approach to RSfM is extremely simple both with respect to its conceptual idea and with respect to implementation. One draw-back is the fundamental assumption of an affine camera. However it is not clear how sensitive this approximation to a real camera is. A possible element of the project would be to clarify this sensitivity. Another serious draw-back is the assumption that all points should be visible in all frames. Usually this assumption is violated. Several approaches exist for handling partial data (among one see Olsen and Bartoli [3]). A possible element of the project could be to compare and evaluate several such approaches. Also robustness to outliers are important in any practical application. A relevant element of the project may be to choose, apply and evaluate a robust estimation method. Finally, upgrading to metric representation is of obvious importance. The standard method described by Quan [2] is simple, but requires a non-linear optimization step. In 3D-Computer Vision such a final improvement step has shown of significant importance in many situations. A experimental evaluation of the importance of a such non-linear improvement could be a relevant element of the project.

The student(s) should clearly describe the chosen focus of the project. The student(s) should argue for and describe the method(s) that are chosen for implementation and should evaluate the implemented method(s) on a set of video sequences. The project includes acquisition of a sufficient number of such sequences. Based on the experiments the performance and limitations of the implemented method(s) should be discussed.

It is a prerequisite that the student(s) have passed the course *Introduktion til digital billedbehandling*. It will be an advantage if the course *3D Computer Vision* has been passed.

[1] C. Tomasi and T. Kanade: *Shape and motion from image streams under orthography: A factorization method.* International Journal of Computer Vision; 9(2):137–154, 1992.
[2] L. Quan: *Self-calibration of an Affine Camera from Multiple Views*. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(10): 1179–1185, 2000
[3] S. Olsen, A. Bartoli: *Implicit Non-Rigid Structure-from-Motion with Priors*. Journal of Mathematical Imaging and Vision, Vol. 31, no.2-3, 233–244, 2008

Projektspecifikke læringsmål:

1. To develop, describe, apply, and evaluate one or several methods for Rigid Structure-from-Motion.

Stereo Tracking

Tracking is a basic task in Computer Vision. If several cameras are used, it becomes possible to track the 3D-shape. There exist a number of commercial systems, eg. VICON. Often such system for *Motion Capture* use many (> 5-10) cameras, special illumination, reflective markers etc. These systems are bulky (not for home usage) and may be very expensive.

This project is aimed at developing a cheap 3D-tracker using a off-the-shelf stereo camera. Using only two cameras severely limits the kind of motion ensuring that the moving shape/body is visible in both images. As example only the front and not the back of the body may be visible. The probability of self-occlusion is large. However, for some applications this limitation may not be serious. In this case the two video streams may be processed to extract (images of) easily visible markers attached to the moving body. By matching the time varying set of image landmarks and by triangulation the 3D-position of the markers can be found. By combining with tracking within the left camera and within the right camera a 3D-tracking can be performed.

The focus of this project is to make a robust real-time cheap Motion Capture system using a Bumblebee stereo camera from Point Grey. The task is to track a person moving/performing about 2 meters in front of the camera. To ease the image processing step it may be assumed that the markers are light emitting diodes (LEDs) attached rigidly to the persons body and limbs. The amount of illumination may be controlled or assumed limited.

The student(s) should argue for and describe the method(s) that are chosen for implementation and should evaluate the implemented method(s) on a set of stereo video sequences. The project includes acquisition of a sufficient number of such sequences. The expected as well as the obtained accuracy of the 3D-tracking should be described and evaluated. Similarly experiments on robustness and sensitivity to eg. body rotations should be evaluated. Based on the experiments the performance and limitations of the implemented method(s) should be discussed and possible future improvements discussed.

It is a prerequisite that the student(s) have passed the course *Introduktion til digital billedbehandling*. It will be an advantage if the course *3D Computer Vision* has been passed.

[1] A. Waxman, J.H. Duncan: *Binocular Image Flows: Steps Towards Stereo-Motion Fusion*. IEEE transaction on PAMI, vol. 8, no. 6, pp. 715–729, 1986.

Projektspecifikke læringsmål:

1. To develop, describe, apply, and evaluate one or several methods for 3D tracking involving a Bumblebee stereo camera.

Visuel Adgangskontrol

Et tænkt institut er plaget af studerende, der efter ordinær lukketid ønsker adgang selv om de har glemt deres nøgle/kort. Instituttet beslutter at installere et visuelt system, der på basis af aftalt fingerkode-sprog kan beslutte om døren skal åbnes. Filosofien er at de studerende altid har højre hånd med sig. Et kamera tænkes opsat så det overvåger mat lys (lysende) plan flade på ca. 25 gange 25 cm. For at simplificere overvågningssystemet monteres en knap (betjent af venstre hånd) der ved aktivering bevirker optagelse af et billede af højre hånd, som holdes i billedfeltet over den lysende flade. Billedet analyseres og det afgøres om et korrekt fingertegn er angivet. I dette tilfælde tændes en grøn lampe, hvorefter studenten kan vise næste fingertegn. I modsat fald tændes en rød lampe kortvarigt til indikation af fejlaflæsning. Hvis en korrekt sekvens af fingertegn genkendes blinkes med den grønne lampe og døren åbnes.

I opgaven skal et fingerkodealfabet samt ordlængden af en korrekt fingerkode fastlægges. Til hjælp kan det oplyses at instituttet har ca. 500 studerende, der hver skal have et personlig kodeord. For at sikre mod uautoriseret adgang bør antallet af mulige kodeord være betydeligt større. De enkelte fingertegn bør vælges således at den visuelle afkodning bliver pålidelig og sådan at sekvensen af tegn ikke bliver urimelig lang.

Der skal udvikles, dokumenteres, implementeres og afprøves (eksperimentelt) et program, der givet et billede kan afgøre om billedet viser et korrekt fingertegn (og i så fald hvilket) eller ej. Dokumentationen for programmet bør redegøre for hvor pålideligt programmet er både mht. korrekt klassifikation af de enkelte tegn samt mht. detektion af ukorrekte tegn. Konsekvenserne af klassifikationssikkerheden for adgangskontrolsystemet skal diskuteres. Dokumentationen af arbejdet bør omfatte en diskussion af og begrundelse for valg af analyse- og klassifikationsmetode. Det udviklede program skal vedlægges som bilag.

Til det eksperimentelle arbejde vil laboratorieafprøvning være tilstrækkelig. Her vil en overheadprojektor samt et kamera monteret på stativ kunne udgøre forsøgsopstillingen. Adgang til DIKUs billedlaboratorium fås efter nærmere aftale. Optagelse af billeder af tilstrækkelig mange testpersoners forsøg på gengivelse af tilstrækkelig mange forskellige koder er en del af opgaven.

Opgaven kan løses individuelt eller i grupper på op til 3 personer. Det forventes at opgaveløserne arbejder pa omtrentlig halvtid med opgaven. Deltagere skal, som del af vejledningen, følge billedgruppens vejledningskursus.

Forudsætninger for løsning af opgaven vil normalt være at kurset *Introduktion til digital billedbehandling* samt kurset *Mønstergenkendelse* eller tilsvarende kurser er bestået.

Projektspecifikke læringsmål:

1. At udvikle, beskrive, anvende og evaluere simple metoder til digital billedbehandling gennem en praktisk „case“.
2. At udvikle, beskrive, anvende og evaluere simple metoder til mønstergenkendelse og klassifikation gennem en praktisk „case“.