Seminar contents

- this seminar includes a selection of the most relevant papers in the field of computer vision in the past 20 years
- papers have been selected to cover different aspects of the topic:
  - denoising and filtering
  - optical flow
  - camera calibration
  - keypoint detection and description
  - image matching and classification
  - segmentation
  - tracking
  - object/face detection
Goals

- You are going to learn:
  - about relevant works in the field of Computer Vision
  - how to read and understand a scientific article
  - how to write a scientific report
  - how to give a talk to an audience, and deal with related questions afterwards
Seminar Schedule

- 4 sessions, 1 every Wednesday, 4pm-6pm
  - November 26
  - December 3
  - December 10
  - December 17
- Each session will include up to three presentations (first come, first served)
- Seminarraum 03.13.010 (except for December 17 in Laborraum 03.13.008)
Presentation

- Each presentation is 25 minutes + 15 minutes for Q&A
- Based on slides (Powerpoints, LaTeX, ..), see website for templates
- The presentation should cover all relevant aspects of the paper
  - Introduction and state of the art
  - Main contribution(s)
  - Experimental results
  - Discussion, summary and future work
- The presentation should be self-contained
- All students are expected to attend all presentations and interact during Q&A (this will influence your final mark)
The report should summarize the paper in the way it has been presented during the talk.

Language: English

Max 8 pages

Template on course website

Once ready, send the report to supervisor, within one week from the day of the presentation.
Evaluation criteria

- Quality of presentation (both regarding slides and speech)
- Quality of the report
- Comprehension of the scientific contents of the presented work
- Interaction and participation during the other talks
Goal is to remove image noise while preserving edges

Idea
- Filter based on *spatial*- and *color*-neighbor relations
- Applicable for N-D images
- Efficient Gaussian kernel

Application
- Data reduction for image transmission and description
- Texture and noise removal
Goal is to find the affine registration between two frames

Idea
- Use image gradients
- Use multi-resolution pyramid to converge in a stable fashion

Application
- Optical Flow
- Stereo Image matching
- Depth map generation

Implemented in OpenCV
Goal is to find intrinsic camera parameters (lens distortion, focal length, principal point)

Idea
- Use only one plane seen from various views
- Refine solution by minimizing projection error

Application
- Structure from motion
- Stereo Image matching

- Goal is to recognize objects in images under various alterations of appearance
- Idea
  - Use models consisting of multiple connected parts
  - Make parts deformable and encode the configuration in a *latent variable*
- Application
  - Contextual scene understanding
  - Surveillance
Goal is to recognize objects in images

Idea
- Foreground segmentation and object classification are tightly coupled
- Learn object shape and use Generalized Hough Transform for segmentation

Application
- Contextual scene understanding
- Surveillance

- SIFT: Local Image Descriptor
  - Highly robust against viewpoint changes

- Algorithmic pipeline
  - Find Interest points
  - Do statistics on local gradient directions
  - Accumulate information

- Application areas
  - Multi-view matching
  - Object recognition

- HOG: Dense grid feature descriptor
- Idea
  - Local object appearance can be described by image gradient
    - Divide image into cells
    - Analyze locally
    - Combine information
- Application areas
  - Object detection
  - Here: Pedestrian detection
D. Comaniciu, P. Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis”, TPAMI 2002

- Non-parametric technique for detecting multiple modes in density functions

- Allows analysis of feature spaces for many tasks including
  - Segmentation
  - Smoothing
  - Clustering
  - Tracking
  - …

- Fast image retrieval and recognition technique
- Based on indexing of local features via quantization and hierarchical clustering
- Robust towards occlusion and background clutter
P. Viola, M. Jones, “Rapid object detection using a boosted cascade of simple features”, CVPR 2001

- Face detection using Haar features
- The most discriminative features for faces are automatically selected by means of AdaBoost classifier
- Efficient since computes the features via incremental schemes (Integral Images)