PhD course on Autonomous Systems
Session 2, Overview:
From Sensing to Artificial Perception
Michael Felsberg
Sensing and Perception in WASP

• autonomous systems with embodiment (robotics) and pure software systems, e.g. for decision support

http://ieiri-lab.jp/it/2014/12/hexh2o.html
https://uk.pinterest.com/pin/299559812692445652/
Case Study: 125 Anniversary
Bertha Mannheim - Pforzheim

Measurements and Interaction

- Both scenarios require measurements and interaction with the environment
- Artificial perception: software system to close the semantic gap between raw sensor output and symbolic representations
Autonomous Driving is Easy?

• According to some experts: Google and Apple have failed with their autonomous car projects [New York Times 9/9, Bloomberg 9/12]

• Tesla’s fatal accident could have been avoided by Lidar instead of computer vision-based perception (Mobileye)? [New York Times 7/12, ArsTechnica 9/16]

https://static01.nyt.com/newsgraphics/2016/07/01/tesla-accident/10c347b26e2d2fb936647182b6b92923cb914729/crash-720.png
Discussion: Regression for Driving

http://deepdriving.cs.princeton.edu/
Overview

• Input from sensors

part I: types and modeling of sensors

http://www.embeddedworks.net/gps-satcom.html

part II


http://www.raymarine.com/view/?id=9860

• Output of relevant information
  • quantitative
  • qualitative

part 2

part 3
Quantitative and Qualitative Streams

sensor data

quantitative stream

qualitative stream

relevant information
Types and Modeling of Sensors

- Often focus on *hardware* (specs, advertisements, discussions)
- “Only” measurement/signal generator
  - *direct* measurements: output from hardware is relevant information itself, e.g. temperature
  - *smart* sensors: software (firmware) on the device produces relevant information, e.g. distance
  - *indirect* measurements: output from hardware needs to be processed to extract relevant information hidden in the signal, e.g. image
Example: Image Data
Example: Image Data
Modalities of Sensing

• Passive sensors rely on influx
  • cameras (incl. thermal and PTZ)
  • microphones

• Active sensors analyze previously sent signal
  • radar and lidar
  • time-of-flight & active illumination (Kinect)

• Sensing by physical interaction
  • tactile and force sensors

• Sensors that rely on physical infrastructure
  • GPS
  • RF and marker-based positioning systems (Vicon)
Camera Types

- monochromatic (grey-scale, IR)
- color and multi-spectral
- photometric sensors (e.g. thermal IR)
- plenoptic (light-field)
- catadioptric (with mirror)
- active vision / articulated cameras (e.g. PTZ)
- event-based cameras

Geometric Camera Models

- affine camera
- perspective camera
- fish-eye camera
- general geometric camera, e.g. catadioptric, rolling shutter
Sound and Air-Pressure Sensors

• Well-known: microphone
  • produces electrical signal from air pressure variations
  • a signal is a function (physical quantity) that
    • varies with time, space, or any other independent variable
    • carries some information.
  • sensitivity may vary depending on orientation, e.g. cardioid

• Ultrasonic sensors

• Barometer

Images courtesy I. Skog

By Lucasbosch - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=31849701

By Nicoguar0 - Own work, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=50230608
Radio-based Sensing

- RFID tags
- GPS
- Pseudolite
- WIFI

Images courtesy I. Skog
Tactile and Force Sensors

• Tactile sensors measure information arising from physical interaction similar to cutaneous touch

• Force sensors produce signals from pressure or force
Acceleration and Rotation Sensors

- gyroscope
- accelerometer
- magnetometer (magnetic field)

Images courtesy I. Skog
Radar and Lidar

• RAdio Detection And Ranging
  • radio or microwaves
  • emitting and receiving antenna

• LIght Detection And Ranging
  • laser light
  • scanner and optics
  • photodetector
  • position/navigation system
Active Cameras

- Time-of-flight (PMD, Photonic Mixer Device)

- Active illumination (Kinect2)
Combination of Sensors

GoPro sequence "RC-car"
Choice and Placement of Sensors

- Scene dependencies
  - update frequency (frame rate)
  - spatial resolution and field of view
  - accuracy and spectrum/modality
  - multi-sensor: synchronization
  - conditions of environment, e.g. light

- System dependencies
  - energy
  - weight and size
  - bandwidth
  - storage
  - computational resources
Choice and Placement of Sensors

• Properties (calibration)
  • delay
  • geometric alignment (6D pose, reference frame)
  • distortion model
  • auto-calibration

• Relations (triangulation)
  • stereo (sound and images)
  • source to receiver (laser triangulation and Kinect1)
  • moving sensor (motion stereo, doppler)

• Example: CENTAUBO head
Choice and Placement of Sensors

- Properties (calibration)
- Delay
- Geometric alignment (6D pose, reference frame)
- Distortion model
- Auto-calibration

Relations (triangulation)
- Stereo (sound and images)
- Source to receiver (laser triangulation and Kinect)
- Moving sensor (motion stereo, doppler)

Example: CENTAUCRO head
Choice and Placement of Sensors

• Example: CENTAURO head
Choice and Placement of Sensors

• Example: CENTAURO head
Observability

1. internal modalities, e.g. fault analysis
2. internal states, e.g. rank of observability matrix for linear system
3. external world – occlusion and multiple view, e.g. next best view planning
Perception, Quantitative Stream

- Often considered *the engineering* approach
  - estimation theory
  - probabilistic modeling
  - geometry
  - energy minimization
- Often signal-based
- Symbols as discrete signals
Filtering and Data Association

- measured data
  - contains outliers or missing data (dropouts)
  - noisy
  - interfering sources

- counter measure: (Bayesian) filtering
  - prediction part: missing data
  - (extended/unscented) Kalman filter: noise
  - particle filters and other MC methods: outliers
Filtering and Data Association

• counter measure to interfering sources: data association
  • multiple hypothesis tracker (MHT)
  • probability hypothesis density (PHD) filter

Mean: 0 (straight ahead)

Steering: -1 (left)
Steering: 1 (right)
Filtering and Data Association

- counter measure to interfering sources: data association
  - multiple hypothesis tracker (MHT)
  - probability hypothesis density (PHD) filter
  - robust estimators (non-convex)
  - MC-methods, e.g. random sample consensus (RANSAC) for perspective-n-point (PnP)

- probabilistic approaches
  - maximum likelihood (ML)
  - (sequential) maximum a-posteriori (MAP)
Localization / Pose Estimation and Mapping

- Landmarks
- Relative pose estimation problem
- Hand-eye calibration problem
- (Visual) odometry problem
- Mapping and SLAM (incl. marginalization)
Sensor Fusion, Multi-modal Input and Prior Information

- Batch and sequential estimation
  - bias
  - covariance and linear estimation
  - optimization (Taylor series, unscented, MC)
- example: sequential SfM, Bundle Adjustment
Sensor Fusion, Multi-modal Input and Prior Information

- Sensor and information fusion
  - early fusion and safe fusion
  - late fusion
  - latent variables
  - example: PC registration
Sensor Fusion, Multi-modal Input and Prior Information

• Feedback and prior information
  • non-informative prior, conjugate prior
  • MAP
  • MC initialization
  • covariance feedback
  • ex: adaptive filtering / steerable filters

By Thomas Schultz, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1201206
Multimodal SLAM and Situation Awareness

• Cross-modality association
  • links to qualitative stream
  • datamining

• Large-scale industrial initiatives
  • maps (Nokia, Apple, Google)
  • Daimler car, Google car

<Networked sensor systems of the S 500 INTELLIGENT DRIVE research vehicle>
Highly autonomous driving based on near-production sensor systems

Cross-Modality Association

Autoliv Instrumented Vehicle
Perception, Qualitative Stream

- Often considered *the* AI/CS approach
  - efficient heuristics
  - procedural methods
  - structural approaches
  - biological motivation
  - dominating: deep learning
- Often: symbols
- Signals as interpolated discrete values
- Note name clashes:
  - background
  - tracking
(Visual) Object Detection and Tracking

• Background modelling
  • anomaly detection
  • probabilistic background model
  • saliency operators

• Object detection
  • generative vs. discriminative
    • generative model
    • distance function
    • discriminative function
  • ex: part-based model
  • optimization vs. learning
(Visual) Object Detection and Tracking

- Region tracking
  - features and landmarks
  - blobs
  - bounding boxes
  - state-space: displacement, scale, rotation
  - perceptual aliasing

- Object tracking
  - adaptivity vs. drift
  - color representations (e.g. color names)
  - descriptors (e.g. HoG)
  - articulation (extended state-space)
(Visual) Recognition / Classification of Objects, Actions, and Poses

- Recognition pipelines
  - Marr’s vision (staged sketches)
  - hierarchical feature extraction
  - classifiers
  - pooling and non-max suppression
(Visual) Recognition / Classification of Objects, Actions, and Poses

• Feature selection and representation
  • biological parallels, Hubel&Wiesel
  • Gabor pyramids and visual gist
  • scale-space and interest points

• PCA, ICA, and curse of dimensionality
• population codes, SIFT, HOG, BoW, …
(Visual) Recognition / Classification of Objects, Actions, and Poses

- Learning approaches
  - classification vs. regression
  - online and incremental vs. offline and batch
  - Hebb learning and PCA
  - kernel-trick and support vector machines (SVM)
  - decision trees
  - ensemble methods; boosting and random forests (RF)
  - activation functions and rectifier / rectified linear unit
  - back-propagation and deep learning
Online Regression Learning
Conclusion and Summary

• Artificial perception: software system to close the semantic gap between raw sensor output and symbolic representations

• One of the most underestimated problems in engineering and computer science

• The Summer Vision Project, MIT CSAIL 1966: “The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system.”

• The project continues … good luck!