Simultaneous Localization and Mapping

Visual Loop-Closure Detection

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

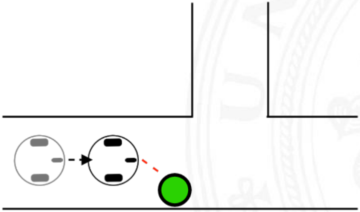
Technical Aspects of Multimodal Systems

04. December 2017

Outline

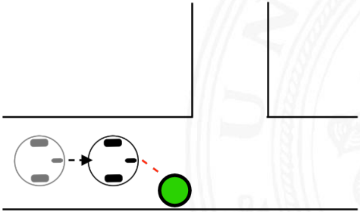
Introduction

- Autonomous robots should be able to navigate in previously unknown environments, without relying on pre-built maps
- Robots must create a map of their environment, while exploring it
- Robots' extrinsic sensors (e.g. laser rangefinders or cameras) capture the environment to build the map
- They either capture landmark locations (feature based) or the environment is divided into grid cells which can either be occupied or not



Introduction

- Always uncertainty in the measurements
- Robot must observe the environment multiple times to find true values by statistical means
- Sensors that track the robot's movements are also noisy -> mapping is difficult
- Without accurate map -> localization is difficult
- SLAM is a 'chicken and egg problem'

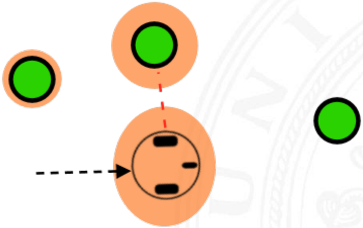


Introduction

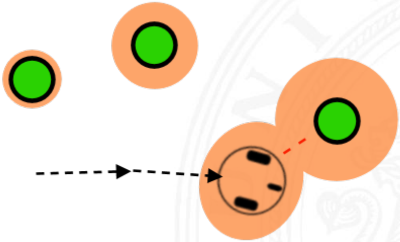
- While robot moves through the environment, uncertainty of mapped landmarks and robot's pose grows
- Produced map is highly inaccurate

Introduction

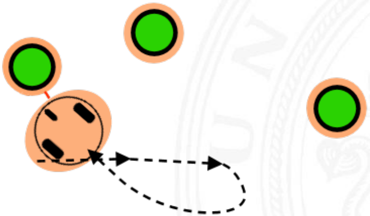
- While robot moves through the environment, uncertainty of mapped landmarks and robot's pose grows
- Produced map is highly inaccurate



- While robot moves through the environment, uncertainty of mapped landmarks and robot's pose grows
- Produced map is highly inaccurate

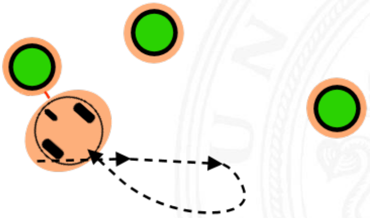


- When robot returns to a known landmark, its own pose can be determined more accurately
- Positions of landmarks can then be updated
- Traversing environment multiple times increases the accuracy of the map



Introduction

- Goal of Full SLAM is to generate map (e.g. locations of all landmarks) and reconstruct the entire path the robot took
- Online SLAM also for generating map but does not keep track of all previous positions
- Only current pose is to be estimated



Introduction

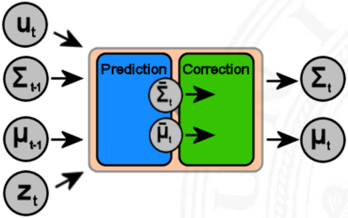
- Several algorithms and their variations that attempt to solve SLAM problem
- They compute the most probable locations of landmarks m_i and robot's pose x_t or path $\{x_0, x_1, ..., x_t\}$
- Estimation is based on readings of robot's movement (odometry or velocity) u_t, observations of the environment z_t and previous poses of the robot

Extended Kalman Filter

- Extended Kalman Filter (for online SLAM) computes probability distributions for each landmark location and robot's pose
- All distributions are Gaussian. Output is a mean value vector and a covariance matrix for the current point in time
- Input is mean value vector and covariance matrix for last time step + executed motion and current observation

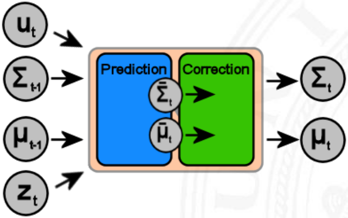
Extended Kalman Filter

- Works in two steps -> prediction and correction step
- Prediction step computes distributions based on the robot's movement
- Motion model is applied to previous location



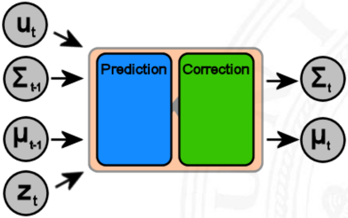
Extended Kalman Filter

- Motion model is matrix for translations and rotations
- Resulting function is locally linearized to maintain Gaussian distributions
- Normally distributed noise is added



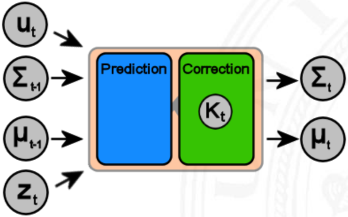
Extended Kalman Filter

- Correction step -> predicted observation is computed by applying observation model
- Again linearized to maintain Gaussian distributions
- Again normally distributed noise is added



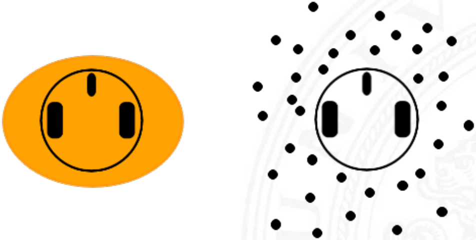
Extended Kalman Filter

- Kalman Gain is computed, weighting factor between predicted observation and actual observation
- Based on uncertainty of motion and sensor observations. If observation is very uncertain, prediction is weighted more and vice versa



Rao-Blackwellized Particle Filters

- Particle filters are not restricted to paramentric distributions, like Gaussians
- Distributions are represented as sets of samples (particles)



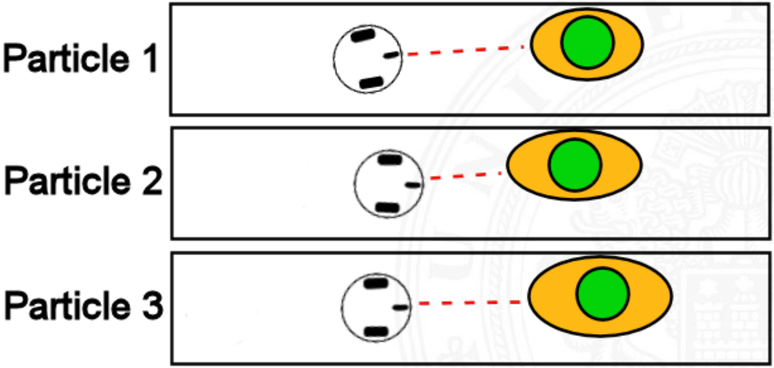
- Each sample represents a possible current pose of the robot and all landmark location estimates as Gaussian Distributions

Rao-Blackwellized Particle Filters

- Particle filter computes robot's pose at current time step by applying motion model based on u_t.
- Particles are 'moved' and noise is added (more particles)

Rao-Blackwellized Particle Filters

- For every particle, predicted landmark locations are computed according to observation model
- Each particle is weighted according to how good its estimates of the landmark locations are by taking into account the current sensor observations z_t



Rao-Blackwellized Particle Filters

- By using one Extended Kalman Filter on each landmark of every particle, map is updated
- Particles are resampled by selecting fixed number of particles with their weight being proportional to probability of selection

Loop-Closure Problem

- Odometry is unreliable for detecting loops in the robot's path
- System must rely on recognizing previsouly seen parts of the environment

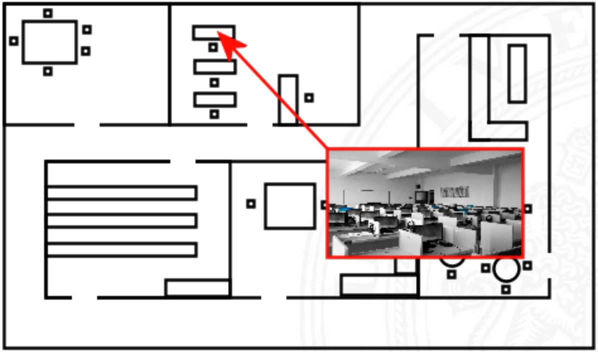
Visual Feature Based Place Detection

- Camera images are potentially feature rich
- Places can be recognized by their visual features

Example of geometric verification [3]

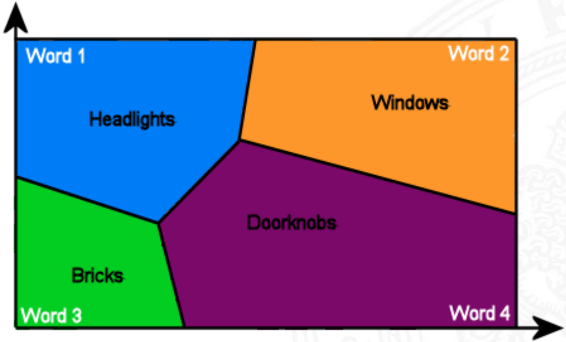
Visual Loop-Closure Detection

- System builds a topological map of places
- Previously unseen places are added, known places are recognized as seen



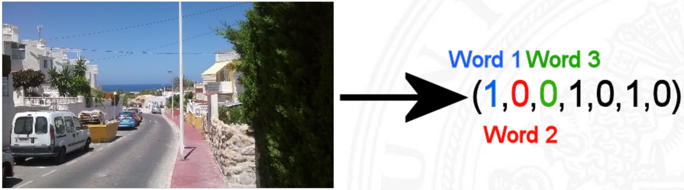
Bag-of-Visual-Words Model

- Many approaches use the Bag-of-Visual-Words model
- Features are extracted from a large set of images and clustered for dimensionaility reduction (e.g. K-Means Clustering)
- Clustered features are called visual words of visual vocabulary



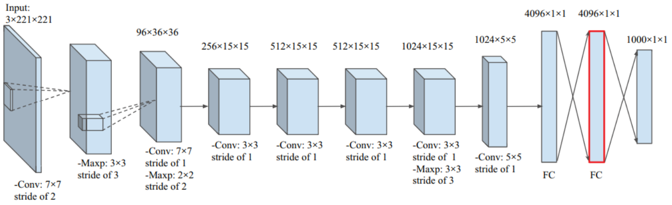
Bag-of-Visual-Words Model

- Images can be represented as word vectors
- Words are not independent -> cooccurrence
- Taking cooccurrence into account leads to better place recognition



CNN for Place Recognition

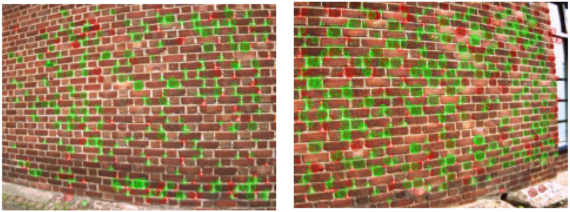
- Recent approach with convolutional neural networks for place recognition [4]
- Image classification network was trained on imagenet dataset
- Feature vectors before fully connected layers were further processed
- Principal component analysis for dimensionality reduction
- Outperformed 'hand-crafted' feature based method



Architecture for Accurate model of OverFeat [4]

Problems with Visual Feature Based Place Recognition

- Features must be invariant towards illumination, scale etc.
- Dynamic objects blocking view
- Features are specifically designed to fit certain environments
- Similar looking images lead to false loop-closure



Similiar looking images [5]

References

[1] R. Siegwart, I. Nourbakhsh, D. Scaramuzza, Introduction to Autonomous Mobile Robots, Cambridge, 2nd Edition, MIT Press, 2011

[2] Giorgio Grisetti, Gian Diego Tipaldi, Cyrill Stachniss, Wolfram Burgard, Daniele Nardi, Fast and accurate SLAM with Rao–Blackwellized particle filters, In Robotics and Autonomous Systems, Volume 55, Issue 1, 2007, Pages 30-38, ISSN 0921-8890

[3] M. Cummins, P. Newman, Appearance-only SLAM at large scale with FAB-MAP 2.0, in The International Journal of Robotics Research, Volume 30, Issue 9, 2011

[4] X. Zhang and Y. Su and X. Zhu, Loop closure detection for visual SLAM systems using convolutional neural network, 2017 23rd International Conference on Automation and Computing (ICAC), 2017

References

[5] Cummins, M. and Newman, P., Invited Applications Paper FAB-MAP: Appearance-Based Place Recognition and Mapping using a Learned Visual Vocabulary Model, 27th Intl Conf. on Machine Learning (ICML2010), 2010