Third Dimension Navigation

Eli Hodges

What is SLAM?

If a mobile robot is placed in a completely unknown environment, how do we give it the ability to contextualize itself in said environment?

History of SLAM

- Started as a conversation between colleagues. 1986 IEEE Robotics and Automation Conference

IEEE Robotics & Automation Society

Notable Conversationers

- Jim Crowley
  - Professor at Grenoble INP

- Peter Cheeseman
  - Ph. D

- Professor Hugh Durrant-Whyte
  - Professor at The University of Sidney

Early Developments

- Early work by Smith and Cheeseman developed the basics of a statistical model for landmark mapping
- Described necessary correlation with landmarks.
  - Common error in the estimation of where the robot actually is.

Estimating Uncertain Spatial Relationships in Robotics*

Randall Smith† Matthew Self‡ Peter Cheeseman§

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

In this paper, we describe a representation for spatial information, called the *stochastic map*, and associated procedures for building it, reading information from it, and revising it incrementally as new information is obtained. The map contains the estimates of relationships among objects in the map, and their uncertainties, given all the available information. The procedures provide a general solution to the problem of estimating uncertain relative spatial relationships. The estimates are probabilistic in nature, an advance over the previous, very conservative, worst-case approaches to the problem. Finally, the procedures are developed in the context of state-estimation and filtering theory, which provides a solid basis for numerous extensions.

1 Introduction

In many applications of robotics, such as industrial automation, and autonomous mobility, there is a need to represent and reason about spatial uncertainty. In the past, this need has been circumvented by special purpose methods such as precision engineering, very accurate sensors and the use of fixtures and calibration points. While these methods sometimes supply sufficient accuracy to avoid the need to represent uncertainty explicitly, they are usually costly. An alternative approach is to use multiple, overlapping, lower resolution sensors and to combine the spatial information (including the uncertainty) from all sources to obtain the best spatial estimate. This integrated information can often supply sufficient accuracy to avoid the need for the hard engineered approach.

In addition to lower hardware cost, the explicit estimation of uncertain spatial information makes it possible to decide *in advance* whether proposed operations are likely to

*The research reported in this paper was supported by the National Science Foundation under Grant ECS-8200615, the Air Force Office of Scientific Research under Contract F49620-84-K-0007, and by General Motors Research Laboratories.

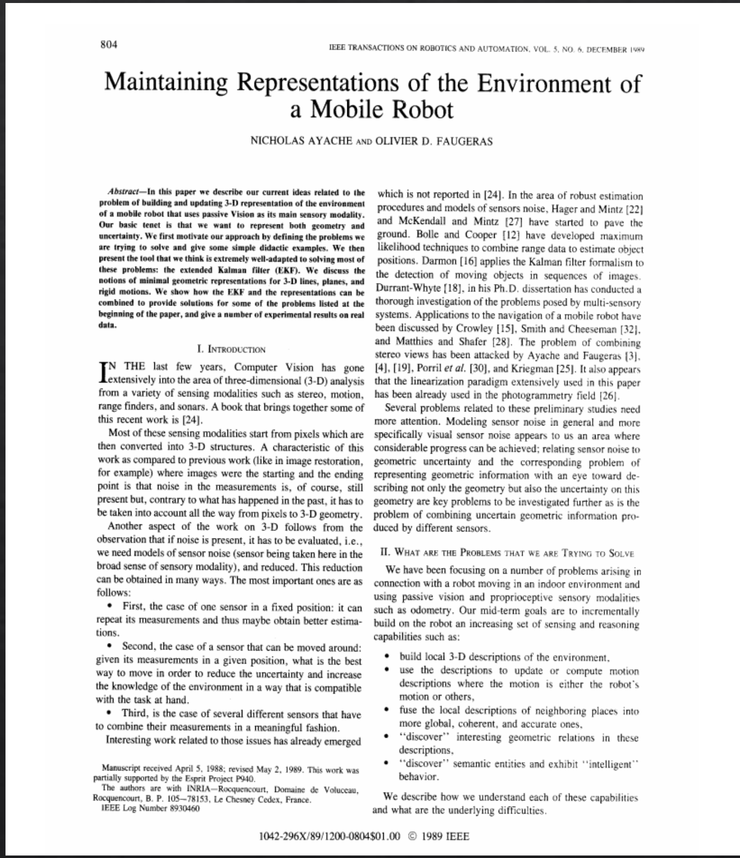
†Currently at General Motors Research Laboratories
Warren, Michigan.

‡Currently at UC Berkeley.

§Currently at NASA Ames Research Center,
Moffett Field, California.

Early Developments – Cont

- Ayache and Faugeras built the foundation of visual navigation using passive vision



Maintaining Representations of the Environment of a Mobile Robot

NICHOLAS AYACHE AND OLIVIER D. FAUGERAS

***Abstract*—In this paper we describe our current ideas related to the problem of building and updating 3-D representation of the environment of a mobile robot that uses passive Vision as its main sensory modality. Our basic tenet is that we want to represent both geometry and uncertainty. We first motivate our approach by defining the problems we are trying to solve and give some simple didactic examples. We then present the tool that we think is extremely well-adapted to solving most of these problems: the extended Kalman filter (EKF). We discuss the notions of minimal geometric representations for 3-D lines, planes, and rigid motions. We show how the EKF and the representations can be combined to provide solutions for some of the problems listed at the beginning of the paper, and give a number of experimental results on real data.**

I. Introduction

IN THE last few years, Computer Vision has gone extensively into the area of three-dimensional (3-D) analysis from a variety of sensing modalities such as stereo, motion, range finders, and sonars. A book that brings together some of this recent work is [24].

Most of these sensing modalities start from pixels which are then converted into 3-D structures. A characteristic of this work as compared to previous work (like in image restoration, for example) where images were the starting and the ending point is that noise in the measurements is, of course, still present but, contrary to what has happened in the past, it has to be taken into account all the way from pixels to 3-D geometry.

Another aspect of the work on 3-D follows from the observation that if noise is present, it has to be evaluated, i.e., we need models of sensor noise (sensor being taken here in the broad sense of sensory modality), and reduced. This reduction can be obtained in many ways. The most important ones are as follows:

- First, the case of one sensor in a fixed position: it can repeat its measurements and thus maybe obtain better estimations.
- Second, the case of a sensor that can be moved around: given its measurements in a given position, what is the best way to move in order to reduce the uncertainty and increase the knowledge of the environment in a way that is compatible with the task at hand.
- Third, is the case of several different sensors that have to combine their measurements in a meaningful fashion.

Interesting work related to those issues has already emerged which is not reported in [24]. In the area of robust estimation procedures and models of sensors noise, Hager and Mintz [22] and McKendall and Mintz [27] have started to pave the ground. Bolle and Cooper [12] have developed maximum likelihood techniques to combine range data to estimate object positions. Darmon [16] applies the Kalman filter formalism to the detection of moving objects in sequences of images. Durrant-Whyte [18], in his Ph.D. dissertation has conducted a thorough investigation of the problems posed by multi-sensory systems. Applications to the navigation of a mobile robot have been discussed by Crowley [15], Smith and Cheeseman [32], and Matthies and Shafer [28]. The problem of combining stereo views has been attacked by Ayache and Faugeras [3], [4], [19], Porril *et al.* [30], and Kriegman [25]. It also appears that the linearization paradigm extensively used in this paper has been already used in the photogrammetry field [26].

Several problems related to these preliminary studies need more attention. Modeling sensor noise in general and more specifically visual sensor noise appears to us an area where considerable progress can be achieved; relating sensor noise to geometric uncertainty and the corresponding problem of representing geometric information with an eye toward describing not only the geometry but also the uncertainty on this geometry are key problems to be investigated further as is the problem of combining uncertain geometric information produced by different sensors.

II. What are the Problems that we are Trying to Solve

We have been focusing on a number of problems arising in connection with a robot moving in an indoor environment and using passive vision and proprioceptive sensory modalities such as odometry. Our mid-term goals are to incrementally build on the robot an increasing set of sensing and reasoning capabilities such as:

- build local 3-D descriptions of the environment,
- use the descriptions to update or compute motion descriptions where the motion is either the robot's motion or others,
- fuse the local descriptions of neighboring places into more global, coherent, and accurate ones,
- ''discover'' interesting geometric relations in these descriptions,
- ''discover'' semantic entities and exhibit ''intelligent'' behavior.

We describe how we understand each of these capabilities and what are the underlying difficulties.

Manuscript received April 5, 1988; revised May 2, 1989. This work was partially supported by the Esprit Project P940.

The authors are with INRIA—Rocquencourt, Domaine de Voluceau, Rocquencourt, B. P. 105—78153, Le Chesney Cedex, France.

IEEE Log Number 8930460

1042-296X/89/1200-0804$01.00

Early Developments – Cont

- Around the same time, Chatila and Laumond used Kalman Filter-type algorithms to develop the foundations of SONAR SLAM

Position referencing and consistent world modeling for mobile robots

Publisher: IEEE

Cite This

PDF

R. Chatila ; J. Laumond **All Authors**

Abstract

Authors

References

Citations

Keywords

Abstract:
In order to understand its environment, a mobile robot should be able to model consistently this environment, and to locate itself correctly. One major difficulty to be solved is the inaccuracies introduced by the sensors. The approach proposed in this paper to cope with this problem relies on 1) defining general principles to deal with uncertainties : the use of a multisensory system, favo ring of the data collected by the more accurate sensor in a given situation, averaging of different but consistent measurements of the same entity weighted with their associated uncertainties, and 2) a methodology enabling a mobile robot to define its own reference landmarks while exploring its environment. These ideas are presented together with an example of their application on the mobile robot HILARE.

Why does this matter?

- This showed that the ‘solution’ to the SLAM problem was going to be computationally daunting.
- Accuracy requires constant updates.
- Additionally, a gigantic state vector would be required.
    - Order: Number of landmarks maintained in the map
    - Computation Behavior: n^2.
        - N = number of landmarks

Continued Research

- The conclusion reached by previous discoveries dissuaded further research on the combined problem.
  - Research split into groups working on the components of the problem.
- Researchers believed that with current information on the topic, no solutions would converge.
  - Behavior was random, with unbounded error growth.

Coining "SLAM"

- In 1995 the approach to the SLAM problem completely changed
  - Reorganized into a *single* estimation problem
  - Datasets emerged showing convergence
  - Correlation between landmarks, previously minimized, became essential to solution generation.

Foundational Theory

In 1997, Michael Csorba formalized SLAM's history into a single source.

Abstract

Michael Csorba
Balliol College

Doctor of Philosophy
Michaelmas Term 1997

Simultaneous Localisation and Map Building

This thesis examines the problem of localising an Autonomous Guided Vehicle (AGV) travelling in an unknown environment. In this problem, the AGV faces the dual task of modeling the environment *and* simultaneously localising its position within it. The Simultaneous Localisation and Map Building (SLAM) problem is currently one of the most important goals of AGV research. Solving this problem would allow an AGV to be deployed easily, with very little initial preparation. The AGV would also be flexible and able to cope with modifications in the environment. A solution to the SLAM problem would enable an AGV to would be truly "autonomous."

The thesis examines the SLAM problem from an estimation theoretic point of view. The estimation approach provides a rigorous framework for the analysis and has also proven to be successful in actual applications.

The most significant contribution of this thesis is to provide, for the first time, a detailed development of the theory of the SLAM problem. It is shown that correlations arise between errors in the vehicle and the map estimates, and these correlations are identified as fundamentally important to the solution of the SLAM problem. It is demonstrated that ignoring these correlations results in the loss of the fundamental structure of the SLAM problem and leads to inconsistency in map and vehicle estimates. The evolution of the map is also examined. It is shown that the map cannot diverge, in particular it is shown that the estimate of individual features in the map, groups of features and relative distances between features cannot diverge. The thesis further examines the theoretical limit to the accuracy of the map and the localisation of the vehicle. The relative nature of the observations are exploited to identify the existence of a limit, and the determination of the limit is also presented.

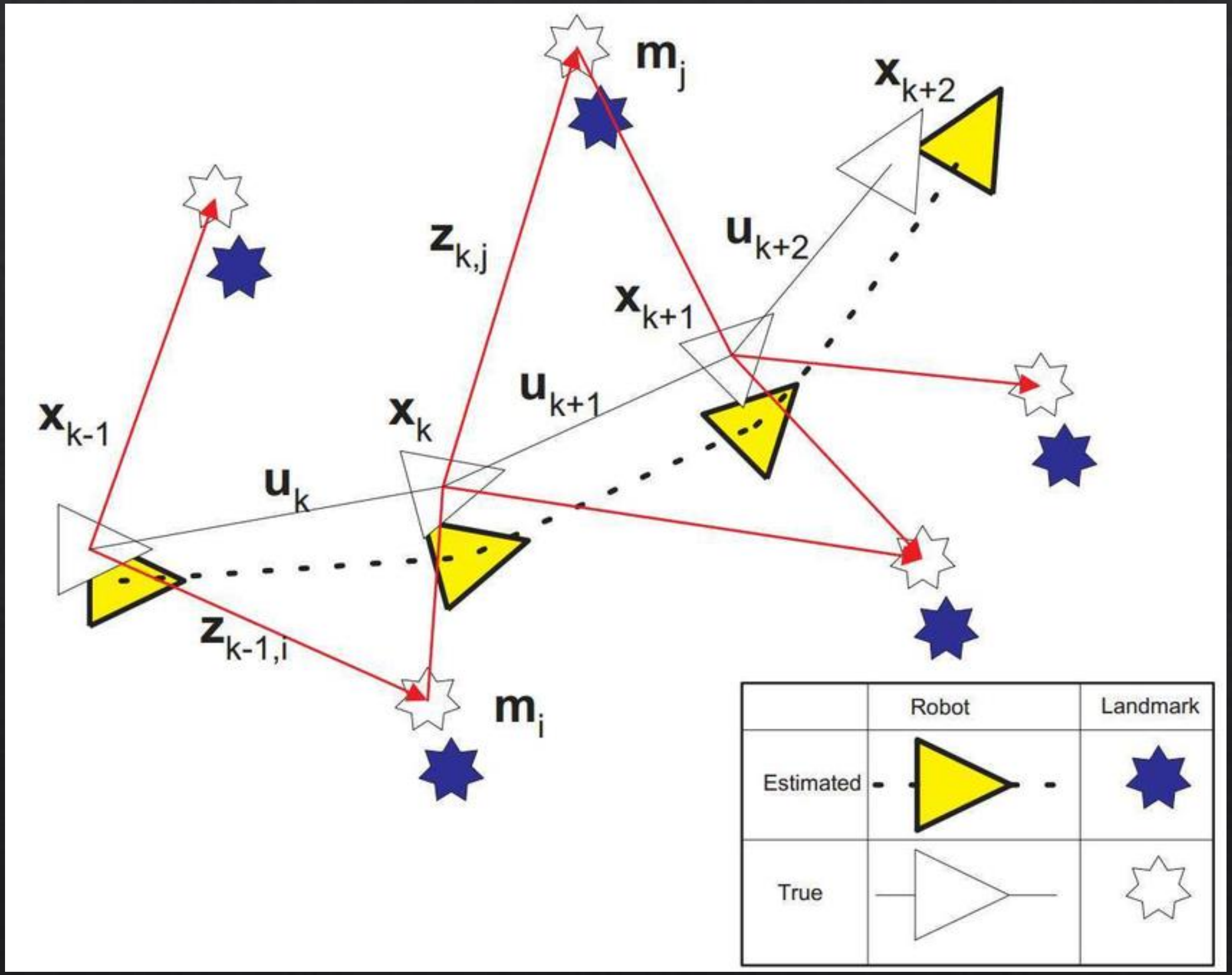
Together, these results show that it is theoretically possible to start at an unknown location in an unknown environment and to build a map by which to navigate.

It is identified in the thesis that the full solution to the SLAM problem results in a large storage requirement and heavy computational burden. This leads to the need to find alternative solutions to the SLAM problem that have both a sound theoretical basis and which also permit practical implementations.

Two such solutions to the SLAM problem are developed in this thesis. The first is the Bounded Region (BOR) filter. The BOR filter makes the assumption of bounded errors in the observation and vehicle models, which permits the filter to be formulated without the need to consider correlations. Consequently the storage and the computational requirement of the BOR filter are very low.

The second alternative solution developed in this thesis is the Relative (REL) filter. The REL filter estimates the relative distance between features and the relative angle formed by groups of three features. Such estimates can be formed without the need to consider correlations between estimates. The REL filter therefore has a low storage requirement and computational burden.

The thesis also investigates the performance of these solutions to the SLAM problem in a real implementation. An AGV equipped with a laser rangefinder builds a map of features in a corridor and simultaneously localises its own position using the map. Each approach is investigated in detail, in particular the evolution of the map is examined and important aspects of the localisation are identified.



- x_k : The state vector describing the location and orientataion of the vehicle
- u_k : The control vector, applied at time k – 1 to drive the vehicle to a state x_k at time k
- m_i : A vector describing the location of the i^{th} landmark whose true location is assumed time invariant
- z_{ik} : An observation of the ith landmark at time k. This informs future decisions

Probabilistic SLAM

$$P(\mathbf{x}_k, \mathbf{m} \mid \mathbf{Z}_{0:k}, \mathbf{U}_{0:k}, \mathbf{x}_0)$$

$$P(\mathbf{z}_k \mid \mathbf{x}_k, \mathbf{m}).$$

$$P(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{u}_k)$$

Probabilistic Form

Represents Overall Confidence

Observation Model

Describes Chance of making a new observation

Motion Model

Markov Process that describes movement behavior

Calculation

Time-Update

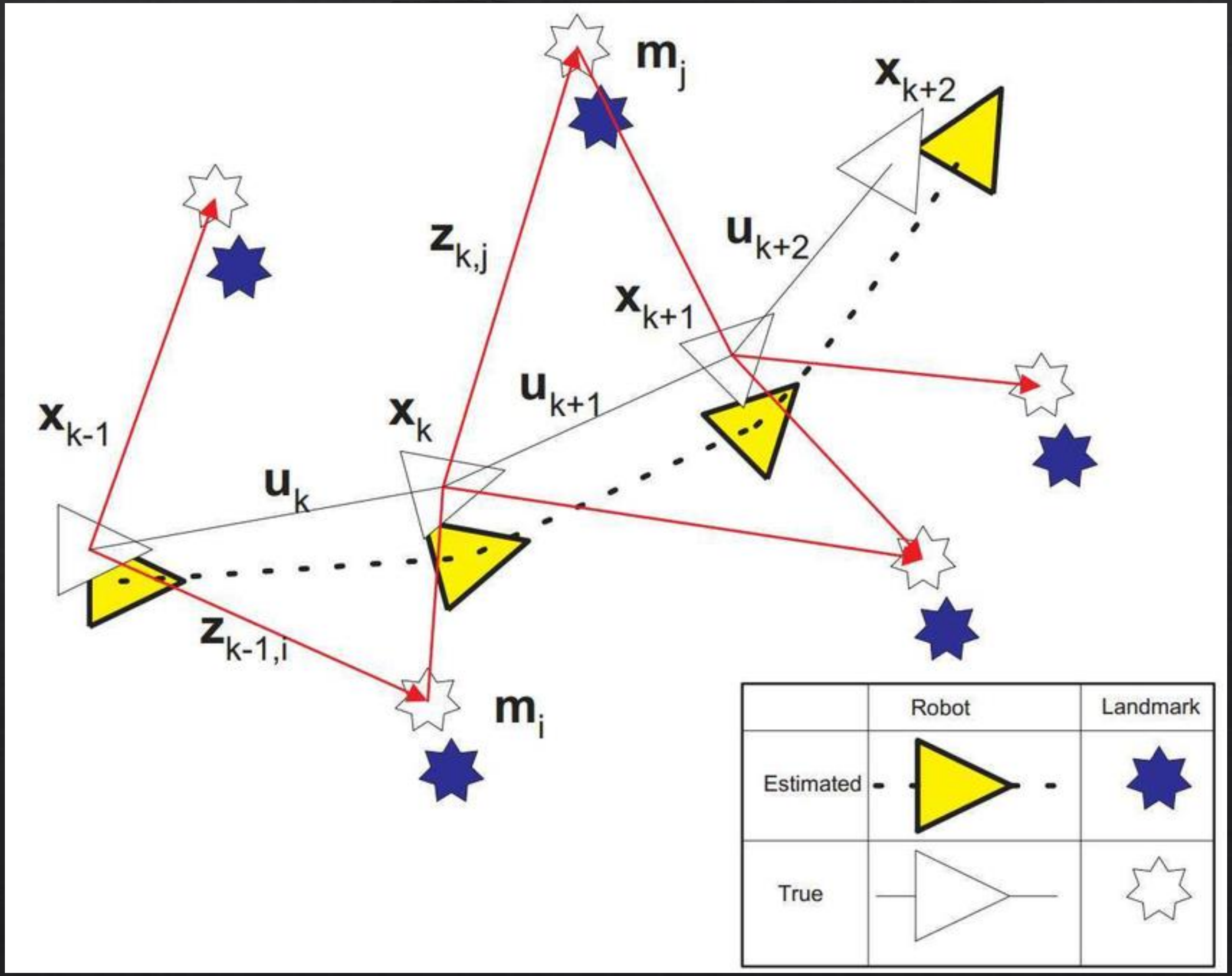
$$
\begin{aligned}
& P(\mathbf{x}_k, \mathbf{m} \mid \mathbf{Z}_{0:k-1}, \mathbf{U}_{0:k}, \mathbf{x}_0) \\
& = \int P(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{u}_k) \\
& \times P(\mathbf{x}_{k-1}, \mathbf{m} \mid \mathbf{Z}_{0:k-1}, \mathbf{U}_{0:k-1}, \mathbf{x}_0) \mathrm{d}\mathbf{x}_{k-1}
\end{aligned}
$$

Measurement Update

$$
\begin{aligned}
& P(\mathbf{x}_k, \mathbf{m} \mid \mathbf{Z}_{0:k}, \mathbf{U}_{0:k}, \mathbf{x}_0) \\
& = \frac{P(\mathbf{z}_k \mid \mathbf{x}_k, \mathbf{m}) P(\mathbf{x}_k, \mathbf{m} \mid \mathbf{Z}_{0:k-1}, \mathbf{U}_{0:k}, \mathbf{x}_0)}{P(\mathbf{z}_k \mid \mathbf{Z}_{0:k-1}, \mathbf{U}_{0:k})}
\end{aligned}
$$

SLAM Structure

Observe the relationship between landmark errors and robot positioning

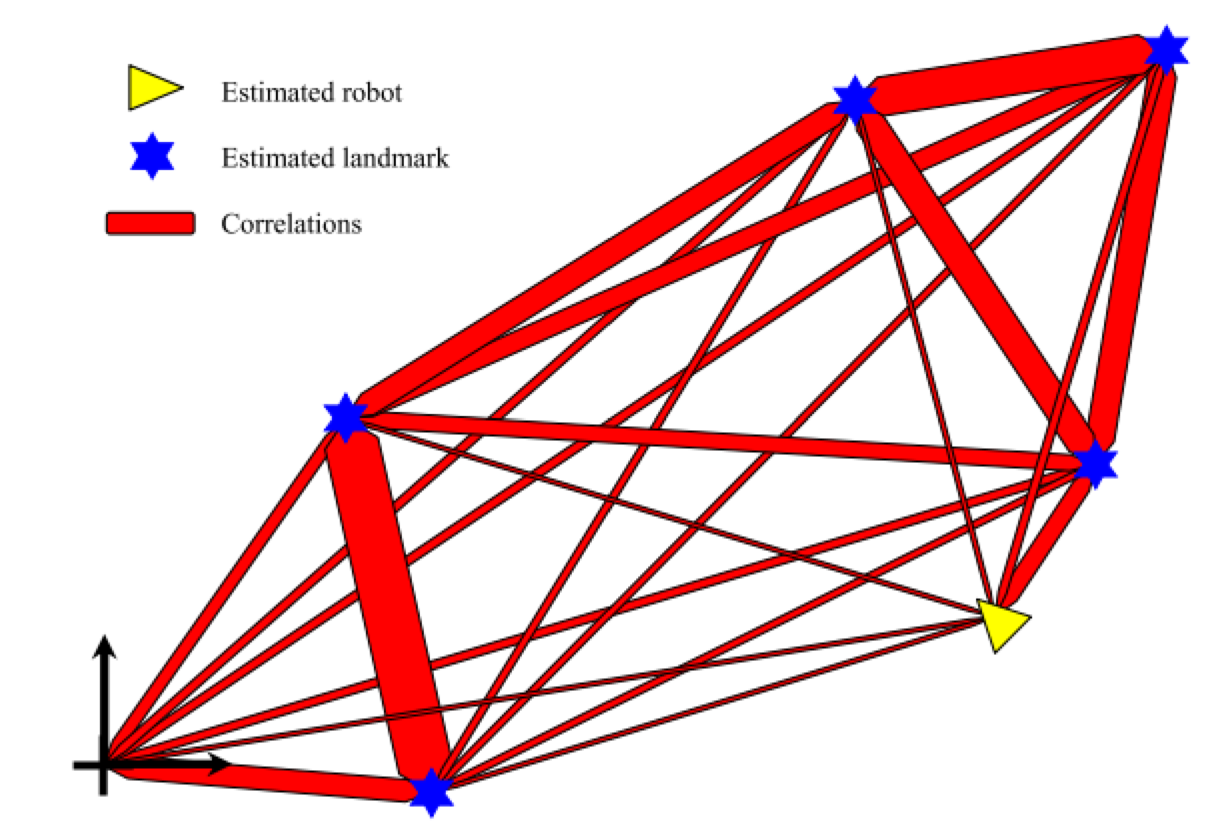


Spring Network

Relative Location of observed landmarks is independent of the robot.

Movement updates observations of landmarks with regards to previous position

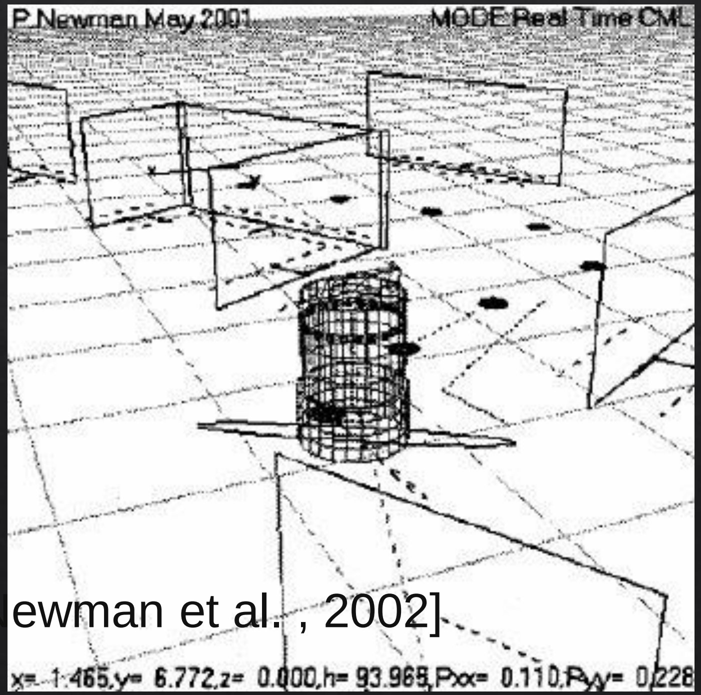
This data propagates backwards, *even updating locations not currently seen by the robot*



Solutions to Localization

Practical Application

Newman et al. 2002

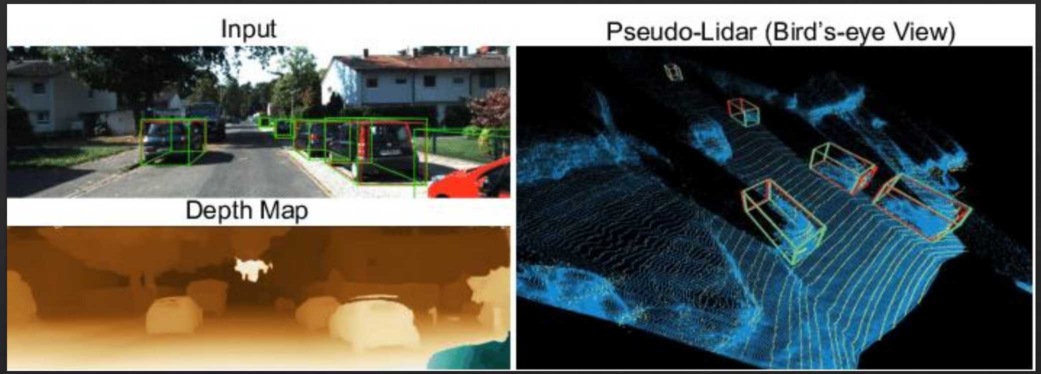


[Newman et al. , 2002]

Modern Solutions

LIDAR

Pseudo-Lidar (Computer Vision)



https://www.youtube.com/watch?v=ibIzthKKZWY

Lidar Location Mapping

https://www.youtube.com/watch?v=yTGFg0euemY&ab_channel=PeRLUM

http://robots.engin.umich.edu/publications/rwolcott-2015a.pdf

Tesla’s Autopilot Mapping

https://twitter.com/tesla/status/1120815737654767616?lang=en

Boston Dynamics’ Spot

https://www.youtube.com/watch?v=Ve9kWX_KXus&ab_channel=BostonDynamics

Things I’ve left out

https://www.youtube.com/watch?v=7Hlb8YX2-W8&ab_channel=georgehotzarchive

References

Hugh Durrant-Whyte., & Tim Bailey. (2006). Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. Retrieved from https://people.eecs.berkeley.edu/~pabbeel/cs287-fa09/readings/Durrant-Whyte_Bailey_SLAM-tutorial-I.pdf

Yan Wang et al. (2019). Pseudo-LiDAR from visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving. Retrieved from https://arxiv.org/abs/1812.07179

Ryan W. Wolcott., & Ryan M. Eustice. Fast LIDAR Localization using Multiresolution Gaussian Mixture maps. Retrieved from http://robots.engin.umich.edu/publications/rwolcott-2015a.pdf

KITTI Vision Benchmark Suite: From http://www.cvlibs.net/datasets/kitti/eval_odometry.php

d