

Bayesian estimation and Kalman filtering: A unified framework for Mobile Robot Localization*

Stergios I. Roumeliotis^{1†} and George A. Bekey^{1,2}

stergios|bekey@robotics.usc.edu

¹Department of Electrical Engineering

²Department of Computer Science

Institute for Robotics and Intelligent Systems

University of Southern California

Los Angeles, CA 90089-0781

Abstract

Decision and estimation theory are closely related topics in applied probability. In this paper, Bayesian hypothesis testing is combined with Kalman filtering to merge two different approaches to map-based mobile robot localization; namely Markov localization and pose tracking. A robot carries proprioceptive sensors that monitor its motion and allow it to estimate its trajectory as it moves away from a known location. A single Kalman filter is used for tracking the pose displacements of the robot inbetween different areas. The robot is also equipped with exteroceptive sensors that seek for landmarks in the environment. Simple feature extraction algorithms process the incoming signals and suggest potential corresponding locations on the map. Bayesian hypothesis testing is applied in order to combine the continuous Kalman filter displacement estimates with the discrete landmark pose measurement events. Within this framework, also known as Multiple Hypothesis Tracking, multi-modal probability distribution functions can be represented and this inherent limitation of the Kalman filter is overcome.

1 Introduction

Mobile robots are fast becoming probably the most significant application of robotics. They have moved from the industry floor and are now being sent on missions to other planets [11], remote areas [26], or dangerous radioactive sites [5]. The potential applications for mobile robots does not only include special missions. They are also being used as guides in museums [24] and for entertainment purposes [10]. In order for a mobile robot to travel from one location to another it has to know its position and orientation (pose) at any given time. In most of the cases today this is achieved by either having a human in the navigation loop who directs the vehicle remotely or by

constraining the robot to operate in a certain area, precisely mapped [7] or suitably engineered; i.e. marked with beacons [13] or other artificial landmarks [16].¹ The level of autonomy depends predominantly on the ability of the robot to know its exact location using a map that contains minimal information for the environment represented in a simple way.

In this paper we present a general localization algorithm, Multiple Hypothesis Tracking [21], that surpasses most of the difficulties encountered by today's localization practices. It can easily adapt to dynamic environments and it is expandable to the case of unknown territory. The key contribution of this algorithm is that it unifies many existing localization techniques under one schema that preserves the benefits of the most commonly used estimation techniques in this field.

The localization problem could be solved if the area that a robot moves could be marked in a unique separable way. That means each place should contain information that will make it recognizable and distinguishable from all the other places in the area. A first approach to the problem would require each location to have a unique "signature". This is already available for some areas where the GPS signals are available. The triangulation of the distances from the satellites in view assigns a unique pair of numbers for each location on the surface of the planet. Longitude and latitude are the geographical coordinates and are used for long distance navigation of airplanes, ships and even cars. This attractive method suffers from a wide variety of problems. The first and most important one is the accuracy of the signals. Even the most expensive commercial GPS devices provide an accuracy of a few tens of centimeters which is more than enough when a car drives from one side of a city to the other but it is not acceptable for robots that move in cluttered environments or handle material located within small distances. Another serious drawback is that the GPS signals are not available indoors or in the vicinity of tall buildings, bridges and structures

*This work is partially supported by NASA-JPL (contract 959816) and DARPA (contracts F04701-97-C-0021 and DAAE07-98-C-L028).

[†]Contact author for correspondence.

¹From now on the words landmark and feature will be used interchangeably.

that occlude the view of the necessary number of satellites. Finally, the cost of a GPS receiver is too high for many robotic applications.

In many scenarios where GPS is not available, the signature of a location has to be determined based on the special characteristics of the area itself. In this case the dissimilarity between signatures of different locations has to be identifiable by the sensors used for this reason. There are many attributes of the environment that can be exploited to gain signature diversity. These can be dynamic or static. The color and the shape of a location are strong candidates for discriminating different areas and tend to remain unchanged over long periods of time. Visual memorization of each place may be sufficient for distinguishing between them. Although the homing capabilities of many types of insects [3], [25] and other animals [9], [8] depend on this type of localization, there are serious implementation issues that have to be resolved before this technique can be applied to mobile robot localization. The requirements for storing and processing the numerous collected images are probably the most difficult to overcome.

Another way to achieve the desired distinguishability of a location of concern is by preparing a detailed volumetric or grid based map of the area and then matching the information from exteroceptive sensors with the information stored in this map. This approach suffers from the following drawbacks: (i) The preparation of the map adds significant overhead to the process. It has to be full, exhaustive, and precise enough to exclude any ambiguities and avoid misinterpretation, (ii) The exteroceptive sensors have to be able to supply the robot with the same level of detail as that imprinted on the map, (iii) The processing requirements to obtain the necessary information and match it to the corresponding locations on the map call for a powerful processor on-board the robot if real-time operation is demanded.

Instead of building elaborate detailed maps of almost every bit of the environment an alternative would be to extract the minimal information that makes a location unique. This could be “What does this location look like” or “How do I get to this location”. If for example a robot is capable of distinguishing two fairly similar doors just by looking at them (as humans do by reading a sign on a door, for example) then the problem is essentially solved. This is one of the prominent ways that mammals recognize their location. Certain aspects of different locations are compiled in a compact form of mnemonic representation. Not all the information is preserved which makes it sometimes unreliable. Our ability to distinguish between one place or another is not unlimited. There are certain cases where the memory is not developed enough to support an unambiguous decision or prevent a mistake. This problem is even more prevalent in artificial machines such as robots equipped with sensing devices like a camera for example. The state of the art pattern recognition techniques are far from being able to imitate the perceptual and processing capabilities of the human eye. The poten-

tial of failing to acquire critical information is such that this method has yet to be applied in real world robots.

The “What does this location look like” methodology, approaches the essence of the landmark based localization. The additional characteristic of the landmark based localization is that only certain locations in the area are marked this way while the rest of the vicinity is marked as “How do I get to this location”. Before going into a detailed examination of the landmark based localization it would be useful to examine the “How do I get to this location” based localization techniques. The main marking of the area in this case is with respect to some known world frame and it is the coordinates (x, y, ϕ) of each location. The fact that these coordinates are with respect to some other location makes them actually $(\Delta x, \Delta y, \Delta \phi)$, i.e. this set of numbers marks the final location by its transition from the initial location there by following a meaningful prespecified way. In this case it is moved by Δx on the x axis and by Δy on the y axis and rotated by $\Delta \phi$. The initial location can be arbitrarily chosen as well as the coordinate system or the units used. The mapping of the area is unique this way. The problem is if the robot can “read” this marking. “Reading this marking” means that the robot is capable of repeating the marking of each location it passes through, i.e. it can determine the coordinates of this location by processing information from odometric or inertial sensors that keep track of the transitions or poses of the robot. A cumulative process that starts from a known initial location and keeps updating every time the robot moves. In practice this is usually done by integrating the wheel encoders’ signals for example. This method suffers from the limited capabilities of the sensors. Noise in the sensor signals, limited precision, potential failures can easily mislead the robot to believe that it is in a different location from where it actually is.

The proposed methodology will be presented for the case of a mobile robot moving into a two dimensional (2 D) office environment. Following the same formulation of the problem, extension to a 3 D environment is straightforward. The main difference is that in the case of an unstructured environment, the exteroceptive sensors have to be capable of extracting natural instead of artificial (structural) features.

2 Kidnapped robot, Absolute Localization, and Pose Tracking

First, we will study the ideal case of a robot equipped with exteroceptive sensors capable of distinguishing between almost identical landmarks. Though this scenario is fictional, it will ease the transition to the more realistic case of imperfect landmark recognition and imperfect odometry.

2.1 Perfect Landmark Recognition, Imperfect Odometry

In the absence of any previous knowledge, the initial pose distribution can be assumed to be uniform and thus the probability density function (pdf) is:

$$f_0(\mathbf{x}) = \frac{1}{2\pi S}, \quad (1)$$

where S is the surface of the area that the robot is allowed to move in and $\mathbf{x} = [x, y, \phi]^T$ is the pose of the robot. In the case of perfect landmark recognition, there is no ambiguity. The decision is binary for each of the possibilities encountered: $P(z_k = A_i) = 1$, $P(z_k = A_j) = 0$, $j = 1, \dots, N$, $j \neq i$. Every time the robot detects and identifies a landmark whose exact pose \mathbf{x}_{A_i} is known, the updated pdf becomes:

$$f(\mathbf{x}/z_k) = f(\mathbf{x}/z_k = A_i) =$$

$$\frac{1}{(2\pi)^{n/2} \det(\mathbf{K}_{A_i})^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_{A_i})^T \mathbf{K}_{A_i}^{-1} (\mathbf{x} - \mathbf{x}_{A_i})\right] \quad (2)$$

in the case that the pose of the landmark is known to follow a Gaussian distribution with mean value \mathbf{x}_{A_i} and covariance² $\mathbf{K}_{A_i} = E[(\mathbf{x} - \mathbf{x}_{A_i})(\mathbf{x} - \mathbf{x}_{A_i})^T]$. The Gaussian assumption for the distribution of the pose of the landmark is often used to compensate for the sensor inaccuracies. The sensors determined to locate the pose of a landmark, have limited precision and their signals are in general noisy. Therefore the calculation of the pose of the landmark with respect to the current pose of the robot has its own inaccuracies. A Gaussian model can be used to describe the uncertainty of the calculated transformation from the egocentric frame to the landmark centered frame.

Every time a landmark is left behind, the robot knows its location precisely. The pdf of its pose is a Gaussian function of its coordinates in the area. This distribution is **unimodal** since we have a 1-to-1 match between encountered landmarks and locations on the map. From then on the robot has to rely on its odometry until it encounters another landmark. In the absence of any external positioning information the uncertainty of the pose estimate will increase continuously as the robot moves. The system is not observable and thus any estimation technique will eventually fail. The accuracy of the pose estimate based on the imprecise odometry will decrease and it will become untrustworthy. The rate of deterioration depends on the number and type of sensors used as well as their noise characteristics. Any form of absolute orientation measurement [20] will reduce the uncertainty while low levels of noise will reduce the rate of increase of the variance of the pose estimate. Substantial improvement is provided by applying Kalman filtering techniques. These techniques have been used successfully in pose estimation problems such

as missile tracking and ship navigation for the last four decades [1] and their field of application has been extended to mobile robots. Relying on a simple kinematic model of the robot [18] or on sensor modeling [19] a Kalman filter can be formulated to estimate the displacement of the robot

$$\widehat{\Delta \mathbf{x}} = [\widehat{\Delta x} \quad \widehat{\Delta y} \quad \widehat{\Delta \phi}]^T \quad (3)$$

and the associated covariance:

$$\mathbf{P}_{\widehat{\Delta \mathbf{x}}} = E[\widehat{\Delta \mathbf{x}} \widehat{\Delta \mathbf{x}}^T] \quad (4)$$

The robot will have to drive towards the pose of another known landmark in order to reduce its uncertainty. When it does so the Kalman filter framework is suitable for fusing the new landmark information (measurement \mathbf{x}_{B_j}) with the current pose estimate $\widehat{\mathbf{x}}_{\text{robot}} = \mathbf{x}_{A_i} + \widehat{\Delta \mathbf{x}}$ as this is calculated based on the previous landmark information \mathbf{x}_{A_i} and the processed odometric data $\widehat{\Delta \mathbf{x}}$. Examples of treatment of similar cases can be found in the related literature [13], [2].

Environments with low landmark density are more difficult to navigate in and require that the path planing module must consider the locations of the landmarks when computing a path in order to facilitate the localization process. It is very possible that a trade off would occur between the total length of the path between the two locations and the required level of accuracy which depends on the frequency of the landmarks within the path.

2.2 Imperfect Landmark Recognition, Imperfect Odometry

This section deals with a more realistic and challenging scenario. In most of the current approaches where both sources of information are considered, either the odometry (eg [4]) or the landmark recognition module (eg [13] or [23]) is assumed perfect. Many times the landmark recognition module is capable of recognizing certain types of landmarks but it cannot distinguish between members of the same type. For example it can distinguish a door from a corner but it is not capable of identifying which door on the map it is facing now, or which corner it is about to turn. The inability of the exteroceptive sensing to determine reliably the identity of a landmark can cost the robot the ability to globally localize in the environment represented on a given a map. To make this more clear, consider the simple case where the pose calculated using the odometry is accurate to a few meters on each direction of motion and the map contains two similar landmarks within this area of a few meters. The appearance of these two landmarks will have no effect on the accuracy of the localization. There is no extra confidence that the encountered landmark is one or another. The uncertainty remains the same as before being able only to categorize the appeared landmark to be of a certain type. Similar problems occur when the landmark recognition module is capable of identifying possible types of landmarks along with the associated probabilities that portray the level of

²For the 2D case $\mathbf{K}_{A_i} = [\sigma_{x_{A_i}}^2 \quad 0 \quad 0; \quad 0 \quad \sigma_{y_{A_i}}^2 \quad 0; \quad 0 \quad 0 \quad \sigma_{\phi_{A_i}}^2]$

confidence. For example a landmark recognition system based on vision and looking for circular or square patches in the environment has a certain hit ratio when categorizing its visual cues. The information provided does not have to be either “square” or “disk”. It can be “square” with probability 70% and “disk” with probability 30%. Although a decision about the robots location cannot be made because the new information is inconclusive, instead of disregarding it, it is better to “store” it and use it later.

A framework that is capable of incorporating uncertain features detected by the exteroceptive sensors and noisy motion measurements recorded by the proprioceptive sensors is presented here. The seemingly un-trustworthy information from the landmark recognition module is compiled in a number of different hypotheses. Each hypothesis is based on certain assumptions for the current location of the robot as well as the possible identity of the encountered landmark and has a level of confidence associated with it. The certainty is portrayed as a probability $P(H_i)$ associated with the particular hypothesis H_i . In the algorithm presented here, both the total pose pdf and the probabilities for each hypothesis are adjusted continuously while the robot is in motion.

As mentioned before, a Kalman filter can be used to fuse information from a variety of proprioceptive sensors onboard a robot (odometric and/or INS), estimate the pose displacement $\widehat{\Delta \mathbf{x}}$, and provide the level of belief to this estimate, i.e. the covariance $\mathbf{P}_{\widehat{\Delta \mathbf{x}}}$. It can also optimally combine the positioning information (from a possible landmark match) with the pose estimate calculated using the odometric information. The uncertainty in the landmark recognition is the main reason for introducing the multiple hypothesis approach. Multiple hypotheses testing for target tracking was first presented in [17] and was applied to mobile robot localization by [4] for distinguishing between structural features inside a building. The main drawback of this application was that a Kalman filter was required for every hypothesis encountered. Thus with the number of hypotheses growing rapidly, this methodology could not be used in real time. Another major drawback was that the belief levels for each hypothesis were updated only when a landmark was encountered. Though it is less obvious, the absence of a landmark also contains information which can eliminate a weak hypothesis and thus keep the number of them considerably smaller. It is also worth mentioning that in [4] the authors assume perfect odometry while our algorithm takes under consideration the odometric uncertainties.

In the approach presented here, only one Kalman filter is used regardless of the number of existing hypotheses. This filter calculates the quantities in Equations (3) and (4) which are the same for each hypothesis and thus reduces considerably the computations required. The confidence levels of the hypotheses change when a new landmark appears. New hypotheses can appear and old ones

can vanish.³ However, the Kalman filter is based on a unimodal distribution and therefore it cannot directly be applied to the multiple hypothesis case. The Bayesian formulation of the problem dictates the competition between the possible localization scenarios and allows for a **multi-modal** pdf. Hereafter we describe the proposed *Multiple Hypothesis Tracking* algorithm (Figures 1 and 2) in more detail.

Step 1[*Kidnapped robot localization*] First the robot finds itself at an unspecified location somewhere in the mapped area. This is the “kidnapped robot” localization problem. The assumption is that the initial pdf is uniform across the space of motion as in Equation (1). The robot starts to move and after a while encounters the first landmark A_i . The feature extraction module assigns probabilities $P(z_1 = A_i)$, $i = 1..N$ to the different choices, where N is the number of possible matches on the map. Due to the uncertainty related to the landmark recognition module, the N hypotheses which assume that different landmarks have been detected, are not mutually exclusive. Each of the probabilities associated with a different landmark A_i receive non zero values. Therefore the new pdf is:

$$f(\mathbf{x}/z_1) = \sum_i P(z_1 = A_i) f(\mathbf{x}/z_1 = A_i) \quad (5)$$

where $i = 1..N$, $\mathbf{x} = [x, y, \phi]^T$, z_1 is the first feature sensed by the exteroceptive sensors and $P(z_1 = A_i)$ is the probability that the sensed feature is A_i .⁴ Here the pdf’s $f(\mathbf{x}/z_1 = A_i)$ are the same ones as in Equation (2) that describe the distribution of the pose for each known landmark. Each of the non zero probabilities defines a new hypothesis H_i with assigned probability $P(H_i) = P(z_1 = A_i)$ that considers the pose of the robot to follow a different Gaussian distribution: $\mathbf{x} \sim \mathbf{N}(\mathbf{x}_{A_i}, \mathbf{K}_{A_i})$, $i = 1..N$. The weighted average of all the pdf’s linked to one of the hypotheses is the total pdf given in Equation (5) and repeated here for the case of the $(k - 1)$ th landmark:

$$f(\mathbf{x}/z_{k-1}) = \sum_i P(H_i) f(\mathbf{x}/H_i) \quad (6)$$

Step k-1[*Kalman filtering*] As the robot leaves the previous landmark $(k - 1)$, it has to rely on its odometry to track its pose. Only *one* Kalman filter is required to process the information from the proprioceptive sensors and calculate the quantities in Equations (3) and (4).⁵ The updated pose estimate for the robot at each time t is given by the following set of equations:

$$\widehat{\mathbf{x}}_{\text{robot}}(t, i) = \mathbf{x}_{A_i} + \widehat{\Delta \mathbf{x}}(t), \quad i = 1..N \quad (7)$$

³The competition between the hypotheses can continue even when there is no landmark at sight. The absence of a landmark carries enough information to terminate some of the hypotheses and thus strengthen others (the total probability of all the hypotheses adds up to one at any given time).

⁴For example, if there exist 7 similar doors in the area then $P(z_1 = A_i) = 1/7$ for each door location.

⁵The formulation of the Kalman filter is omitted due to space limitations. The interested reader is referred to [15] and [14] for a detailed description of the propagation and update equations. Some simple implementations can be found in [12] and [6].

Each equation corresponds to a hypothesis H_i with probability $P(H_i)$. \mathbf{x}_{A_i} is the pose of the last visited landmark and $\widehat{\Delta \mathbf{x}}(t)$ is the displacement estimated in the Kalman filter. The pose covariance related to each hypothesis H_i is calculated by adding the covariance of the pose of the corresponding landmark A_i to the covariance of the displacement $\widehat{\Delta \mathbf{x}}(t)$ which is computed at each time step in the Kalman filter.

$$\mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}}(t, i) = \mathbf{K}_{A_i} + \mathbf{P}_{\widehat{\Delta \mathbf{x}}}(t) \quad (8)$$

As the robot travels in-between landmarks the uncertainty of its displacement grows. Thus, the pose covariance for each hypothesis is constantly increasing. The new total pdf at time t is the weighted average of N pdf's each described by a Gaussian $\mathbf{N}(\widehat{\mathbf{x}}_{\text{robot}}(t, i), \mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}}(t, i))$:

$$\begin{aligned} f(\mathbf{x}/z_{k-1}, \widehat{\Delta \mathbf{x}}(t)) &= \sum_i P(H_i) f(\mathbf{x}/H_i) = \\ & \sum_i P(H_i) \frac{1}{(2\pi)^{n/2} \det(\mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}}(t, i))^{1/2}} \times \\ & \exp\left[-\frac{1}{2}(\mathbf{x} - \widehat{\mathbf{x}}_{\text{robot}}(t, i))^T \mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}}(t, i)^{-1} (\mathbf{x} - \widehat{\mathbf{x}}_{\text{robot}}(t, i))\right] \quad (9) \end{aligned}$$

in accordance with Equation (6).

Step k[*Bayesian estimation*] After the robot encounters the k th landmark, the distribution of its pose is described by the following pdf:

$$\begin{aligned} f(\mathbf{x}/z_k) &= f(\mathbf{x}/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k) = \\ & \sum_{j=1}^M f(\mathbf{x}/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) P(z_k = B_j) = \\ & \sum_{j=1}^M P(z_k = B_j) \sum_{i=1}^N f(\mathbf{x}, H_i/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) \times \\ & P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) = \\ & \sum_{j=1}^M \sum_{i=1}^N f(\mathbf{x}, H_i/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) \times \\ & P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) P(z_k = B_j) \quad (10) \end{aligned}$$

where N is the number of hypotheses in the previous step and M is the number of the possible matches on the map for the k th landmark. $P(z_k = B_j)$ is the probability that the k th landmark is B_j . The first quantity in Equation (10) is the updated pdf for the pose estimate after taking into consideration the new k th landmark. This quantity

is calculated in the Kalman filter as the latest estimates (Equations (7) and (8)) are combined with the new information for the pose of the new landmark $B_j, j = 1..M$. Each *new* hypothesis H_j assumes that the pose of B_j follows a Gaussian distribution with pdf as in Equation (2). The fusion of the latest Kalman filter estimates $\widehat{\mathbf{x}}_{\text{robot}}(t, i)$ and $\mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}}(t, i)$ with the pose information $\mathbf{x}_{B_j}, \mathbf{K}_{B_j}$ for each of the possible landmarks B_j is a *one step* process and it can be implemented in parallel. The resulting $f(\mathbf{x}/z_k)$ is a *multi-modal* Gaussian distribution composed of 1-to- $N \times M$ functionals. The second quantity in Equation (10) is the result of the Bayesian estimation step (Multiple Hypothesis Testing) and it is calculated for each hypothesis H_i as follows:

$$\begin{aligned} P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}}, z_k = B_j) &= \\ & \frac{P(z_k = B_j/z_{k-1}, \widehat{\Delta \mathbf{x}}, H_i) P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}})}{\sum_{i=1}^N P(z_k = B_j/z_{k-1}, \widehat{\Delta \mathbf{x}}, H_i) P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}})} \quad (11) \end{aligned}$$

where $P(H_i/z_{k-1}, \widehat{\Delta \mathbf{x}})$ is the a priori probability for each hypothesis H_i available from the previous step and $P(z_k = B_j/z_{k-1}, \widehat{\Delta \mathbf{x}}, H_i)$ is the probability to detect landmark B_j given that the current location of the robot is $\widehat{\mathbf{x}}_{\text{robot}}(t, i)$:

$$\begin{aligned} P(z_k = B_j/z_{k-1}, \widehat{\Delta \mathbf{x}}, H_i) &= \frac{1}{(2\pi)^{n/2} \det(\mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}} + \mathbf{K}_{B_j})^{1/2}} \times \\ & \exp\left[-\frac{1}{2}(\mathbf{x}_{B_j} - \widehat{\mathbf{x}}_{\text{robot}}(t, i))^T (\mathbf{P}_{\widehat{\mathbf{x}}_{\text{robot}}} + \mathbf{K}_{B_j})^{-1} (\mathbf{x}_{B_j} - \widehat{\mathbf{x}}_{\text{robot}}(t, i))\right] \quad (12) \end{aligned}$$

for $j = 1..M$. Each of the N previously existing hypotheses H_i are now tested in light of the new evidence. The hypotheses whose assumption for the estimated pose of the robot (given in Equation (7)) is closer to the location of any of the new possibly encountered landmarks $B_j, j = 1..M$ will be strengthened. The rest will be weakened. The total probability calculated by summing the probabilities in Equation (11) equals to 1. Due to the additional constraints imposed from the information associated with the pose tracking, a few only hypotheses will sustain probabilities over a preset threshold and continue to the next round, i.e. when the next landmark is encountered.

After a sufficient number of repetitions of **Step k-1** and **Step k** of this algorithm the robot will be left with one only choice for possible location. This is demonstrated in detail in the next section.

3 Experimental Results

In this section we demonstrate the efficiency of the presented algorithm in the test case of an office environment shown in Figure 3a. The position of the robot is marked with a small triangle.

In order to support dead-reckoning based position estimation, the vehicle carries shaft encoders and a gyroscope.

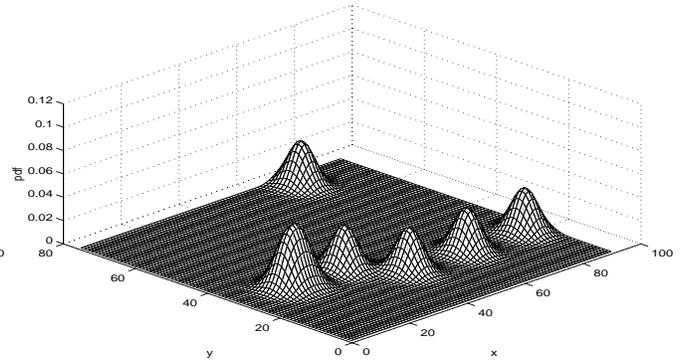
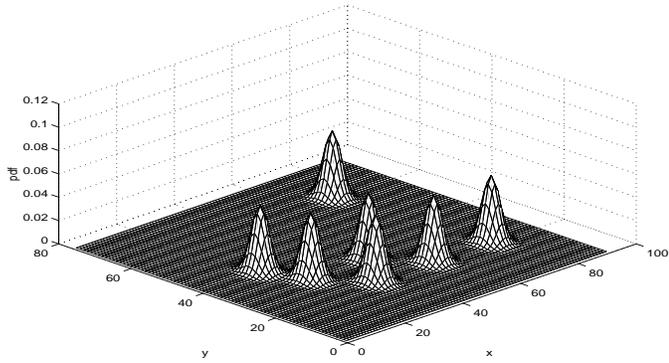
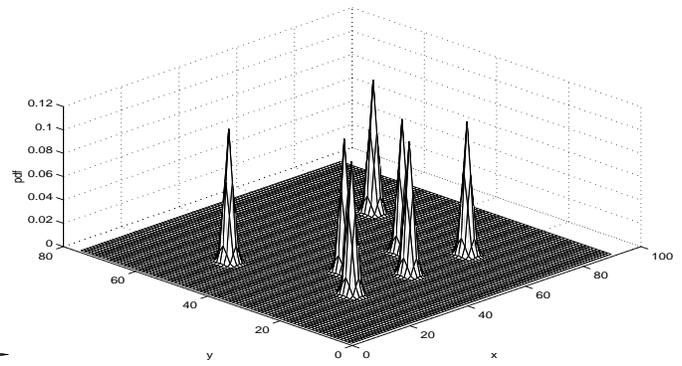
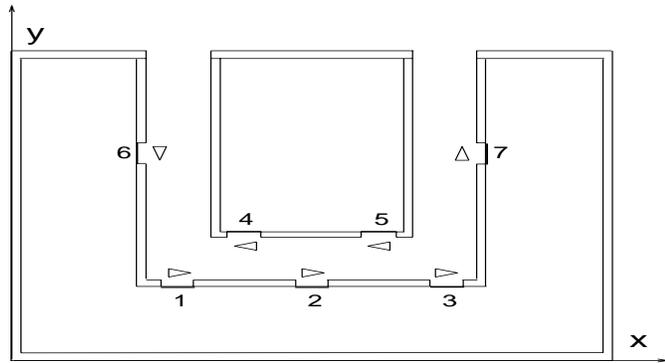


Figure 3: a. The environment of the robot, b. The initial pdf for the position of the robot when it senses one door on its right, c. The pdf after the robot moves for 2 meters, d. The pdf after the robot moves for another 2 meters and before sensing the second door on its right. Notice that 6 instead of 7 Gaussians are shown in this figure. The explanation is that 2 of the 7 hypotheses point to the same location and thus their pdf's are combined to a single Gaussian which is taller than any of the other 5. The units on these plots are 0.2m/div.

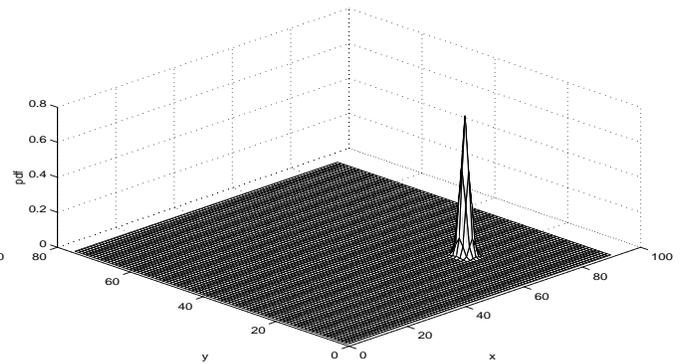
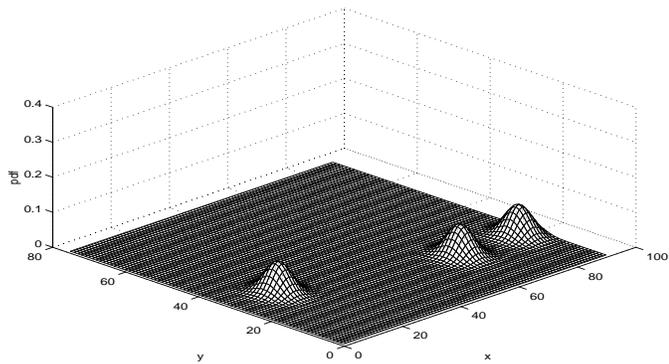
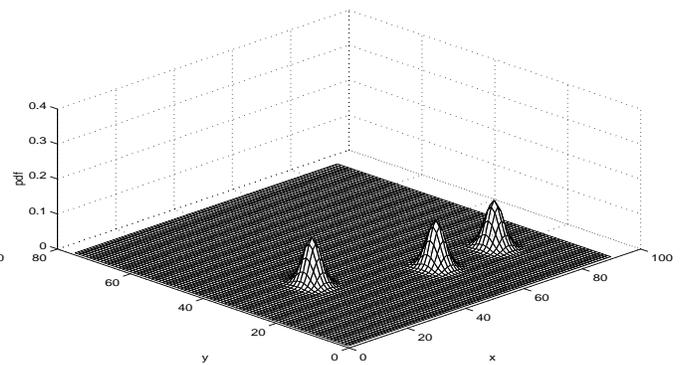
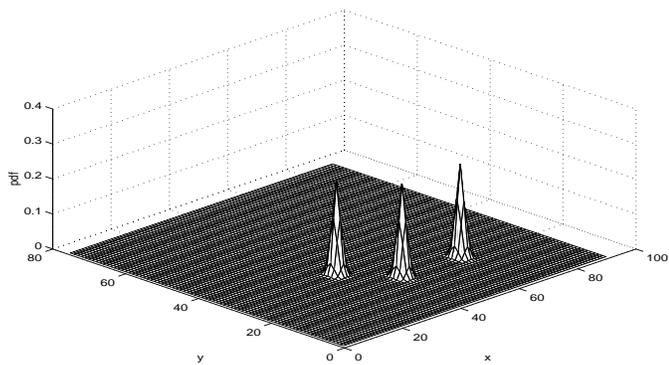


Figure 4: a. The pdf right after the robot has sensed the second door on its right, b. The pdf after the robot moves for another 2 meters, c. The pdf after the robot moves for another 2 meters and before sensing the third door on its right, d. The pdf right after the robot has sensed the second door on its right. The units on these plots are 0.2m/div.

These signals are fed to a Kalman filter implemented for this set of sensors. The computed position displacement estimate $\widehat{\Delta \mathbf{x}}(t)$ is available to the localization module at any given time. In addition the robot is equipped with exteroceptive sensors; a laser range finder. A feature extraction module capable of detecting doors is fed with these signals [22]. Although a door can be detected, it can not be identified. There are seven doors facing this U shaped corridor. A number is assigned to each of them. The robot has stored a map of the floor containing the positions of each of the doors in some coordinate frame.

In the beginning, the robot is brought and set at an arbitrary location in the corridor. In the absence of any initial positioning information, the localization module assumes a uniform position pdf along the corridor and sets the robot in motion until the first landmark is in sight. When the robot encounters the first door on its right, say number 1, it does not know which door this is. There are in fact 7 equally likely hypotheses of where it is. These are depicted in Figure 3a. The pdf's associated with each of these hypotheses are fairly sharp and are shown in Figure 3b.

As the robot begins to move again following the wall on its right, **Step k-1** is applied. That is, the new information from the odometric sensors is processed in the Kalman filter and an estimate of its position is computed along with the uncertainty associated with it. After the robot has covered 2 meters the position uncertainty has increased and thus each pdf related to each of the 7 hypotheses is spread around a larger area as shown in Figure 3c. As the robot covers another 2 meters this phenomenon becomes more prominent. This can be seen in Figure 3d.

After the robot has covered a total of 4 meters, it detects another door on its right (number 2). Application of **Step k** of the algorithm immediately reduces the number of hypotheses to 3. Out of the 7 previous hypotheses only 3 are possible. These are the ones that assumed that the previous landmark was door number 1, door number 2, or door number 5. For these hypotheses there is a new door on the right side of the robot after the vehicle has moved for 4 meters. In Figure 4a we see again that the precise localization information related to a potential landmark makes the pdf's for each of the surviving hypotheses sharp again.

The robot has to travel a longer distance before it is able to determine its position uniquely. **Step k-1** [Kalman filtering] is applied again and thus the position estimates related to the 3 hypotheses deteriorate. As the robot moves for another 2 + 2 meters, the uncertainty increase is shown in Figures 4b and 4c. Finally the robot senses a third door on its right (number 3) and by applying **Step k** of the algorithm again, it concludes with the position estimate described by the pdf of Figure 4d.

4 Conclusions

A new approach to the mobile robot localization problem was presented in detail in this paper. The two main methods applied to this problem, namely Bayesian esti-

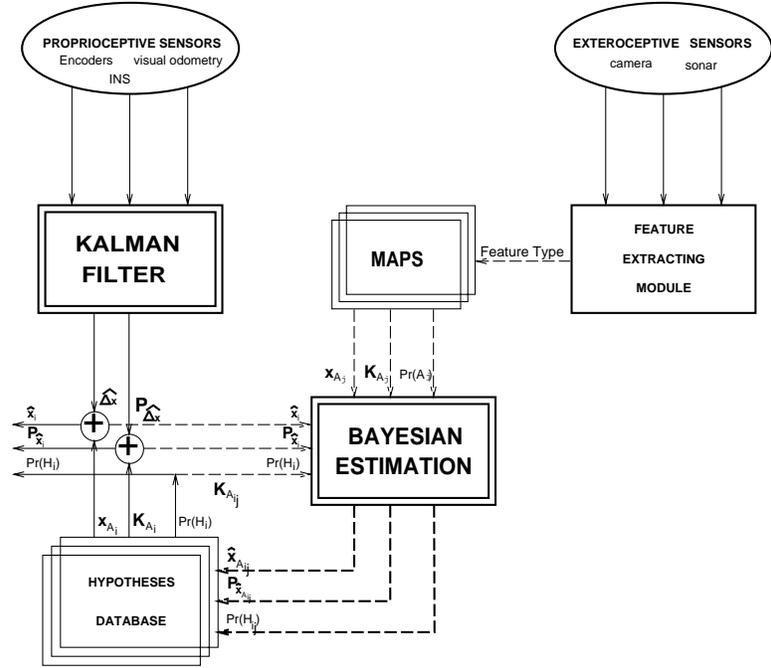


Figure 1: Multiple Hypotheses Tracking Architecture

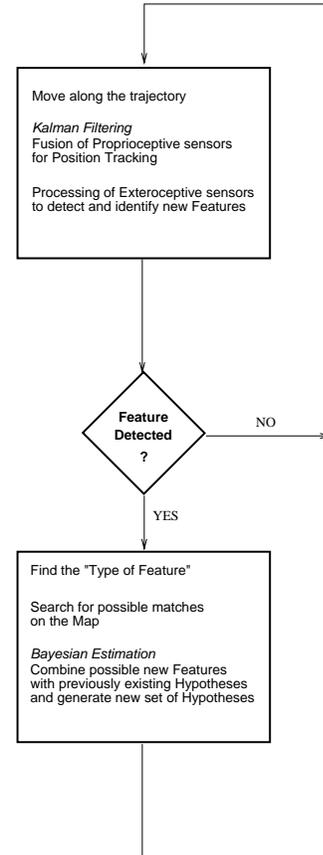


Figure 2: Multiple Hypotheses Tracking Algorithm

mation and Kalman filtering were combined within this unified framework. A mobile robot equipped with noisy odometric sensors is capable of localizing itself very precisely. A map of the environment with rough details is used for this reason. The features of the environment represented on this map can be doors, corners, columns etc. The robot also carries exteroceptive sensors that help it sense its surroundings. Simple featurizing extraction algorithms process the signals from these sensors and detect features equivalent to the ones stored in the map. A multiple hypothesis approach allows the localization algorithm to carry along all the information about the detected but not uniquely identified landmarks. Combination of consecutive landmark detections along with the odometric information processed in a Kalman filter results first in a small number of hypotheses and finally in a single location where the robot is located.

Acknowledgments

The authors would like to thank Dr. D. Bayard for several illuminating discussions and Dr. P. Pirjanian for his constructive criticism.

References

- [1] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, New York, 1988.
- [2] E. T. Baumgartner and S. B. Skaar. An autonomous vision-based mobile robot. *IEEE Transactions on Automatic Control*, 39(3):493–501, March 1994.
- [3] M. Collett, T. S. Collett, S. Bisch, and R. Wehner. Local and global vectors in desert ant navigation. *Nature*, 394(6690):269–272, 1998.
- [4] I. J. Cox and J. J. Leonard. Modeling a dynamic environment using a bayesian multiple hypothesis approach. *Artificial Intelligence*, 66(2):311–344, April 1994.
- [5] T. Denmeade. A pioneer’s journey into the sarcophagus. *Nuclear Engineering International*, 43(524):18–20, March 1998.
- [6] H. F. Durrant-Whyte. Where am i? a tutorial on mobile vehicle localization. *Industrial Robot*, 21(2):11–16, 1994.
- [7] A. Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, June 1989.
- [8] R. Epstein and N. Kanwisher. A cortical representation of the local visual environment. *Nature*, 392(6676):598–601, 1998.
- [9] C. H. Greene and R. G. Cook. Landmark geometry and identity controls spatial navigation in rats. *Animal Learning & Behavior*, 25(3):312–323, 1997.
- [10] N. Hager, A. Cypher, and D. C. Smith. Cocoa at the visual programming challenge 1997. *Computing*, 9(2):151–169, April 1998.
- [11] S. Hayati, R. Volpe, P. Backes, J. Balaram, and R. Welch. Microrover research for exploration of mars. In *Proceedings of the 1996 AIAA Forum on Advanced Developments in Space Robotics*, University of Wisconsin, Madison, August 1-2 1996.
- [12] A. Kelly. A 3d state space formulation of a navigation kalman filter for autonomous vehicles. Technical report, CMU, CMU-RI-TR-94-19.
- [13] J. J. Leonard and H. F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(3):376–382, June 1991.
- [14] P. S. Maybeck. *Stochastic Models, Estimation and Control*, volume 141-1 of *Mathematics in Science and Engineering*, chapter 6. Academic Press, 1979.
- [15] J. M. Mendel. *Lessons in Digital Estimation Theory*, chapter 17. Prentice-Hall, 1987.
- [16] R. R. Murphy, D. Hershberger, and G. R. Blauvelt. Learning landmark triples by experimentation. *Robotics and Autonomous Systems*, 22(3-4):377–392, Dec. 1997.
- [17] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, AC-24(6):843–854, Dec. 1979.
- [18] S. I. Roumeliotis and G. A. Bekey. An extended kalman filter for frequent local and infrequent global sensor data fusion. In *Proceedings of the SPIE (Sensor Fusion and Decentralized Control in Autonomous Robotic Systems)*, pages 11–22, Pittsburgh, PA, Oct. 1997.
- [19] S. I. Roumeliotis, G. S. Sukhatme, and G. A. Bekey. Circumventing dynamic modeling: Evaluation of the error-state kalman filter applied to mobile robot localization. In *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, Detroit, MI, May 10-15 1999.
- [20] S. I. Roumeliotis, G. S. Sukhatme, and G. A. Bekey. Smoother-based 3-d attitude estimation for mobile robot localization. In *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, Detroit, MI, May 10-15 1999.
- [21] S.I. Roumeliotis. Reliable mobile robot localization. Technical Report IRIS-99-374, University of Southern California, April 1999. (<http://iris.usc.edu/irislib/>).
- [22] S.I. Roumeliotis and G.A. Bekey. Segments: A layered, dual-kalman filter algorithm for indoor feature extraction. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2000. (submitted).
- [23] S. Thrun. Bayesian landmark learning for mobile robot localization. *Machine Learning*, 33(1):41–76, Oct. 1998.
- [24] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hhnel, C. Rosenberg, N. Roy, J. Schulte, and D. Shulz. Minerva: A second-generation museum tour-guide robot. In *Proceedings of the 1999 IEEE International Conference in Robotics and Automation*, Detroit, MI, 1999.
- [25] R. Wehner, B. Michel, and P. Antonsen. Visual navigation in insects: Coupling of egocentric and geocentric information. *Journal of Experimental Biology*, 199(1):129–140, 1996.
- [26] D. Wettergreen, H. Pangels, and J. Bares. Behavior-based gait execution for the dante ii walking robot. In *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 274–279, Pittsburgh, PA, 5-9 Aug. 1995.