

# Simultaneous Localization and Mapping with Multi Robot Map Joining

John Downs

Feb 2013

# Contents

<b>1</b>	<b>Introduction: What is SLAM?</b>	<b>1</b>
1.1	Motion . . . . .	3
1.2	Observation . . . . .	6
1.3	probablistic models . . . . .	7
1.4	EKF . . . . .	10
<b>2</b>	<b>Literature Review: State of the Art</b>	<b>17</b>
2.1	Least Squares Estimates . . . . .	20
2.2	Map Joining . . . . .	20
2.3	Spherical Matrices, One Step SLAM and Map Joining . . . . .	22
2.4	SLAM and Machine Learning . . . . .	23
<b>3</b>	<b>Methods</b>	<b>25</b>
3.1	Robotic Data . . . . .	25
3.2	Simulator . . . . .	26
3.3	Experiments . . . . .	27

<b>4</b>	<b>Results</b>	<b>29</b>
<b>5</b>	<b>Discussions</b>	<b>33</b>
<b>6</b>	<b>Conclusions</b>	<b>35</b>
<b>7</b>	<b>Future Work</b>	<b>37</b>

## Abstract

This paper presents an overview of Simultaneous Localization and Mapping. The focus is on getting the reader familiar with the essential concepts of the topic. Key ideas include the robot model, probabilistic interpretations of motion and observation, and map representation. There is a discussion of the Extended Kalman Filter with an aim towards getting the reader ready to make a simple implementation of SLAM. Finally there is a description of the state of the art in SLAM.

The experimental portion of this work focuses on map joining techniques. First, maps are created with an Extended Kalman Filter to provide a baseline of accuracy and performance. Then, a single step map joining approach based on Iterated Sparse Local Submap Joining Filter by Dr. Shoudong Huang, et al is employed and extended to the multi robot case. The multi robot case is shown to be as effective as the single robot case and significantly better than the traditional Extended Kalman Filter algorithm.

# Chapter 1

## Introduction: What is SLAM?

A mobile robot in an unknown environment has two tasks that precede anything else. First, it needs a map of where it has been. The robot must ask 'what does my environment look like.' Second, it must localize itself using that map. The question becomes 'where am I?' In any partially observable environment, these two tasks are inseparable. The two tasks, localization and mapping, are combined into a single problem known as Simultaneous Localization and Mapping, or SLAM. Along with motion planning, these tasks form the basis for all mobile robot navigation. While not discussed in detail here, localization, mapping and planning are applicable to manipulator arm robots as well as wheeled mobile robots.

Robot navigation may seem simple at first glance. Simply measure the location of objects in the environment and move between a pair of points. The problem with this is sensor noise. Every measurement has a limit to its accu-

racy. This introduces error. All sensors can suffer from some systemic errors as well; wheels can slip, cross talk can affect sonar, and so on. The consequence is uncertainty in the position of environmental features and a robot's position. If left unchecked, that uncertainty will grow without bounds and invariably results in the catastrophic failure of any navigation scheme. Effectively dealing with this uncertainty requires some sort of probabilistic model.

Such a model is used to create an estimate of the map and the robot's position. As the robot continues to operate in the environment, it becomes more confident about the shape of that environment. The movement and observation are highly correlated because the observations depend on the position of the robot. Due to that correlation, the certainty about the robot's position also improves over time. The end result is an estimate of the map and a statement about its certainty.

Prior to a discussion about a probabilistic model, a discussion of deterministic robotics is necessary to provide a foundation. Two models will be developed that describe the essential behavior of the robot: the motion model and the observation model. These two models will give a framework for creating a probabilistic interpretation.

## 1.1 Motion

A robot's pose is a combination of the position and orientation, represented as a vector. In the two dimension case, this is represented by the column vector  $[xy\theta]^T$  with position  $(x, y)$  and heading  $\theta$ . The motion model describes the state transition a prior pose at time  $t - 1$  and a new pose at time  $t$  given a control input  $u$ . Several simplifying assumptions are often made. Most models use kinematics, a simple mathematical description of motion, rather than dynamics, which focuses on the forces responsible for motion. In it's simplest form, robot motion can use a particle like model consisting of linear and angular velocity in a plane. This extends to three dimensions without difficulty.

A derivation of the model begins with decomposing the components of motion. There are three components:  $x$  position,  $y$  position and heading  $\theta$ . The control signal  $u$  consists of a linear velocity  $v \frac{m}{s}$  and angular velocity  $\omega \frac{rad}{s}$ . Consider linear velocity  $v$  with a heading  $\theta$ , with an angular velocity of 0 from an initial position  $(x, y)$ . Velocity is the first derivative of position  $r$  with respect to time,  $v = \frac{dr}{dt}$ . The the  $x'$  component is given by  $\int_0^t v \cos \theta dt$ , the  $y'$  component is  $\int_0^t v \sin \theta dt$ . The heading remains unchanged. If the angular velocity  $\omega$  is non-zero and the linear velocity is zero, the change in position is also zero, but the change in heading is  $\theta' = \int_0^t \omega dt$ .

Combining the translational and angular velocities complicates the model. Assume that the robot travels with a constant velocity for a time interval. This

is a realistic assumption for almost all scenarios where the interval is small enough. Rather than moving along a line or on a point, it now moves along an arc  $A$  with a radius  $r$ . The new radius is calculated by:

$$\vec{x}_{t+1} = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} + \begin{pmatrix} -\frac{v}{\omega} \sin(\theta) + \frac{v}{\omega} \sin(\theta + \omega \delta t) \\ \frac{v}{\omega} \cos(\theta) - \frac{v}{\omega} \cos(\theta + \omega \delta t) \\ \omega \delta t \end{pmatrix} \quad (1.1)$$

citeThrunPR

Special consideration must be given to the case of motion in a straight line. If the motion were strictly linear, the angular velocity will be 0, so the formula given in 1.1 would involve division by 0. In this case, the formula is instead:

$$\vec{x}_{t+1} = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} + \begin{pmatrix} x + v \sin \theta \\ y + v \cos \theta \\ \omega + 0 \end{pmatrix} \quad (1.2)$$

Robotic motion can take on a variety of forms, from wheels to legs to flying rotors. The most common drive systems are either car-like or differential drives. Car like systems are very effective for outdoor experiments with uneven terrain, but are more difficult to model. Differential drives are better for indoor experiments and is remarkably simple to model, but has weaknesses that make it difficult to model in sloped environments. The differential drives consist of two wheels separated by an axle. Each wheel is controlled by a different motor



and turns independently on the axle. This configuration allows the robot to turn in a circle either centered on a wheel or the axle.

The translation between the particle model and differential drive model is important. When we think of how something moves, we think of how fast it is moving in a direction and how fast it is turning. It also provides a common format for controls that can be used independently of the wheel configuration. But there needs to be a correspondence between this idea and the actual motion of the wheels on the robot.

The motion of a differential drive comes from the velocity of its two wheels and their relative distance. Consider a differential drive robot traveling along an arc. Rotation is caused by a difference in speed between the right and left wheel that is proportional to the distance between those wheels  $l$ . Linear velocity is simply the average of the two wheel velocities.

$$\omega = \frac{v_r - v_l}{l} \quad (1.3)$$

citeDUDEK

$$v = \frac{v_r + v_l}{2} \quad (1.4)$$

## 1.2 Observation

The observation model describes how the robot's sensors are used to measure its environment. There are many kinds of sensors, but they can be divided first into two categories: those that measure an aspect of the robot state and those that measure an aspect of the environment. We are concerned here with those that measure the environment. Some examples include sonar, laser range finders and cameras. For mapping, it is best to abstract the details of perception into range and bearing measurements. If the robot knows the relative distance and orientation of an object, it is relatively easy to put that into a global frame of reference. Features in the environment are usually represented as a set of points. Other representations, notably geometric and topological, are sometimes used, but will not be considered here [?].

A range sensor is usually fixed on a robot such that we know its orientation. When used, such a sensor will return the distance to the nearest object directly in front of it. The robot then has an estimated range and bearing of an object relative to itself. This is precisely the definition of a polar coordinate. Converting to Cartesian coordinates from a range  $r$  and bearing  $\theta$  is done with:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix} \tag{1.5}$$

We will also frequently need to use the inverse of this model as well. This is

just the conversion from Cartesian coordinates to polar form:

$$\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan(\frac{x}{y}) \end{pmatrix} \quad (1.6)$$

While not strictly a property of the sensor, a correspondence variable  $s$  is usually added to the tail of the observation vector. This is used to identify it as either a landmark that has been seen previously or a new landmark. If the landmarks are not easily identifiable, there are many classification algorithms that will match them. A characteristic algorithm, Maximum Likelihood, will be covered later.

### 1.3 probabilistic models

A naive mapping algorithm would use the motion model to keep track of the robot's position, and the observation model to locate landmarks. The flaw in this approach is that no matter how accurate the sensors are, the error will be additive, and thus grow without bounds. An additional model is needed: a model of the uncertainty the robot has about its position and the position of environment features. The structure developed is known as a stochastic map [?]. It consists of an estimate of the robot's pose and surrounding landmarks along with a covariance matrix describing the confidence in that estimate. As a robot travels through an environment making observations, the correlation

between landmarks increases, converging towards an accurate map.

Consider a pair of features, each some uncertain distance from the robot. The sensors will measure the distance as some value with a degree of error. The robot then moves to another position and again observes these two landmarks. This second observation reinforces the relationship between the landmarks, reducing the uncertainty of their relative distance. This reduction in uncertainty comes because the measurement error is inherent in the robot, so all the landmarks share the same error. As a network of relationships between landmarks and robot poses grows, the map tends towards an accurate relative representation [?]. This increasing certainty in landmark position allows the robot to accurately determine its own position given an observation of a feature it has seen before.

Movement is described by:

$$p(x_t | x_{t-1}, u_t) \tag{1.7}$$

This gives the probability of the robot being at a particular pose  $x_t$  given its last pose at time  $t - 1$  and a control input,  $u$ . A possible source of confusion is the use of  $x$  as the pose variable;  $x$  is a vector describing the position and heading angle,  $[x, y]$ . This is separate from the  $x$  of the Cartesian coordinates of the robot and should be clear from context.

Observation is described by:

$$p(z_t|x_t) \tag{1.8}$$

This model gives the probability that there is something at position  $z$ , that was correctly observed given the current pose of the robot,  $x$ , at some time  $t$ . Usually the robot will want to also keep track of landmarks it has seen, so it is common to add a map  $m$ .

$$p(z_t|x_t, m_t) \tag{1.9}$$

These probabilistic models allow us to use a predictor-updater style algorithm known as the Recursive Bayesian Estimation. Based on a prior state estimation, plus new information from observation and a control signal, a new state estimate is formed.

---

**Algorithm 1** Recursive Bayesian Estimation

---

**Require:** Prior estimate  $prior(x_{t-1})$   
Control signal  $u_t$   
Observation  $z_t$   
**for**  $i = 0 : t$  **do**  
     $prediction(x_t) := \int p(x_t|u_t, x_{t-1})prior(x_{t-1})dx_{t-1}$  ▷ Prediction Step  
     $correction(x_t) := \eta p(z_t|x_t)prediction(x_t)$  ▷ Correction Step  
**end for**  
**return**  $correction(x_t)$

---

Intuitively, the Bayes Filter involves the robot moving to a position and guessing where it is. It then looks around and used that information to improve the guess. More rigorously, the filter begins with an initial estimate of the state

$x_0$ , which is used as the base prior. The state is represented by a probability density function. The filter also requires a control signal  $u_t$  and observation  $z_t$ . The first step of the for loop predicts the new state of the environment based on the robot's motion. In a static environment, only the pose should change. This estimate includes a belief about the position of landmarks relative to the robot. It's inclusion in the next step tells us how much information is provided by the new observations, which is the probability that an observation  $z$  was made given the likely state  $x_t$ . The new information is used to create a better estimate of the state. This is repeated until the system stops.

Because the two steps in the for loop of the Bayes Filter do not usually have a closed form, it's not very useful in practice [?]. A number of algorithms approximate the Bayes filter. The most commonly used is the Extended Kalman Filter or EKF. The EKF is similar to the simpler Kalman Filter, with the addition of a linearization step that makes it applicable to non-linear systems.

## 1.4 EKF

In order to implement the EKF, we need to begin with data structures. First, we assume that the noise in the motion and observation models is Gaussian. While this is not necessarily true, it does allow for the relative simplicity of the EKF and a tractable estimate is often preferable to a more precise estimate. We will also assume discrete time steps with a constant interval. The mean state

estimate  $x$  is a column vector that combines the robot's pose at the current time  $t$   $R_t = \langle x_r, y_r, \theta_r \rangle^T$  and the position of all the observed features  $f_i \in F$ ,  $f_i = \langle f_i x, f_i y \rangle^T$ . It's size is  $1 \times N$ , where  $N$  is the length of the pose and the combined length of the features. The covariance  $P$  describes the confidence in each landmark's position and has a size  $N \times N$ .

The EKF is built on the Kalman Filter (KF) with an additional linearization step to allow for estimation involving non-linear motion and observation models. We will begin with a derivation of the Kalman Filter and then the changes necessary to create an Extended Kalman Filter.

Kalman Filters are an approximation of the Bayes Filter. A pair of models are used; the process model and the observation model. These correspond to the motion and measurement models for a mobile robot. The generalized process model is:

$$x_{t+1} = Fx_t + Gu_t + \epsilon \quad (1.10)$$

$$P_{t+1} = FP_tF^T + Q \quad (1.11)$$

Here, the state at time  $t+1$ ,  $x_{t+1}$  is determined by the state transition matrix  $F$  applied to the prior state estimate  $x_t$  plus the control signal  $u$  modified by some gain matrix  $G$  along with some Gaussian noise  $\epsilon$  with a mean of 0 and covariance  $Q$ . Just like the Bayes Filter, this step makes the best initial guess of

the state at the next instance. This is then corrected by using the observation model. That model uses an observation  $z$  at time  $t + 1$ :

$$z_{t+1} = Hx_{t+1} + \gamma \quad (1.12)$$

Here,  $H$  is a matrix describing the observation process and  $\gamma$  is Gaussian noise with a mean of 0 and covariance  $R$ .

Prior to the update step, two new values are needed. The first is the innovation covariance  $S$  and the second is the Kalman gain  $K$ . Innovation is a distribution indicating the difference between the expected observation in EQUATION and the actual measurement. It is how much information a new observation provides for the filter, modified by a gain which is proportional to the current confidence in the estimate of a feature. The formulas for  $S$  and  $K$  respectively are:

$$S = HP_{t+1}H^T + R \quad (1.13)$$

and

$$K = P_{t+1}H^TS^{-1} \quad (1.14)$$

The update step uses this observation value to make a new state estimate:



$$\hat{x}_{t+1} = x_{t+1} + K(z_{t+1} - Hx_t + 1) \quad (1.15)$$

$$\hat{P}_{t+1} = P_{t+1} - KSK^T \quad (1.16)$$

The trouble with the Kalman Filter is that it is only applicable to linear Gaussian processes, and the motion and observation models for a mobile robot are certainly not linear, due to the presence of transcendental functions. A modification is necessary to use the Kalman Filter for SLAM. Rather than a state transition matrix and an observation matrix, the Extended Kalman Filter uses two differentiable non-linear functions.

$$x_{t+1} = f(u_t, x_t) + \epsilon \quad (1.17)$$

$$z_{t+1} = h(x_{t+1}) + \gamma \quad (1.18)$$

In the prediction step, the initial state estimate is simply the result of  $f$ . But in order to compute the covariance  $P_{t+1}$ ,  $f(u, x)$  needs to be linearized. This is done by computing the Jacobian matrix  $\nabla f$ , which is the total derivative of  $f$  evaluated at  $x_t, u_t$ . The resulting function is a plane tangential to  $f$  that is a reasonable approximation. The estimate for  $P_{t+1}$  is then calculated by:

$$P_{t+1} = \nabla f P_t \nabla f^T + R \quad (1.19)$$

The observation model is linearized in the same way, with the resulting Jacobian matrix  $\nabla h$ .

The Innovation matrix and Kalman gain are calculated by:

$$S = \nabla h P_{t+1} \nabla h^T + R \quad (1.20)$$

$$K = P_{t+1} \nabla h^T S^{-1} \quad (1.21)$$

The final state estimate and covariance are the same as the Kalman Filter.

The complete EKF algorithm is as follows in Matlab.

```
function [x P] = EKF(x0, P0, u, z, R, Q)
    x_est = f(u, x0);
    P_est = F * P0 * F' + R;
    K = P_est * H' * S^-1 + Q;
    x = x_est + K * (z - h(x_est));
    P = P_est - K * S * K';
end
```

Some additional work is necessary to do SLAM with the EKF. To begin, odometry and sensor data is necessary. It is convenient to have a simulator to create this data.

```
% Config
landmarks = 10;
waypoints = 100;
map_radius = 30;
```

```

dt = 1;
close = 1;

motion_noise = .001;
observation_noise = .01;

sensor_arc = pi/4;
sensor_range = 10;

% Initialize
start_pose = [0;0;0];
start_time = 0;
map_true = create_landmarks(map_radius, landmarks);
waypoint_list = create_waypoints(map_radius, landmarks);

mapFig = figure(1);
axis([-map_radius map_radius -map_radius map_radius])
axis square
LG = line('parent',gca,...
    'linestyle','none',...
    'marker','o',...
    'color','r',...
    'xdata',map_true(:,2),...
    'ydata',map_true(:,3));

WG = line('parent',gca,...
    'linestyle','none',...
    'marker','x',...
    'color','b',...
    'xdata',waypoint_list(:,2),...
    'ydata',waypoint_list(:,3));

R = line('parent',gca,...
    'linestyle','none',...
    'marker','+',...
    'color','r',...
    'xdata',[],...
    'ydata',[]);

pose = start_pose;

odometry = [0 0 0]; % t v w

```

```

time = 0;
observations = [get_visible_landmarks(map_true, pose, sensor_range, sensor_arc, ob

for ii = 1:waypoints
    wpt = waypoint_list(ii,2:3);
    pose_old = pose;

    while pdist([wpt;pose(1:2)'], 'euclidean') > close;
        delta_a = atan2(wpt(2) - pose(2), wpt(1) - pose(1)) - pose(3);
        if abs(delta_a) > 10 * eps
            u = [0 rotate(pose, wpt(1:2), dt)];
        else
            u = [velocity(pose(1:2)', wpt(1:2), 1) 0];
        end
        time = time + dt;
        [odo, pose] = move_robot(pose, u, dt, motion_noise);
        obs = get_visible_landmarks(map_true, pose, sensor_range, sensor_arc, obse
        odometry = [odometry; time odo'];
        observations = [observations; obs];
        set(R, 'xdata', pose(1), 'ydata', pose(2));

        drawnow;
    end
end
end

```

TODO: Show EKF SLAM TODO: Show EKF Inconsistency

## Chapter 2

# Literature Review: State of the Art

Square Root Smoothing and Mapping ( $\sqrt{SAM}$ ) is a full SLAM algorithm developed by Frank Dellaert and Michael Kaess that represents a break from earlier filtering approaches. Rather than filtering, it uses statistical smoothing methods over the robot trajectory. Smoothing over the entire trajectory turns out to be more efficient in practice than the EKF once the number of features exceeds 600 [?]. Where filtering assumes that the current state is complete at each step when it is actually a linearized estimate, smoothing uses each pose in the trajectory to help calculate the most likely state.

Incremental Smoothing and Mapping (iSAM) is an extension of  $\sqrt{SAM}$  that takes advantage of sparseness inherent in the full SLAM formulation to make only incremental updates as necessary [?]. The incremental nature of the algorithm means it can handle data as it comes in, which allows it to be run on an actual robot in service. iSAM does SLAM by formulating each step as a least squares problem. The result is the error in the map is a linear function that can be solved by the Gauss-Newton optimization method.

iSAM is often used as a map optimization tool for graph based SLAM ap-

plications [?]. A mobile robot that is operating in an environment can add nodes and edges to a graph, just as in GraphSLAM. Once the map is needed for planning, iSAM can be run on the graph, even if the graph is extremely large, and return a result in a matter of seconds. While this is still not quite fast enough to make some time critical decisions, it is a great leap forward for SLAM algorithms.

Iterated Sparse Local Submap Joining (I-SLSJF) is yet another approach that tries to reduce the complexity of SLAM. In this case, the reduction comes from dividing the problem into a series of overlapping submaps [?]. Where graph based SLAM first estimates a graph of poses, submap joining works more like traditional SLAM in that it focuses on the estimation of feature locations. As a side effect of the submap joining process, a coarse estimate of the robot trajectory is produced. But because the entire trajectory is not maintained, the number of dimensions in the problem is greatly reduced when compared to other batch SLAM solutions.

Submaps are usually produced with an EKF or other simple estimation technique and consist of a single step with a start and end pose in a local frame of reference. The end pose of a map is always the start pose of the next map in order to facilitate the fusing process. The initial version of the algorithm uses an information filter to match observations shared between maps. This filtering can still suffer from some linearization errors, so if an inconsistency is detected at any step, it can apply smoothing to the global map to recover from this inconsistency.

Another area of very active research is cooperative multi-robot mapping [?][wang2007multi) [?][multiSEIF), [?][4399142), [?][4543634), [?][5509154). Outside of SLAM research, robot groups are popular because of improved redun-

dancy and the obvious benefits of being able to execute a task in a distributed manner. Often multi-robot systems consist of lower cost components because the group as a whole has a higher fault tolerance than any individual member. If an individual fails, the impact on the completion of the overall goal is mitigated. This can be useful in situations where the individual chance of failure is relatively high, or the cost of failure is high.

There are two challenges associated with multi-robot SLAM: distinguishing robots from landmarks and the determination of a shared frame of reference. Mistaking a robot for a landmark and adding it to the map can lead to an extremely inconsistent landmark. If a robot in the team is observed and classified as a landmark in one location and then observed again at another location, the observer may conclude that a loop has been closed, when this has in fact not occurred. The simplest way to cope with this is for the team to have a priori knowledge of the other members and devise some way to uniquely distinguish them from the environment, such as a barcode or unique pattern of infrared flashes. The other alternative is to make the SLAM implementation robust in a dynamic environment, but this comes with unique data association challenges.

Determination of a shared reference frame can be done by sharing maps when a mutual pose observation occurs between two or more robots. Once a robot determines that it has observed another robot, it can communicate with the other team member and share its map and current pose estimate. In the map thus shared, there will be an estimate of the position of the team member. This point can then be used like the shared start end end poses with I-SLSJF SLAM. Once the location of the two robots is determined, it becomes possible to calculate the correct rotation and translation vectors to align the shared map with the robot's local map.

## 2.1 Least Squares Estimates

The key to the performance of full SLAM solutions is the least squares formulation. The objective is to minimize the Mahalanobis distance between the predicted state and combined prediction and observation. Mahalanobis distance is a measure of similarity for two data sets. Applied to SLAM, this estimate uses the entire robot trajectory and observations to date to calculate a total state estimate. Because it uses the entirety of the data available, it is called smoothing, to contrast it with filtering, which only uses the current estimate, rather than all historical data. Smoothing finds a single state estimate that best fits all the available data. The Least Squares formulation of SLAM is to minimize:  $f(x_t - 1, u)^T P f(x_t - 1, u) + h(z)^T R h(z)$  where  $f()$  is the motion model with covariance  $P$ , and  $h()$  is the observation model with covariance  $R$ ,  $u$  is the set of control inputs and  $z$  is the set of observations. Many methods exist for solving linear least squares problems, but the motion and observation models are almost never linear in practice. When a linear least squares problem is at hand, there is always a closed form solution, thus it can be solved in a single step. However, because SLAM is non-linear, it requires iterative methods to solve. The reason for this is that no methods exist for solving these types of problems, so we must search for a solution. A number of machine learning algorithms exist to solve these problems, both deterministic and stochastic.

## 2.2 Map Joining

The map joining approach to SLAM relies on the creation of submaps: local maps focused only on a subset of a trajectory and the immediately related observations. The creation of local maps is usually accomplished by either Ex-



tended Kalman or Information Filters. Markov Chain Monte Carlo methods have also been suggested . While these methods prove to be inconsistent, this inconsistency is only significant in large sets of observations. By limiting their scope, linearization errors, inconsistency and complexity can be held in check. The earliest paper I have found on map joining is [?]. This describes the fundamental operations of map joining: transformation from a common observation into a global coordinate frame and feature association. More recent formulations [?] have eliminated the need for explicit transformation through the use of a graph theoretic formulation of SLAM. In this case, the more general term “map alignment” is used over transformation. No matter how the maps are aligned, the second step is data association. Because of possible noise in the observation of a common landmark, alignment may not place all landmarks at the same point. This requires a classification algorithm to be run over the map, such as k-Nearest Neighbor or Joint Compatibility Branch and Bound [?]. Other classifiers can be used, but are not common in the literature. Classification can be eliminated if noiseless identification of features is possible, such as when using cameras and unambiguous barcodes. If there are common landmarks between the two submaps, after classification, a new state estimate is required to make sense of the matched but non-coincident features. This is accomplished by calculating a least squares estimate of the new global map, to create a best fit “curve” describing all the observations.

## 2.3 Spherical Matrices, One Step SLAM and Map Joining

A key observation to the reduction of dimensionality in single step SLAM is the use of spherical covariance matrices. A spherical matrix is defined as any matrix that is commutative with a rotation matrix. A rotation matrix  $R(\theta)$  is  $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ . For all  $\theta$  and any spherical matrix  $A$ ,  $AR = RA$ .

It is important that the covariance matrices are also positive definite. A matrix is positive definite if for all positive, non-zero column vectors  $v$ ,  $v^T M v > 0$ . This is to allow for methods analogous to finding the square-root of a matrix, such as Cholesky or QR factorization to be used in the solution of the least squares estimate. In the Automatica preprint, Dr. Huang claims that this objective function, when applied to single step SLAM, is equivalent to a one dimensional optimization problem. Through repeated application of single step SLAM for each timestep, an approximate solution to the previous objective function can be easily found. It is approximate because a spherical covariance matrix is used, rather than the actual covariance associated with observation uncertainty. In On The Num of Local Minima, Dr. Huang shows that single step SLAM and map joining SLAM share the same property of having only 1 or 2 minima, one of which is global, when covariances are approximated by spherical matrices. This provides the benefits of the reduction in linearization error due to map joining, along with the low dimensionality of single step SLAM. Additionally, if multi-robot map joining belongs to this class of problems, map sharing can possibly become quite efficient. In a remark from [?], Huang notes that the covariance matrices must only be spherical, but not identical. That is, the covariance can differ for each odometry measurement and observation. This

lends some home to the possibility that multi-robot map joining is in this class of problems.

## 2.4 SLAM and Machine Learning

There are two sub-problems within SLAM that call for the application of machine learning techniques, landmark association and least squares optimization. These are two separate types of problems, the former being unsupervised classification, the latter is convex optimization.

Landmark association, also known as data association, is a classification problem that uses a pair of feature maps or a map and set of observations and tries to match known landmarks with new observations. In most cases, this is an unsupervised learning problem because the set of features can vary so greatly from map to map, it is not possible to provide examples for a supervised approach. This is necessary for any environment where there can be ambiguity in landmarks. While it might be possible in a lab to put barcodes on landmarks that can be recognized with a camera, in scenarios where a camera might not be available or barcoding landmarks unfeasible, landmark association is required. The least squares portion of SLAM is non-linear and of a very high dimension. Because of this, single step techniques such as linear regression are not applicable. Other algorithms such as Levenberg-Marquardt, Gauss-Newton and Stochastic Gradient Descent must be used instead. These three methods are the most popular among SLAM researchers.

The Gauss-Newton method is used effectively in iSAM [?]. It is a variation on Newton's method for finding the minima of a function taught in elementary calculus courses but is modified to solve least squares problems. It begins with

an initial guess  $x_0$  of a possible solution. For an objective function  $f(x)$  and a residual function  $r(x)$ , the jacobian at  $r(x_0)$  is calculated. The jacobian is a linearization estimate of  $r(x)$  and is used for the next guess. This process is repeated until it converges on a solution. It is possible in certain situations to overshoot the optimum and fail to converge. This algorithm can also fail to find a global optimum and instead converge on a local optimum.

Gradient descent is an optimization algorithm similar to hill climbing, but rather than selecting a single element to improve, it follows the slope (or gradient) of the function towards a solution. Like Gauss-Newton, gradient descent can get stuck in local optima. A modification known as stochastic gradient descent can overcome this limitation.

# Chapter 3

## Methods

### 3.1 Robotic Data

The experiment uses the UTIAS multi-robot data set. The data consists of odometry as linear and angular velocities, and feature observations as range and bearing relative to the center of the robot along its line of motion. This was produced by six iRobot Creates, each with a monocular camera mounted on the center of the robot along the local x-axis. Each robot was marked by a vertical barcode to allow for easy identification of other agents.

The landmarks were posts with barcodes similar to those on the robot chassis. The number of landmarks varied between  $N$  and  $M$ , and were arranged in different configurations for each system. A total of nine sets are provided.

Groundtruth for landmarks was provided, along with groundtruth for the robots at each measured time. Groundtruth was gathered by an accurate object tracking system with an approximate accuracy of 1 mm.

The data set provided a script for sampling from the data sets and interpolating to provide a consistent interval. Early work with this data showed that

very small intervals too much data for the available disk space and processing power.

The data as available was preprocessed to allow for easy integration with backend slam algorithms. Range and bearing information was extracted from the images. Images were provided but not used. Odometry information was converted from differential drive form to the particle model.

Prior to experimentation, some additional work was necessary. In some data sets there were erroneous measurements that needed to be removed. The data was sampled and interpolated into one second intervals.

## 3.2 Simulator

In order to test the map joining algorithm with a large collection of features, a point mass robot simulator is used. The simulator creates a set of uniformly random landmarks and uniformly random waypoints within a particular radius. The size of the sets and radius are configurable. The sensor is a range and bearing sensor with a range of 10 meters and covers an arc of  $\pi/4$  radians.

The simulator starts with an initial observation before any motion. The robot rotates towards the first waypoint and then heads towards it with a maximum velocity of one meter per second. Once the robot reaches the waypoint, the process repeats until the robot reaches the final landmark, when the simulation ends. Observations occur every second, immediately after the motion step. The resulting odometry and observation data is formatted identically to that used in the UTIAS data set. The simulator was run thirty times each for landmark counts of 10 and 100.

### 3.3 Experiments

The first experiment was to create single robot maps using just the EKF algorithm from THRUN. The algorithm assumes known and correct, noise free data association. while not true in most realistic operating environments, the simplifying assumption allows map quality to be assessed without confounding factors. The results from this experiment are the baseline for those that follow. Accuracy of the estimation is calculated for both the map and the trajectory using the root mean squared error (RMSE) technique to measure how close the estimate was to the available groundtruth. The average RMSE of all runs in both the UTIAS data set and all the runs in the simulated data set is calculated and used as the benchmark.

The next experiment was single robot map joining. The purpose of this experiment is to record the accuracy improvements of the map joining approach over the EKF alone. Using the EKF, a submap was saved at each time step. The submap is just a snapshot of the current state estimate. Prior to prediction phase, the state estimate is cleared of all landmarks to preserve the independence of each submap. Only the mean estimate of features and the ending pose are saved. The covariance is discarded because it will be estimated by an identity matrix during the map joining phase. This results in a space savings of 144 Kb for each submap with a state vector of 48 elements, assuming 64 bit doubles. Map joining was done offline after all submaps were collected.

Offline map joining was accomplished by the algorithm described in cite-HUANG. The algorithm keeps track of the global map state and a single robot's trajectory. Map joining is accomplished by a by finding the state  $X_{join}$  that satisfies the least squares problem

$$\underset{X_{join}}{\operatorname{argmin}} \sum_{j=1}^k (\hat{X} - H(X_{join})^T) P^{-1} (\hat{X} - H(X_{join})) \quad (3.1)$$

Optimization is done with the Matlab optimization toolbox Cite, using the Quasi Gauss Newton algorithm Cite. This algorithm was selected because it is the simplest optimization method for a function with a single minimum. While this may fail in the case of two local minima, that will only occur if data association is very poor. Because the maps all have completely accurate data association, that condition will not occur. Accuracy for the trajectory and map are determined by comparing the resulting estimates to groundtruth and computing the RMSE, just as with the EKF.

The final experiment was multi robot map joining. It followed the lines of single robot map joining, except when another robot was observed. At that time step, the local maps for each robot were exchanged. Singular value decomposition was used to determine the ideal rotation matrix and translation vector. After applying that transformation, the two maps are joined using the same function as single robot map joining, except trajectory is ignored. Optimization is also done with the Quasi Gauss Newton algorithm. Additional, optimization is verified using a genetic algorithm as a global optimization technique. If the shared map joining least squares problem has only one local minima, the resulting estimate will be similar.



# Chapter 4

## Results

The first experiment, single robot EKF, had nine data sets with five robots each, for a total of forty-five maps that were created. The aggregate Root Mean Squared Error between the map and groundtruth is 0.4004 m with a standard deviation of 0.0806 m. The details for each run are shown in table 4.1. The aggregate RMSE of the trajectory estimate and groundtruth is 0.0462 m with a standard deviation of 0.0074 m.

The second experiment, single robot map joining has an aggregate RMSE of 0.0347 meters and a standard deviation of 0.0164 meters. As with the EKF experiment, this was done with nine data sets each with five robots, for a total of forty-five maps. The mean RMSE of the trajectory compared to groundtruth is  $7.7\text{e-}5$  meters with a standard deviation of  $4.9\text{e-}5$  meters. A paired T test of the hypothesis that the mean RMSE of the first experiment's map is equal to the RMSE of the second experiment's map with 95% confidence gives a p-value of  $3.04\text{e-}4$ .

The third experiment, multi-robot map joining has an aggregate mean RMSE of 0.0467 meters and aggregate standard deviation of 0.0240 meters. Trajectory

was not considered in this experiment. A paired T test of the hypothesis that the mean RMSE of the single robot map joining experiment is equal to the mean RMSE of the multi-robot map joining experiment at a 95% confidence interval gives a p-value of 0.1018.

Table 4.1: Root Mean Square Error of Single Robot EKF Maps in Meters

Set	Robot 1	Robot 2	Robot 3	Robot 4	Robot 5	Average RMSE	Std Dev
1	0.3304	0.4240	0.2255	0.3733	0.2437	0.3194	0.0844
2	0.3990	0.1655	0.6744	0.3071	0.5666	0.4225	0.2026
3	0.0599	0.4383	0.7593	0.5151	0.4371	0.4419	0.2510
4	0.2483	0.3933	0.5499	0.3501	0.2237	0.3531	0.1305
5	0.1555	0.3135	0.3689	0.5330	0.3344	0.3411	0.1350
6	0.6738	0.5618	0.5167	0.5235	0.5974	0.5746	0.0642
7	0.4676	0.1768	0.2946	0.5613	0.3348	0.3670	0.1503
8	0.2980	0.2922	0.1912	0.7909	0.1180	0.3381	0.2640
9	0.1874	0.7517	0.5520	0.5976	0.1405	0.4458	0.2683
RMSE	0.3133	0.3908	0.4592	0.5058	0.3329	-	-
Std Dev	0.1840	0.1854	0.1994	0.1480	0.1727	-	-

Table 4.2: Root Mean Square Error of Single Robot EKF Trajectory

Set	Robot 1	Robot 2	Robot 3	Robot 4	Robot 5	Average RMSE	Std Dev
1	0.0393	0.0353	0.0399	0.0450	0.0419	0.0403	0.0036
2	0.0526	0.0352	0.0467	0.0362	0.0455	0.0433	0.0074
3	0.0391	0.0381	0.0620	0.0602	0.0347	0.0468	0.0132
4	0.0437	0.0412	0.0412	0.0516	0.0479	0.0451	0.0045
5	0.0281	0.0300	0.0327	0.0318	0.0368	0.0319	0.0033
6	0.0563	0.0557	0.0634	0.0609	0.0474	0.0567	0.0061
7	0.0563	0.0452	0.0418	0.0755	0.0466	0.0531	0.0136
8	0.0391	0.0381	0.0620	0.0602	0.0347	0.0468	0.0132
9	0.0415	0.0672	0.0471	0.0595	0.0420	0.0514	0.0114
RMSE	0.0440	0.0428	0.0485	0.0534	0.0422	-	-
Std Dev	0.0094	0.0117	0.0013	0.0137	0.0050	-	-

Table 4.3: Root Mean Square Error of Single Robot Joined Maps in Meters

Set	Robot 1	Robot 2	Robot 3	Robot 4	Robot 5	Average RMSE	Std Dev
1	0.0501	0.0055	0.0110	0.0164	0.0146	0.0195	0.0176
2	0.0268	0.0031	0.1410	0.0312	0.1102	0.0625	0.0596
3	0.0044	0.0296	0.0647	0.1097	0.0123	0.0441	0.0434
4	0.0345	0.0212	0.0046	0.0008	0.0302	0.0183	0.0150
5	0.0308	0.0497	0.0680	0.0140	0.0609	0.0447	0.0222
6	0.0337	0.0220	0.0054	0.0000	0.0310	0.0184	0.0151
7	0.0337	0.0220	0.0054	0.0000	0.0310	0.0184	0.0151
8	0.0419	0.0484	0.0004	0.1284	0.0114	0.0461	0.0502
9	0.0246	0.0723	0.0910	0.0049	0.0088	0.0403	0.0390
RMSE	0.0312	0.0304	0.0435	0.0339	0.0345	0.0347	0.0308
Std Dev	0.0119	0.0212	0.0473	0.0467	0.0308	0.0155	0.0164

Table 4.4: Root Mean Square Error of Single Robot Smoothed Trajectory in Meters

Set	Robot 1	Robot 2	Robot 3	Robot 4	Robot 5	Average RMSE	Std Dev
1	1.3e-4	2.8e-5	3.5e-5	1.4e-4	6.1e-6	6.7e-5	6.2e-5
2	1.1e-4	5.1e-5	7.1e-6	1.2e-4	3.7e-5	6.5e-5	4.8e-5
3	1.4e-4	6.2e-6	2.0e-5	2.1e-4	2.4e-5	8.1e-5	9.1e-5
4	1.5e-4	1.5e-5	7.0e-5	1.6e-4	2.2e-5	8.4e-5	7.0e-5
5	5.8e-6	9.1e-6	6.2e-6	2.3e-5	2.6e-5	2.4e-5	2.1e-5
6	3.6e-5	7.5e-5	1.8e-4	3.0e-5	1.8e-4	2.2e-4	1.1e-4
7	2.0e-5	9.4e-5	1.7e-5	1.5e-4	6.0e-6	9.2e-5	8.2e-5
8	6.8e-7	1.5e-5	2.2e-5	2.8e-6	2.3e-5	1.4e-5	9.0e-6
9	1.6e-6	1.4e-5	1.0e-4	1.3e-5	7.6e-5	4.4e-5	4.2e-5
RMSE	6.6e-5	3.4e-5	5.1e-5	9.4e-5	4.5e-5	7.7e-5	6.0e-5
Std Dev	4.7e-5	3.7e-5	6.8e-5	5.2e-5	8.3e-6	7.0e-5	4.9e-5

Table 4.5: Root Mean Square Error of Multi Robot Map Joining in Meters

Set 1	Set 2	Set 3	Set 4	Set 5
0.0480	0.0617	0.0254	0.0392	0.0726
Set 6	Set 7	Set 8	Set 9	-
0.0103	0.0335	0.0423	0.0877	-
Mean	StdDev			
0.0467	0.0240			

# Chapter 5

## Discussions

The performance of the EKF for single robot mapping is poor relative to the map joining approach. For comparison, the diameter of an iRobot Create is approximately 0.35 meters and the EKF error was 0.40 meters. Upon visual inspection, it appears that the maps may be similar but require a two-dimensional transformation to accurately represent groundtruth. Of course, this is not possible to determine that transformation if groundtruth is not available.

The two possible explanations for this are an inaccurate initial pose and linearization errors. In the case of the experiment using the UTIAS data set, an accurate initial pose for each robot was available. The discrepancy must be due to linearization errors then. This, combined with the overconfidence of the EKF means the robot cannot recover from this error and will continue to converge on an inaccurate map.

The accuracy is acceptable, considering the very slow monocular camera and imprecise odometry on the Create platform, for tasks where a Create might be used. But it turns out that drastic improvements are realized when using the map joining approach. For these experiments, the resulting accuracy improves

by an order of magnitude. This result is within the radius of the robot's chassis, which is a nice first pass heuristic for the expected accuracy. The Quasi-Newton algorithm has a complexity of  $O(n^2)$  [?] for  $n$  dimensions. This is less than the  $O(n^3)$  of the EKF, so it does not add any additional complexity to the process. Additionally, the optimization step only took 2.98 iterations on average and never more than 3 iterations.

Multi-robot map joining is not significantly different from single robot map joining. The lack of accuracy improvement is likely to be due to the map joining being a smoothing approach and thus already an ideal estimator [?]. While there is no improvement in mapping accuracy, there is also no degradation when map data is shared in a homogenous swarm. Trajectory is not considered when sharing maps because tracking is not considered. Without modification, this algorithm cannot improve on trajectory estimation because it relies on landmarks being stationary over time.

# Chapter 6

## Conclusions

Multi Robot map joining using least squares optimization works very well in a homogenous swarm of robots. While some additional processing is required over the traditional EKF SLAM approaches, it is a small constant factor and does not add any complexity. That small factor gives strong gains in mapping accuracy. Compared to the time required for the robot to make observations, the method shown here is fast enough to be used in real time, even in an interpreted language such as Matlab.

Another feature of this approach is the uniform method used for local smoothing and map sharing. This method's dual use means the robot's navigation system has, metaphorically, less moving parts to fail. This is unique in the literature involving map joining SLAM [?] Other map joining schemes use a separate algorithm for creating individual maps and joining shared maps.

Finally, multi robot map joining is an iterative process that can be picked up any time as resources are available. It can also be paused and picked up again at any time, without adversely affecting the results. The tradeoff is the space required for submap storage. While non-submap techniques only require

a single point for each map feature, submapping requires a point to be stored for each observation. However, even the Mars rover, Curiosity has 2GB flash memory [?], which is sufficient for the storage of several large submaps prior to joining. The need for space is tempered by the estimate of the covariance matrix. Because of this the submaps need only store the estimate and not the covariance.



# Chapter 7

## Future Work

Numerous tasks exist for future work implied by this thesis. The most obvious is to extend this system to handle unknown data association. While data association is a well studied subject, it adds considerable complexity to the system. Three dimensional mapping is also important for the usefulness of this system. Work is in progress showing that map joining in three dimensions has at most two local minima [?]. This would require an efficient stochastic optimization method to replace the Gauss-Newton method currently used.

Trajectory tracking is another important area. Having a multi robot system provides an opportunity to add additional trajectory data, but this also would add a great deal of complexity to the system. It is not clear that this can be done within the current framework and may require additional algorithms.

It is unknown whether multi-robot map joining requires homogenous robots. Work on sensor fusion and learning the accuracy of various robots on a heterogeneous swarm could help answer this question.

Finally, an actual on-line system needs to be developed. The current system works with the data after it has been collected and does not work with the data

as it arrives in real time streams. While I believe this is an accurate model of the problem, it is not currently usable on robots in real time.