

# Object Localization and Tracking in 3D

Vidya Viswanathan

A Thesis Submitted to  
Indian Institute of Technology Hyderabad  
In Partial Fulfillment of the Requirements for  
The Degree of Master of Technology



भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad

Department of Electrical Engineering

July, 2015

## Declaration

I declare that this written submission represents my ideas in my own words, and where ideas or words of others have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

Vidya

(Signature)

\_\_\_\_\_  
(Vidya Viswanathan)

EE13M1007

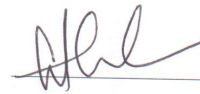
(Roll No.)

## Approval Sheet

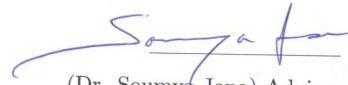
This Thesis entitled Object Localization and Tracking in 3D by Vidya Viswanathan is approved for the degree of Master of Technology from IIT Hyderabad



(Dr. Ketan P. Detroja) Examiner  
Department of Electrical Engineering  
IIT Hyderabad



(Dr. Sumohana Channappayya) Examiner  
Department of Electrical Engineering  
IIT Hyderabad



(Dr. Soumya Jana) Adviser  
Department of Electrical Engineering  
IIT Hyderabad

## Acknowledgements

I would like to place on record, profound gratitude for those who made the successful completion of this project possible. I would like to thank my advisor Dr. Soumya Jana for his able guidance, constant encouragement, suggestions and comments. Without his guidance this project could not have been completed successfully.

I would like to thank my peers who have been a constant support and have provided an enormous amount of help in the completion of this project. Specifically, I would like to thank Kiran Kumar for his insightful suggestions and comments on various topics related to the thesis. I would also like to thank Ranjith Kumar, G. Mamatha and Roopak R. Tamboli for helping me with the generation of data required for the thesis .

I would like to express my sincere gratitude to Dr. Sumohana Channappayya, Dr. Ketan P. Detroja and Dr. Shanti Swarup from Uurmi Systems Pvt. Ltd., for sharing their invaluable suggestions and views from time to time.

# Dedication

To my family and friends who have been part of my learning experience

## Abstract

The field of Computer Vision has repeatedly been recognized as an intellectual frontier whose boundaries of applicability are yet to be stipulated. The work attempts to demonstrate that vision can achieve an automatic localization and tracking of targets in a 3D space. Localization of targets has gained importance in the recent past due to the myriad of applications it plays a significant role in. It is analogous to detection of objects in a video sequence in the image processing domain. This work aims to localize a target based on range measurements obtained using a network of sensors scattered in the 3D continuum. To this end, the use of the biologically inspired particle swarm optimization (PSO) algorithm is motivated. In this context, a novel modification of PSO algorithm is proposed that leads to faster convergence, and eliminates the flip ambiguity encountered by coplanar sensors. The initial results over several simulation runs highlight the accuracy and speed of the proposed approach.

Among various factors influencing the performance of localization system, including quality of sensing devices and strength of the signal, the relative position of the sensors has a significant effect. In this paper, our goal is to optimize the positions of the sensors in order to localize an unknown target with maximum accuracy. In order to minimize the variance of the localization error, a statistical measure provided by Cramer Rao Lower Bound which relates the variance of the error to the Fisher Information is incorporated in this framework. The computational advantage of Fisher Information method combined with the accuracy of the use of variance based objective function helps in providing the optimal position of sensors. The proposed technique is validated using practical experiments that provide the optimal position of sensors on a terrain to localize a target present anywhere on the road.

Localization serves as a precursor to all the tracking applications. Tracking involves prediction, data association and estimation of motion vectors based on the observation. The problem of tracking is dealt with from the image processing front and to illustrate 3D tracking of object, a network of stereo cameras is considered. Common feature points are identified in both the views of the stereo camera and the motion vectors are estimated using Kanade-Lucas-Tomasi (KLT) feature point tracker. These points are then projected in 3D using the camera parameters estimated during the calibration. Experimental results illustrate the scope of the tracking algorithm and motivates the need for a multiple-stereo network.

# Contents

Declaration . . . . .	iv
Approval Sheet . . . . .	v
Acknowledgements . . . . .	vi
Abstract . . . . .	viii
<b>Nomenclature</b>	<b>x</b>
<b>1 Introduction</b>	<b>4</b>
1.1 Scope of the Thesis . . . . .	4
1.2 Literature Survey and Motivation . . . . .	6
1.3 Primary Contributions . . . . .	8
<b>2 Object Localization in 3D</b>	<b>10</b>
2.1 Introduction . . . . .	10
2.2 Related Work . . . . .	10
2.3 Target localization in 3D with sensors at known locations . . . . .	11
2.4 Particle Swarm Optimization . . . . .	12
2.5 PSO based Localization . . . . .	13
2.5.1 Flip ambiguity . . . . .	14
2.6 Experimental Results . . . . .	15
<b>3 Sensor Placement Problem</b>	<b>17</b>
3.1 Optimal sensor placement for known target location . . . . .	17
3.2 Optimal sensor placement for known target location set . . . . .	18
3.3 FIM determinant in closed form . . . . .	18
3.4 Multiple Maxima: Illustrative Example . . . . .	21
3.5 Experimental Results . . . . .	25
3.5.1 Stationary target and sensors located on the ground terrain . . . . .	26
3.5.2 Moving target and sensors located on the ground terrain . . . . .	28
<b>4 Vision-based tracking of objects in 3D</b>	<b>31</b>
4.1 Introduction . . . . .	31
4.2 Sparse point-cloud generation using Stereo pair . . . . .	32
4.3 Kanade-Lucas-Tomasi Point tracking algorithm . . . . .	33
4.4 Experimental Results . . . . .	34

<b>5</b>	<b>Summary and Discussion</b>	<b>36</b>
5.1	Object Localization in 3D . . . . .	36
5.2	Sensor Placement Problem . . . . .	36
5.3	Vision-based tracking of objects in 3D . . . . .	37
<b>A</b>	<b>Theoretical Background</b>	<b>38</b>
A.1	Camera Model and Need for Calibration . . . . .	38
A.2	Auto Calibration . . . . .	40
	<b>References</b>	<b>42</b>



# List of Figures

1.1	Application of localization and tracking . . . . .	5
1.2	Multiple maxima of FIM determinant . . . . .	8
2.1	Trilateration in 2D. . . . .	11
2.2	Graphical representation of localization . . . . .	11
2.3	Flowchart for PSO. . . . .	13
2.4	Cluster-based initialization in 2D for collinear sensors. . . . .	14
2.5	Localization in 3D on California terrain obtained from Google Maps. . . . .	15
2.6	Variation of localization error with measurement error. . . . .	15
2.7	Elimination of flip ambiguity. . . . .	16
3.1	Variation of FIM determinant with $\theta_{ij}$ and indication of multiple maxima. . . . .	23
3.2	Orientations of the sensors corresponding to the maxima of FIM determinant. . . . .	24
3.3	Plot indicating the symmetry of FIM determinant. . . . .	24
3.4	RMSE of different sensor orientations. The CRLB and the RMSE of the optimal positions are shown. . . . .	25
3.5	Section of Anathagiri Hills consider for practical experiments . . . . .	26
3.6	Figure depicting the position of sensors that maximize FIM determinant and provide lowest localization error when the target is stationary and located on the road. . . . .	27
3.7	MSE comparison of the two sets of sensor positions that localize a target on the road shown in Figure 3.6 . . . . .	28
3.8	Figure depicting the position of sensors that maximize FIM determinant and position of sensors that provide lowest localization error when the target could be located anywhere on a small stretch of the road. . . . .	29
3.9	MSE comparison of the two sets of sensor positions that localize a target on the trajectory shown in Figure 3.8 . . . . .	29
3.10	Figure depicting the position of sensors that maximize FIM determinant and position of sensors that provide lowest localization error when the target could be located anywhere on a small stretch of the road. . . . .	30
3.11	MSE comparison of the two sets of sensor positions that localize a target on the trajectory shown in Figure 3.10 . . . . .	30
4.1	Graphical representation of multi-view tracking . . . . .	31
4.2	Triangulation using stereo pair . . . . .	33
4.3	3D point cloud using a pair of calibrated stereo cameras . . . . .	34

4.4	Point cloud tracking from a stereo pair: Frames 1, 20 and 40 (Left to Right) . . . . .	35
A.1	Pinhole model of a camera . . . . .	38
A.2	Camera calibration using extraneous objects . . . . .	39
A.3	Euclidean Autocalibration . . . . .	40

# List of Tables

2.1	Comparison of Localization performace of PSO, BFGS and LS optimization methods.	16
3.1	Maxima of $ I(s, \mathbf{q}) $ and performance comparison . . . . .	27

# Chapter 1

## Introduction

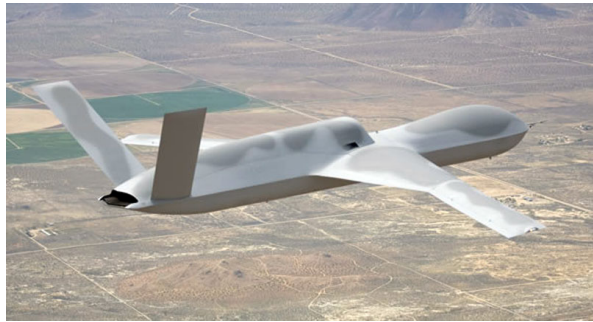
### 1.1 Scope of the Thesis

In a variety of applications, including target tracking and wireless cellular networks, location information plays a crucial role. In order to properly route calls between two mobile users and provide a seamless connection, the base stations at the respective geographical locations must be aware of the position of the users [1]. The design of handoff strategies and allocation of bandwidth are developed with the help of localization. In addition, the locations of the base stations are important in providing an efficient communication system [2], [3] and to provide maximum coverage while minimizing the localization error [4], [5]. Localization is not restricted to outdoor scenario but is also used in indoor applications like security systems [6] [7]. In the area of target tracking, localization plays a vital role in ensuring successful navigation of unmanned aerial vehicles (UAVs) and terrestrial objects. UAVs play vital role in several scientific and commercial applications including monitoring wildlife and vegetation, surveillance, search and rescue operations and so on. The quality of the intelligence gathered is often determined by the accuracy of the estimated position at any time instant [8]. Location information is also central to terrestrial vehicle tracking systems [9], [10]. Vehicle tracking systems are often coupled with security systems to notify unauthorized access of an area. It is also used in recovering stolen vehicles, managing e-commerce, traffic control and surveillance. Even though localization of objects does not directly lead to tracking the object, it serves as a precursor to several tracking algorithms [11], [12]. The first step in determining the trajectory of a moving object is to localize the object in the 3D continuum to reduce the search space. Using the estimated location of the object as initialization, its location in subsequent time instants can be estimated with maximum accuracy. However, the localization error has to be maintained below a certain threshold. Among various factors affecting the localization error, including the quality of sensing devices and the transmitted signal, one of the most important factors that has a direct influence on the accuracy of localization is the placement of sensors. The geometry and position of the sensors play a key role in determining the ability of the system to estimate the location of the target with minimum uncertainty [13], [14]. Therefore, it is necessary to optimize the positions of the sensors [15], [16]. This work deals with a systematic approach to determine the optimal position of sensors that can be used in any of the applications specified above. In addition to localization, tracking is also demonstrated using visual fiducials and a near-accurate track of the

object is provided in 3D.



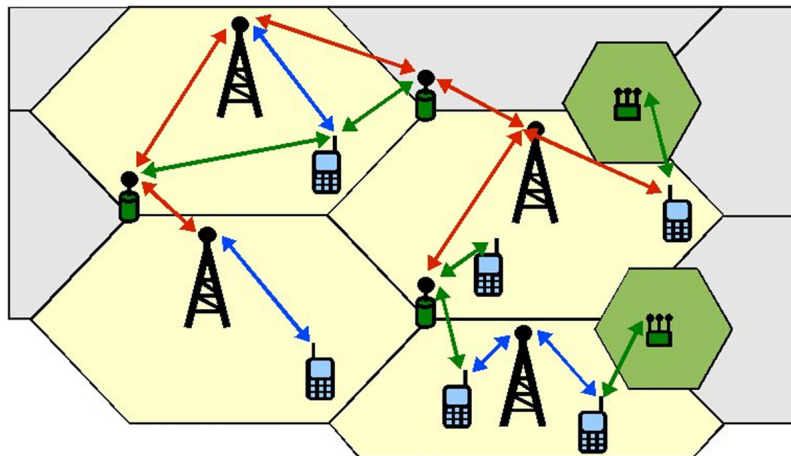
(a) Missile localization



(b) UAV tracking



(c) Vehicular tracking



(d) Wireless networks

Figure 1.1: Application of localization and tracking

## 1.2 Literature Survey and Motivation

Localization techniques can be broadly classified into the following two categories: (i) Range-free localization, and (ii) Range-based localization. Range-free algorithms estimate the location of the target using the geometry of reference points and do not require distance measurement. This technique proves to be simpler, however, compromising on the accuracy [17], [18], [19]. In contrast, applications requiring high accuracy employ range-based algorithms, where a network of sensors are deployed at certain reference points with known locations. Based on the sensing modalities, the distance between the sensors and the target is estimated based on various measured quantities including the received signal strength (RSS) [20], [21], [22], time of arrival (TOA) [23], time difference of arrival (TDOA) [24], [25], [26] or angle of arrival (AOA) [27], [28]. One of the simplest range-based localization techniques is Multilateration [29] which is motivated by geometry. Yuang Zhang *et al.* [30] proposed a 3D localization technique that combines both distance as well as angle measurements to localize the object. Integrating sensors with additional hardware has also been incorporated by Modar Ibraheem [31] to improve the localization accuracy. Despite the fact that the problem of range-based localization has been addressed to a great extent [17], [31], [32], this work will provide an improved solution to localization with a significantly higher accuracy and lower computational overhead. In particular, this paper deals with systems that estimate the range measurements based on received signal strength (RSS) and localizes an unknown target using a network of sensors.

The need for precise estimation of location has lead researchers to focus more on optimal placement of sensors. An obvious inference is that clustering of sensors results in poor ability to provide an accurate location estimate. When the target is stationary, it can be seen intuitively that the optimal position of sensors is the vertices of a regular polygon with the target at the center. However, this situation serves no practical significance because the purpose of localization is moot when the position of the target is known apriori. A direct extension of the intuition is also not possible when the target is not stationary. Therefore, the urge for a rigorous mathematical framework to optimize the sensor placement stems from the complexity of the problem. As a first step to an optimization problem, it is necessary to define an objective function. The objective function must be a reflection of the accuracy of the estimated location. For several years, researchers in the GPS community used Geometric Dilution of Precision (GDOP) as a metric to assess the quality of satellite geometry [33], [34]. GDOP is an indication of how the error in the measured data results in inaccuracy in the estimated location. Despite the simplicity of the metric, it is not used extensively because of its inability in capturing both the stochastic nature of the measurement and the geometry of sensors individually. Several information theoretic measures have also been used to select a set of optimal candidate sensors [35]. Fisher *et al.* [36] consider a Bayesian filtering strategy and use the principle of mutual information to determine the optimal set of sensors. An entropy-based selection heuristic was used by Wang *et al.* [37] that selects an informative sensor which when combined with existing sensors results in the maximum reduction of entropy of the distribution of target location. The authors claim that this heuristics is computationally less expensive and accurate compared to mutual information based techniques. A more natural selection of a measure of inaccuracy is the variance of the location estimate. The computation of variance is complex and an exhaustive enumeration is computationally infeasible. Therefore, an alternate, yet computationally simple, representation of variance is required. Information inequality provides a statistical bound on the variance of an estimated variable. According to Cramer Rao lower bound, the variance of an unbiased estimate is

lower bounded by the inverse of a statistical measure called Fisher Information. Fisher Information often admits closed form expression and therefore, maximization of Fisher Information is practical. Instead of minimizing the variance of the location estimate, the lower bound on the variance is minimized which essentially provides a balance between computational cost and accuracy. Fisher Information based approach has been used by several researchers to resolve the issue of sensor placement [16], [38], [39]. However, this work proposes a novel technique that fuses the computational advantage of Fisher Information based methods as well as the accuracy of direct minimization of variance.

Fisher Information is a statistical measure of the amount of information present in the observable data about the underlying unknown parameter. When the information content is high, the unknown variable can be estimated with higher confidence. Using the noise models of range measurements proposed by Cassioli *et al.* [40], Borah *et al.* proposed an algorithm that determines the optimal position of sensors based on maximization of Fisher Information for a stationary as well as a moving target [38]. The work is primarily based on the closed form expressions for Fisher Information derived in [41] and later used in [16] and [39]. The paper provides a solid mathematical foundation for placement of heterogeneous sensors for accurate localization. The authors have proposed an efficient algorithm that solves the optimization problem using exponential smoothing. The key features of the algorithm include a convex search space and the use of a projected gradient search technique to restrict the search space. The authors have illustrated with numerical results how the proposed technique enables accurate localization of target over large volume of search space as compared to random sensor locations. The effect of the heterogeneity of the sensors is incorporated in the framework and illustrated using simple examples. Given a network of sensors, the algorithm provides the optimal location of an additional sensor to localize a fixed or moving target. The optimal position of sensors is obtained by maximizing the determinant of Fisher Information matrix (FIM) defined with respect to the target location. When the target is not stationary, the cost function is averaged over the search space of the target to incorporate the prior information about the target location. However, Fisher information fails to capture the variation in the variance of the estimated location. Maximizing Fisher Information only lowers the bound on the error covariance, it does not translate to lowest expected localization error which is described in terms of root mean squared error (RMSE). Therefore, the solution to the sensor placement problem is not just to find the maxima of the determinant of FIM but to also examine if the corresponding sensor placement results in the lowest RMSE. This complexity of the problem is further increased when the determinant of FIM is multimodal. A multimodal objective could consist of either multiple global maxima or a combination of global and local stationary points. The RMSE is locally minimum around each of the stationary points of FIM determinant. However, the global maxima of FIM determinant does not necessarily correspond to the global minimum of RMSE. Figures 1.2a and 1.2b give a graphical representation of the same. In Figure 1.2a, all the stationary points of FIM determinant have the same value but the behavior of each in a localization system is significantly different. This is reflected in the variation of RMSE plotted in blue. Only one of the global maxima has the lowest RMSE and this therefore, the optimal configuration of sensors. In a practical scenario, there could be several maxima of FIM determinant with varying values as shown in Figure 1.2b. However, the RMSE does not follow the same pattern. The global minima of RMSE could lie around one of the local maxima of FIM determinant. The primary contribution of the thesis is to highlight the existence of

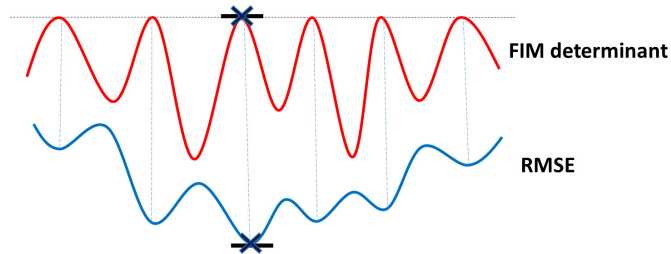
multiple maxima of FIM and compare the performance of each in a localization system in order to provide an optimal sensor position. This is supported with the help of an illustrative toy example where the existence of multiple maxima is proved analytically and an experimental comparison of the performance of each maxima is provided.

### 1.3 Primary Contributions

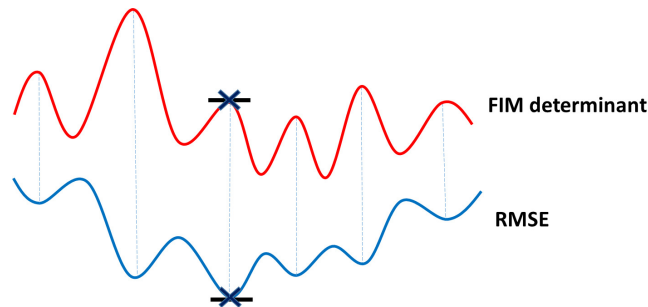
The main areas of focus in this thesis are:

- Object Localization in 3D
- Optimal Placement of sensors for an improved localization
- Tracking of an object in 3D using visual data

In the field of localization, the thesis proposes a modification of the bio-inspired optimization algorithm called Particle Swarm Optimization that leads to faster convergence and eliminates the flip ambiguity encountered by coplanar sensors. This thesis also proposes a novel optimal sensor placement technique that first determines all the local maxima of Fisher Information. An exhaustive evaluation of each of the local maxima is performed to determine the solution with minimum variance. To motivate the need for a two step optimization, a toy example is considered. The location of a stationary target is assumed to be known. The sensors are further constrained to be located along the circumference of a circle with the target located at its center. This example would clearly highlight the fact that maximization of Fisher Information merely serves as an initial step to finding the optimal sensor positions and is not the final solution. A real life situation of tracking a vehicle on a



(a) Plot showing that all the global maxima of FIM determinant do not have the same RMSE



(b) Plot showing that a local maxima (instead of a global maxima) of FIM determinant has the lowest RMSE

Figure 1.2: Multiple maxima of FIM determinant



road is considered for corroboration. The stationary target is first assumed to be located on a road with sensors to be deployed along the road. As an extension of the above, the algorithm is demonstrated for a moving target whose location is not known precisely but a set of probable target locations is known. The target can, therefore, be located anywhere on a known section of the road and sensors, subject to practical constraints, are placed on either side of the road to localize the target efficiently. As can be seen from the results obtained, the sensor orientation with better localization performance has a significantly lower value of Fisher Information compared to its global maxima. A comparison of the global maxima of FIM determinant and the local maxima resulting in minimum Root Mean Squared Error (RMSE) shows that the local maxima results in approximately 50% improvement in the RMSE over the global maxima of FIM determinant. The results obtained conclusively prove that the global maxima of Fisher Information does not always result in the lowest localization error. This work also makes use of a real data set that was obtained for practical illustrations.

Although the problem of localization has been addressed from the point of view of wireless sensor networks, the tracking of object is considered using visual data to highlight the effects of occlusions that hinder the efficiency of tracking algorithms. This effect is eliminated by using a network of cameras (stereo cameras in particular) instead of a single camera because the notion of occlusion is obsolete in 3D. This work provides a foundation for 3D object tracking highlighting the several vital components of the process and illustrates how a point cloud generated using a stereo pair can be tracked in 3D.

## Chapter 2

# Object Localization in 3D

### 2.1 Introduction

Location information is crucial in several applications such as navigation, target tracking and environment monitoring. For instance, value of intelligence collected by unmanned aerial vehicles (UAVs) often crucially depends on accurately finding their locations in 3D, and enabling successful navigation. Accurate location information is also central in efficient design of mobile networks. In particular, accurate localization of a mobile user enables one to design efficient handoff strategies and hence optimize energy budgeting. Localization error is often kept below a threshold by using a network of sensors. A majority of localization techniques fall under one of the two categories: (i) Range-free localization, and (ii) Range-based localization. Range-free algorithms make use of only sensor geometry, do not require distance measurement, and tend to be simpler but less accurate [17]. In contrast, applications requiring high accuracy often employ range-based algorithms, where a network of sensors, capable of recording and interpreting range-finding signals, are deployed at certain reference points with known locations. Practically, the range is not directly measured, but estimated based on various measured quantities including the received signal strength (RSS), time of arrival (TOA), time difference of arrival (TDOA) or angle of arrival (AOA). It is worthwhile to mention that measurement of such quantities has become increasingly accurate and efficient with recent advancements in the fields of wireless sensor network and wireless communication. Typical SLAM-based localization methods use visual fiducials and operate on indoor geometries of smaller scale. However, we aim to provide a solution to the localization problem can even be extended to using RF signals spanning several kilometers.

### 2.2 Related Work

Several attempts have been made to accurately localize an object using range measurements. One of the simplest techniques proposed in the literature is Multilateration [29]. Fig 3.1 depicts how trilateration (number of sensors equal to 3) is used to localize an object in 2D. The location of the target is the point of intersection of circles drawn with the sensors at the center and radii equal to the distances measured. The presence of noise introduces ambiguity in the radius of the circles which results in a region of intersection in contrast to a point. The complexity increases when the number of

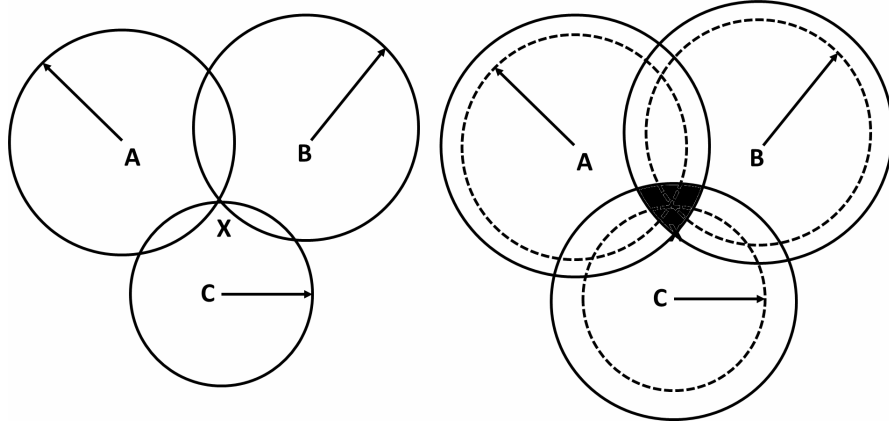


Figure 2.1: Trilateration in 2D.

sensors becomes large. Yuang Zhang et al. [30] proposed a 3D localization technique that combines both distance as well as angle measurements to localize the object. Several authors have proposed techniques that require additional hardware in addition to the sensors deployed. 3D pose estimation by Modar Ibraheem [31] integrates data from odometer and gyro meter to minimize localization error. The following work is based on the use of Particle Swarm Optimization to estimate the location of an unknown target using range measurements derived from the strength of the received signal. The following section describes the mathematical framework for the localization problem and motivates the need for an evolutionary search algorithm.

### 2.3 Target localization in 3D with sensors at known locations

Consider a network of  $N$  sensors at locations  $r_i = (x_i, y_i, z_i)^T$ , where  $i = 1, \dots, N$ . If the position of the target in 3D space is denoted as  $s = (x, y, z)^T$ , the distance between the  $i^{th}$  sensor and the target is given by

$$d_i = \|r_i - s\|_2. \quad (2.1)$$

Due to inaccuracies in the sensing devices, the measured distance  $m_i$  deviates from the true distance

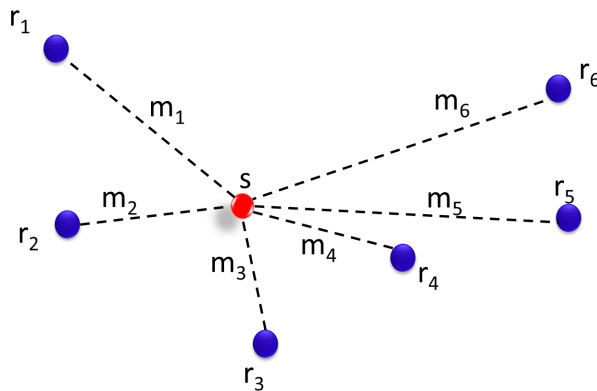


Figure 2.2: Graphical representation of localization

$d_i$  by a factor  $n_i$  given by

$$m_i = d_i + n_i. \quad (2.2)$$

If  $\mathbf{m} = [m_1, m_2, \dots, m_N]^T$  represents the vector of all the measured distances,  $\mathbf{d} = [d_1, d_2, \dots, d_N]^T$  is the vector of true Euclidean distances and  $\mathbf{n} = [n_1, n_2, \dots, n_N]^T$  is the vector of all measurement noise components, Equation 2.2 can be represented in vector form as

$$\mathbf{m} = \mathbf{d} + \mathbf{n}. \quad (2.3)$$

Therefore, given the measured distances  $\mathbf{m}$  and the position of the sensors  $r_i = (x_i, y_i, z_i)^T, \forall i \in \{1, \dots, N\}$ , the problem of localization is defined as estimation of the vector  $\mathbf{d}$  (which in turn gives the position of the target  $s$ ) such that the effect of measurement noise is minimized. In other words, the true Euclidean distances are estimated such that the norm of the measurement error  $\|\mathbf{n}\|_2 = \|\mathbf{m} - \mathbf{d}\|_2$  is minimized.

This deterministic framework results in a trivial solution of  $\hat{\mathbf{d}} = \mathbf{m}$ . However, the practical constraints impels the need for a physical model of the measurement noise. We adopt one such noise model proposed in literature [40] that models the noise vector to jointly follow Gaussian distribution with mean 0 and covariance matrix that is a function of the Euclidean distance between the target and the sensor. Therefore, the measurements are said to follow a Gaussian distribution  $\mathcal{N}(\mathbf{d}, \mathbf{A})$  given by

$$p(\mathbf{m}; s, \mathbf{q}) = \frac{1}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{d})^T \mathbf{A}^{-1}(\mathbf{m} - \mathbf{d})\right). \quad (2.4)$$

Localization is a multidimensional non-linear optimization problem which can be solved using an iterative approach only. Most of the techniques address the localization problem in 2D and cannot be directly upgraded to 3D. Equation 3.4 involves a nonconvex objective function (with a Hessian matrix not positive definite). Therefore, L2 optimization methods are not generally optimal. A PSO-based heuristics that balances accuracy and exploration at low computational cost is adopted. PSO is a widely used iterative algorithm that has proved to be robust in many situations. The entire search space is scanned to find the global extremum. This can be cumbersome when the dimensions of the search space is large. Therefore, it is necessary to provide the right initialization of the particles so that the time taken for convergence is reduced drastically. Before delving in to the details of the proposed PSO based localization algorithm, a brief overview of PSO is provided.

## 2.4 Particle Swarm Optimization

Particle Swarm Optimization is a global optimization technique that is inspired by fish schooling and swarm behavior of birds. It was first proposed by Kennedy et al. in 1995 [42]. The algorithm iteratively scans the search space to find the optimal solution. Each particle in the algorithm represents a potential solution in the search space. The evolution of the particles is guided by a fitness function (a measure of its quality). The particles keep track of its own best position and is aware of the best particle in the group. It is influenced by its own experience and also by the social behavior of the group. The velocity  $v_i$  of the  $i^{th}$  particle is updated using the following equation

$$v_i^{k+1} = \omega \times v_i^k + c_1 \times r_1 \times (p_{i\_best} - X_i^k) + c_2 \times r_2 \times (g_{best} - X_i^k), \quad (2.5)$$

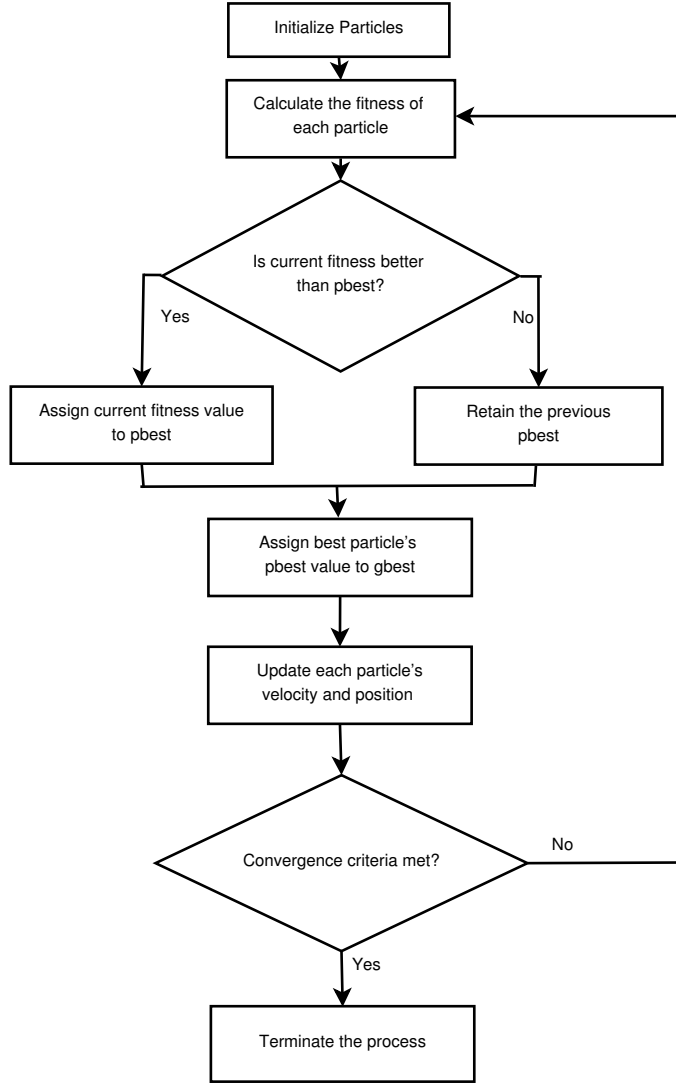


Figure 2.3: Flowchart for PSO.

where  $v_i^k$  represents the velocity of the  $i^{th}$  particle at the  $k^{th}$  iteration,  $X_i^k$ , the position of the  $i^{th}$  particle,  $c_1$  and  $c_2$  are correction factors,  $\omega$  is inertia and  $r_1$  and  $r_2$  are random numbers between 0 and 1. Fig 2.3 explains how PSO is used for optimization. The position of the  $i^{th}$  particle is updated by the equation

$$X_i^{k+1} = v_i^{k+1} + X_i^k. \quad (2.6)$$

## 2.5 PSO based Localization

As explained in section 2.4, PSO finds the extremum using particles. In this context, a particle refers to the coordinates of the target. The fitness function  $F$ , given by

$$F = \frac{1}{N} \sum_i \left( m_i - \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2} \right)^2 \quad (2.7)$$

represents the mean squared difference between the measured and the actual distance between the target and each of the sensors. The particle with the minimum fitness value is said to have a higher quality. All particles evolve in this search space till they converge at the minimum fitness value. Instead of initializing the particles randomly in the entire space, the search space of PSO is restricted. This is inspired by trilateration technique. The search space is now restricted to the shaded region in Fig. 3.1 thereby reducing the the time required to compute the target location. A similar method is incorporated in case of 3D. The circles in 2D are replaced by spheres in 3D and the region of intersection is a volume in 3D. If the particles are initialized around the centroid of the region, the algorithm converges at a faster rate.

### 2.5.1 Flip ambiguity

Flip ambiguity is an phenomenon that is commonly observed when the sensors are collinear or coplanar. This occurs because of the existance of two minima of  $F$ . Instead of the original location, the object is estimated to be present at its reflection about the line joining the sensors in case of 2D or the plane of the sensors in 3D.

Consider a set of three collinear sensors. For a given set of range measurements, there are two points satisfying the distance constraint. If the initial value supplied to the estimator is incorrect, it can converge at either of the points. Fig 3.2 shows how a set of three sensors with erroneous range measurements results in 6 points of intersection of circles. These points are classified into two clusters based on the Euclidean distance. The cluster with lower volume is chosen and the particles of PSO are initialized around the centroid of the cluster.

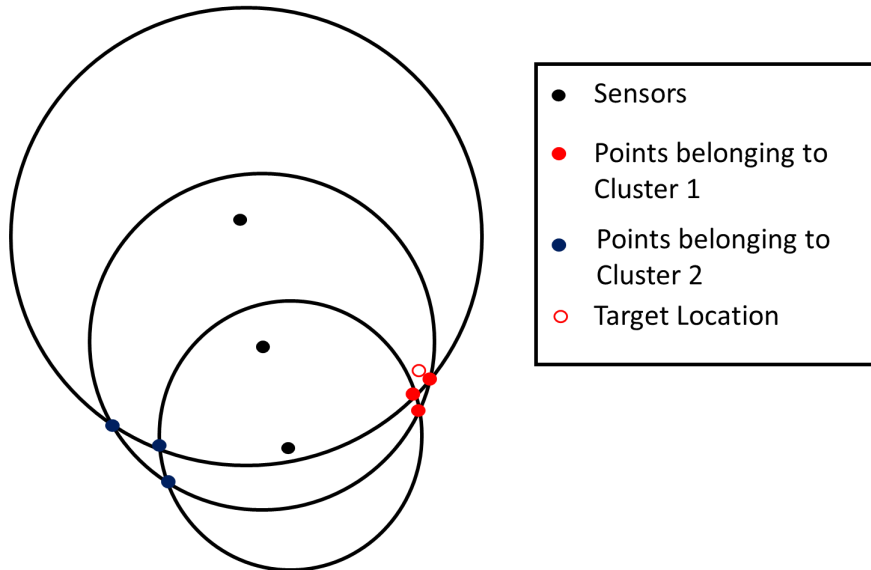


Figure 2.4: Cluster-based initialization in 2D for collinear sensors.

## 2.6 Experimental Results

This section describes the experiments conducted to validate the theories proposed. The results of the experiments reflect the performance of the algorithm in terms of computational time and accuracy. The sensors are placed arbitrarily on a given terrain and the trajectory of the target is marked in the 3D continuum. A section of California terrain obtained from Google Earth is considered to show that even with minimal variation in the heights of the sensors, a target flying high above the surface can be localized with maximum accuracy. As shown in Fig. 2.5, the maximum variation in height of the sensors is only around 500m whereas the target is at a height of 2km from the sea level. Error in the range measurements has an impact on the accuracy of the location estimate. Fig. 2.6 shows the variation of the localization error with the range measurement error. When the distances measured are accurate, the localization error is zero. It can be observed that the variation of location estimate error with measurement error does not assume a linear pattern.

The proposed cluster-based initialization has a significant advantage in comparison to the standard PSO in terms of speed of computation. While standard PSO with random initialization requires about 120 function evaluations to converge, cluster-based initialization needs only 60 function evaluations for convergence. This method not only increases the speed of the localization, but it also

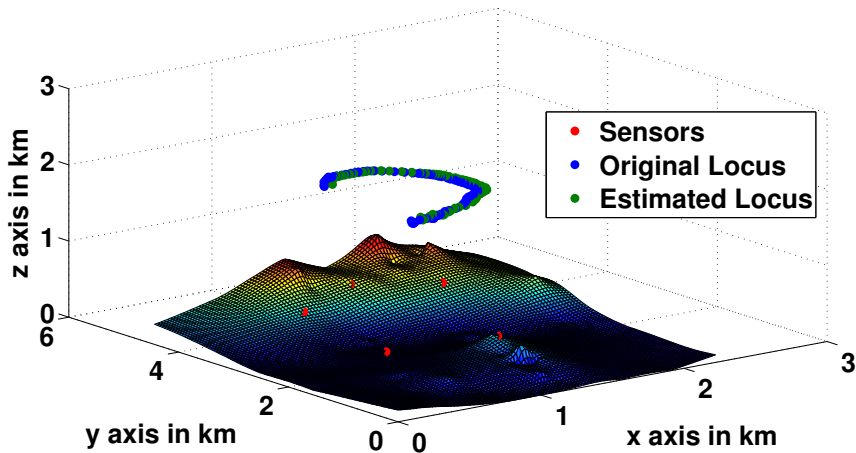


Figure 2.5: Localization in 3D on California terrain obtained from Google Maps.

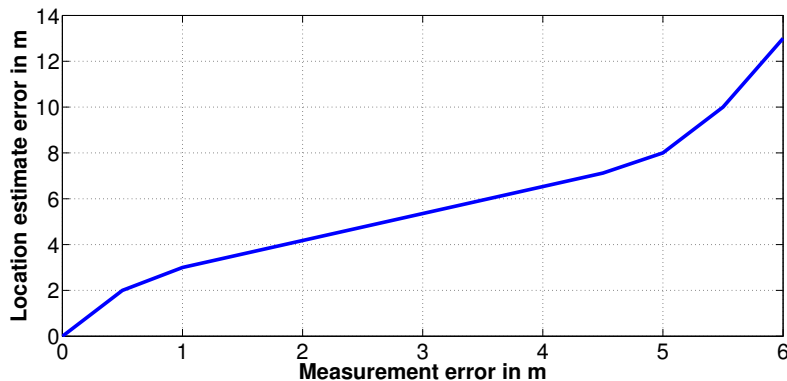


Figure 2.6: Variation of localization error with measurement error.

eliminates flip ambiguity as illustrated in Fig. 2.7. Random initialization can lead to convergence at the reflection of the target’s original location about the line joining the sensors causing large localization errors. PSO based localization is compared with the standard L2 optimization techniques like Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm and Least Squares (LS) method. Table 2.1 depicts how this method outperforms the other two by eliminating flip ambiguity. Therefore, the proposed PSO based localization not only solves the complex optimization problem but also eliminates flip ambiguity and is computationally faster. This is vital because location information is required in real-time applications. A detailed explanation of the proposed algorithm and the results obtained are available in the paper proposed by Vidya *et al.* [43].

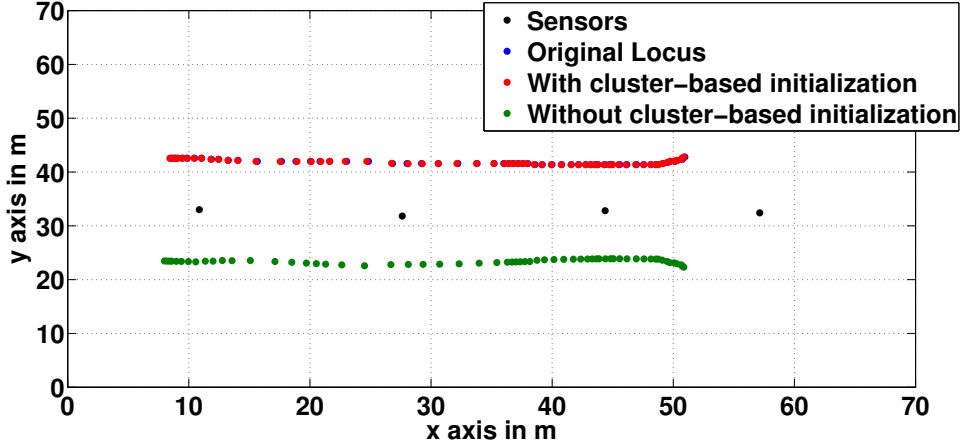


Figure 2.7: Elimination of flip ambiguity.

Table 2.1: Comparison of Localization performance of PSO, BFGS and LS optimization methods.

Method	$x$ error	$y$ error	$z$ error	MSE
Proposed PSO	0.0052	0.0022	0.0086	0.0114
BFGS	1.9388	0.6364	4.9221	5.356
LS	3.56	1.33	10.37	11.05



# Chapter 3

## Sensor Placement Problem

This work is aimed at providing an optimal network of sensors that can localize any unknown target with the best average performance. The positions of the sensors, subject to some practical constraints, are estimated so that the localization error incurred is minimum. This section describes the mathematical framework for localization, followed by the optimal sensor placement problem.

### 3.1 Optimal sensor placement for known target location

Localization, as described previously, is used to estimate the location of a target using a set of sensors at pre-defined locations. These sensors, however, do not localize every target location in a given search space with the same accuracy. This is due to the difference in the relative positions of the sensors with respect to the each of the target locations. For the target location which is localized with maximum accuracy, the given network of sensors is said to be optimally placed. Optimal sensor placement is a process of determining such orientations that has the ability to localize a target with minimum error. It can also be viewed as the inverse of localization. A statistical measure of accuracy is the variance of the estimated variable. Localization error is expressed in terms of the variance of the estimated location. For a fixed target location, the set of sensor locations that provides the minimum variance is said to be optimal [38] [40]. If  $\hat{s}$  is the estimate of the target location  $s$ , the covariance matrix is given by  $\mathbf{R} = E \left[ (s - \hat{s})(s - \hat{s})^T \right]$ .

For a fixed target location, the optimal position of the sensors is the one that minimizes the determinant of the covariance matrix given by

$$\hat{q} = \arg \min_q \left| E \left[ (s - \hat{s})(s - \hat{s})^T \right] \right|, \quad (3.1)$$

where  $\mathbf{q} = [x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N]^T$  is a vector of the position of all the sensors. Owing to the complexity in computation of covariance matrix, the lower bound on the determinant of covariance matrix defined by Cramer Rao Lower Bound is minimized. If  $s$  is an unknown scalar variable, the variance of an unbiased estimator  $\hat{s}$  of  $s$  ( $E[\hat{s}] = s$ ) is bounded by the reciprocal of Fisher Information  $I(s)$ :

$$\text{var}(\hat{s}) \geq \frac{1}{I(s)}, \quad (3.2)$$

where the Fisher Information  $I(s)$  is defined by

$$I(s) = E \left[ \left( \frac{\partial \log(p(m; s))}{\partial s} \right)^2 \right].$$

In vector form, a bound on the covariance matrix implies that  $\mathbf{R}$  succeeds the Fisher Information, i.e.

$$\mathbf{R} = E \left[ (s - \hat{s})(s - \hat{s})^T \right] \succeq [\mathbf{I}(s, \mathbf{q})]^{-1}, \quad (3.3)$$

where  $\mathbf{I}(s, \mathbf{q})$  is the Fisher Information Matrix (FIM). This also implies that  $\mathbf{R} - [\mathbf{I}(s, \mathbf{q})]^{-1}$  is positive semidefinite. The Fisher Information matrix is represented in vector form as

$$\mathbf{I}(s, \mathbf{q}) = E \left[ (\nabla_s \ln p(\mathbf{m}; s, \mathbf{q}))^T (\nabla_s \ln p(\mathbf{m}; s, \mathbf{q})) \right]. \quad (3.4)$$

Using Cramer Rao Lower Bound, the optimal sensor positions for a fixed target location are obtained by minimizing the lower bound of the determinant of covariance matrix, or maximizing the determinant of Fisher Information matrix. Therefore, Equation 3.1 can be written as

$$\hat{q} = \arg \max_q |\mathbf{I}(s, \mathbf{q})|. \quad (3.5)$$

This provides a framework for optimally placing sensors to localize a known target. Although this does not have practical significance, the exercise was vital to develop a systematic framework that is applicable in a practical setting. The sensor placement problem for a moving target is proposed in the subsequent section.

## 3.2 Optimal sensor placement for known target location set

In practical situations, the location of the target is not fixed. However, it is possible to know the set of locations the target is most likely to be present in. Using the prior distribution on the set of target locations, it is possible to determine where the sensors are to be placed in order to provide the lowest average localization error. The framework described previously is extended by averaging the determinant of FIM over the set of target locations and then maximizing with respect to  $\mathbf{q}$ . Equation 3.5 is written as

$$\hat{q} = \arg \max_q E_s [|\mathbf{I}(s, \mathbf{q})|]. \quad (3.6)$$

It is often observed that Fisher Information Matrix admits a closed form expression. Section 3.3 provides the analytical expression for Fisher Information Matrix which is then used in the sensor placement problem.

## 3.3 FIM determinant in closed form

Consider the distribution of the measured distance  $p(\mathbf{m}; s, \mathbf{q})$  defined by Equation 2.4. The distances are measured independent of each sensor and therefore it is justified to assume that the measured distances have an independent distribution. We further assume that the measurements are identically distributed. This assumption is valid and helps in obtaining a closed form expression for FIM

determinant. Each of the measured distances  $m_i$  follows a Gaussian distribution with mean  $d_i$  and a variance of  $\sigma_i^2 \left(\frac{d_i}{d_0}\right)^{\alpha_i}$ , where  $\alpha_i$  is the path-loss exponent of the  $i^{\text{th}}$  sensor,  $d_0$  is the path-loss reference distance which is assumed to be 1 for simplicity, and  $\sigma_i^2$  is a constant for the  $i^{\text{th}}$  sensor. In this regard, Equation 2.4 is first rewritten as

$$p(\mathbf{m}; s, \mathbf{q}) = \frac{1}{(2\pi)^{N/2}} \prod_{i=1}^N \frac{1}{(\sigma_i^2 d_i^{\alpha_i})^{1/2}} \exp\left(-\frac{(m_i - d_i)^2}{2\sigma_i^2 d_i^{\alpha_i}}\right). \quad (3.7)$$

Using the above to determine the Fisher Information Matrix, we obtain

$$\nabla_s \ln p(\mathbf{m}; s, \mathbf{q}) = \mathbf{k}^T \nabla_s \mathbf{d}, \quad (3.8)$$

where  $\mathbf{k}^T$  is  $1 \times N$  vector with the  $i^{\text{th}}$  element given by

$$k_i = \left[ \frac{(m_i - d_i)}{\sigma_i^2 d_i^{\alpha_i}} + \frac{(m_i - d_i)^2 \alpha_i}{\sigma_i^2 d_i^{\alpha_i + 1}} - \frac{\alpha_i}{d_i} \right]. \quad (3.9)$$

Substituting the above in Equation 3.4, FIM is given by

$$I(s, \mathbf{q}) = (\nabla_s \mathbf{d})^T E(\mathbf{k}\mathbf{k}^T) \nabla_s \mathbf{d}. \quad (3.10)$$

On simplification, it can be observed that  $E(\mathbf{k}\mathbf{k}^T)$  is a diagonal matrix with  $i^{\text{th}}$  diagonal element given by

$$E(k_i^2) = \frac{\alpha_i^2}{2d_i^2} + \frac{1}{\sigma_i^2 d_i^{\alpha_i}} = \beta_i^2. \quad (3.11)$$

Since  $m_i$  is modeled to follow Gaussian distribution with mean  $d_i$ , the central moments of the variable are given as follows:

$$\begin{aligned} E[(m_i - d_i)] &= 0 \\ E[(m_i - d_i)^2] &= \sigma_i^2 d_i^{\alpha_i} \\ E[(m_i - d_i)^3] &= 0 \\ E[(m_i - d_i)^4] &= 3\sigma_i^4 d_i^{2\alpha_i}. \end{aligned} \quad (3.12)$$

Also,  $E(\mathbf{k}\mathbf{k}^T)$  is an  $N \times N$  matrix given by

$$E(kk^T) = E \begin{pmatrix} k_1^2 & \dots & k_1 k_N \\ \vdots & \ddots & \vdots \\ k_N k_1 & \dots & k_N^2 \end{pmatrix}. \quad (3.13)$$

The  $(i, j)^{\text{th}}$  element of the matrix is given by

$$E(k_i k_j) = E \left[ \left( \frac{(m_i - d_i)}{\sigma_i^2 d_i^{\alpha_i}} + \frac{(m_i - d_i)^2 \alpha_i}{\sigma_i^2 d_i^{\alpha_i + 1}} - \frac{\alpha_i}{d_i} \right) \times \left( \frac{(m_j - d_j)}{\sigma_j^2 d_j^{\alpha_j}} + \frac{(m_j - d_j)^2 \alpha_j}{\sigma_j^2 d_j^{\alpha_j + 1}} - \frac{\alpha_j}{d_j} \right) \right]. \quad (3.14)$$

The above expression can be simplified using Equation 3.12 and the fact that  $m_i$ 's are independent and identically distributed. When  $i \neq j$ ,  $E(kk^T)$  reduces to

$$E(k_i k_j) = \frac{1}{4} \left[ \frac{2\alpha_i \alpha_j}{d_i d_j} - \frac{2\alpha_i \alpha_j}{d_i d_j} \right] = 0. \quad (3.15)$$

When  $i = j$ ,

$$E(k_i^2) = \frac{1}{\sigma_j^2 d_j^{\alpha_j}} + \frac{\alpha_i^2}{2d_i^2} = \gamma_i^2. \quad (3.16)$$

Therefore,  $E(\mathbf{k}\mathbf{k}^T)$  is a diagonal matrix with diagonal entries given by

$$E(k_i^2) = \frac{\alpha_i^2}{2d_i^2} + \frac{1}{\sigma_i^2 d_i^{\alpha_i}} = \beta_i^2. \quad (3.17)$$

where  $i$  varies from 1 to  $N$ . The determinant of FIM is given by

$$|I(s, \mathbf{q})| = \frac{1}{6} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \beta_i^2 \beta_j^2 \beta_k^2 \|((\nabla_s d_i) \times (\nabla_s d_j)) \cdot (\nabla_s d_k)\|^2, \quad (3.18)$$

where  $s = (x, y, z)^T$  refers to the location of the target in 3D and  $\mathbf{q}$  is a  $3N$  dimensional vector denoting the position of the sensors. However, when all the sensors are located on a plane, the  $3^{rd}$  dimension of the sensor locations would be insignificant. Therefore, the expression for FIM determinant simplifies to

$$|I(s, \mathbf{q})| = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \beta_i^2 \beta_j^2 \|(\nabla_s d_i) \times (\nabla_s d_j)\|^2, \quad (3.19)$$

where  $s = (x, y)^T$  refers to the location of the target in 2D and  $\mathbf{q}$  is a  $2N$  dimensional vector denoting the position of the sensors. For a fixed target location  $s$ , the optimal position of the sensors is obtained by maximizing  $|I(s, \mathbf{q})|$  given by Equations 3.18 and 3.19.

Although a closed form expression is obtained for the objective function, the optimization problem is not simplistic. It is observed that multiple maxima are obtained while optimizing the determinant of FIM. Two sensors having the same value of FIM determinant do not necessarily localize the target with the same accuracy. Therefore, among the set of all maxima, it is necessary to identify the one that is optimal in terms of average localization error. If the global maxima of FIM determinant coincides with the minima of the actual covariance, then the sensor placement problem simplifies to determining the global maxima. However, it is not necessary that the above situation is valid. Instead of the global maxima, the minima of covariance could correspond to one of the local maxima of FIM determinant. Alternately, the minima of covariance could lie in the neighborhood of the maxima of FIM determinant. Therefore, reaching the global maxima of FIM determinant does not ensure that the optimal sensor locations are obtained. A further insight about the complexity of the optimization problem is developed by considering an illustrative example in Section 3.4. Although the example considered might seem simplistic, it portrays clearly the existence of multiple maxima of FIM determinant and the difference in the behavior of each of the maxima in localizing a target. Section 3.5 shows experimentally how the proposed technique can be applied to practical scenarios.

### 3.4 Multiple Maxima: Illustrative Example

Consider the following toy example in 2D where the target location is known and sensors are placed at the same distance  $d$  from the target. This implies that the sensors are aligned along the circumference of a circle with the target located at its center. Incorporating this, the expression for  $|I(s, \mathbf{q})|$  simplifies to

$$|I(s, \mathbf{q})| = \frac{1}{2} \beta^4 \sum_{i=1}^N \sum_{j=1}^N \sin^2 \theta_{ij}, \quad (3.20)$$

with the assumption that all the sensors are identical, i.e.  $\beta_i = \beta_j = \beta$ . In Equation 3.20,  $\theta_{ij}$  represents the angle between the  $i^{\text{th}}$  and  $j^{\text{th}}$  sensors with respect to the center. Since the distance between the sensors and the target is fixed, the only parameter influencing FIM determinant is the angle between the sensors.

If  $\theta_i$  represents the azimuthal angle of the  $i^{\text{th}}$  sensor (angle subtended by the  $i^{\text{th}}$  sensor with respect to the positive x-axis), Equation 3.20 can be written as

$$|I(s, q)| = \frac{1}{2} \beta^4 \sum_{i=1}^N \sum_{j=1}^N \sin^2 (\theta_i - \theta_j). \quad (3.21)$$

**Proposition 1:** When sensors are placed along the circumference of a circle to localize a target at the center, the following are observed-

1. The maximum value of  $|I(s, \mathbf{q})|$  is  $\frac{N^2}{4} \beta^4$ , that is obtained when the following conditions are satisfied:

$$\sum_{i=1}^N \cos 2\theta_i = 0 \text{ and } \sum_{i=1}^N \sin 2\theta_i = 0.$$

2. The value of  $\theta_{ij} = |\theta_i - \theta_j|$  that maximizes  $|I(s, \mathbf{q})|$  is equal to  $k\pi/N$  (module  $\pi$ ), where  $k$  is an integer and  $N$  represents the number of sensors.

**Proof:** Using trigonometric identity  $\sin^2 x = \frac{1 - \cos 2x}{2}$ , the expression for FIM determinant reduces to

$$|I(s, \mathbf{q})| = \frac{1}{2} \beta^4 \sum_{i=1}^N \sum_{j=1}^N \frac{1 - \cos 2(\theta_i - \theta_j)}{2}. \quad (3.22)$$

However,  $\cos(A - B) = \cos A \cos B + \sin A \sin B$ . Therefore,

$$|I(s, \mathbf{q})| = \frac{N^2}{4} \beta^4 - \frac{1}{4} \beta^4 \left[ \sum_{i=1}^N \sum_{j=1}^N (\cos 2\theta_i \cos 2\theta_j + \sin 2\theta_i \sin 2\theta_j) \right]. \quad (3.23)$$

Rearranging the terms in Equation 3.23, we obtain

$$\begin{aligned} |I(s, \mathbf{q})| = & \frac{N^2}{4} \beta^4 - \frac{1}{4} \beta^4 \sum_{i=1}^N \cos 2\theta_i \sum_{j=1}^N \cos 2\theta_j \\ & - \frac{1}{4} \beta^4 \sum_{i=1}^N \sin 2\theta_i \sum_{j=1}^N \sin 2\theta_j. \end{aligned} \quad (3.24)$$

The above expression helps in obtaining the maximum value of  $|I(s, \mathbf{q})|$ .

**Proposition 2:** In particular,

1. When the number of sensors  $N$  is odd, the exhaustive set of sensor positions maximizing  $|I(s, \mathbf{q})|$  is given by Proposition 1.
2. When the number of sensors  $N$  is even,  $|I(s, \mathbf{q})|$  is maximum when
  - the angle between adjacent sensors is equal to  $k\pi/N$
  - the angle between  $N/2$  distinct pairs of sensors is equal to  $\pi/2$ .

**Proof:** (a) From equation 3.24, it can be observed that the maximum value of  $|I(s, \mathbf{q})|$  is obtained when

$$\sum_{i=1}^N \cos 2\theta_i \sum_{j=1}^N \cos 2\theta_j + \sum_{i=1}^N \sin 2\theta_i \sum_{j=1}^N \sin 2\theta_j = 0, \quad (3.25)$$

resulting in a value of  $|I(s, \mathbf{q})| = \frac{N^2}{4} \beta^4$ .

Further, since  $\sum_{i=1}^N \cos 2\theta_i \sum_{j=1}^N \cos 2\theta_j$  can be written as  $\left(\sum_{i=1}^N \cos 2\theta_i\right)^2$ , Equation 3.25 can be simplified as

$$\left(\sum_{i=1}^N \cos 2\theta_i\right)^2 + \left(\sum_{i=1}^N \sin 2\theta_i\right)^2 = 0.$$

Both the terms in the above expression are positive which implies that each of the terms has to be equal to 0 to satisfy the condition specified above. Therefore,  $|I(s, \mathbf{q})|$  is maximized when

$$\sum_{i=1}^N \cos 2\theta_i = 0 \text{ and } \sum_{i=1}^N \sin 2\theta_i = 0.$$

(b) Consider the following set of equations

$$\begin{aligned} \sum_{i=1}^N \cos 2\theta_i &= 0 \\ \sum_{i=1}^N \sin 2\theta_i &= 0. \end{aligned} \quad (3.26)$$

On simplification, Equation 3.26 reduces to

$$\begin{aligned} \sum_{i=1}^N \cos 2(\theta_i \sim \theta_j) &= 0 \\ \sum_{i=1}^N \sin 2(\theta_i \sim \theta_j) &= 0, \quad \forall j \in \{1, 2, \dots, N\}. \end{aligned} \quad (3.27)$$

The equation above can be represented in Euler's form resulting in

$$\sum_{i=1}^N e^{\sqrt{-1} \times 2(\theta_i \sim \theta_j)} = 0, \quad \forall j \in \{1, 2, \dots, N\}. \quad (3.28)$$

Using trigonometric identities, the value of  $\theta_i \sim \theta_j$  satisfying Equation 3.28 is given by

$$\begin{aligned} 2(\theta_i \sim \theta_j) &= \frac{2\pi k}{N}, \quad k \in Z \\ \Rightarrow (\theta_i \sim \theta_j) &= \frac{\pi k}{N}. \end{aligned} \quad (3.29)$$

Therefore, the angle between adjacent pair of sensors is equal to  $k\pi/N$  ( $k \in \mathbb{Z}$ ) when  $|I(s, \mathbf{q})|$  is maximum. For example, when  $N = 4$ ,  $|I(s, \mathbf{q})|$  is maximum when  $\theta_{21} = \theta_{32} = \theta_{43} = \frac{k\pi}{4}$  (From Proposition 1.(2)). If we partition the set of sensors in to two groups such that sensors 1 and 3 belong to one group, then  $|I(s, \mathbf{q})|$  is maximized if  $\theta_{31} = \theta_{42} = \frac{\pi}{2}$ . This proves that there are several orientations that result in the same FIM determinant.

**Proposition 3:** All the maxima of FIM determinant are equal, but do not necessarily result in the same localization error.

**Proof:** Consider the case when  $N$  is even. The sensors can be partitioned into  $N/2$  groups of 2 sensors such that the choice of group is mutually exclusive and collectively exhaustive. If  $g_1, g_2, \dots, g_{N/2}$  represents the partition of sensors where each  $g_i$  is a set containing the indices of the two sensors belonging to the group, then by definition,

$$\begin{aligned} \bigcup_{i=1}^{N/2} g_i &= \{1, 2, \dots, N\}, \\ \bigcap_{i=1}^{N/2} g_i &= \Phi. \end{aligned} \quad (3.30)$$

If  $\theta_{g_i}$  represents the angle between the two sensors in the  $i^{\text{th}}$  group  $g_i$ , Equation 3.26 can be written as

$$\begin{aligned} \sum_{i=1}^{N/2} (\cos 2\theta_{g_i,1} + \cos 2\theta_{g_i,2}) &= 0, \\ \sum_{i=1}^{N/2} (\sin 2\theta_{g_i,1} + \sin 2\theta_{g_i,2}) &= 0, \end{aligned} \quad (3.31)$$

where  $\theta_{g_i,1}$  represents the azimuthal angle of the first sensor in group  $g_i$  and  $\theta_{g_i,2}$  represents the azimuthal angle of the second sensor in group  $g_i$ . When  $\theta_{g_i} = \theta_{g_i,1} \sim \theta_{g_i,2} = \pi/2, \forall i \in 1, 2, \dots, N/2$ , the condition specified by Equation 3.31 is satisfied.

To corroborate the propositions, consider a network with three sensors constrained to be located at a distance of 2 units from the target. Figure 3.1 shows the variation of FIM determinant with the angle between two sensors and clearly marks the 8 different maxima obtained. For ease of illustration, the first sensor is assumed to subtend an angle of 0 with respect to the positive  $x$ -axis. If  $\theta_{21}$  represents the angle between sensors 2 and 1 and  $\theta_{31}$  represents the angle between sensors 3 and 1, it can be observed from Figure 3.1 that the value of FIM determinant is maximum for 8 different combinations of  $\{\theta_{21}, \theta_{31}\}$ . The orientation of the sensors corresponding to the 8 stationary points

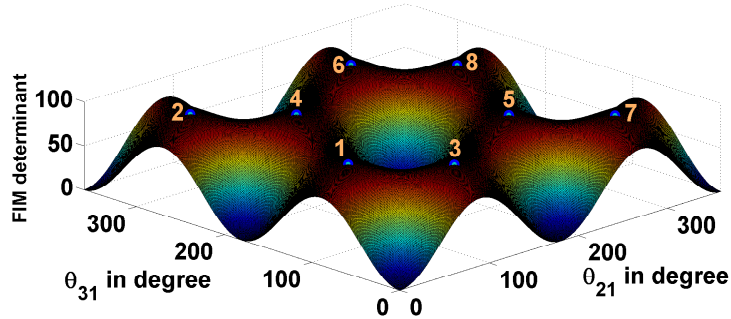


Figure 3.1: Variation of FIM determinant with  $\theta_{ij}$  and indication of multiple maxima.

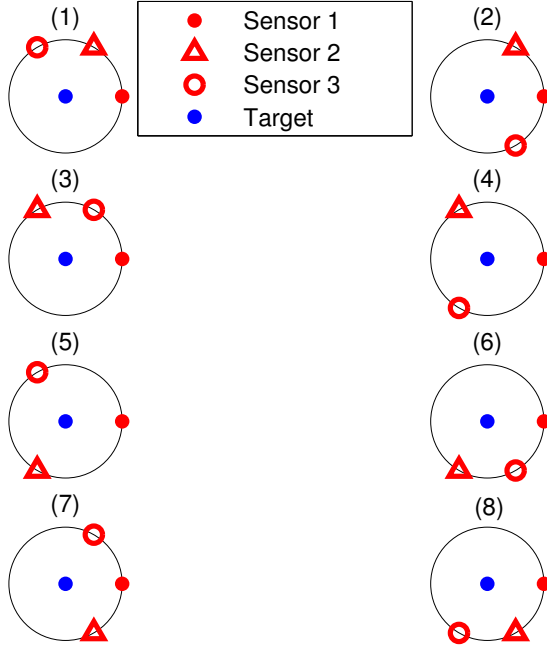


Figure 3.2: Orientations of the sensors corresponding to the maxima of FIM determinant.

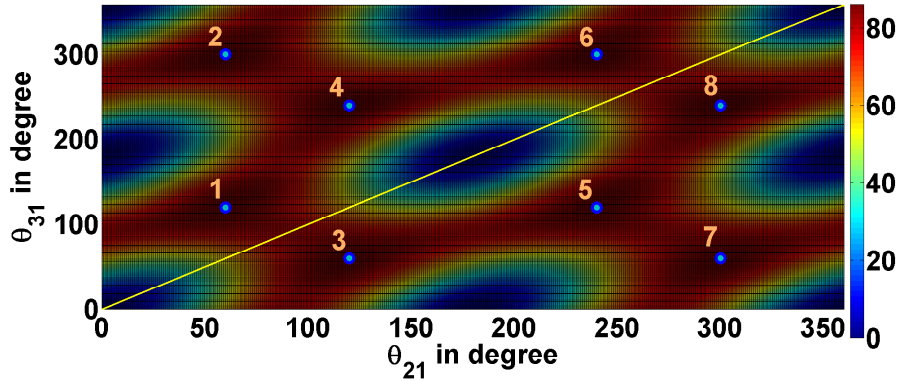


Figure 3.3: Plot indicating the symmetry of FIM determinant.

is shown in Figure 3.2. For convenience, the azimuthal angle of the first sensor is assumed to be 0. The values of  $\theta_{21}$  and  $\theta_{31}$  that maximize  $|I(s, \mathbf{q})|$  are equal to 60 or 120 which is in agreement with the claims of Proposition 1. Figure 3.2 shows the orientation of the sensors corresponding to each of the maxima of FIM determinant. However, there are some orientations that are equivalent i.e. the positions of the sensors are the same except that the labels of the sensors are interchanged. This is a direct consequence of the symmetry of the function  $|I(s, \mathbf{q})|$  about the line  $\theta_{21} = \theta_{31}$  as shown in Figure 3.3. Kindly note that the labeling of different orientations of sensors is consistent throughout this section. Equivalent orientations have the same performance in terms of localization error which is intuitive because the positions of sensors in the network are the same in both the orientations and



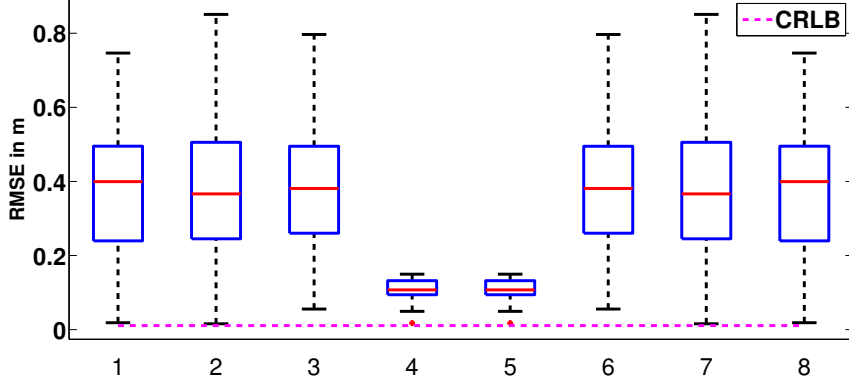


Figure 3.4: RMSE of different sensor orientations. The CRLB and the RMSE of the optimal positions are shown.

all sensors are assumed to be identical (same pathloss exponent). When each of these orientations is used in localizing the target, the average localization error expressed in terms of Mean Squared Error (MSE) is significantly different despite the fact that all these orientations have the same value of FIM determinant. This is because maximizing FIM determinant only implies that the lower bound on the localization error is minimized. This does not mean that the average localization error is the same for all the orientations. Figure 3.4 depicts graphically how the performance of different orientations vary on an average. The figure indicates the CRLB on the localization error and how the average MSE of each orientation varies in comparison. The average MSE of orientations (4) and (5) are comparable to the lower bound as opposed to the other orientations. In these two orientations, the angle between consecutive sensors is equal to 120. This proves that the multiple maxima of FIM determinant behave differently in a practical setup. It is also observed that a network with well distributed sensors is more efficient compared to a network with clustered sensors.

In a real life situation, the likelihood of the existence of several global maxima of FIM determinant is low. However, several local maxima of FIM determinant could perform as good as or better than the global maxima. This paper proves experimentally that the sensor position resulting in the lowest localization error is not necessarily a global maxima of FIM determinant.

### 3.5 Experimental Results

Optimal position of sensors is determined on a section of the terrain of Ananthagiri Hills located near the city of Vikarabad, Telangana in India. The terrain and GPS information of Ananthagiri Hills was collected on site and using the information collected at discrete points in the region, the terrain map of the entire region was estimated by Kriging based interpolation technique. The target is assumed to be present anywhere on a particular section of the road located in Ananthgiri Hills and the sensors are not allowed to be placed on the road. In order to experimentally validate the claims of the the paper, the following two scenarios are considered-

1. The target is stationary and the sensors are located on the ground terrain (3D)
2. The target is moving along the road and the sensors are located on the ground terrain (3D)

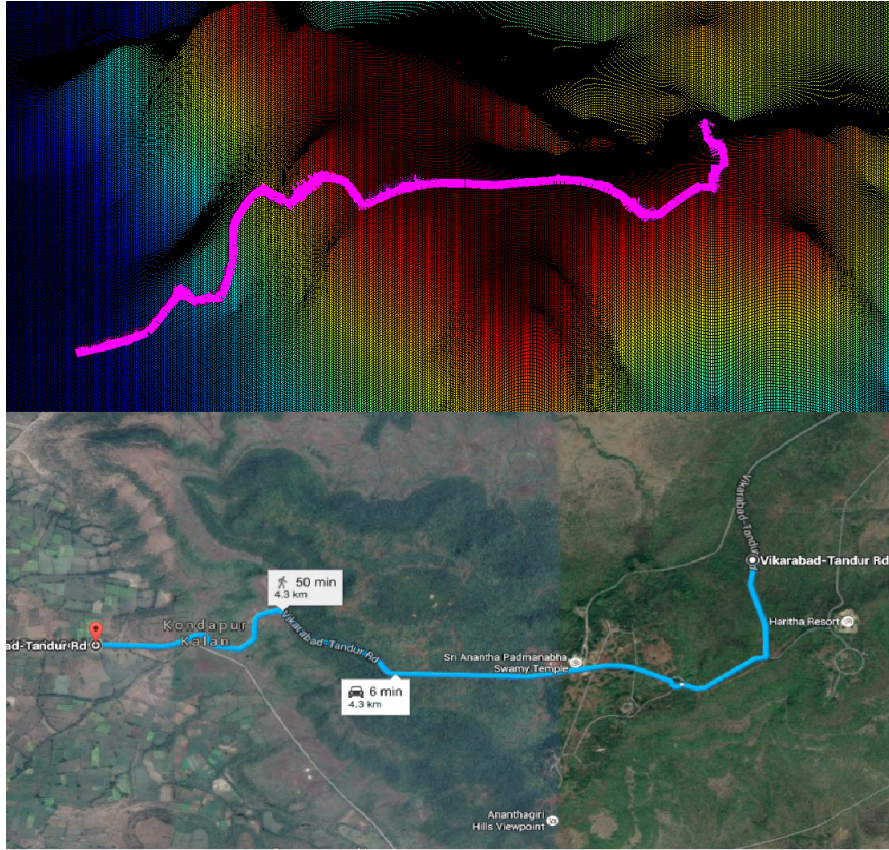


Figure 3.5: Section of Anathagiri Hills consider for practical experiments

The global and local maxima of FIM determinant in both the cases are determined using genetic algorithm and among the set of maxima, the one that provides the lowest average localization error is identified.

### 3.5.1 Stationary target and sensors located on the ground terrain

A section of a road located at Ananthagiri hills along with the surrounding terrain is considered. The target is assumed to be located on the road and the sensors are placed on a 3 dimensional terrain. The position of the target is fixed and known apriori and the optimal network of sensors corresponding to this target location is determined. Owing to the irregularities in the terrain, it is not possible to determine the optimal sensor positions analytically. The possibility of existence of multiple maxima of FIM determinant is also high. Therefore, several runs of genetic algorithm based optimization is conducted and the results are recorded. Table 3.1 shows some of the results obtained using genetic algorithm along with the corresponding sensor orientations. This table shows how different local maxima of  $|I(s, \mathbf{q})|$  result in a different sensor network compares the localization performance of each. Although, one might expect the localization error to be minimum for the global maxima of FIM determinant, there are some orientations of sensors with lower FIM determinant that result in a more accurate localization system. Figure 3.6 shows the difference in distribution of sensors with maximum FIM determinant and one that corresponds to minimum average localization error. Clearly, the two sets of sensors have different orientations. This proves experimentally that

the best network of sensors is not necessarily the global maximum of FIM determinant. Figure 3.7 shows graphically the difference in the performance of the two networks shown in Figure 3.6. The variation in mean squared error (MSE) of the two networks over several iterations is shown in the figure. The plot in blue corresponds to the network with maximum FIM determinant while the plot in red corresponds to the network with the lowest average localization error.

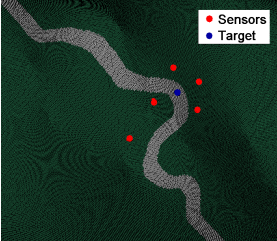
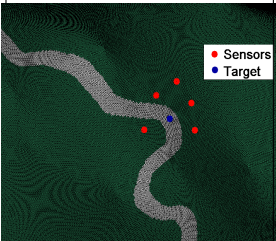
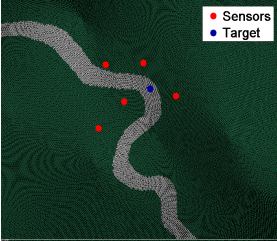
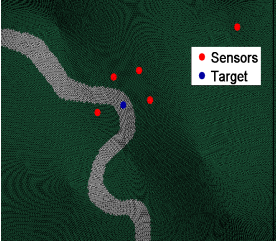
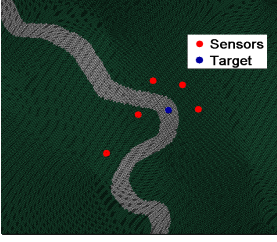
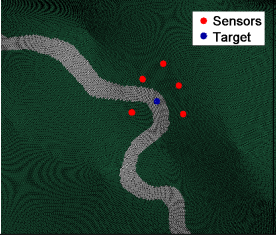
Sensors	Attributes	Sensor locations	Attributes
	$FIM_d = 1.2 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 3.4m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 1.6m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 13.4m$ <b>RMSE = 8.06 m</b>		$FIM_d = 1.1 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 3.5 m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 1.7m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 13.1 m$ <b>RMSE = 7.9113 m</b>
	$FIM_d = 0.5 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 65.31 m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 68.29 m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 24.28 m$ <b>RMSE = 56.3295 m</b>		$FIM_d = 0.6 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 12.27 m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 9.58 m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 26.8 m$ <b>RMSE = 17.8943 m</b>
	$FIM_d = 0.9 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 3.03 m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 2.28m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 14.7 m$ <b>RMSE = 8.7627 m</b>		$FIM_d = 0.8 \times 10^{-10} m^{-6}$ $E[ \mathbf{x}_s - \hat{\mathbf{x}}_s ] = 1.56 m$ $E[ \mathbf{y}_s - \hat{\mathbf{y}}_s ] = 1.96 m$ $E[ \mathbf{z}_s - \hat{\mathbf{z}}_s ] = 7.35m$ <b>RMSE = 4.489 m</b>

Table 3.1: Maxima of  $|I(s, \mathbf{q})|$  and performance comparison

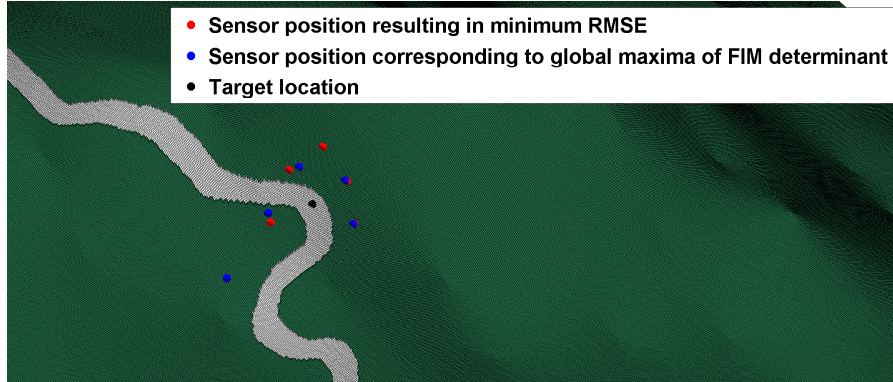


Figure 3.6: Figure depicting the position of sensors that maximize FIM determinant and provide lowest localization error when the target is stationary and located on the road.

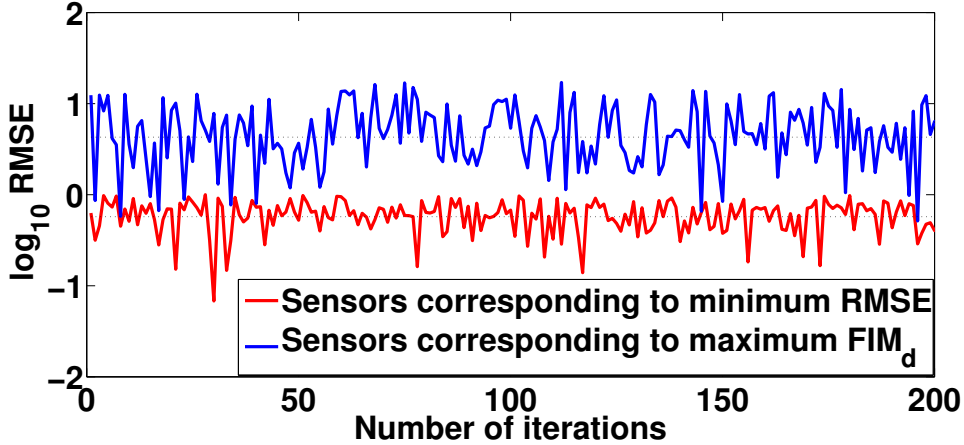


Figure 3.7: MSE comparison of the two sets of sensor positions that localize a target on the road shown in Figure 3.6

### 3.5.2 Moving target and sensors located on the ground terrain

The following practical scenario is considered where a target is moving along the road and the sensors are placed on either side of the road to localize the target at any time instant. Although the target location is not precisely known, using the knowledge about the probable locations of the target on the road, the optimal sensor positions are now determined for that section of the road. In this situation, Equation 3.5 cannot be maximized directly since the location of the target is unknown. Hence, we assume a prior distribution on the position of the target and use Equation 3.6 to determine the optimal positions of the sensors. When a uniform prior is assumed, the target is equally likely to be present anywhere on that section of the road. The objective function for optimization is the expected value of  $|I(s, \mathbf{q})|$  (the expectation taken over the prior distribution on location of the target).

Similar to the previous situation, the network of sensors resulting in the lowest localization error is not the global maxima of  $E_s [|I(s, \mathbf{q})|]$ . Figure 3.8 and 3.10 show how the optimal sensor network differs from the network with maximum  $E_s [|I(s, \mathbf{q})|]$  for two different sections of the road. Despite the fact that the network shown in red is not the global maxima of the objective function, the average localization error given by the root of the expected mean squared error obtained using that network is the lowest. Figure 3.9 and 3.11 provides the graphical comparison of the two networks in terms of localization error for the two sections of the road in Figures 3.8 and 3.10 respectively. About 100 points on the section of the road were picked randomly and localized using both the networks.

Although directly maximizing FIM determinant is computationally simplistic, the local maxima provides a solution that performs nearly 50% better than the global maxima. This, however, comes with a computational overhead which is negligible because the process of sensor placement is conventionally offline. The proposed algorithm brings out the computational elegance of FIM based optimization as well as the accuracy of exhaustive enumeration.

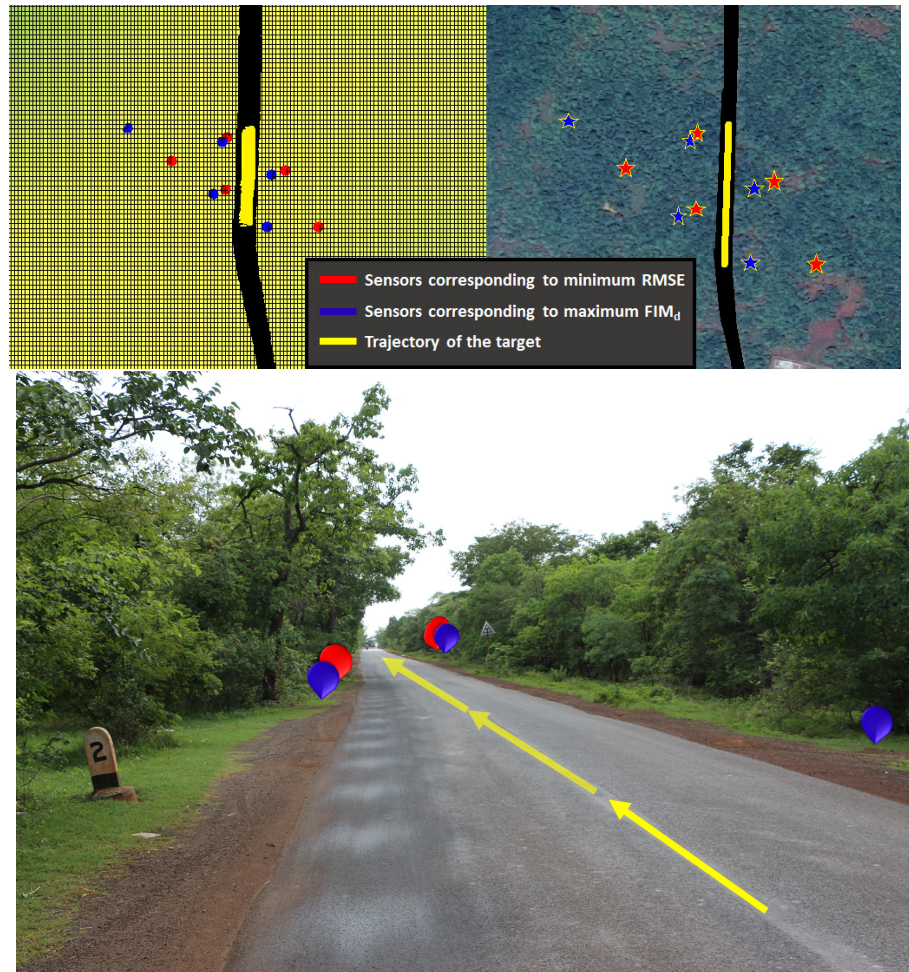


Figure 3.8: Figure depicting the position of sensors that maximize FIM determinant and position of sensors that provide lowest localization error when the target could be located anywhere on a small stretch of the road.

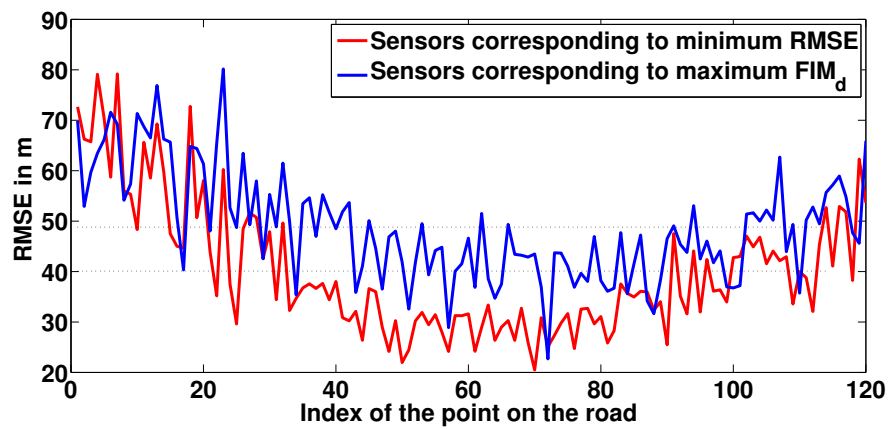


Figure 3.9: MSE comparison of the two sets of sensor positions that localize a target on the trajectory shown in Figure 3.8

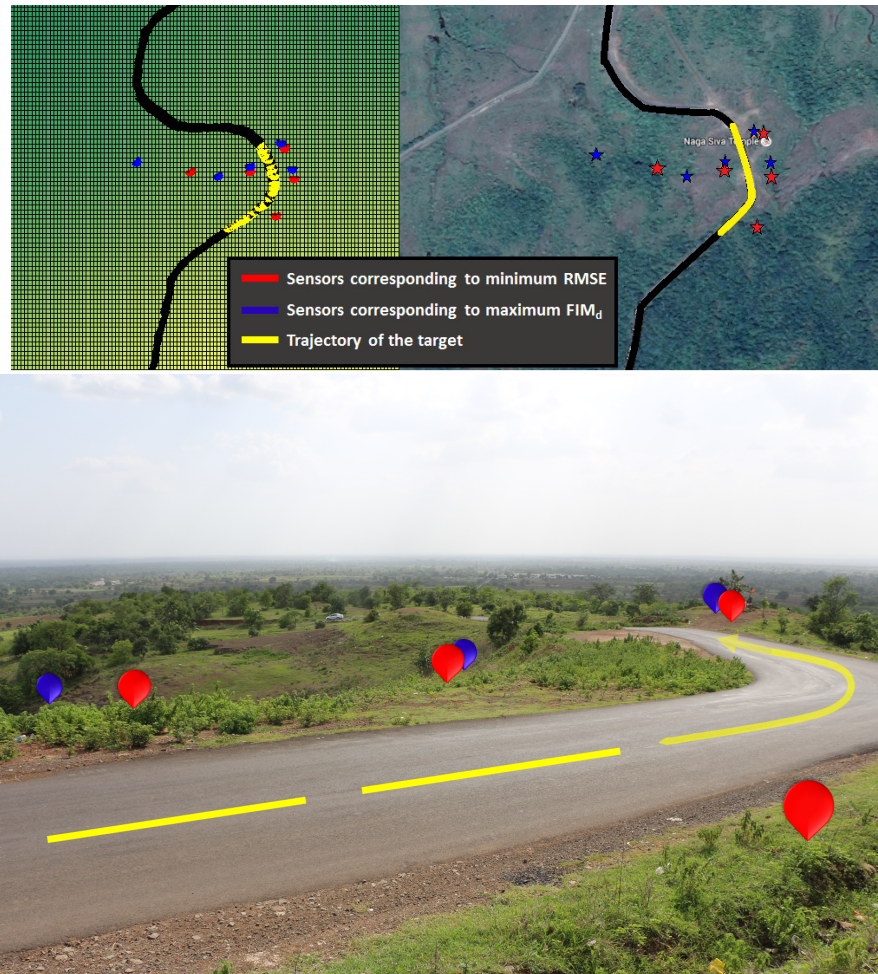


Figure 3.10: Figure depicting the position of sensors that maximize FIM determinant and position of sensors that provide lowest localization error when the target could be located anywhere on a small stretch of the road.

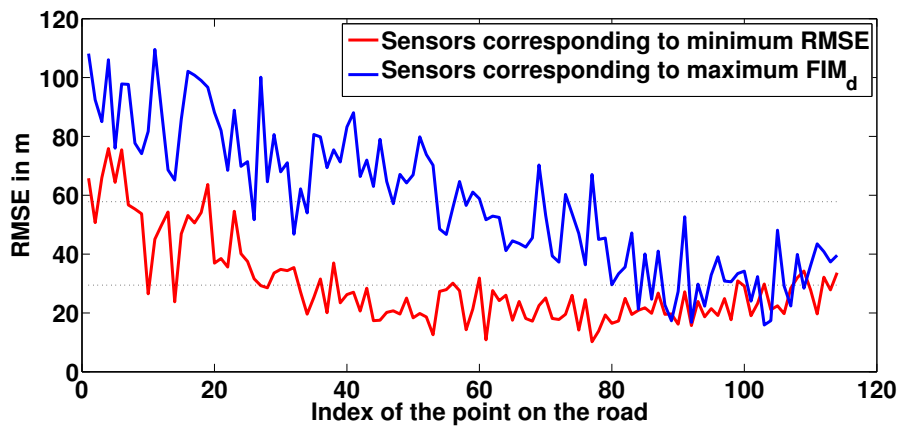


Figure 3.11: MSE comparison of the two sets of sensor positions that localize a target on the trajectory shown in Figure 3.10

# Chapter 4

## Vision-based tracking of objects in 3D

### 4.1 Introduction

Object tracking is a very important technique that is often applied in robotics vision, surveillance systems and other industrial applications. Basically object tracking is conducted through the comparison of the characteristics of the target object and video images in the search region. Most of the tracking algorithms are some form of optimization which are tackled using either deterministic or stochastic methods. Deterministic methods usually involve a gradient descent search to minimize a cost function. Compared with the deterministic counterparts, stochastic methods are usually more robust, but they suffer a large computational load, especially in high-dimensional state space.

The main challenge to robust tracking is the lack of visibility of the object in crowded scenes. This problem can be overcome to an extent by the use of multiple cameras that monitor the same

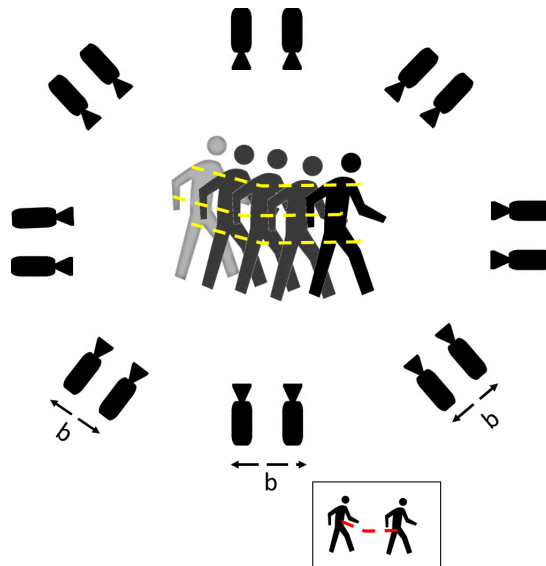


Figure 4.1: Graphical representation of multi-view tracking

scene at different angles. The probability that the object is not visible in all the cameras is low and the effect of occlusion will therefore be reduced to a great extent. This can help in tracking the object seamlessly. Figure 4.1 shows a typical multi-camera network that is used to track the object in 3D. Each of these cameras do not function independently in providing an accurate 3D track. Therefore, the knowledge of multiple camera networks and camera geometry is necessary to locate the object in all the views and fuse the total information.

Calibration of the camera network is a vital step towards 3D representation of objects and tracking. Calibration is a process of determining the intrinsic parameters of the camera including its local length, aspect ratio and skew and the extrinsic parameters like the rotation and translation of the camera center with respect to a reference coordinate system. The details of camera geometry and calibration is provided in Appendix A.

This work considers a network of stereo cameras monitoring the same object from different angles. The use of stereo cameras exhibits two fold advantage: Firstly, the process of calibration is made independent of extraneous objects due to the baseline restriction in the stereo pair. The second advantage is that the use of stereo cameras provides information that can be correlated with human perception because the human eyes can be modeled as a stereo camera. For ease of illustration, this work demonstrates the tracking of object in 3D by the use of a single pair of stereo cameras. Once the track of the object due to a single stereo pair is obtained, it is possible to fuse the information from the other stereo pairs to obtain a complete 3D track of the target of interest. The basics of stereo geometry and the autocalibration technique proposed by Kiran *et al.* [44] adopted is explained in Appendix A. Using a set of calibrated cameras, the object is tracked using Kanade-Lucas-Tomasi (KLT) point tracking algorithm. The details of the tracking algorithm is described in the subsequent section with the corresponding experimental results.

## 4.2 Sparse point-cloud generation using Stereo pair

Once the camera matrices  $P_1$  and  $P_2$  of the stereo pair are computed, it is possible to obtain a 3D point cloud of certain feature points that are common across both the views. The feature points are identified using ASIFT (Affine Scale-Invariant Feature Transform) [45] and are projected back on to the 3D space. If a set of point correspondences in two views determine the fundamental matrix uniquely, then the scene and cameras may be reconstructed from these correspondences alone. This process is known as triangulation.

Figure 4.2 shows pictorially how triangulation is used to obtain a sparse 3D point cloud. If  $v_L$  and  $v_R$  represent the images of a 3D point  $P$ , it is possible to obtain the coordinates of the point  $P$  with the knowledge of the focal lengths of the two cameras and the baseline separation between the stereo cameras.



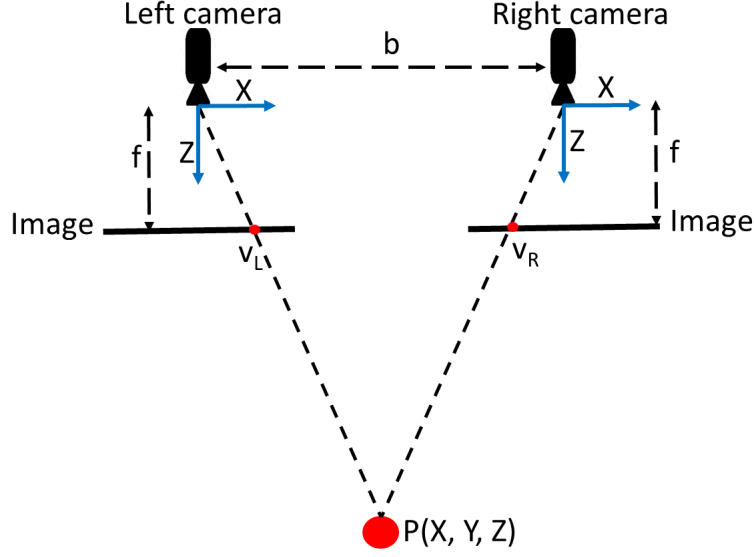


Figure 4.2: Triangulation using stereo pair

### 4.3 Kanade-Lucas-Tomasi Point tracking algorithm

The point tracker algorithm tracks the set of detected feature points using the Kanade-Lucas-Tomasi (KLT), that is basically inspired by the concept of optical flow. It was originally designed to address the issue of image registration [46]. It was later used for video stabilization, camera motion estimation, and object tracking. It works particularly well for tracking objects that do not change shape and for those that exhibit visual texture.

If  $F(x)$  and  $G(x)$  represent the image intensity of a point in subsequent frames, this algorithm aims to estimate the disparity  $h$  that matches the two points with maximum accuracy. In other words, the disparity is calculated such that the difference in the some measure between  $F(x+h)$  and  $G(x)$  is minimized for all the feature points  $x$  in region of interest  $R$ . The following are the most commonly used measures of difference:

- $L_1$  norm =  $\sum_{x \in R} |F(x+h) - G(x)|$
- $L_2$  norm =  $\sqrt{\sum_{x \in R} [F(x+h) - G(x)]^2}$

KLT tracker is based on two papers: The first one uses a gradient-based approximation to determine the disparity  $h$ . The second paper provides a generalized extension to the first by not restricting the kind of displacement to just translation.

If  $h$  is the displacement between the two images  $F(x)$  and  $G(x) = F(x+h)$ , the following approximation is made

$$F'(x) \approx \frac{F(x+h) - F(x)}{h} = \frac{G(x) - F(x)}{h}. \quad (4.1)$$

Therefore,

$$h \approx \frac{G(x) - F(x)}{F'(x)} \quad (4.2)$$

Considering all the feature points, an average estimate of disparity is obtained as follows

$$h \approx \sum_x \frac{G(x) - F(x)}{F'(x)}. \quad (4.3)$$

The accuracy of the estimate is further improved by incorporating a weighting function that is derived from the second derivative of  $F(x)$ . If the weighting function  $w(x)$  is defined as

$$w(x) = \frac{1}{|G'(x) - F'(x)|}, \quad (4.4)$$

the average weighted disparity is given by

$$h \approx \frac{\sum_x \frac{w(x)[G(x) - F(x)]}{F'(x)}}{\sum_x w(x)}. \quad (4.5)$$

As the point tracker algorithm progresses over time, points can be lost due to lighting variation, out of plane rotation, or articulated motion. To track an object over a long period of time, you may need to reacquire points periodically. The second paper proposed by Tomasi and Kanade [47] provides a framework for identifying the feature points that are suitable for tracking. The detected features are selected if both the eigen values of the gradient matrix are larger than some threshold.

## 4.4 Experimental Results

In this work, the set of point correspondences are obtained using Affine SIFT in both the views of the stereo camera. Using the camera parameters and the point correspondences, a sparse 3D point cloud is obtained. Figure 4.3 shows the common feature points that are detected in both the views

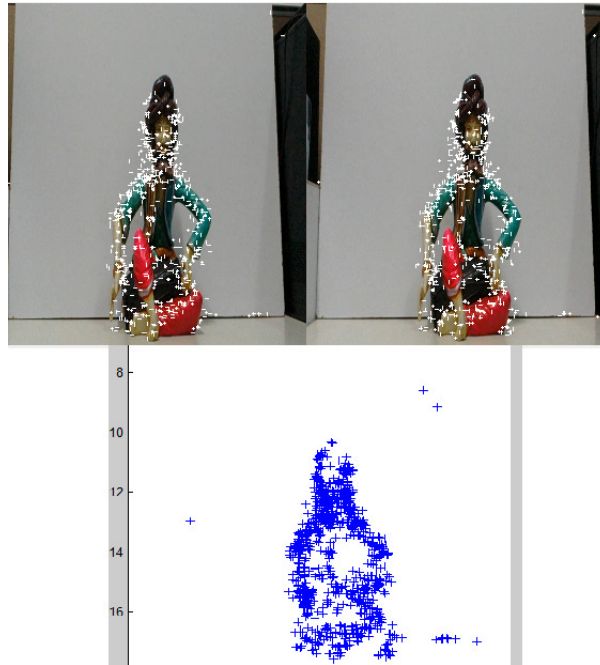


Figure 4.3: 3D point cloud using a pair of calibrated stereo cameras

and the point cloud generated using triangulation.

Using the feature points detected, the KLT tracker is run on the video sequence of the left camera. The corresponding location of the points in the other view is mapped using Fundamental matrix. The tracked locations are further projected on to the 3D continuum using triangulation. Figure 4.4 shows the results of the algorithm both in 2D as well as 3D.

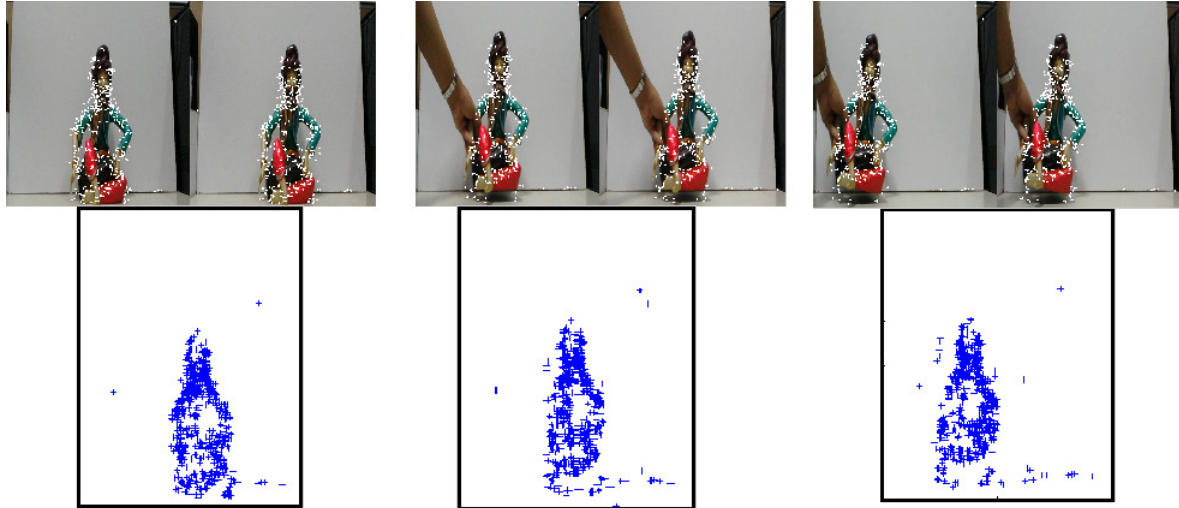


Figure 4.4: Point cloud tracking from a stereo pair: Frames 1, 20 and 40 (Left to Right)

# Chapter 5

## Summary and Discussion

The two main areas of this work include localization and tracking of objects in 3 dimensions. Localization involves estimating the location of an object with respect to some reference points and is analogous to detection in case of videos. In this regard, a novel algorithm is proposed that has a computational advantage over traditional algorithms and provides significantly better results. Additionally, the problem of optimal placement of sensors is considered in order to improve the efficiency of the proposed localization system. Finally, the problem of object tracking is addressed using visual data from a stereo camera network. In the course of these studies, several interesting insights were obtained. This chapter summarizes the work in each field and discusses limitations and future scope in this direction.

### 5.1 Object Localization in 3D

In this thesis, PSO-based localization is proposed to estimate the location of a target in 3D. We introduce a novel modification that is demonstrated to be computationally faster compared to standard PSO. The proposed modification also eliminates the flip ambiguity inherent in the standard version. A cluster-based approach is adopted to ensure that the initialization of PSO is accurate enough to eliminate flip ambiguity. The cluster with lower volume is identified to be the right cluster and therefore, the particles of the PSO algorithm are initialized around the centroid of that cluster. The notion of the minimum volume of cluster stems from experimental observations and it lacks a mathematical explanation. Therefore, the next step would be to provide an analytical explanation for the observation to highlight the need for a cluster-based initialization.

### 5.2 Sensor Placement Problem

A mathematical framework to determine the optimal position of sensors to localize an unknown target in 3D is proposed. In order to ensure that maximum localization accuracy is obtained, a statistical measure of inaccuracy i.e. covariance is used. Due to the computational complexity of covariance, the bound on covariance, provided by Cramer Rao Lower Bound, is computed instead. CRLB relates the covariance of the estimated location to Fisher Information matrix that exhibits a closed form expression. Maximizing the determinant of Fisher Information matrix reduces the

volume of the error ellipsoid which in turn minimizes the variance of the estimated location. A keen analysis of the expression of FIM determinant reveals that there are several orientations that maximize FIM determinant. This is experimentally verified by considering a toy example where the location of the target is known and the sensors are restricted to lie on the circumference of a circle with the target at the center. This example also reveals that the localization performance of each of the maxima of FIM determinant is significantly different. This is because maximizing FIM determinant only ensures that the lower bound on the variance is minimized, not the actual variance. Therefore, a two-step optimization algorithm is proposed that first determines all the local and the global maxima of FIM determinant and then exhaustively determines the sensor orientation with the lowest average localization error. The proposed algorithm is corroborated using practical examples. When the optimal position of the sensors to localize a specific target on the road was determined, it was observed that the global maxima of FIM determinant (value=  $1.237 \times 10^{-10} m^{-6}$ ) resulted in an average RMSE of 8.08 *m* whereas the sensor positions corresponding to the FIM determinant value of  $0.8701 \times 10^{-10} m^{-6}$  provided the lowest average RMSE of 4.489 *m*. Therefore, our proposed algorithm provides improved results that comes with an additional computational cost. However, considering the fact that the sensor placement problem is usually addressed offline and the improvement in the results obtained is about 50%, the additional computational overhead is admissible. In this thesis, the objective function is considered to be isotropic i.e., the errors along the *x*, *y* and *z*-direction are considered to be the same. In several applications, especially terrestrial tracking, the error along the *z*-direction could be neglected. This means that a weighted objective function could be considered instead so that the effect of the *z*-error could be neglected. In addition, this framework is for deploying range-based sensors. It could be extended to video cameras because similar to sensor-based tracking, the position of the cameras play a significant role in the accuracy of visual tracking. Therefore, this work can serve as a platform for a systematic framework to determine the optimal position of cameras with suitable modification to the objective function considered. The details will be available in my forthcoming paper.

### 5.3 Vision-based tracking of objects in 3D

Stereoscopic vision has helped in reconstruction and tracking of an object in a multicamera system. This thesis aims to provide a platform for 3D object tracking. KLT feature tracker is used to identify and track certain feature points present on the object. Using a set of calibrated stereo cameras, the feature points are mapped in 3D and also projected on to the other view. It also paves way for several improvements. Most of the cameras are equipped with Autofocus facility. In order to incorporate the effects of autofocus, it is necessary to update the camera parameters with time along with the motion vectors. In addition, the information from multiple stereo pairs has to be fused in order to obtain a more accurate location information and faithful reconstruction. Finally, the object can be tracked in 3D directly using 3D variation of algorithms like Particle Filter.

# Appendix A

## Theoretical Background

Multi view geometry is used to map the points (location of the object) in one view to the other views. This can be done by considering a pair of cameras at a time. One of the cameras (where the object remains unoccluded throughout) is fixed as a reference and the mapping to the other views is done with respect to that. Therefore the multi view network can be interpreted as a series of pairs of cameras. Therefore we will first consider the geometry of two perspective views only. The following steps are involved in mapping points from one view to another.

- Calibration of the individual cameras
- Finding the homography matrix that maps pixel locations in one view to the other

### A.1 Camera Model and Need for Calibration

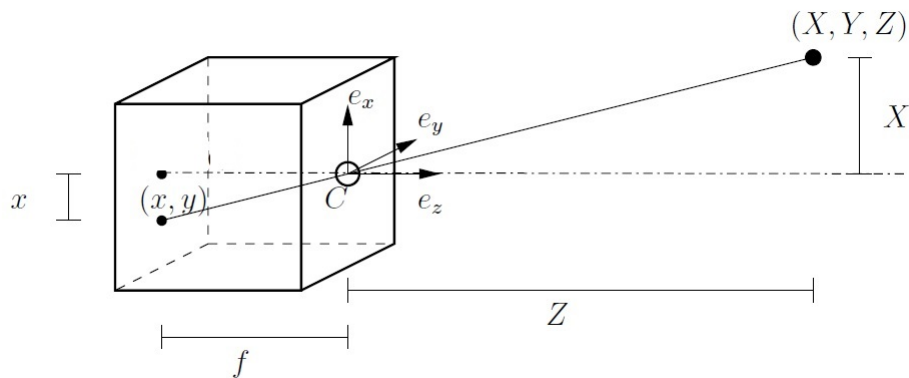


Figure A.1: Pinhole model of a camera

Figure A.1 represents the process of image formation in a camera using a pin hole model. Using

the laws of optics, the process of image formation is represented mathematically as,

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_K \underbrace{\begin{pmatrix} r_{11} & r_{12} & r_{13} & | & t_1 \\ r_{21} & r_{22} & r_{23} & | & t_2 \\ r_{31} & r_{32} & r_{33} & | & t_3 \end{pmatrix}}_{[R|t]} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (\text{A.1})$$

In order to invert the process of image formation to obtain a 3D representation of the object from 2D images, it is mandatory to know the parameters of the camera matrix. Camera calibration is a necessary step in 3D computer vision in order to extract metric information from 2D images. Image is nothing but the projection of a 3 dimensional scene on to a 2d plane. Camera calibration helps us identify the transformations that map the 3d world coordinates to the 2d image coordinates. Conventional techniques use an extraneous object for calibration. These technique requires to observe a planar pattern shown at a few different orientations to estimate the projection matrix. The simplest would be the chess board pattern where the dimensions of each square is known. Using this knowledge and the images of the chess board pattern in different orientation, the projection matrix can be obtained. The assumption made in this type of calibration is that the model plane is on  $Z=0$  of the world coordinate system. To overcome the burden of using an external object for calibration, Kiran *e al.* proposed a calibration technique that eliminates the need for a calibration object. In the proposed method, calibration is performed without imposing any restriction on the scene.

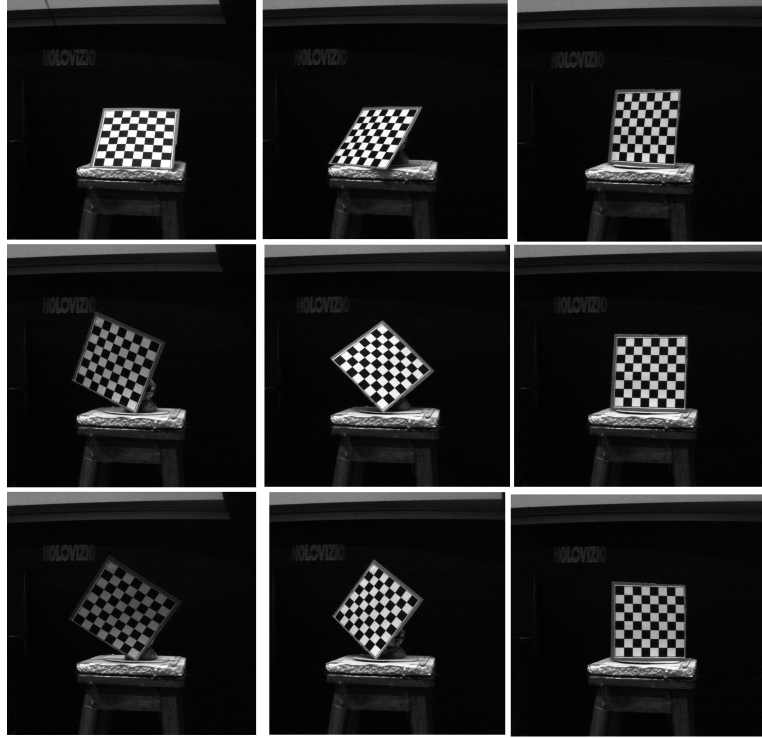


Figure A.2: Camera calibration using extraneous objects

## A.2 Auto Calibration

An non-intrusive calibration technique proposed by Kiran *et al.* eliminates the need for manual intervention and provides a faithful Euclidean reconstruction of the scene. The authors propose a three camera network which comprises of a stereo pair. The fixed baseline separation between the stereo cameras facilitates the unique estimation of the camera parameters, both intrinsic and extrinsic. This technique could be extended to a network consisting of more than 3 cameras with ease. In this framework, the authors consider a three-camera configuration, where two of the cameras form a stereo pair with parallel optical axes with known baseline separation. Consider  $N$  points on the object in 3D which are visible in all the three cameras. The camera model equations are written for all the three cameras with the left camera of the stereo pair as reference. The rotation matrices of the stereo pair is hence identity and the translation vector of the stereo right with respect to left is equal to the baseline separation. If the  $j^{th}$  3D point is given by  $\bar{X}_j = [X_j, Y_j, Z_j, 1]^T$  and the corresponding image on the  $i^{th}$  camera is denoted as  $\bar{x}_{ij} = [x_{ij}, y_{ij}, 1]^T$ , the camera equations of the stereo cameras are given as follows:

$$s_{1j} \begin{bmatrix} x_{1j} \\ y_{1j} \\ 1 \end{bmatrix} = \begin{pmatrix} f_1 & 0 & 0 \\ 0 & f_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0|0 \\ 0 & 1 & 0|0 \\ 0 & 0 & 1|0 \end{pmatrix} \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix} \quad (\text{A.2})$$

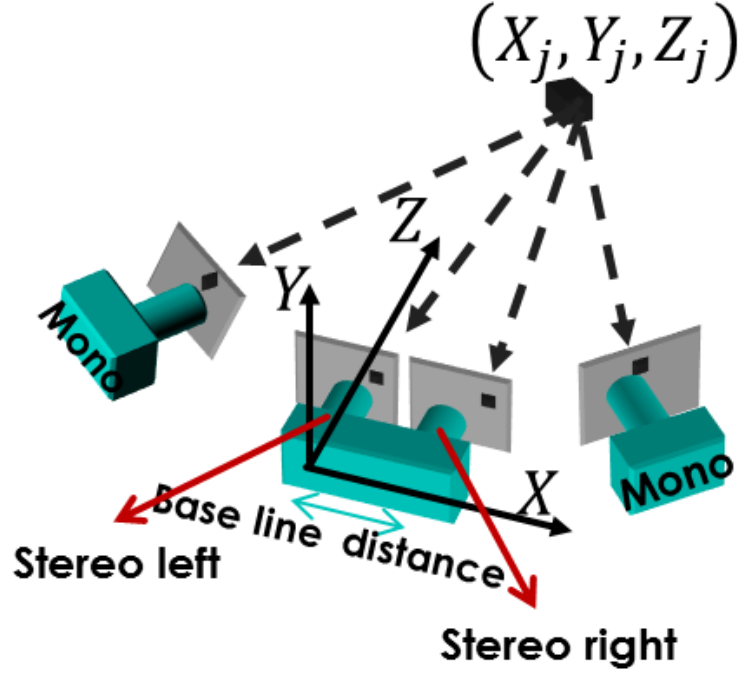


Figure A.3: Euclidean Autocalibration



$$s_{2j} \begin{bmatrix} x_{2j} \\ y_{2j} \\ 1 \end{bmatrix} = \begin{pmatrix} f_2 & 0 & 0 \\ 0 & f_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0|-l \\ 0 & 1 & 0|0 \\ 0 & 0 & 1|0 \end{pmatrix} \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix} \quad (\text{A.3})$$

The equation of the mono cameras can be written as

$$s_{ij} \begin{bmatrix} x_{ij} \\ y_{ij} \\ 1 \end{bmatrix} = \begin{pmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1^i & r_2^i & r_3^i|t_x \\ r_4^i & r_5^i & r_6^i|t_y \\ r_7^i & r_8^i & r_9^i|t_z \end{pmatrix} \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix} \quad (\text{A.4})$$

Using the above, the authors determine the parameters of the camera network by means of a multi-objective optimization.

# References

- [1] L. Xiao, L. J. Greenstein, and N. B. Mandayam, “Sensor-assisted localization in cellular systems,” *Wireless Communications, IEEE Transactions on*, vol. 6, no. 12, pp. 4244–4248, 2007.
- [2] Y. Shi and Y. T. Hou, “Optimal base station placement in wireless sensor networks,” *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, no. 4, p. 32, 2009.
- [3] S. S. Dhillon and K. Chakrabarty, *Sensor placement for effective coverage and surveillance in distributed sensor networks*. IEEE, 2003, vol. 3.
- [4] G. Wang, G. Cao, and T. La Porta, “Movement-assisted sensor deployment,” *Mobile Computing, IEEE Transactions on*, vol. 5, no. 6, pp. 640–652, 2006.
- [5] P. González-Brevis, J. Gondzio, Y. Fan, H. V. Poor, J. Thompson, I. Krikidis, and P.-J. Chung, “Base station location optimization for minimal energy consumption in wireless networks,” in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*. IEEE, 2011, pp. 1–5.
- [6] J. Zhou, J. Shi, and X. Qu, “Landmark placement for wireless localization in rectangular-shaped industrial facilities,” *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 6, pp. 3081–3090, 2010.
- [7] Y. Chen, J.-A. Francisco, W. Trappe, and R. P. Martin, “A practical approach to landmark deployment for indoor localization,” in *Sensor and Ad Hoc Communications and Networks, 2006. SECON’06. 2006 3rd Annual IEEE Communications Society on*, vol. 1. IEEE, 2006, pp. 365–373.
- [8] S. S. Ponda, R. M. Kolacinski, and E. Frazzoli, “Trajectory optimization for target localization using small unmanned aerial vehicles,” Ph.D. dissertation, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, 2008.
- [9] C. Tessier, C. Cariou, C. Debain, F. Chausse, R. Chapuis, and C. Rousset, “A real-time, multi-sensor architecture for fusion of delayed observations: application to vehicle localization,” in *Intelligent Transportation Systems Conference, 2006. ITSC’06. IEEE*. IEEE, 2006, pp. 1316–1321.
- [10] J. Levinson, M. Montemerlo, and S. Thrun, “Map-based precision vehicle localization in urban environments.” in *Robotics: Science and Systems*, vol. 4. Citeseer, 2007, p. 1.
- [11] X. Sheng, Y.-H. Hu, and P. Ramanathan, “Distributed particle filter with gmm approximation for multiple targets localization and tracking in wireless sensor network,” in *Proceedings of the*

- 4th international symposium on Information processing in sensor networks.* IEEE Press, 2005, p. 24.
- [12] J. Teng, H. Snoussi, C. Richard, and R. Zhou, “Distributed variational filtering for simultaneous sensor localization and target tracking in wireless sensor networks,” *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 5, pp. 2305–2318, 2012.
- [13] J. R. Spletzer and C. J. Taylor, “Dynamic sensor planning and control for optimally tracking targets,” *The International Journal of Robotics Research*, vol. 22, no. 1, pp. 7–20, 2003.
- [14] V. Isler, S. Khanna, J. Spletzer, and C. J. Taylor, “Target tracking with distributed sensors: The focus of attention problem,” *Computer Vision and Image Understanding*, vol. 100, no. 1, pp. 225–247, 2005.
- [15] W.-P. Chen, J. C. Hou, and L. Sha, “Dynamic clustering for acoustic target tracking in wireless sensor networks,” *Mobile Computing, IEEE Transactions on*, vol. 3, no. 3, pp. 258–271, 2004.
- [16] S. Martínez and F. Bullo, “Optimal sensor placement and motion coordination for target tracking,” *Automatica*, vol. 42, no. 4, pp. 661–668, 2006.
- [17] S. Mihail and V. Ramadurai, “Localization of wireless sensor networks with a mobile beacon,” *Mobile Ad Hoc and Sensor Systems, IEEE*, pp. 174–183, 2004.
- [18] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, “Range-free localization schemes for large scale sensor networks,” in *Proceedings of the 9th annual international conference on Mobile computing and networking.* ACM, 2003, pp. 81–95.
- [19] S. Lee and K. Kim, “3d localization with a mobile beacon in wireless sensor networks,” in *2012 IEEE Sensors*, 2012.
- [20] A. N. Bishop and P. Jensfelt, “An optimality analysis of sensor-target geometries for signal strength based localization,” in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2009 5th International Conference on.* IEEE, 2009, pp. 127–132.
- [21] V. Seshadri, G. V. Zaruba, and M. Huber, “A bayesian sampling approach to in-door localization of wireless devices using received signal strength indication,” in *Pervasive Computing and Communications, 2005. PerCom 2005. Third IEEE International Conference on.* IEEE, 2005, pp. 75–84.
- [22] X. Li, “Collaborative localization with received-signal strength in wireless sensor networks,” *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 6, pp. 3807–3817, 2007.
- [23] Y.-T. Chan, W.-Y. Tsui, H.-C. So, and P.-c. Ching, “Time-of-arrival based localization under nlos conditions,” *Vehicular Technology, IEEE Transactions on*, vol. 55, no. 1, pp. 17–24, 2006.
- [24] J. T. Isaacs, D. J. Klein, and J. P. Hespanha, “Optimal sensor placement for time difference of arrival localization,” in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on.* IEEE, 2009, pp. 7878–7884.

- [25] L. Lin, H.-C. So, F. K. Chan, Y. Chan, and K. Ho, “A new constrained weighted least squares algorithm for tdoa-based localization,” *Signal Processing*, vol. 93, no. 11, pp. 2872–2878, 2013.
- [26] B. Yang, “Different sensor placement strategies for tdoa based localization,” in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II–1093.
- [27] R. Peng and M. L. Sichitiu, “Angle of arrival localization for wireless sensor networks,” in *Sensor and Ad Hoc Communications and Networks, 2006. SECON’06. 2006 3rd Annual IEEE Communications Society on*, vol. 1. IEEE, 2006, pp. 374–382.
- [28] A. N. Bishop, B. Fidan, K. Doğançay, B. D. Anderson, and P. N. Pathirana, “Exploiting geometry for improved hybrid aoa/tdoa-based localization,” *Signal Processing*, vol. 88, no. 7, pp. 1775–1791, 2008.
- [29] F. Thomas and L. Ros, “Revisiting trilateration for robot localization,” *IEEE Transaction on Robotics*, vol. 21, no. 1, February 2005.
- [30] Y. Zhang, S. Liu, and Z. Jia, “Localization using joint distance and angle information for 3d wireless sensor networks,” *IEEE Communications Letters*, vol. 16, no. 6, pp. 809–811, June 2012.
- [31] M. Ibraheem, “Towards a 3d localization system for an intelligent walker in uneven outdoor environments,” *Methods and Models in Automation and Robotics, IEEE*, vol. 4, pp. 1942–1948, 1995.
- [32] Q. Zhang, J. Huang, J. Wang, C. Jin, J. Ye, W. Zhang, and J. Hu, “A two-phase localization algorithm for wireless sensor network,” in *Information and Automation, 2008. ICIA 2008. International Conference on*. IEEE, 2008, pp. 59–64.
- [33] R. Yarlagadda, I. Ali, N. Al-Dhahir, and J. Hershey, “Gps gdop metric,” *IEE Proceedings-radar, sonar and navigation*, vol. 147, no. 5, pp. 259–264, 2000.
- [34] N. Levanon, “Lowest gdop in 2-d scenarios,” *IEE Proceedings-radar, sonar and navigation*, vol. 147, no. 3, pp. 149–155, 2000.
- [35] H. Wang, K. Yao, and D. Estrin, “Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking,” *Communications and Networks, Journal of*, vol. 7, no. 4, pp. 438–449, 2005.
- [36] E. Emre, J. Fisher, and L. C. Potter, “Maximum mutual information principle for dynamic sensor query problems,” *Information processing in sensor networks, Springer Berlin Heidelberg*, 2003.
- [37] H. Wang, K. Yao, G. Pottie, and D. Estrin, “Entropy-based sensor selection heuristic for target localization,” *Proceedings of the 3rd international symposium on Information processing in sensor networks. ACM*, 2004.
- [38] J. Perez-Ramirez, D. K. Borah, and D. G. Voelz, “Optimal 3-d landmark placement for vehicle localization using heterogeneous sensors,” *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 2987–2999, September 2013.

- [39] K. Doğançay and H. Hmam, “Optimal angular sensor separation for aoa localization,” *Signal Processing*, vol. 88, no. 5, pp. 1248–1260, 2008.
- [40] D. Cassioli, M. Z. Win, and A. F. Molisch, “The ultra-wide bandwidth indoor channel: from statistical model to simulations,” *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 6, pp. 1247–1257, 2002.
- [41] S. A. S. M. F. Bullo, “On optimal sensor placement and motion coordination for target tracking.”
- [42] J. Kennedy and R. Eberhart, “Particle swarm optimization,” *Neural Networks Proceedings, IEEE*, pp. 357–362, August 2011.
- [43] Vidya Viswanathan, Soumya Jana, Shati Swarup, “Fast Range-based Localization of Targets using Particle Swarm Optimization,” *International Conference on Control, Automation and Robotics, IEEE*, May 2015
- [44] Vupparaboina, Kiran Kumar, Karthika Raghavan, and Soumya Jana, “Smart camera networks: An analytical framework for auto calibration without ambiguity.” *2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, IEEE, 2013.
- [45] Morel, Jean-Michel, and Guoshen Yu, “ASIFT: A new framework for fully affine invariant image comparison.” *SIAM Journal on Imaging Sciences*, pp. 438-469, 2009
- [46] Lucas, Bruce D., and Takeo Kanade, “An iterative image registration technique with an application to stereo vision.” —*emphIJCAI*, Vol. 81. 1981.
- [47] Tomasi, Carlo, and Takeo Kanade, “Detection and tracking of point features”, *Pittsburgh: School of Computer Science, Carnegie Mellon University*, 1991.