Active Sensing for Collaborative Localization in Swarm Robotics

Shengsong Yang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science in Electrical and Computer Engineering

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

ABSTRACT

Localization is one of the most important capabilities of mobile robots. Thanks to the fast development of embedded computing hardware in recent years, many localization solutions, such as simultaneous localization and mapping (SLAM), have been vastly investigated. However, popular localization solutions rely heavily on the working environment and are not applicable to scenarios such as search and rescue in the wild, where the working environment is not accessible before the localization operation or where the environment lacks information on features and textures. The thesis thus proposes a design for an innovative localization sensor and a collaborative pose estimation scheme using the localization sensor in order to alleviate the reliance on information from the environment, while providing reliable and accurate pose estimates for mobile robots.

The proposed collaborative pose estimation scheme is comprised of individual and collaborative landmark position estimation, localization sensor inter-calibration, and collaborative sensor pose estimation, among which the inter-calibration process ensures that the sensor provides capability to also estimate orientations. In the collaborative scheme, multiple instances of the proposed sensor collaborate to estimate their respective poses by measuring the relative distance and angle among them, where the measurement errors are characterized as Gaussian white noise. Two instances of the proposed localization sensor are implemented, and the collaborative scheme is tested with the instances in the thesis. Both sensor instances reliably and accurately estimate the position of a stationary landmark, and it is demonstrated that the collaboratively estimated position estimate is more accurate than its individual counterpart. Additionally, the two instances also demonstrate their ability to track and estimate the position of a moving landmark. Lastly, the inter-calibration is experimentally validated with the instances with satisfactory performance. The experimental results presented in this work confirm the feasibility and usability of the proposed collaborative pose estimation scheme in a wide range of robotic applications.

ACKNOWLEDGMENTS

I would like to thank my academic supervisor, Professor Pierre Payeur, for accepting me as a student. I am genuinely grateful to him for his continuous support and guidance throughout the last three years.

I am also thankful to all my friends and colleagues at the University of Ottawa. And I would like to express my special thanks to Omar Al-Buraiki, Wenbo Wu, and Eric Maia for their valuable help and insightful suggestions.

Finally, I would like to express my sincerest gratitude to my family for their unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 - INTRODUCTION

Mobile robots have become more and more prevalent over the years, boosted by the fast development of embedded system solutions. The size of the robots has become smaller, while their computational capabilities remain powerful. Among the many capabilities of mobile robots, localization is the most essential and intriguing of all. From a sweeping robot to self-driving vehicles, localization is an indispensable component that can be found in nearly every application.

The localization problem is to estimate the position and orientation of mobile robots by taking measurements. The measurements, however, are most likely corrupted by noise, sometimes even errors. For example, the robots using odometry sensors are subject to non-systematic errors resulting from the environment, such as skidding. Therefore, methodologies that reduce the effects of such undesired behavior in the measurements must be researched and developed. Merging data from other sources, for instance, is a widely applied solution that can be found in the literature. The uncertainty on the measurements is evaluated and propagated to corresponding individual localization estimates, which are often modelled as probability distributions. The reliability of each individual result is then reflected by its estimated uncertainty, and a weighted average with merged uncertainties is usually taken as the combined result.

Localization using a group of mobile robots has been widely researched in recent years. Each robot in the group is localized in a collaborative fashion, where the relative location information among multiple robots is utilized in the estimation process. In this way, the collaborative scheme is more robust to individual estimation failure and potentially more accurate. But challenges remain in reducing the reliance on the working environment and reducing the effects of noise in the measurement efficiently.

1.1 Motivation

The Global Positioning System (GPS) has been widely applied to cellphones and vehicles for outdoor localization. The estimation errors can be at least a few meters using only one receiver, but the errors could be reduced to about 0.01 meters using differential GPS with another receiver at a known position, according to [1]. The accuracy degrades to about ten meters or more when the receiver is indoors, subject to effects such as multipath and occlusion [2].

Some commercial solutions with high accuracy for localization could be found on the market. A commonly used motion tracking system, the Vicon-460 system [3], has a general accuracy of about 63±5 μm. An alternative consumer-oriented solution, the HTC VIVE [4], is reported to have an accuracy within millimeters. Besides the fact that markers or tracking devices are required to be attached to the tracking target, both systems require extra installation of devices at specific locations in the environment. Though such motion tracking systems provide high-performance localization, they are only operational within a limited working area, and the systems are not suitable when the working environment is inaccessible before the localization operation.

While there are several solutions found in the literature, the dominant ones rely heavily on the surrounding environment. Additionally, popular solutions like simultaneous localization and mapping (SLAM) can be too computationally intensive for devices with limited resources, such as a mobile robot. A demand for innovative localization solutions with little requirements on the environment thus still exists, especially for mobile robots. Bearing in mind such a demand, the thesis proposes an innovative active localization sensory approach for mobile robotic applications.

1.2 Objectives

The thesis first aims to design and implement a localization sensor, which takes relative measurements to a type of specific landmark, each associated with an individual ground mobile robotic agent. The proposed sensor should be sufficiently compact to permit usage with mobile robotic agents and should not suffer from limitations regarding the environment it is operating in. When using such a sensor, it should be easily mounted on a mobile robot without interfering with its original routines. In addition, the localization sensor should have an appropriate tracking and servo mechanism. As robots may be moving consistently during each run, the sensor must have the capability of actively searching for and tracking the landmark it attempts to take distance measurements to. Most important of all, the sensor should have reliable performance when working individually.

The second objective of the thesis is to provide a mathematical localization model of each implemented sensor, which can be readily merged with information from other sources. Because of the existence of noise and estimation errors, the sensor must first be carefully configured and calibrated to counter these effects. Besides, due to the nature of the relative measurements acquired between a sensor and a given landmark, the representation of the uncertainties in the localization results has a complex form, which is not suitable for merging data. A proper approximation is formulated to represent the uncertainties in a coherent form.

The thesis then aims to develop a collaborative pose estimation scheme, where the individual measurements from multiple instances of the sensor are used in the pose estimation process. By merging the individual pose estimates based on the developed localization model, more reliable results can be obtained through the proposed collaborative pose estimation process.

1.3 Organization

The thesis starts with a literature review on both non-collaborative localization and collaborative localization solutions grouped by their methods in Chapter 2. Details of the design and implementation of the proposed localization framework are documented in Chapter 3. In Chapter 4, the estimation methods used by the sensor are presented where approximated Gaussian models of the estimation results are developed. The implemented sensors are tested and experimentally validated with multiple use cases in Chapter 5. The thesis finishes with a conclusion provided in Chapter 6.

CHAPTER 2 - LITERATURE REVIEW

Localization methods can be sorted, depending on the sources of information utilized in the process, into two categories: non-collaborative localization and collaborative localization. In non-collaborative localization solutions, a target to be localized often requires self-contained measurements or useful information from the environment. Previously deployed sensors or known map features are often used as reference points in these localization solutions. This type of localization solution is well suited for scenarios where the targets being located do not need to be aware of the existence of other targets. Collaborative localization solutions, on the contrary, usually determine the positions of multiple targets at the same time, with the information on relative location among unlocated targets contributing to the estimation process. By nature, this method would partially, or even entirely reveal to the targets their relationships to others.

In this chapter, non-collaborative localization solutions are reviewed in Section 2.1 and collaborative localization solutions are surveyed in Section 2.2.

2.1 Non-collaborative localization

Non-collaborative localization solutions are usually comprised of two stages, sensor data acquisition and estimation with respect to the information extracted from the acquired data. In the first stage, different signal characteristics, such as received signal strength indicator (RSSI), are measured to interpret valuable information for localization. According to the types of information extracted, various techniques would be engaged in the second stage to localize the target. In this section, non-collaborative localization solutions are discussed in detail grouped by the modalities of the utilized information and thus solutions within each group depend on similar techniques on estimation.

2.1.1 Geometric Localization Solutions

Geometric information, such as distance and bearing, can be extracted from many signal characteristics, and such information is beneficial towards localization of a specific target. Solutions involving geometric information are grouped and discussed in this subsection.

2.1.1.1 Trilateration

Trilateration makes use of direct distance information between a target and a group of reference points. By measuring the distance between the target and one reference point, the position of the target can be located on a circle, of which the centroid is located at the reference and the radius is the distance obtained from measurement. Therefore, with enough measurements between different reference points and the target, the target could then be easily located at the intersection of the circles. In a general case of 2-D localization, let $\boldsymbol{x} = [x, y]^T$ be the position vector of the target and $\boldsymbol{x}_i = [x_i, y_i]^T$ be the position vector of the i^{th} reference point, where $i = 1,2, \dots, N$. The Euclidean distance ρ_i between the target and the i^{th} reference can be expressed by the following equation:

$$\|\boldsymbol{x} - \boldsymbol{x}_i\| = \rho_i, i = 1,2,3, \dots, N, \tag{2.1}$$

where $\| \cdot \|$ is the l^2-norm, and $\|\boldsymbol{x} - \boldsymbol{x}_i\| = \sqrt{(x - x_i)^2 + (y - y_i)^2}$. It is straightforward that at least three non-collinear reference points are required to achieve a unique intersection point. Similarly, in three-dimensional cases, where the position vector is defined as $\boldsymbol{x} = [x, y, z]^T$ with additional z coordinates, the circles become spheres, and at least four non-coplanar reference points are required.

A geometric solution is presented in [5]. The geometric solution uses lines of positions (LOPs) in two-dimension instead of solving the non-linear equations in (2.1). When squared on both sides of the equations, the quadratic terms of x and y can be eliminated by subtracting the i^{th} squared equation from the $i + 1^{th}$ equation, and the following equations could be obtained:

$$(x_{i+1} - x_i)x + (y_{i+1} - y_i)y = \frac{1}{2}\left(\rho_i^2 - \rho_{i+1}^2 - \|\boldsymbol{x}_i\|^2 + \|\boldsymbol{x}_{i+1}\|^2\right), \tag{2.2}$$

when written in matrix form:

$$\begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_2 & y_3 - y_2 \\ \cdots & \cdots \\ x_N - x_{N-1} & y_N - y_{N-1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \rho_1^2 - \rho_2^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2 \\ \rho_2^2 - \rho_3^2 - \|\boldsymbol{x}_2\|^2 + \|\boldsymbol{x}_3\|^2 \\ \cdots \\ \rho_{N-1}^2 - \rho_N^2 - \|\boldsymbol{x}_{N-1}\|^2 + \|\boldsymbol{x}_N\|^2 \end{bmatrix}, \tag{2.3}$$

$$\text{or } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

When using three non-collinear reference points, $\boldsymbol{A}$ becomes a 2×2 non-singular matrix, and the position estimate could be given by $\hat{\boldsymbol{x}} = \boldsymbol{A}^{-1}\boldsymbol{b}$. When there are more than three reference points available, each combination of three reference points could generate an estimate and the mean of these estimates is taken as the final estimate. Noticeably, the LOPs are actually the radical lines among the circles, thus they can be generated even though the circles cannot intersect with each other due to measurement noise, and the LOPs of any three would intersect at one unique point. The intuitive solution, however, becomes infeasible in real-world scenarios where multiple reference points are available, as it becomes more complicated to solve the geometric equations. Thus, the authors have provided a least-squares (LS) solution to Equation (2.3) with N available reference points: $\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$.

A similar 3-D geometric solution for localizing aircraft or submarines is provided in [6]. Though with an extra unknown variable on the z-axis, it only requires three reference points for calculation. The ambiguity is removed by selecting from the operating altitude of the target, either above or below the surface formed by the reference points. Similar to Equation (2.1) and (2.2), the 3-D solution uses the following equations:

$$\rho_i^2 = x_i^2 + y_i^2 + z_i^2 - 2x_i x - 2y_i y - 2z_i z + x^2 + y^2 + z^2, \quad i = 1, 2 \text{ or } 3, \tag{2.4}$$

and

$$(x_i - x_1)x + (y_i - y_1)y = \frac{1}{2}\left(\rho_1^2 - \rho_i^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_i\|^2\right) - (z_i - z_1)z, \quad i = 2 \text{ or } 3. \tag{2.5}$$

The above equation could then be written in matrix form similar to Equation (2.3) as:

$$\begin{aligned} \boldsymbol{W}\boldsymbol{r}_h &= (\boldsymbol{b} - \boldsymbol{d}z), \\ \text{or } \boldsymbol{r}_h &= \boldsymbol{W}^{-1}(\boldsymbol{b} - \boldsymbol{d}z), \end{aligned} \tag{2.6}$$

where $\boldsymbol{r}_h = [x, y]^T$, $\boldsymbol{W} = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix}$, $\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} \rho_1^2 - \rho_2^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2 \\ \rho_1^2 - \rho_3^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_3\|^2 \end{bmatrix}$, and $\boldsymbol{d} = [z_2 - z_1, z_3 - z_1]^T$. Substituting Equation (2.6) in (2.4) when $i = 1$, the operating altitude, z, could be solved using the quadratic formula. By taking into consideration the operating altitude, only one possible solution of z is maintained, and the 3-D position vector, $\boldsymbol{x}$, could then be derived from Equation (2.6) as $\boldsymbol{x} = [\boldsymbol{r}_h^T, z]^T$. Though using less reference points to acquire the desired result, the noise in the 1st reference point has a major effect on the accuracy and larger noise may lead to inaccurate altitude. In addition, it is incapable of dealing with measurements from more reference points efficiently. A similar solution using Cayley-Menger determinants is proposed in [7]. The effect of noise could be minimized, but the solution remains heavy for multiple-reference cases since repetitive processes on different combinations of reference points are required to get a more precise result.

Foy [8] has solved the localization problem using an iterative LS algorithm by applying the Taylor-Series linearization with multiple reference points, and each reference point may have multiple measurements. In the Taylor-Series linearization-based LS algorithm, the left side of Equation (2.1) is first linearized at a guess of the true position with its first-order derivatives. Let $f(\boldsymbol{x}, \boldsymbol{x}_i) = \|\boldsymbol{x} - \boldsymbol{x}_i\|$ and $\boldsymbol{x}_g = \left[x_g, y_g\right]^T$ be the guess, the function is linearized as:

$$f(\boldsymbol{x}, \boldsymbol{x}_i) \approx f(\boldsymbol{x}_g, \boldsymbol{x}_i) + \left[\frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_i)}{\partial x}\Big|_{\boldsymbol{x}_g} \quad \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_i)}{\partial y}\Big|_{\boldsymbol{x}_g}\right]\boldsymbol{\delta}, \quad i = 1,2,\dots,N, \tag{2.7}$$

where $\boldsymbol{\delta} = [\delta_x, \delta_y]^T$ and $\boldsymbol{\delta} = \boldsymbol{x} - \boldsymbol{x}_g$. According to Equation (2.1), the above equation could be written as:

$$\left[\frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_i)}{\partial x}\Big|_{\boldsymbol{x}_g} \quad \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_i)}{\partial y}\Big|_{\boldsymbol{x}_g}\right]\boldsymbol{\delta} = \rho_i - f(\boldsymbol{x}_g, \boldsymbol{x}_i),\ i = 1,2,\dots,N, \tag{2.8}$$

when written in matrix form:

$$\begin{bmatrix} \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_1)}{\partial x}\Big|_{\boldsymbol{x}_g} & \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_1)}{\partial y}\Big|_{\boldsymbol{x}_g} \\ \vdots & \vdots \\ \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_N)}{\partial x}\Big|_{\boldsymbol{x}_g} & \frac{\partial f(\boldsymbol{x}, \boldsymbol{x}_N)}{\partial y}\Big|_{\boldsymbol{x}_g} \end{bmatrix}\boldsymbol{\delta} = \begin{bmatrix} \rho_1 - f(\boldsymbol{x}_g, \boldsymbol{x}_1) \\ \vdots \\ \rho_N - f(\boldsymbol{x}_g, \boldsymbol{x}_N) \end{bmatrix}, \quad \text{or } \boldsymbol{A}_{TS}\boldsymbol{\delta} = \boldsymbol{b}_{TS}. \tag{2.9}$$

The estimated result is given by $\widehat{\boldsymbol{\delta}} = (\boldsymbol{A}_{TS}^T\boldsymbol{A}_{TS})^{-1}\boldsymbol{A}_{TS}^T\boldsymbol{b}_{TS}$. Based on the result, the new guesses $\boldsymbol{x}_g' = \boldsymbol{x}_g + \widehat{\boldsymbol{\delta}}$ are formed and then fed into the same estimation process iteratively until the estimated results are small enough. The method requires an initial guess close to the true position and estimates the errors iteratively, which could be computationally complex in practice. Nevertheless, the convergence of the estimation is not guaranteed.

Probabilistic methods are also proposed to solve the problem. A confidence-based iterative localization (CIL) is presented in [9]. A metric called quality of trilateration (QoT) is introduced to achieve CIL and it is defined as a circular function centered at the true location of the estimated target:

$$Q(t) = P(\boldsymbol{x} \in D(\boldsymbol{x}_0, R)), \tag{2.10}$$

where $D(\boldsymbol{x}_0, R)$ is a circular area centered at the true location $\boldsymbol{x}_0$ with a radius of R. Apparently, the QoT function is part of the probability density function (PDF) of a point being the estimation result, which is given by:

$$P(\boldsymbol{x}) = \prod_{i=1}^{3} P_i(\|\boldsymbol{x} - \boldsymbol{x}_i\|), \tag{2.11}$$

where $P_i(\|\boldsymbol{x} - \boldsymbol{x}_i\|)$ is the PDF function of the distance measurement to the i^{th} reference point. While the true position is unknown, the point with maximum probability is used instead. The confidence of the trilateration is defined as:

$$C_t(\boldsymbol{x}) = Q(t) \prod_{i=1}^{3} C(\boldsymbol{x}_i), i = 1,2, \dots, N, \tag{2.12}$$

where $C_t(\boldsymbol{x})$ is the confidence of the trilateration process, t, at point $\boldsymbol{x}$, $Q(t)$ is the quality of the trilateration, and $C(\boldsymbol{x}_i)$ is the confidence of reference point $\boldsymbol{x}_i$. The confidences of the reference points are defined as 1, since their positions are previously known, and the confidence of the targets are defined as the maximum confidence from the best trilateration results. In this way, the method is able to locate multiple points albeit iterations are necessary. Although only three reference points are used for two-dimensional cases in the work, it could certainly be expanded to the case where multiple reference points are available. Nevertheless, the analytical form of the quality of trilateration is not given explicitly, though the PDF of the trilateration is given as the product of the measurement PDFs, and it is likely to be a 2-D nonlinear function, which could be computationally expensive to solve.

A range-based maximum likelihood estimation (MLE) is proposed in [10]. The method is proposed for wireless sensor networks (WSN) that communicate through ultra-wideband (UWB) radio. The range is then estimated by the received signal strength (RSS) using a propagation path loss model:

$$\begin{aligned} r_r &= r_t - a, \\ a &= a_0 + 10\alpha \log_{10} d + n_a, \end{aligned} \tag{2.13}$$

where r_r and r_t is the received and transmitted logarithmic RSS values in decibel (dB) respectively, which are calculated by $r = 10 \log_{10} R$ and R is the non-logarithmic RSS in Watt (W), a is the attenuation or the path-loss, a_0 is the reference attenuation at 1 meter, d is the distance between the receiver and

transmitter, and α is the path loss exponent, which is related to the propagation conditions. n_a is Gaussian white noise with a variance of σ_a^2, thus $n_a \sim N(0, \sigma_a^2)$. Therefore, the distance estimates could be expressed as $\hat{d} = c \cdot 10^{\alpha \log_{10} d + n_a}$, where c is a proportional factor to cancel the bias in the estimation. Using the same vector $\boldsymbol{b}$ as defined in Equation (2.3), where $\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} \rho_1^2 - \rho_2^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2 \\ \rho_1^2 - \rho_3^2 - \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_3\|^2 \end{bmatrix}$, the MLE then could be defined as the following form:

$$\begin{aligned} \hat{\boldsymbol{x}}_{MLE} &= \arg\max_{x}\{p(\boldsymbol{b}|\boldsymbol{x})\}, \\ p(\boldsymbol{b}) &= \frac{1}{\sqrt{(2\pi)^N \|\boldsymbol{\Sigma}\|}} \exp\left(-\frac{(\boldsymbol{b}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{b}-\boldsymbol{\mu})}{2}\right), \end{aligned} \tag{2.14}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the expectation and covariance matrix of $\boldsymbol{b}$. As the PDF is a multidimensional Gaussian distribution, the MLE condition could be obtained at the only one extreme value of $\boldsymbol{x}$ by solving the following equation:

$$\nabla_{\boldsymbol{x}} \log\big(p(\boldsymbol{b}|\boldsymbol{x})\big)\big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{MLE}} = 0. \tag{2.15}$$

The method is simulated to have a lower root mean squared error (RMSE) than the LS method and have similar performance compared to fingerprinting techniques, which will be described in Section 2.1.2.

2.1.1.2 Multilateration

Multilateration, or hyperbolic localization, is another method that makes use of distance information to localize targets. However, instead of using direct distances measured from the reference points, this method takes advantage of the distance differences between a target and the reference points. Let ρ_{ij} be the distance difference between the target to each reference point, the multilateration problem could be expressed as:

$$\|\boldsymbol{x}-\boldsymbol{x}_i\|-\left\|\boldsymbol{x}-\boldsymbol{x}_j\right\|=\rho_{ij},\ i,j=1,2,\dots N,\text{and } i\neq j, \tag{2.16}$$

where $\boldsymbol{x}$ is the position vector of the target, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are the position vectors of the i^{th} and j^{th} reference points. In 2-D cases, where $\boldsymbol{x}=[x,y]^T$, $\boldsymbol{x}_i=[x_i,y_i]^T$, and $\boldsymbol{x}_j=\left[x_j,y_j\right]^T$, the possible locations of the target comprise one branch of a hyperbola, while in 3-D cases, a sheet of hyperboloid is formed instead, where $\boldsymbol{x}=[x,y,z]^T$, $\boldsymbol{x}_i=[x_i,y_i,z_i]^T$ and $\boldsymbol{x}_j=\left[x_j,y_j,z_j\right]^T$. With enough reference points, the one-branch hyperbolae or one-sheet hyperboloids would intersect at the position of the target.

Time difference of arrival (TDOA) is a common means of measurement when using multilateration, and the distance difference d_{ij} could be inferred from the time difference by multiplying the signal propagation speed. TDOA methods usually require the signal to be transmitted from the unlocated target to the reference points to measure the time difference. Therefore, the synchronization between the target and reference points can be relaxed, while it remains necessary to synchronize the reference points that are receiving signals.

The Taylor-Series linearization-based LS method [8] is also applicable to the multilateration localization. Given Equation (2.16), the corresponding Taylor-Series expansion with first-order derivatives at $\boldsymbol{x}_g=\left[x_g,y_g\right]^T$ is:

$$\left[\frac{x_g-x_i}{\rho_{gi}}-\frac{x_g-x_j}{\rho_{gj}}\quad\frac{y_g-y_i}{\rho_{gi}}-\frac{y_g-y_j}{\rho_{gj}}\right]\boldsymbol{\delta}\approx\rho_{ij}-\left(\rho_{gi}-\rho_{gj}\right)-e_{ij}, \tag{2.17}$$

where $\rho_{gi}=\|\boldsymbol{x}_g-\boldsymbol{x}_i\|$, $\rho_{gj}=\left\|\boldsymbol{x}_g-\boldsymbol{x}_j\right\|$, $\boldsymbol{\delta}=\left[\delta_x,\delta_y\right]^T$, and $\boldsymbol{\delta}=\boldsymbol{x}-\boldsymbol{x}_g$. The estimation then follows a similar iterative process as that mentioned in trilateration, hence the same issues. An MLE method utilizing Taylor-series linearization is provided in [11], which is equivalent to the linearized LS method. First, two equations of ρ_{12} and ρ_{23} from Equation (2.16) are linearized the same way using Taylor-series

expansion, which yields two equations in the form of Equation (2.17). Rewritten in matrix form, the noise, in vector form, could be expressed as:

$$\boldsymbol{e} = \boldsymbol{\rho} - \boldsymbol{\rho}_g - \boldsymbol{G\delta}, \tag{2.18}$$

where $\boldsymbol{e} = [e_{12}, e_{23}]^T$, $\boldsymbol{\rho} = [\rho_{12}, \rho_{23}]^T$, $\boldsymbol{\rho}_g = \left[\rho_{g1} - \rho_{g2}, \rho_{g2} - \rho_{g3}\right]^T$, $\boldsymbol{\delta} = \left[\delta_x, \delta_y\right]^T = \boldsymbol{x} - \boldsymbol{x}_g$, and $\boldsymbol{G} = \begin{bmatrix} \frac{(x_g - x_1)}{\rho_{g1}} - \frac{(x_g - x_2)}{\rho_{g2}} & \frac{(y_g - y_1)}{\rho_{g1}} - \frac{(y_g - y_2)}{\rho_{g2}} \\ \frac{(x_g - x_2)}{\rho_{g2}} - \frac{(x_g - x_3)}{\rho_{g3}} & \frac{(y_g - y_2)}{\rho_{g2}} - \frac{(y_g - y_3)}{\rho_{g3}} \end{bmatrix}$.

By assuming the noise to be independent Gaussian white noise, the likelihood function of the position, $\boldsymbol{x}$, given the measurement vector, $\boldsymbol{\rho}$, is then presented as $p(\boldsymbol{\rho}|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\|\boldsymbol{\Sigma}\|}} \exp\left(-\frac{1}{2}\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e}\right)$, where $\boldsymbol{\Sigma}$ is the covariance matrix of $\boldsymbol{e}$. By maximizing on the likelihood function, the estimator minimizes on a cost function, $J(\boldsymbol{x}) = \boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e}$. The estimated value of $\boldsymbol{x}$, $\hat{\boldsymbol{x}}$, is then given by setting the gradient of $J(\boldsymbol{x})$ to zero, $\nabla_{\boldsymbol{x}} J(\boldsymbol{x}) = 0$. As $\boldsymbol{\Sigma}$ is a positive-define symmetric matrix, $\boldsymbol{\Sigma}^{-T} = \boldsymbol{\Sigma}^{-1}$ and the gradient can be expressed as $\nabla_{\boldsymbol{x}} J(\boldsymbol{x}) = -2\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e} = -2\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\rho} - \boldsymbol{\rho}_g + \boldsymbol{G}\boldsymbol{x}_g\right) + 2\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{G}\boldsymbol{x}$. The estimated value is given by $\hat{\boldsymbol{x}} = \boldsymbol{x}_g + (\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\rho} - \boldsymbol{\rho}_g\right)$. The point of interest used in the linearization, $\boldsymbol{x}_g$, should be close enough to $\boldsymbol{x}$, otherwise heavy iterations would be inevitable, and convergence is not guaranteed.

Besides LS estimation and MLE solutions, geometric solutions have also been proposed in the literature. Fang [12] presents a simple 3-D solution that employs three reference points to estimate the position of the target. It is assumed that all the three reference points are located on the XOY plane, therefore $z_i = 0$, $i = 1,2, and\ 3$. This solution uses a method similar to lines of positions (LOPs) in trilateration; a plane of possible locations at the intersection of two hyperboloids is obtained first. By re-organizing the equations from Equation (2.16) and squaring on both sides:

$$\begin{aligned} \|\boldsymbol{x}-\boldsymbol{x}_1\|^2 &= (\rho_{12}+\|\boldsymbol{x}-\boldsymbol{x}_2\|)^2, \\ \|\boldsymbol{x}-\boldsymbol{x}_3\|^2 &= (\rho_{32}+\|\boldsymbol{x}-\boldsymbol{x}_2\|)^2. \end{aligned} \tag{2.19}$$

Then by canceling the remaining first-order term of $\|\boldsymbol{x}-\boldsymbol{x}_2\|$, a plane perpendicular to the XOY plane is derived:

$$Ax + By + C = 0, \tag{2.20}$$

where $A = 2\rho_{12}(x_3 - x_2) - 2\rho_{32}(x_1 - x_2)$, $B = 2\rho_{12}(y_3 - y_2) - 2\rho_{32}(y_1 - y_2)$, and $C = \rho_{12}\rho_{32}(\rho_{32} - \rho_{12}) - \rho_{12}(x_3^2 + y_3^2) + \rho_{32}(x_1^2 + y_1^2) + (\rho_{12} - \rho_{32})(x_2^2 + y_2^2)$.

Instead of calculating another plane from different combinations similar to Equation (2.19), the solution combines the plane from Equation (2.19) with one of the hyperboloids from Equation (2.16) to get the projection of the intersection on the XOZ plane, which is either a hyperbola or an ellipse. In order to localize the target, additional information, such as operating altitude and measurements from other reference points, is required. This need for additional information, along with the requirement on reference points installation, has made the solution less applicable to three-dimensional scenarios. Besides, the solution is cumbersome to solve when more measurements are available.

2.1.1.3 Triangulation

Unlike trilateration and multilateration solutions, triangulation solutions leverage information of angles rather than distances. Triangulation has been widely used throughout history, as it is intuitive to measure the angle of a target from reference points by using the line of sight. In 2-D cases, by measuring an unknown target from two known reference points, the two lines of sight form a triangle along with the line connecting the two reference points, and the position of the unknown target could then be obtained. Let $\boldsymbol{x}_i = [x_i, y_i]^T$ be the i^{th} reference point, and $\boldsymbol{x} = [x, y]^T$ be the unknown position of the target. The measured bearing angle, θ_i, at the reference point then could be expressed as:

$$\theta_i = \arctan\frac{y-y_i}{x-x_i}. \tag{2.21}$$

The Taylor-series linearization LS estimation [8], being a flexible approach, is also applicable to this method. By applying the Taylor-Series expansion with first-order derivatives at the vicinity of $\boldsymbol{x}_g = \left[x_g, y_g\right]^T$, Equation (2.21) could be linearized as:

$$\theta_i - \arctan\frac{y_g-y_i}{x_g-x_i} = \frac{-(y_g-y_i)\delta_x+(x_g-x_i)\delta_y}{(x_g-x_i)^2+(y_g-y_i)^2}. \tag{2.22}$$

With measurements from more than two reference points, the above equation can be rewritten in matrix form:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\delta} \tag{2.23}$$

where $\boldsymbol{b} = \left[\theta_1 - \arctan\frac{y_g-y_1}{x_g-x_1}, \dots, \theta_N - \arctan\frac{y_g-y_N}{x_g-x_N}\right]^T$, $\boldsymbol{A} = \frac{1}{(x_v-x_i)^2+(y_v-y_i)^2}\begin{bmatrix} -(y_g-y_1) & x_g-x_1 \\ \dots & \dots \\ -(y_g-y_N) & x_g-x_N \end{bmatrix}$, and $\boldsymbol{\delta} = \left[\delta_x, \delta_y\right]^T$. The estimate is given by $\widehat{\boldsymbol{\delta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$. And the new $\boldsymbol{x}_g' = \boldsymbol{x}_g + \widehat{\boldsymbol{\delta}}$ is then plugged into the same process iteratively until the estimate becomes small enough.

An MLE solution to triangulation is given using small-angle assumptions in [13]. The angular errors are represented by the distance from the true position of the target to the line of sight. Let ϵ_i be the error in the i^{th} angular measurement, θ_i, and the distance be $p_i = d_i\epsilon_i$, where d_i is the distance from the target to the sensor. By assuming that the angular error is a zero-mean Gaussian error and its variance is $\sigma_{\epsilon_i}^2$, the PDF of the distance could be expressed as:

$$p(p_i) = \frac{1}{\sqrt{2\pi\sigma_{p_i}^2}} \exp\left(-\frac{p_i^2}{2\sigma_{p_i}^2}\right), \tag{2.24}$$

where $\sigma_{p_i} = d_i\sigma_{\epsilon_i}$. Assume the estimate is $\hat{\boldsymbol{x}} = [\hat{x}, \hat{y}]$, the distance between the estimated position of the target to the line of sight is then expressed by $q_i = p_i + \hat{x}\sin\theta_i - \hat{y}\cos\theta_i$, where the line of sight could be expressed as $x\sin\theta_i + y\cos\theta_i = 0$. The PDF of the distance q_i is then described as:

$$p(q_i) = \frac{1}{\sqrt{2\pi\sigma_{p_i}^2}} \exp\left(-\frac{(p_i+\hat{x}\sin\theta_i-\hat{y}\cos\theta_i)^2}{2\sigma_{p_i}^2}\right). \tag{2.25}$$

Suppose the measurements from the N reference points are independent of each other, the PDF of all such distances is given as:

$$p(q_1, q_2, \dots, q_N) = \frac{1}{\sqrt{(2\pi)^N \prod_{i=1}^N \sigma_{p_i}^2}} \exp\left(-\sum_{i=1}^N \frac{(p_i+\hat{x}\sin\theta_i-\hat{y}\cos\theta_i)^2}{2\sigma_{p_i}^2}\right), \tag{2.26}$$

when written in matrix form:

$$p(q_1, q_2, \dots, q_N) = \frac{1}{\sqrt{(2\pi)^N \prod_{i=1}^N \left(d_i^2\sigma_{\phi_i}^2\right)}} \exp\left(-\frac{1}{2}(\boldsymbol{Ax}-\boldsymbol{b})^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{-1}(\boldsymbol{Ax}-\boldsymbol{b})\right), \tag{2.27}$$

where $\boldsymbol{A} = \begin{bmatrix} \sin\theta_1 & -\cos\theta_1 \\ \dots & \dots \\ \sin\theta_N & -\cos\theta_N \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -p_1 \\ \dots \\ -p_N \end{bmatrix}$, $\boldsymbol{R} = diag(d_1^2, \dots, d_N^2)$, and $\boldsymbol{Q} = diag(\sigma_{\epsilon_1}^2, \dots, \sigma_{\epsilon_N}^2)$. The estimated result maximizes on the probability and minimizes on the function $f(\boldsymbol{x}) = (\boldsymbol{Ax}-\boldsymbol{b})^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{-1}(\boldsymbol{Ax}-\boldsymbol{b})$, and it is given as:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} f(\boldsymbol{x}) = (\boldsymbol{A}^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{-1}\boldsymbol{b}. \tag{2.28}$$

Gavish and Weiss [14] provided another MLE method and analyzed its performance against the work provided in [13]. The method uses a series of observation functions of the angular measurements from the position of the target:

$$\boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\epsilon_\theta}, \tag{2.29}$$

where $\boldsymbol{x} = [x, y]^T$ is the true position of the target, $\boldsymbol{\theta}$ is a vector of N angular measurements, $\boldsymbol{\epsilon_\theta}$ is the corresponding noise vector, and $\boldsymbol{g}(\boldsymbol{x}) = [g_1(\boldsymbol{x}), \dots, g_N(\boldsymbol{x})]^T$. Apparently, $g_i(\boldsymbol{x}) = \arctan\frac{y-y_i}{x-x_i}$, where $[x_i, y_i]^T$ is the position of the i^{th} sensor that takes the angular measurement. Assuming that $\boldsymbol{\epsilon_\theta}$ consists of independent zero-mean Gaussian noise and its covariance matrix is $\boldsymbol{\Sigma}$, the estimator minimizes on a cost function which has the following form:

$$J(\boldsymbol{x}, \boldsymbol{\theta}) = \frac{1}{2}(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\theta}). \tag{2.30}$$

The result could then be solved by the gradient-based iteration method:

$$\hat{\boldsymbol{x}}' = \hat{\boldsymbol{x}} + (\boldsymbol{G}_{\boldsymbol{x}}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{G}_{\boldsymbol{x}})^{-1} \boldsymbol{G}_{\boldsymbol{x}}^T \boldsymbol{\Sigma}^{-1}\big(\boldsymbol{\theta} - \boldsymbol{G}(\hat{\boldsymbol{x}})\big), \tag{2.31}$$

where $\boldsymbol{G}_{\boldsymbol{x}} = \left[\frac{\partial \boldsymbol{g}(\boldsymbol{x})}{\partial x} \; \frac{\partial \boldsymbol{g}(\boldsymbol{x})}{\partial y}\right]$, $\hat{\boldsymbol{x}}'$ is the iteration result, and $\hat{\boldsymbol{x}}$ is the previous iteration result as a starting point. This MLE method [14] is stated to have lower mean squared errors (MSE) than the previous one [13].

In the work provided by Spingarn [15], a stationary emitter is located by a moving observer by using an Extended Kalman Filter (EKF) on a 2-D plane. The estimated position of the observer at time step k, $\boldsymbol{x}_k = [x_k, y_k]^T$, is given at each time step, which is similar to the previously covered cases but not all measurements are available at the time of estimation. Since the emitter is stationary, the state transition matrix is constant and the extrapolation of the state estimate at timestep k has a simple form of:

$$\begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{y}_{k|k-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1|k-1} \\ \hat{y}_{k-1|k-1} \end{bmatrix}, \quad \text{or } \widehat{\boldsymbol{X}}_{k|k-1} = \boldsymbol{\Phi}\widehat{\boldsymbol{X}}_{k-1|k-1}, \tag{2.32}$$

where $\widehat{\boldsymbol{X}}_{k|k-1}$ stands for the predicted position at timestep k given $k-1$ observations and $\widehat{\boldsymbol{X}}_{k-1|k-1}$ is the smoothed estimate at time $k-1$ given $k-1$ observations. Similar to Equation (2.29), the observation model is expressed by:

$$\begin{aligned} \theta_k &= h\left(\widehat{\boldsymbol{X}}_{k|k-1}\right) + \epsilon_k \\ &= \arctan\frac{\hat{y}_{k|k-1} - y_k}{\hat{x}_{k|k-1} - x_k} + \epsilon_k, \end{aligned} \tag{2.33}$$

where ϵ_k is assumed to be zero-mean Gaussian noise with variance R_k. The predicted position is then corrected at timestep k given k observations:

$$\widehat{\boldsymbol{X}}_{k|k} = \widehat{\boldsymbol{X}}_{k|k-1} + \boldsymbol{K}_k\left(\theta_k - h\left(\widehat{\boldsymbol{X}}_{k|k-1}\right)\right), \tag{2.34}$$

where θ_k is the measurement taken by the observer and $\boldsymbol{K}_k$ is the Kalman gain matrix at timestep k. The Kalman gain matrix is updated by the predicted state and the predicted error covariance matrix, $\boldsymbol{P}_{k|k-1}$, at timestep k:

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}^T|_{\widehat{\boldsymbol{X}}_{k|k-1}} S^{-1}, \tag{2.35}$$

where $S = \boldsymbol{H}|_{\widehat{\boldsymbol{X}}_{k|k-1}} \boldsymbol{P}_{k|k-1}\boldsymbol{H}^T|_{\widehat{\boldsymbol{X}}_{k|k-1}} + R_k$ and $\boldsymbol{H}$ is the Jacobian matrix of the observation model, and in this case, $\boldsymbol{H}$ is also a 1×2 vector. The error covariance matrix is then updated and extrapolated by:

$$\begin{aligned} \boldsymbol{P}_{k|k} &= \left(I - \boldsymbol{K}_k\, \boldsymbol{H}|_{\widehat{\boldsymbol{X}}_{k|k-1}}\right)\boldsymbol{P}_{k|k-1}, \\ \boldsymbol{P}_{k+1|k} &= \boldsymbol{\Phi}\boldsymbol{P}_{k|k}\boldsymbol{\Phi}^T + \boldsymbol{Q}, \end{aligned} \tag{2.36}$$

where $\boldsymbol{Q} = diag(q_{11}, q_{22})$ is the covariance matrix of the process noise, and by assuming the nonexistence of the process noise, q_{11} is set to zero and q_{22} is set to a small positive number to prevent $\boldsymbol{P}$ from becoming singular. The EKF is initialized by prior estimates and corresponding covariances. The author also added that the estimation result can be improved by repeatedly updating the Kalman gain at the corrected state in Equation (2.34), but it proves to be not significantly different from the EKF method. In order to start the EKF process, an initial estimate on the state of position along with its corresponding variance matrix is required. The position estimate is obtained from the first two angular measurements, θ_1 and θ_2, using a basic triangulation,

$$\begin{bmatrix} x_{init} \\ y_{init} \end{bmatrix} = \begin{bmatrix} \frac{\lambda \sin(\alpha+\theta_1)\cos\theta_2}{\sin(\theta_1-\theta_2)} + x_2 \\ \frac{\lambda \sin(\alpha+\theta_1)\sin\theta_2}{\sin(\theta_1-\theta_2)} + y_2 \end{bmatrix}, \tag{2.37}$$

where $\boldsymbol{x}_1 = [x_1, y_1]$ and $\boldsymbol{x}_2 = [x_2, y_2]$ are the corresponding positions of the observer, λ is the distance between the two, and $= \arctan\frac{y_2-y_1}{x_2-x_1}$. Let the above transformation be as $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = [\theta_1, \theta_2]^T$, the Jacobian matrix of the transformation function is $\boldsymbol{J} = \left[\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \theta_1}, \frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \theta_2}\right]$. The covariance matrix of $\boldsymbol{x}$ is then approximated as

$$\boldsymbol{\Sigma}_{\boldsymbol{x}} = \boldsymbol{J}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{J}^T, \tag{2.38}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is the covariance matrix of the two measurements. Though only one moving observer is demonstrated in the work, the solution can be easily extended to the case where multiple independent observers are available to be applied in an arbitrary sequence equivalent to the time steps. The uncertainties of the observer are not considered in this work, which could have improved the estimation accuracy. Besides, the application uses an identity matrix for the state transition model because of the stationary target, while a more complicated model is usually required in general robotic applications.

A 3-D triangulation method is presented in [16]. The method has taken into consideration the noise that exists in the locations of the reference points by finding the effects of noise in the measurements and reference points. Let the true position of the target be $\boldsymbol{x} = [x, y, z]^T$ and the position of the i^{th} reference point be $\boldsymbol{x}_i = [x_i, y_i, z_i]^T$, the measurement made at the i^{th} reference point can be expressed by using its corresponding azimuth angle θ_i and elevation angle ϕ_i, and thus

$$\begin{aligned} \theta_i &= \arctan\frac{y-y_i}{x-x_i}, \\ \phi_i &= \arctan\frac{z-z_i}{(x-x_i)\cos\theta_i + (y-y_i)\sin\theta_i}. \end{aligned} \tag{2.39}$$

Without the effects of noise, the following equations can be obtained:

$$0 = (x - x_i)\sin\theta_i - (y - y_i)\cos\theta_i,$$
$$0 = (x - x_i)\sin\phi_i\cos\theta_i + (y - y_i)\sin\phi_i\sin\theta_i - (z - z_i)\cos\phi_i. \tag{2.40}$$

When noise is added, the above equations no longer hold. Therefore, define the residuals η_{θ_i} and η_{ϕ_i} as:

$$\eta_{\theta_i} = (x - x_{mi})\sin\theta_{mi} - (y - y_{mi})\cos\theta_{mi} = \boldsymbol{g}_{\theta_{mi}}^T(\boldsymbol{x}_{mi} - \boldsymbol{x}),$$
$$\eta_{\phi_i} = \big((x - x_{mi})\cos\theta_{mi} + (y - y_{mi})\sin\theta_{mi}\big)\sin\phi_{mi} - (z - z_{mi})\cos\phi_{mi} = \boldsymbol{g}_{\phi_{mi}}^T(\boldsymbol{x}_{mi} - \boldsymbol{x}), \tag{2.41}$$

where $\boldsymbol{x}_{mi} = [x_{mi}, y_{mi}, z_{mi}]^T$ are the positions of the references corrupted by noise, $\boldsymbol{g}_{\phi_{mi}} = [\sin\phi_{mi}\cos\theta_{mi}, \sin\phi_{mi}\sin\theta_{mi}, \cos\phi_{mi}]^T$, $\boldsymbol{g}_{\theta_{mi}} = [-\sin\theta_{mi}, \cos\theta_{mi}, 0]^T$, and ϕ_{mi} and θ_{mi} are the corrupted measurements. Let the distance between the target and the i^{th} reference point be d_i, the angular errors be $\epsilon_{\theta_{mi}}$ and $\epsilon_{\phi_{mi}}$, and the position errors be $\boldsymbol{\epsilon}_{\boldsymbol{x}_{mi}} = \boldsymbol{x}_{mi} - \boldsymbol{x}_i = \left[\epsilon_{x_{mi}}, \epsilon_{y_{mi}}, \epsilon_{z_{mi}}\right]^T$, and assume that the errors are small enough, the residuals could be written as follows:

$$\eta_{\theta_i} = \epsilon_{\theta_{mi}} d_i\cos\phi_i - \epsilon_{x_{mi}}\sin\theta_i + \epsilon_{y_{mi}}\cos\theta_i = \epsilon_{\theta_{mi}} d_i\cos\phi_i + \boldsymbol{g}_{\theta_i}^T\boldsymbol{\epsilon}_{\boldsymbol{x}_{mi}},$$
$$\eta_{\phi_i} = \epsilon_{\phi_{mi}} d_i - \epsilon_{x_{mi}}\sin\phi_i\cos\theta_i - \epsilon_{y_{mi}}\sin\phi_i\sin\theta_i + \epsilon_{z_i}\cos\phi_i = \epsilon_{\phi_{mi}} d_i + \boldsymbol{g}_{\phi_i}^T\boldsymbol{\epsilon}_{\boldsymbol{x}_{mi}}. \tag{2.42}$$

By cancelling the residual terms in Equation (2.41) and Equation (2.42), the above equations could then be written in matrix form:

$$\boldsymbol{B}\boldsymbol{e} = \boldsymbol{h} - \boldsymbol{G}\boldsymbol{x}, \tag{2.43}$$

where $\boldsymbol{B} = \begin{bmatrix} diag(d_i\cos\phi_i) & \boldsymbol{0} & diag(\boldsymbol{g}_{\theta_i}) \\ \boldsymbol{0} & diag(d_i) & diag(\boldsymbol{g}_{\phi_i}) \end{bmatrix}$, $\boldsymbol{e} = \left[\epsilon_{\theta_{m1}}, \dots, \epsilon_{\theta_{mN}}, \epsilon_{\phi_{m1}}, \dots, \epsilon_{\phi_{mN}}, \boldsymbol{\epsilon}_{\boldsymbol{x}_{m1}}^T, \dots, \boldsymbol{\epsilon}_{\boldsymbol{x}_{mN}}^T\right]^T$, $\boldsymbol{G} = \left[\boldsymbol{g}_{\theta_{m1}}, \dots, \boldsymbol{g}_{\theta_{mN}}, \boldsymbol{g}_{\phi_{m1}}, \dots, \boldsymbol{g}_{\phi_{mN}}\right]^T$, and $\boldsymbol{h} = \left[\boldsymbol{g}_{\theta_{m1}}^T\boldsymbol{x}_{m1}, \dots, \boldsymbol{g}_{\theta_{mN}}^T\boldsymbol{x}_{mN}, \boldsymbol{g}_{\phi_{m1}}^T\boldsymbol{x}_{m1}, \dots, \boldsymbol{g}_{\phi_{mN}}^T\boldsymbol{x}_{mN}\right]^T$

$$\hat{\boldsymbol{x}} = (\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{h}, \tag{2.44}$$

where $\boldsymbol{W} = E[\boldsymbol{B}\boldsymbol{e}\boldsymbol{e}^T\boldsymbol{B}^T]$ is the weight matrix. Though the ground truths of θ, ϕ, and d, are involved for the calculation of the weight matrix, $\boldsymbol{W}$, the authors state that the approximation using corresponding measurements has little performance loss. The presented method is reported to achieve better performance

but at the expense of heavy computations when noise corrupts the positions of the reference points. However, the residual models are dedicated only to the angular measurements and it may turn out to be completely different for other types of measurements.

Triangulation could be integrated with other solutions to estimate the position of the target given other forms of information such as distances. In [17], a moving mobile terminal is located by multiple base stations in a cellular network by a hybrid trilateration and triangulation method in 2-D. All the base stations can estimate their distances to the mobile terminal by measuring the time of arrival (TOA) of signals transmitted by the terminal, while some base stations are equipped with additional antenna arrays and can take angular measurements towards the mobile terminal. The method consists of two estimators, an ML estimator for rough estimates and an iterative LS estimator to refine the estimates. The trilateration part of the ML estimator is similar to those mentioned previously, but it selects a different variable, $\boldsymbol{z} = [x, y, x^2 + y^2]^T$, to be estimated in re-arranged Equation (2.1):

$$\begin{gathered}\left(\rho_{mi} - \epsilon_{\rho_{mi}}\right)^2 = (x - x_i)^2 + (y - y_i)^2, \\ \text{or } \rho_{mi}\epsilon_{\rho_{mi}} - \frac{1}{2}\epsilon_{\rho_{mi}}^2 = \frac{1}{2}\left(\rho_{mi}^2 - x_i^2 - y_i^2\right) + [-x_i, -y_i, 1/2]\boldsymbol{z},\end{gathered} \tag{2.45}$$

where ρ_{mi} is the noisy distance measurement from the i^{th} base station and $i = 1, \dots, N$. For the triangulation part, the method employs a similar error expression to Equation (2.40):

$$\rho_{mj}\epsilon_{\theta_{mj}} = \left(x - x_j\right)\sin\theta_{mj} - \left(y - y_j\right)\cos\theta_{mj}, \tag{2.46}$$

where θ_{mj} is the angular measurement from a base station, associated with the number j, to the mobile terminal. Note that the base station j is one of the stations that are equipped with additional antenna arrays and thus capable of taking angular measurements. Let $S = \{j_1, \dots, j_M\}$ $(M \leq N)$ be the set of such base stations, thereby $S \subseteq \{1, \dots, N\}$ and $j \in S$. By stacking Equations (2.45) and (2.46), the error vector of all measurements is expressed as:

$$\boldsymbol{e} = \begin{bmatrix} \frac{1}{2}(\rho_{m1}^2 - x_1^2 - y_1^2) \\ \cdots \\ \frac{1}{2}(\rho_{mN}^2 - x_N^2 - y_N^2) \\ -x_{j_1}\sin\theta_{mj_1} + y_{j_1}\cos\theta_{mj_1} \\ \cdots \\ -x_{j_M}\sin\theta_{mj_M} + y_{j_M}\cos\theta_{mj_M} \end{bmatrix} - \begin{bmatrix} -x_1 & -y_1 & 1/2 \\ \cdots & \cdots & \cdots \\ -x_N & -y_N & 1/2 \\ -\sin\theta_{mj_1} & \cos\theta_{mj_1} & 0 \\ \cdots & \cdots & \cdots \\ -\sin\theta_{mj_M} & \cos\theta_{mj_M} & 0 \end{bmatrix} \boldsymbol{z}, \quad (2.47)$$

$$\text{or } \boldsymbol{e} = \boldsymbol{h} - \boldsymbol{G}\boldsymbol{z},$$

where $\boldsymbol{e} = \left[\rho_{m1}\epsilon_{\rho_{m1}} - \frac{1}{2}\epsilon_{\rho_{m1}}^2, \ldots, \rho_{mN}\epsilon_{\rho_{mN}} - \frac{1}{2}\epsilon_{\rho_{mN}}^2, \rho_{mj_1}\epsilon_{\theta_{mj_1}}, \ldots, \rho_{mj_M}\epsilon_{\theta_{mj_M}}\right]^T$. The estimate is given as $\boldsymbol{z} = (\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{h}$, where $\boldsymbol{\Sigma}$ is the covariance matrix of the error vector approximated at the measurements. Since the variables to be estimated in $\boldsymbol{z}$ are not all independent of each other, the method uses an iterative LS estimator to refine the result. The second estimator then expands Equation (2.46) using Taylor-Series with first-order derivatives at a guess $\boldsymbol{x}_g = \left[x_g, y_g\right]^T$:

$$\rho_{mj}\epsilon_{\theta_{mj}} = (x_g - x_j)\sin\theta_{mj} - (y_g - y_j)\cos\theta_{mj} - \left[-\sin\theta_{mj}, \cos\theta_{mj}\right]\boldsymbol{\delta}, \quad (2.48)$$

where $\boldsymbol{\delta} = \left[x - x_g, y - y_g\right]^T$. Again, by stacking all M equations from Equation (2.48), $\boldsymbol{\delta}$ can be obtained. And by repeating the LS estimator from a new guess, $\boldsymbol{x}_g^{'} = \boldsymbol{x}_g + \boldsymbol{\delta}$, the position estimate will be refined iteratively.

A 3-D multilateration/triangulation method is presented in [18], which follows the work provided in [16]. The vector from the reference points, $\boldsymbol{x}_i$ $(i = 1\ or\ 2)$, to the target, $\boldsymbol{x}$, could be expressed as:

$$\boldsymbol{x} - \boldsymbol{x}_i = \rho_i\boldsymbol{b}_i,\ i = 1\ or\ 2, \quad (2.49)$$

where $\boldsymbol{b}_i = [\cos\phi_i\cos\theta_i, \cos\phi_i\sin\theta_i, \sin\phi_i]^T$, θ_i and ϕ_i are the azimuth angle and elevation angle from the reference points to the target, respectively. Noticeably, $\boldsymbol{b}_i$ is a unit norm vector and $(\boldsymbol{b}_2 - \boldsymbol{b}_1)^T(\boldsymbol{b}_2 + \boldsymbol{b}_1) = 0$. Given the multilateration measurements as defined by Equation (2.16), $\rho_2 - \rho_1 = \rho_{21}$, the following equation could be derived:

$$\rho_2(\boldsymbol{b}_2 - \boldsymbol{b}_1)^T(\boldsymbol{b}_2 + \boldsymbol{b}_1) = (\boldsymbol{b}_2 - \boldsymbol{b}_1)^T(\rho_2\boldsymbol{b}_2 + \rho_1\boldsymbol{b}_1 + \rho_{21}\boldsymbol{b}_1) = 0, \tag{2.50}$$

and substituting in Equation (2.49),

$$2(\boldsymbol{b}_2 - \boldsymbol{b}_1)^T\boldsymbol{x} = (\boldsymbol{b}_2 - \boldsymbol{b}_1)^T(\boldsymbol{x}_1 + \boldsymbol{x}_2 - \rho_{21}\boldsymbol{b}_1). \tag{2.51}$$

The equations from (2.40) could be rewritten as:

$$\boldsymbol{G}_i\boldsymbol{x} = \boldsymbol{G}_i\boldsymbol{x}_i, i = 1 \; or \; 2, \tag{2.52}$$

where $\boldsymbol{G}_i = \begin{bmatrix} \sin\theta_i & -\cos\theta_i & 0 \\ \sin\phi_i\cos\theta_i & \sin\phi_i\sin\theta_i & -\cos\phi_i \end{bmatrix}$. Similar to the work provided in [16], residuals appear when the measurements are substituted into Equation (2.51) and (2.52). By approximating the residuals using Taylor-Series expansion, the position vector, $\boldsymbol{x}$, could be obtained by a weighted LS estimator. Though the solution comes with a simpler form than the previous work in [16], the uncertainties in the reference points are no longer taken into consideration.

In applications such as [14][15][17], angular measurements often require to be taken at multiple positions, where specialized receivers such as antenna arrays are being used at each position to acquire the angles. There are also applications that have employed specialized cameras for measuring angles, especially in the field of robotics. Drocourt *et al.* [19] use a stereoscopic omnidirectional vision system, which consists of two omnidirectional cameras. Each camera is comprised of a CCD camera and a conical mirror that reflects the surroundings into the CCD camera. The vision system uses vertical lines as features and matches the features against a previously given map. The system reportedly has a precision of about 10 cm in position estimation with a maximum error of 16 cm in 11 setups. The orientation estimation error of the system ranges from 0 to 3.5° with an average of about 1°. However, the size of each setup is not disclosed in the work. In the work presented by Cao *et al.* [20], a fisheye lens with a field of view of 185° is mounted on top of a camera, which points up vertically. The camera can estimate the elevation angle of a detected landmark by applying the Equal Distance Projection Theorem. Given a number of

known features in the surrounding environment, the pose of the camera could be estimated. In a testing area of 1.53×2.32 m^2 with two known landmarks, the camera obtained an average position estimation error of 0.156 m with a maximum of 0.25 m. The orientation estimation error, except for its maximum value of 21°, ranges between 0 and 9°. Both vision-based systems require the camera to be properly installed and adjusted, as the angular measurements are closely correlated to the installation errors. The angular errors of both systems are not sufficiently small for long-distance localization, as with an angular error of 2° and accurate distance measurements, at 5 m away, a position estimation error of $\frac{2^\circ}{180^\circ} \times \pi \times 5 \approx 0.175$ m is expected, and since both use only angular measurements, the error could actually be worse than the expectation.

2.1.1.4 Commercial Motion Tracking Systems

Commercial motion tracking systems are common solutions used in robotic research. Because of the outstanding performance of these systems, the tracking results are often taken as ground truth for analyzing the results of many other localization solutions. Motion tracking systems often require installation of fixed devices, and a tracking device or markers need to be attached to the tracked target as well. However, due to the proprietary nature of the commercial solutions, the details of the localization approaches are not usually disclosed. As reported in the literature [3][4], some motion tracking systems use methods similar to triangulation described in Section 2.1.1.3. Apart from geometric information, inertial measurements from IMUs are also involved in some systems [4].

The Vicon-460 [3] is a commonly used motion tracking system. It requires multiple specialized high-resolution cameras to be set up and fixed in the working environment, and the work presented in [3] analyzes setups with at least three cameras. Retroreflective markers are required to be attached on the target to be tracked, and the markers are captured by the specialized cameras when reflecting the light

emitted by the system. The motion tracking solution has been tested to have a general accuracy of 63±5 μm.

Another consumer-oriented solution is the HTC VIVE [4], which is comprised of two lighthouses and a tracking device. The lighthouses are fixed in the working environment and emit synchronized infrared light sweeping through the working area vertically and horizontally. The tracking device is pre-installed with photodiodes, which receive infrared signals and record corresponding timestamps, and an IMU is also integrated into the tracking device for estimating its orientation. According to the work presented in [4], the system reaches accuracy within millimeters.

As mentioned earlier, motion tracking systems require some components to be fixed in the working environment, and localization operation can only be performed within the workspace covered by the systems, which is rather limited for general robotic applications. Moreover, in some real life operational contexts the working environment may not be accessible before the localization operation, and motion tracking systems are not applicable to such scenarios.

2.1.2 Fingerprinting

Fingerprinting, or scene analysis, is a popular indoor localization solution for mobile devices in recent years. Enabled by the ubiquitous Wi-Fi networks, fingerprinting solutions have the potentials to locate devices in very complex indoor environments [21][22]. The measurements in the previous techniques usually are assumed to satisfy the line-of-sight (LOS) condition, where the target being located is measured directly from the reference points. However, the surrounding environment of the target is often complicated indoors due to the nature of constructions, and those measurements suffer from the non-line-of-sight (NLOS) condition, such as multipath effects and occlusions. Applications like GPS that works outdoors often perform poorly when located indoors. Fingerprinting solutions, by comparison, do

not depend on the LOS condition, as the measurements of received signal strength (RSS) are robust to the aforementioned effects.

Fingerprinting solutions usually are composed of two phases, an offline survey phase and an online query phase. During the offline survey phase, a site survey is carried out to record the RSS measurements at many selected reference points spread all over the indoor space. At each reference point, the values of RSS from all the Wi-Fi access points are recorded in the database. Then during the online query phase, the user device measures the RSS values of each Wi-Fi access point and queries the localization system, which determines the location of the user device by matching the pre-recorded data and returns a possible location result.

One problem with the fingerprinting solutions is that different receiving devices may have different receiving characteristics acquiring the signals. When the receiving device in the query phase is different than that used in the survey phase, it is likely that the localization accuracy will be affected. He *et al.* [22] provide a simultaneous localization and calibration (SLAC) method to overcome this problem by fusing data from pedestrian dead-reckoning. The discrepancy of the characteristics of the recording and receiving devices and the characteristics of the user strides are identified and adjusted in the process of estimation.

As mentioned above, the technique depends on the exhaustive offline survey process, and thus it cannot be used for cases where the environment has not been surveyed previously. Besides, once the surveyed environment has changed, such as a wall has been torn down, the RSS map in the environment is then changed and a new survey is necessary to update the database to keep the estimation results reliable. In addition, the technique only works within the coverage of Wi-Fi access points, and in order to identify the user's location, RSS values from multiple access points should be available at each location, which requires more access points than usual to be installed.

2.1.3 Dead Reckoning

The aforementioned techniques all take measurements relative to the environment, therefore, targets are located with absolute locations or relative locations with respect to a map. Dead reckoning, on the contrary, often utilizes proprioceptive measurements to locate moving targets. Since dead reckoning uses a motion model of the target for estimation, the orientation of the target is indispensable, and the estimation results are usually relative to their previous ones. This technique is extensively engaged in 2-D applications such as navigation, especially for mobile robotics, but it is often combined with other techniques because of its own limitations. The localization process, by nature, is subject to errors accumulated through each timestep, and it becomes impossible eventually to tell whether the target is at a visited position due to the unbounded growth of errors and lacking absolute position information to the environment. These errors either reside in the proprioceptive measurements of displacement and orientation or rise from the unpredictable environment. According to [23], errors in dead reckoning are generally categorized into two types based on their sources: systematic errors and non-systematic errors. For ground mobile robotics, systematic errors are mainly caused by the imprecise modeling of the robots and the accuracy limitations of the proprioceptive sensors, and non-systematic errors raise when interacting with the environment, such as skidding wheels on a slippery floor.

Let $\boldsymbol{p}_t = [x_t, y_t, \theta_t]^T$, $t = 0,1,2, \dots, N$ be the pose of the target at timestep t, where $\boldsymbol{x} = [x_t, y_t]^T$ and θ_t are the position vector and the orientation of the target at timestep t, respectively. Let $\Delta\boldsymbol{p}_t = [\Delta x_t, \Delta y_t, \Delta\theta_t]^T$ be the measured pose difference between two consecutive poses, $\boldsymbol{p}_t$ and $\boldsymbol{p}_{t-1}$, then the poses have the following relationship:

$$\boldsymbol{p}_t = \boldsymbol{p}_{t-1} + \Delta\boldsymbol{p}_t. \tag{2.53}$$

A widely used and fairly simple way of dead reckoning is to use odometry sensors on differential-drive vehicles as shown in [23][24] to analyze the pose difference. The method uses the assumption that

the traces of the robot in between every two consecutive time steps are approximately straight, which requires that the measurements should be taken over a very short period of time. Suppose the displacements of the left and right wheels between two consecutive time steps are measured and calculated as $\Delta d_{L,t}$ and $\Delta d_{R,t}$ respectively, the displacement of the robot is then given as the average of the two displacements, $\Delta d_t = \left(\Delta d_{L,t} + \Delta d_{R,t}\right)/2$, and the change in the robot's orientation is given as $\Delta\theta_t = \left(\Delta d_{R,t} - \Delta d_{L,t}\right)/l$, where l is the distance between the two contact points where the wheels touch the ground. Equation (2.53) could then be written as:

$$\boldsymbol{p}_t = \begin{bmatrix} x_{t-1} + \Delta d_t \cos\theta_t \\ y_{t-1} + \Delta d_t \sin\theta_t \\ \theta_{t-1} + \Delta\theta_t \end{bmatrix}. \tag{2.54}$$

Even though calibration methods exist to minimize the effect of systematic errors [24], it is very difficult to deal with the non-systematic errors using only the odometry sensors. The Gyrodometry introduced in [25] makes use of a gyroscope to compensate for such errors. The method employs a simple strategy that monitors the discrepancy between the measured orientation changing rates from the gyroscope and the odometry sensor. The method switches to the gyroscope when the discrepancy is larger than a pre-set threshold, and the rate from the odometry sensor is used otherwise. The data from the gyroscope is not used constantly because of its inherent drifts, and the drifts will affect minimally the orientation estimation if the gyroscope is used in the described way. The method manages to keep the orientation errors within ±0.5° in a two-minute test, but the accumulation of the errors remains unknown for longer period of time. In addition, the robot used in the test runs at a speed of 10 cm/s, which covers about 12 m during the test, and the effects of higher speed and longer distance should be examined as well.

Inertial navigation system (INS) [26][27] is another important solution for dead reckoning. INS solutions employ inertial measurement units (IMU) such as gyroscopes and accelerometers, which measure angular velocities and linear accelerations, respectively. Some other sensors, such as tilt sensors,

may contribute to a more precise result, but may not participate in the estimation directly. The measurements are then fed to an EKF to estimate the state of the system. The data from odometry sensors is also fused with INS in [27] to overcome the non-systematic errors due to uneven surface for lunar rovers. Though INS solutions show great performance in these methods, accurate sensors are required for such methods to minimize the effects of drifts, and such sensors can be costly. In addition, the growth of errors due to drifts is not eliminated and usually requires other sources of sensory data to correct the errors.

2.1.4 Exteroceptive Odometry and SLAM

Instead of using classic odometry sensors, there are also solutions that estimate the motions of a moving robot by making use of information from the environment. In this way, the effects of non-systematic errors in dead reckoning between two consecutive pose estimates can be reduced. Using 2D laser scanners [28] or cameras [29][30][31] are two well-known solutions for exteroceptive odometry, which are often referred to as laser odometry or visual odometry, respectively.

Lingemann *et al.* [28] demonstrated a laser odometry solution for fast moving robots. The solution achieves the goal of estimating the transformation by matching the features in two consecutive laser scans. The features are first generated by filtering scan data points in the polar coordinate system, which reduces potential computational complexity from matching 2-D features. In the filtering process, the scan data points are sharpened on the radial distances, and extrema found at the zero crossing points of the corresponding smoothed gradient are used as required features. The obtained features are then classified into three types, which are maximums, minimums and inflection points. Only those assigned within the same type are considered as potential matches and the matching is completed by evaluating a predefined distance of the potential matches. Assume $\boldsymbol{m}_i$ and $\boldsymbol{n}_j$ are two feature points from two consecutive scans

in a Cartesian coordinate system, and their coordinates in a polar system are denoted by $\left[\rho_{m_i}, \theta_{m_i}\right]^T$ and $\left[\rho_{n_j}, \theta_{n_j}\right]^T$, respectively. The distance is defined as follows:

$$d_{ij} = \sqrt{\omega_1 \left(\rho_{m_i} - \rho_{n_j}\right)^2 + \omega_2 \left(\theta_{m_i} - \theta_{n_j}\right)^2} + \omega_3 \|\boldsymbol{m}_i - \boldsymbol{n}_j\|, \tag{2.55}$$

where ω_1, ω_2, and ω_3 are corresponding weights. Assuming the topology of the features in the two scans remains approximately the same during the movement of the robot, by centering the matched features of each scan on their geometric center, the rotation between the two sets could be easily achieved by maximizing an error function:

$$E(\theta) = \sum_{i=1}^{\mathrm{N}} \hat{m}_i \boldsymbol{R}(\theta) \hat{n}_i, \tag{2.56}$$

where $\hat{m}_i$ and $\hat{n}_i$ are shifted feature points, and $\boldsymbol{R}(\theta)$ is the rotation matrix. The translation between the two sets of features then can be easily derived from the geometric centers of the original feature sets.

Laser scanners have introduced more information to the laser odometry method than the classic ones to accomplish higher accuracy in relative motion. However, similar to the assumption for classic odometry that the transitional movement is straight, laser odometry assumes that the topologies in the two consecutive scans are approximately the same, which is subject to the effects of moving objects and requires the speed of the robot to be reasonably slow to maintain the consistency. In addition, the method for extracting features relies on the surrounding environments of the laser scanner, and it is possible that in a feature-less environment not enough features can be generated to match the scans.

Similar approaches have been applied to visual odometry (VO) solutions since images have the advantage of containing much more information. Scaramuzza and Fraundorfer [30][31] provided a thorough tutorial on feature-based visual odometry solutions. In general, feature-based visual odometry has two stages, feature detection and feature matching. In the first stage, various features have been

developed, including SIFT [32], SURF [33], and ORB [34]. Some of the features, such as SIFT, have excellent performance in accuracy but are computationally intensive. As reported in [35], ORB has the fastest time in feature detection and matching time, but it has also the most drifts over time. The detected features in two consecutive image frames are then matched in the second stage. An exhaustive search can apply for matching the features in consecutive image frames, but it can be inefficient for online applications. An approximated motion model can improve the searching efficiency by predicting the areas where the features in the first frame are likely to be in the second frame. The epipolar constraints between any consecutive frames are also helpful for reducing the matching computations. Once the features are matched, the relative motion between the image frames can then be estimated based on the camera model being used. Optimization methods such as pose-graph optimization [36] and windowed bundle adjustment [37] can also be applied on the estimated motion to reduce drifts and improve local consistency. Feature-based VO solutions require proper lighting condition for image processing and that the environment should have a reasonable number of features. Also due to the distortion caused by the camera, the images are usually rectified in order to accurately estimate the motion.

Appearance-based, or direct, VO solutions have also been developed over the years. Instead of extracting and matching features, direct VO solutions match on photometric information of the image. Direct sparse odometry (DSO) [29] is a recent example of direct solutions that offer great performance. DSO is based on a photometric camera model developed in [38]. A photometric residual is defined as the weighted sum of squared errors on a designated pattern of a point and its back-projected point in the next image, and a cost function is defined as the sum of all the photometric errors over all frames of all points. Since the photometric errors are dependent on the camera poses in the consecutive frames, DSO obtains these poses by optimizing on the photometric cost function. Noticeably, the method uses a sliding window of frames instead of all the frames in order to trade-off for computational efficiency. Direct VO solutions

like DSO do not have the requirement of the number of features in the environment, but they may suffer from significant environmental lighting change. Besides, they do require delicate calibrations of the camera not only for classic intrinsics but also photometrics, such as exposure time, vignette effects of the camera, and even the effects of rolling shutters.

Similar to the exteroceptive odometry techniques, simultaneous localization and mapping (SLAM) solutions [39][40][41][42] have been developed and became the most popular approaches in recent years. SLAM solutions not only provide pose estimation of the sensor but also build a globally consistent map of its surrounding environment. In order to build such a map, SLAM solutions are equipped with a distinct functionality from the odometry solutions, loop-closure, which helps recognize a previously visited location. With the help of loop-closure, SLAM solutions produce results relative to the generated map reducing the accumulated errors found in odometry solutions. However, SLAM solutions become highly computationally intensive and not acceptable for devices with limited computational power. Besides, SLAM solutions based on vision and LiDAR are also faced with the restrictions found in LiDAR- and vision-odometry solutions such as requirements on number of features and delicate camera calibration.

2.2 Collaborative Localization

As presented previously in non-collaborative localization, each target is being located independently regardless of others, while collaborative localization solutions leverage the relative information among the targets and locate multiple targets collaboratively at a time. Through this relative information, collaborative localization solutions are more robust to failures of individual localization and may require less information from the environment compared to non-collaborative localization methods. The relative information participates in different stages during the localization process, therefore collaborative methods could be divided roughly into two categories.

One category of methods uses the relative information directly to locate the targets along with the measurements taken from a known reference. Dardari and Conti [43] provide an MLE solution based on RSS measurements. The method employs a similar propagation model as that in Equation (2.13), though the distance is not used for the estimation but rather the path-loss. The PDF of the path-loss measurements given the positions of the targets are then used to be maximized, which is given as:

$$\widehat{\boldsymbol{P}}_{ML} = \arg\max_{\boldsymbol{P}} \prod_{i=1}^{\mathrm{N}} \exp\left(-\frac{(\tilde{a}_i - a_i)^2}{2\sigma^2}\right), \tag{2.57}$$

where $\boldsymbol{P}$ is the position matrix for all targets, $\tilde{a}_1$ is the measured path-loss, and a_i is the path-loss calculated given the position matrix. The MLE solution requires an exhaustive search through numerous possible subdivisions, thus the authors provide an iterative square grid search as a sub-optimal solution. In each iteration, the targets are first set in the center of the grid that they belong to, and the grid is then divided equally to smaller squares, where the MLE algorithm is run to decide which targets correspond to the new squares. The smaller squares will be divided in the next iteration until the size of the square is smaller than a pre-set threshold. A similar MLE solution is presented in [44], where the estimated distances from the propagation model are used in the estimation instead.

An optimal solution to MLE is presented in [45], where a brand-and-bound searching strategy is applied to the linearized residual function by the reformulation linearization technique (RLT). The residual function of the estimated positions is defined as:

$$
\begin{aligned}
f(\boldsymbol{P}) &= \sum_{i=1}^{n}\sum_{j=i+1}^{n+m}\frac{1}{\sigma_{ij}^2}\left(d_{ij}-\sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2}\right)^2 \\
&= \sum_{i=1}^{n}\sum_{j=i+1}^{n+m}\frac{1}{\sigma_{ij}^2}\left(d_{ij}^2-2d_{ij}Z_{ij}+W_{ij}\right) \\
&= L(\boldsymbol{Z},\boldsymbol{W}),
\end{aligned}
\tag{2.58}
$$

where $\boldsymbol{P}$ is a $2\times N$ position matrix comprised of n targets to be located and m reference points whose position is known, d_{ij} is the measured distance between the i^{th} and the j^{th} target, $\boldsymbol{Z}=\{Z_{ij}\}$, $\boldsymbol{W}=\{W_{ij}\}$, $Z_{ij}=\sqrt{W_{ij}}$, and $W_{ij}=\sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2}$. The relationship between Z_{ij} and W_{ij} is then linearized by the RLT, where the non-linear constraint is then transferred into multiple linear constraints that form a polygonal region on the plot of Z_{ij} and W_{ij}. Given the upper and lower bounds of W_{ij} as W_{ij}^U and W_{ij}^L, and a chosen point between the two, W_{ij}^B, the corresponding points on the $W_{ij}-Z_{ij}$ could be presented as $\left(W_{ij}^L,\sqrt{W_{ij}^L}\right)$, $\left(W_{ij}^B,\sqrt{W_{ij}^B}\right)$, and $\left(W_{ij}^U,\sqrt{W_{ij}^U}\right)$, and the linear constraints are selected as the tangents at these points, plus the line connecting the two points at the bounds. The polygonal region then could be expressed as:

$$
\begin{aligned}
Z_{ij} &\le \frac{1}{2\sqrt{W_{ij}^L}}\left(W_{ij}+W_{ij}^L\right) \\
Z_{ij} &\le \frac{1}{s\sqrt{W_{ij}^B}}\left(W_{ij}+W_{ij}^B\right) \\
Z_{ij} &\le \frac{1}{s\sqrt{W_{ij}^U}}\left(W_{ij}+W_{ij}^U\right) \\
Z_{ij} &\le \left(W_{ij}+\sqrt{W_{ij}^L W_{ij}^U}\right)/\left(\sqrt{W_{ij}^L}+\sqrt{W_{ij}^U}\right).
\end{aligned}
\tag{2.59}
$$

$\boldsymbol{Z}$ and $\boldsymbol{W}$ then take part in the new estimation using the linearized residual function $L(\boldsymbol{Z},\boldsymbol{W})$. Similar linearization is applied towards $W_{ij}=\sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2}$ between W_{ij} and the j^{th} target if it is to

be located. The initial bounds are set according to the relaxed constraints, for example in the constraint of $Z_{ij} = \sqrt{W_{ij}}$, the bounds of W_{ij} are set to 0 and +∞. By applying the brand-and-bound searching strategy, the polygonal region is bi-partitioned in each iteration at the variables that lead to the maximum relaxation error between the relaxed and unrelaxed values and then tightened by using the function $L(\boldsymbol{Z}, \boldsymbol{W})$ until the discrepancy between the upper and lower bounds is small enough. Although optimal solutions could be obtained by this solution, the number of iterations remains high according to the sub-optimal solutions. As shown in the work, a noise-free case with three reference points and one target may take approximately 20 iterations to get the bounds to converge.

The second category of collaborative methods does not use the relative information directly to estimate the position of the targets but to improve the lower accuracy that some targets may get. Chan *et al.* [46] provide a solution for scenarios where multiple mobile devices form stable clusters. The solution consists of three steps: neighborhood detection, confidence estimation, and collaborative error collection. The neighborhood is first detected by using ZigBee signals since the protocol covers a short propagation distance of about 2 meters and the distance could be estimated through RSS. Therefore, once the devices are connected using ZigBee, the distance between them is estimated and they are taken as members in the same group if the distance is smaller than 2 meters. The confidence of the location estimates is given by an independent location system that localizes the targets individually, where higher confidence stands for higher accuracy of the estimate, and vice versa. The confidences are at last used to collaboratively correct the errors among the members of a group in a way similar to magnetic forces, where more accurate estimates pull less accurate estimates towards them. The authors state that using such collaborative error correction achieves slight improvement on the accuracy and the cases with more devices outperform the cases with fewer devices. But due to the loose constraint of determining a neighbor and complex movements of the mobile devices, noise may be introduced into the estimation and may degrade the

performance through the collaborative correction process. A similar solution using a spring model is given in [47], where the position of each mobile device is located using fingerprinting techniques and the relative information is measured by a propagation model of Bluetooth signal. However, few improvements are made on the neighbor detection method, and the same problems may exist as well. Higuchi *et al.* [48] impose more strict constraints on neighbor detection, where two targets must stay within range for a period of time and their traces must be similar enough to each other. Pedestrian dead reckoning is used for recording the trace, and Bluetooth is used for detecting neighboring targets. The similarity of the traces along with the distribution of the location estimates and the relative distances acquired from Bluetooth RSS is used for collaborative error correction. However, this method does not use information measured from the environment, therefore the error will be accumulating though it could be improved by the collaborative correction.

2.3 Summary

In this chapter, several localization techniques are reviewed according to the types of measured information. It is noticeable that most of the literature is focused on localization solutions for individual entities, while limited work has been dedicated to the case where the entities are located collaboratively.

Measurements related to distances and angles in the localization techniques are mostly acquired by using the propagation characteristics of the transmitted signal, such as time difference of arrival (TDOA) and received signal strength (RSS), while only a few solutions extract angular measurements from omni-directional cameras. The iterative LS algorithm is often used with the distance and angular measurements for position estimation. However, the iterative LS algorithm requires a close guess to initiate the estimation and reference points whose positions are known. In addition, its iterative nature is unsuitable for online estimation.

Some signal characteristics, such as RSS, can also be used for localization when compared to a survey database of the working environment. The offline surveying process is exhaustive and yet the database easily becomes outdated once the environment is changed. Thus, such solutions cannot be used in scenarios where the access to the working environment is unavailable until the localization operation is initiated.

Odometry and inertial measurements represent the relative motion of the sensors in between time steps, and thus only relative poses can be estimated by accumulating all the measurements over a period of time. The measurement errors also build up through accumulation and eventually lead to unreliable estimation results. But if the systematic errors in the measurements are minimized, the relative measurements over a short period could be considered reliable and may contribute to the estimation results of solutions using other types of measurements.

Information acquired by visual sensors and LiDAR scanners can be applied to the exteroceptive odometry and SLAM approaches. Most of these approaches rely heavily on the features already existing in the environment, while others are based on the assumption of consistent environmental lighting along with delicate sensor calibration. Besides, these approaches generally are computationally intensive and are not acceptable to devices with limited computational power, such as a mobile robot.

CHAPTER 3 - LOCALIZATION SENSOR DESIGN AND IMPLEMENTATION

Most predominant solutions in the literature obtain estimates of pose or position by extracting information from the working environment, which inevitably increases the dependence of the estimation accuracy on the environment. Such dependence limits the application scenarios of those solutions and may deteriorate their performance when the working environment is not ideal. Some solutions such as visual odometry and SLAM are usually computationally expensive and demand adequate resources. However, mobile robots are often equipped with limited resources and cannot afford to execute such solutions. Therefore, an innovative localization sensor is proposed in this thesis, aiming at providing pose estimation for ground mobile robotic agents with limited computational power and at reducing the dependency on the working environment characteristics through collaboration between the agents.

The design and implementation of the proposed localization sensor are documented in this chapter. Section 3.1 presents the conceptual design. Details about the physical implementation and data acquisition are presented in Section 3.2 and Section 3.3, respectively. In Section 3.4, details are provided about the configuration of the implemented localization sensor instances.

3.1 Conceptual Design

Using the proposed localization sensor requires collaboration among a number of instances of the device. The sensor units will actively monitor each other, forming a collaborative localization sensor network. In this network, each sensor unit carries a cylindrical landmark with a unique color-coded pattern and serves as a visible feature that can be identified by other units. As a result, pose estimation can be robustly performed without any form of dependency on information existing in the environment.

Actual pose estimation with the localization sensor is achieved by measuring the relative locations among the sensor units, which are acquired as relative distance and angular measurements collected from

one unit to another. An example with five localization sensor units is illustrated in Figure 3.1, in which the sensor units are represented by corresponding and uniquely color-coded landmarks, while the orientation of the respective units is illustrated as large arrows. The relative distances and angles among the units are denoted by $\rho_{<i,j>}$ and $\theta_{<i,j>}$, respectively, the subscript of which indicates that the corresponding measurements are made by *Sensor i* to *Sensor j*.

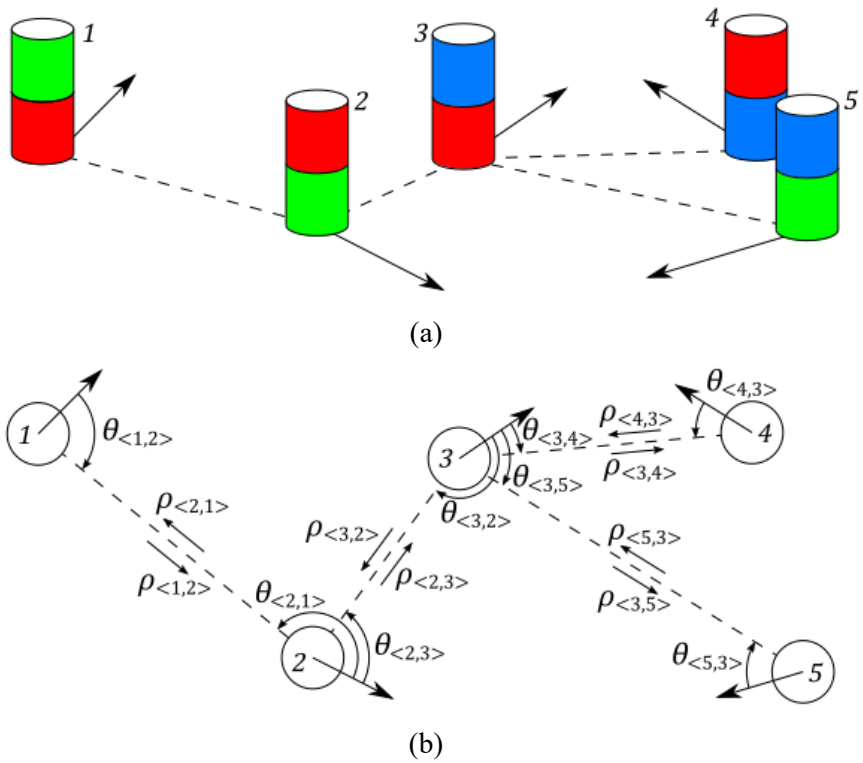
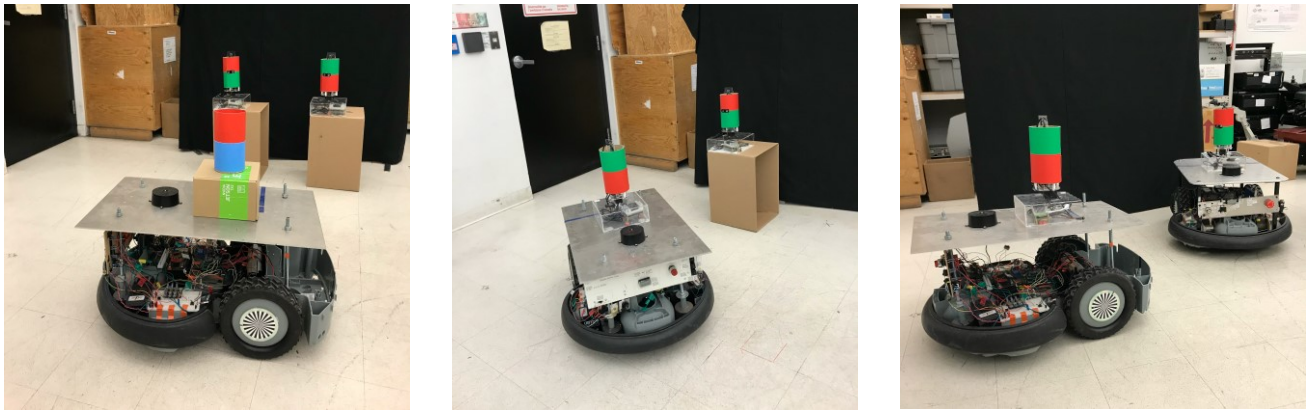


Figure 3.1 An example of collaboration among five units of the localization sensor: (a) an orthographic view of the units, (b) a corresponding top view.

In the example of Figure 3.1, given the absolute pose of *Sensor 1* and the relative measurements made by it towards *Sensor 2*, the position of *Sensor 2* can be readily estimated. The orientation of *Sensor 2* can be determined after it takes relative angular measurements back to *Sensor 1*. In the meantime, the relative distance measurements taken by *Sensor 2* contribute to improving the accuracy of its estimated position. When estimating the pose of multiple sensor units while only the pose of some is known, the pose of all sensor units can be estimated sequentially and collaboratively. For instance, if the pose of

Sensor 1 and *Sensor 4* are known, given all the measured relative locations shown in Figure 3.1(b), the pose of *Sensor 2* and *Sensor 3* are first estimated and the pose of *Sensor 3* can be further improved in collaboration with *Sensor 2*. The pose of *Sensor 5* can then be estimated based on the estimated pose of *Sensor 3* and the relative location between the two measured by both.

When using the localization sensor, multiple units will be attached to individual ground mobile robots with known relative pose with respect to the robots or be installed at fixed locations at approximately the same elevation. If the pose of some sensor units is accessible with respect to a given world reference frame, the pose of other units can be estimated by these units with respect to the same reference frame. Otherwise, the pose of all the units can be estimated with respect to the local reference frame of a chosen sensor unit. Some application scenarios using the proposed localization sensor units are shown in Figure 3.2. By estimating the position of the landmark or the pose of the sensor units attached to the robots, the position or pose of the robots can be estimated. Since only two instances were implemented for validating the proposed localization sensor design, the collaborative sensor pose estimation using the proposed localization sensor is not shown in the scenarios. However, in Figure 3.2(a), the red and blue landmark simulates a third localization sensor unit and showcases the collaborative landmark position estimation, which is an essential part of the collaborative sensor pose estimation using the proposed localization sensor. Details of the developed estimation models and methods will be further discussed in Chapter 5.



(a) (b) (c)

Figure 3.2 Some application scenarios using two instances of the proposed localization sensor. (a) The units individually and collaboratively estimate the position of the landmark with red and blue sections installed on a mobile robot. (b)(c) The units inter-calibrate with each other to estimate the pose of one with respect to the other. In (b), one of the sensor units is fixed and remains static.

Multiple instances of the proposed localization sensor collaborate by wirelessly sending their collected relative measurements to a base station, on which the estimation process takes place, and the base station will send back the estimated pose of each unit. Then, the pose of a robot with a localization sensor unit installed onboard can be estimated by making use of the relative pose between the robot and the sensor unit. In order to deal with the varying relative locations among the robots, a color camera is integrated into each sensor unit for visually detecting and tracking landmark patterns carried by other units in the environment, and the orientation of the camera is rotated around a vertical axis by a full-revolution servo motor. A laser rangefinder is installed along with the camera for collecting relative distance measurements, and therefore its orientation is also rotated by the servo motor. The laser rangefinder is mechanically aligned with the camera, and the servo keeps the detected pattern centered on the image so that the rangefinder is pointed towards a sensor unit corresponding to the given landmark pattern to collect distance measurements. Relative angular measurements are collected from the servo motor feedback simultaneously with the relative distance measurements.

3.2 Physical Implementation of Localization Sensor

In order to fit on a mobile robot, the proposed localization sensor should be of appropriate size while having little interference in the original routines of the robot. The individual components of the sensor should be compact in size, but its processing module, as the most important component, should also have sufficient processing power to be competent for processing camera images. Moreover, as the localization sensor will be mounted on mobile robots, all the components need to be energy efficient for long-term operations.

A prototype of the localization sensor is shown in Figure 3.3. The prototype sensor has a dimension of 281 mm×306 mm×141 mm and it is mounted on a Turtlebot3 Waffle Pi robot. The sensor shown in the figure is currently connected to power adapters but can be easily modified to use the same 12V power as that of the mobile robot. The cylindrical landmark is installed right on top of the axis of rotation of the servo motor, which is also considered the center of the sensor unit, and thus the actual relative distances and angles among the units are directly reflected by the relative measurements.

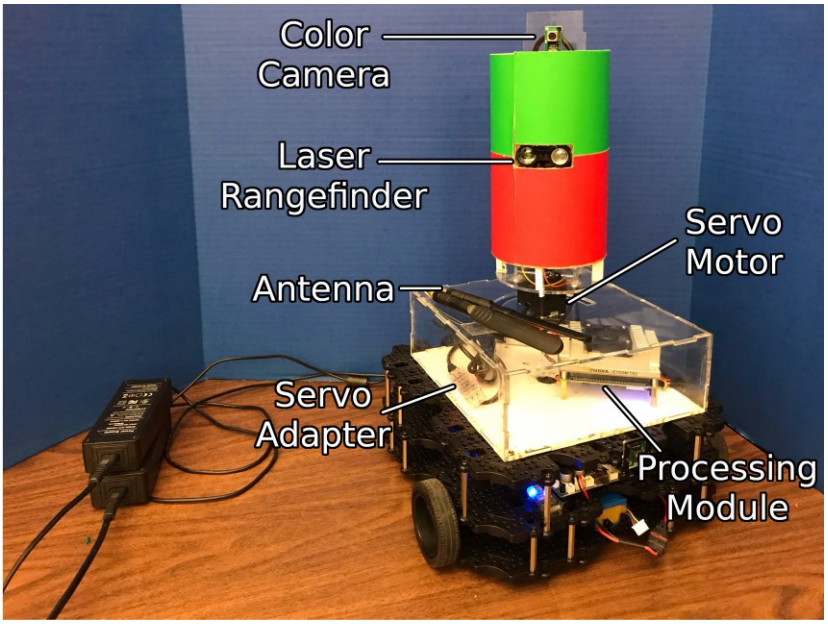


Figure 3.3 A prototype of the localization sensor mounted on a Turtlebot3 Waffle Pi robot.

In this thesis, two instances of the proposed localization sensor are implemented to experimentally validate the collaborative pose estimation with the sensor. Both instances use the same components except for the processing module. In the following subsections, details about the components of the localization sensor are provided.

3.2.1 Camera

Cameras have been widely used in the field of mobile robotic applications in recent years, as they provide abundant information about the surrounding environment of the robots, yet their price can be inexpensive. The camera used for the localization sensor must have accurate color response and resolution in order to detect and track a landmark pattern robustly. Each implemented sensor is equipped with a See3CAM_130 USB 3.1 camera [49] by E-Con Systems, which offers 13 megapixels with a horizontal field of view of 60.2°. Though the cameras are compactly designed on a printed circuit board of 80 mm×15 mm, they are able to run 9 frames per second (fps) at the highest resolution of 4224×3156. In addition,

useful functions such as auto white balance and autofocus are offered by the cameras, which are helpful towards tracking the landmark by providing consistent images.

At the highest resolution, the camera provides rich information about the surrounding environment. However, the computational power required for processing the camera images is very high as well. Besides, the framerate at higher resolutions is slower than that at lower resolutions. To ensure a fast response of the landmark detection and tracking algorithm, the operating resolution of the camera is then set at 1920×1080 with a framerate of about 30 Hz. The details of the landmark detection and tracking algorithm will be discussed in Section 3.3.2.

3.2.2 Laser Rangefinder

High performance laser rangefinders are often costly. Although cheaper rangefinders are available on the market, most do not fit the design because of either short nominal range or large beam divergence. Among the many choices, the selected Garmin LiDAR Lite v3 [50] measures distance over 0.5 to 40 meters and has a beam divergence of 0.008 rad. The LiDAR offers an accuracy of 0.025 m within five meters and 0.1 m at a longer range. Along with its small form factor of 20×48×40 mm and versatile I2C interface, the LiDAR is well-suited for the design of the proposed localization sensor. Additionally, the laser emitted by the LiDAR is rated as class 1, which is safe under normal use conditions. The accuracy of measurements varies according to the measured distance, but deviations can be properly compensated by modeling the characteristics of the LiDAR, which is documented in Section 3.4.2. The LiDAR is said to be able to measure distances at a frequency up to 500 Hz, but in reality, with a frequency higher than 50 Hz, measurements become less consistent. The frequency is therefore set to 20 Hz to make sure the distance measurements are reliable.

Since the laser beam diverges and its dispersion increases over distance, the size of the landmark and the working range should be determined accordingly. A simplified illustration of the beam divergence, θ, is shown in Figure 3.4, where L is the distance from a LiDAR placed at S facing a flat panel on the right. According to the manufacturer, the beam diameter, d, can be estimated as $L/100$ at distances over 1 m. For instance, at a working distance of 10 meters, the beam diameter is approximately 0.1 m. On the implemented localization sensor, a cylindrical landmark with a diameter of 0.106 m and a height of 0.2 m is implemented, which fully reflects the laser beam emitted by the LiDAR at a working distance up to 10 meters.

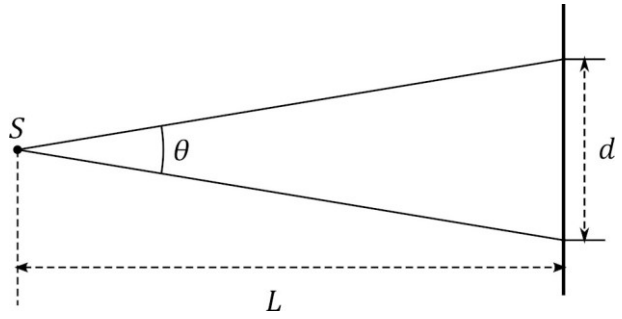


Figure 3.4 An illustration for beam divergence.

3.2.3 Servo Motor

In the implementation of the proposed localization sensor, a Dynamixel XL430-W250 [51] motor is selected as the servo motor that directly drives the rotary part, which contains the camera and the LiDAR. The motor provides accurate feedback on motor states at 100 Hz and a USB interface for simple integration. The feedback of the servo motor has negligible delays, and therefore it is considered reliable to be directly used for angular measurements indicating the relative direction of a detected landmark. Several working modes, such as velocity control and position control, are offered by the motor. To achieve better accuracy on angular measurements, the implemented localization sensor relies on the position control mode, which takes desired positions as control commands. The parameters of the mode are chosen

empirically so that the servo controller has a fast and accurate response but insignificant oscillations. A screen capture of the R+ Manager provided by Robotis, which is used for tuning the motor parameters, is shown in Figure 3.5.

Figure 3.5 The control table in Robotis R+ Manager.

In position control mode, the motor travels within 0° and 360°, while the 12-bit contactless magnetic absolute encoder generates 4096 steps for each revolution. As angles are defined in $[-\pi, \pi)$ on the localization sensor, the encoder readings are mapped to this range, where 2048 serves as the zero reading of the motor. The motor resolution is about $360°/4096 \approx 0.088° \approx 0.00154$ rad, which is smaller than the beam divergence of the LiDAR. Based on similar calculations to that depicted in Figure 3.4, the motor is considered suitable for the chosen landmark design over a working range of up to 10 meters.

However, the motor cannot rotate unlimitedly in one direction due to the limitation of the encoder in position control mode. Once the motor has reached an endpoint of the revolution, the motor rotates to the other endpoint at a fast speed in the reverse direction and then continues the previous rotation. Fortunately, this feature does not pose a threat to the detection and tracking of the landmark, and instead,

it is helpful to prevent entangling the wires connecting between the camera and the processing module on the localization sensor.

3.2.4 Processing module

Since the proposed localization sensor is developed for mobile robotic applications, the embedded processing module is limited by the size and power consumption of the sensor, while providing sufficient computational power. To facilitate prototyping, the NVIDIA Jetson TX2 [52] module is chosen for the localization sensor. The module has a compact size of 50×90 mm and features a GPU that provides 256 CUDA cores, which is capable of complicated image processing tasks. It also features a 2GHz hex-core CPU complex, which consists of a Quad-Core ARM Cortex-A57 CPU and a Dual-Core NVIDIA Denver 2 64-bit CPU. Along with its 8GB 128Bit LPDDR4 memory, the module is competent enough for the implementation of the localization sensor. To fully leverage the small size and power of the module, a CTI Orbitty carrier board [53] is used, which provides helpful breakout ports, such as USB 3.0, Gigabytes Ethernet and GPIO pins, for integration with the abovementioned components. The Jetson TX2 module comes with a wireless communication module, and a maximum communication speed of 867 Mbps can be achieved with additional antennae installed, which is capable of communicating with the base station for exchanging information on measurements and estimated poses.

After the first implementation of the proposed localization sensor, NVIDIA has released the Jetson Nano kit [54] in April 2019. Though the kit has only one ARM Cortex-A57 CPU clocked at 1.43GHz with half the memory and CUDA cores as the Jetson TX2 module, it has been tested to be competent for the localization sensor. The Jetson Nano kit does not support wireless communication natively, but it comes with an M.2 Key E for installing an additional Wi-Fi network adapter card. With a proper Wi-Fi card and antenna installed, a Jetson Nano kit is used in a second instance of the proposed sensor.

3.3 Data Acquisition with Localization Sensor

The collaborative pose estimation of the localization sensor is implemented under a centralized communication scheme where all sensor units send their information to a central base station. The base station estimates the pose of each sensor unit based on the received information and sends back the estimated pose of the unit. It is assumed that all sensor units are within the communication range to connect to the base station, and all sensor units and the base station are connected to the same Wi-Fi during operation.

3.3.1 Data Acquisition Software

The software has been developed as packages leveraging Robotic Operating Systems (ROS) [55] provided by the Open Robotics Foundation. ROS is one of the most popular open-source platforms for robotics and provides abundant tools from its community. Packages developed using ROS are usually highly modularized and reusable, and the communication among threads is well encapsulated in simplified messaging interfaces. Therefore, the use of ROS accelerates the process of development and implementation. For instance, the packages developed initially for the first sensor unit could be easily deployed to the following ones.

A graph illustrating the packages running on each localization sensor unit is shown in Figure 3.6. In the figure, except for packages developed by or adopted from the ROS community, which are grayed out, all other packages are dedicatedly developed for the proposed localization sensor unit in this work. The drivers for the camera, servo motor and LiDAR are shown on the left side of the graph. The visual detection and tracking package developed in this work receives the camera image and outputs the position of the landmark center on the image. Upon detection of a landmark, the servo controller calculates the control command accordingly and adjusts the orientation of both the camera and the LiDAR. Due to the

time spent on image processing, the detected position of a moving landmark is always slightly delayed with respect to the most recent feedback of the servo motor state and distance measurement from the LiDAR. As the motor may be still rotating after an image is taken, this delay causes inaccurate servo control commands and imprecise distance estimates, degrading the performance of the sensor. By introducing a dedicated package for buffering data from the LiDAR driver and servo driver, each measurement is stored with a timestamp, and a service is provided by the data buffer to respond to measurement queries at given timestamps. In the implemented data buffer, only measurements within a certain period of time are stored in order to save onboard memories and the queried measurements are always linearly interpolated between two consecutive timestamps. If the interpolation fails, a negative large number is returned to indicate the situation.

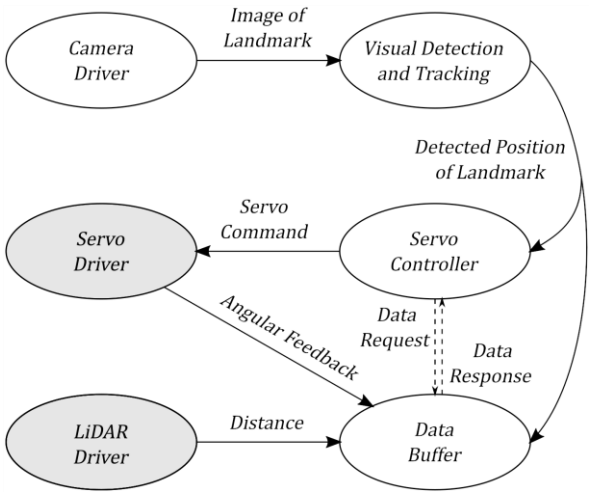


Figure 3.6 A graph of the ROS packages on each sensor unit. The solid arrows represent information from one package to another, while the dashed arrow represents a query for measurements. The grayed out packages are developed by third parties and not originally for this work.

Though the proposed localization sensor works under a centralized communication scheme, each sensor runs the same ROS packages independently without sharing the ROS message server. A centralized communication utility based on RabbitMQ [56] is developed for sending and receiving information between the sensor units and a base station, and the communication server is deployed on the base station.

The communication utility makes use of the messaging system of ROS, as the ROS messages can be easily serialized and deserialized with respect to their message types, and the utility can be easily modified to send and receive any required messages. The protocol used in the communication utility is also interchangeable with many other options, which enables the possibility to exchange information between only the sensors for a decentralized communication scheme. Using such communication utility offers the advantage to reduce communication load between the sensor units and the base station when compared to the original ROS message server. The sensor units no longer need to send the camera images to the base station to be processed and only a small amount of data, such as the measurements and detected landmark positions, is required to be sent to the base station. Since most processing tasks are performed locally at each localization sensor unit, the specifically developed software for the sensor is robust to delays caused by communication through the wireless network.

3.3.2 Landmark Detection and Tracking

As mentioned in Section 3.1, each localization sensor unit carries a cylindrical landmark with a unique color-coded pattern. The cylindrical shape of the landmark has a horizontal cross-section of a circle, and thus it is invariant to any horizontal angle of observation when observed by another sensor unit at a similar elevation. On the implemented sensor units, the landmarks are made from a Uline kraft mailing tube with a diameter of 106 mm, and the height of the landmark is about twice its diameter.

The unique pattern of each landmark consists of a combination of multiple vertically neighboring sections with distinct colors. The use of colored sections helps increase robustness to motion blurs. Using more colored sections can increase the contrast between the landmark and its surroundings. More color combinations would be available, but in this work, only two sections are used for the landmark, as it proves sufficient for the implementation and experimental validation of the localization sensor. The

different landmarks used in this work are shown in Figure 3.7. The landmarks with green and red are to be installed on the implemented sensor units, and the cut-outs on the landmarks are for the laser rangefinders to take measurements to other landmarks.

Figure 3.7 The landmarks used in this work.

The landmark detection and tracking algorithm is developed by improving a sampling-based algorithm presented in [57], where similar but stationary color-coded landmarks are detected and tracked as a mobile robot moves in the environment. In the original algorithm, a landmark depicting a color pattern is detected by many samples with weights spread all over the input images. The position of the landmark center on the image is considered as a random variable with distribution and represented by the weighted samples. At each time step, the samples are processed in three phases as follows. First, the positions of all the samples are predicted according to the motion model of the mobile robot. Second, the weight of each sample is updated by a landmark measurement model, which compares the actual colors in the image to the known colors on the landmark to be detected. To simplify the comparison, only the colors at two pixels vertically and symmetrically located above and below the position of each sample are actually compared, and the positions of the pixels correspond to the positions of the colored sections used on the target

landmark. The measurement model is illustrated using one sample on a landmark pattern in Figure 3.8, where the black dot represents the position of the sample on the image and the white circles represent the two pixels. A color difference between the observed values and those expected on the landmark is determined at each pixel using red and blue chromaticities, but the details are not disclosed in the original work. The weight of each sample is then updated as the inverse product of the color differences at the two pixels. In the last phase, a resampling process is carried out according to the weights of the samples so that the samples will propagate towards those with higher weights, which are usually located at the actual position of the landmark center. Noticeably in the procedure, the distance between the two pixels of the measurement model is also considered a random variable, and its value varies randomly along with the sample position at each time step.

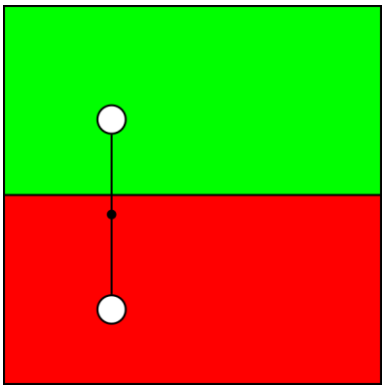


Figure 3.8 The original landmark detection model for each sample.

The original sampling-based algorithm has the advantage of reducing the computational burden by computing over only samples instead of the whole image. But several problems remain in its application. The chromaticity color representation in RGB color space is not robust enough under varying lighting conditions. Using only two pixels for each sample fails easily in complicated environments. Also, the calculation for weights using color differences generates non-constant weights on the landmark due to the existence of noise, leading to off-center detection results.

In the solution developed in this thesis, the first problem is solved by using representations in Hue-Saturation-Value (HSV) color space, which are more robust to varying lighting conditions, and colors could be distinguished solely by their hues if properly saturated and lightened. An image of an HSV color bar with fully saturated and lightened colors organized by hue is shown in Figure 3.9. The range of hue is adopted from OpenCV as [0,179] inclusive, and the color plotted at 179 is almost as red as that at 0. Ideally, there are 180 colors that can be used for coding the landmarks, but due to the existence of noise in the camera, the selected pair of colors should be appropriately distinctive from each other with evident hue difference.

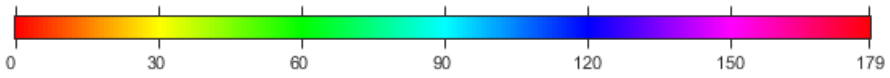


Figure 3.9 A color bar with fully lightened and saturated colors plotted by hue.

For the second problem, an improved measurement model is formulated for calculating the weight for each sample, where two lines of pixels are used instead of a single pixel on each side of the sample, as shown in Figure 3.10. The length of the lines is set proportional to the distance between the upper and lower pixels, and the proportion is chosen empirically to trade off fast computational speed for robustness to the noise in the imaging process. As mentioned earlier, assigning weights according to color is subject to image noise, which would lead to off-center landmark detection results. In order to better obtain the center position of the landmark on the image, predefined constant weights are assigned to the samples instead of calculated values. A large constant weight is assigned to a sample if the HSV components of all colors at its pixels are within an empirically chosen range from the colors of the landmark. Otherwise, a very small weight will be assigned. Therefore, when a landmark is detected, the samples will eventually be symmetrically centered on the color transition part of the landmark instead of spreading as a unimodal distribution, and the position of the landmark center is given by the average of the positions of all samples.

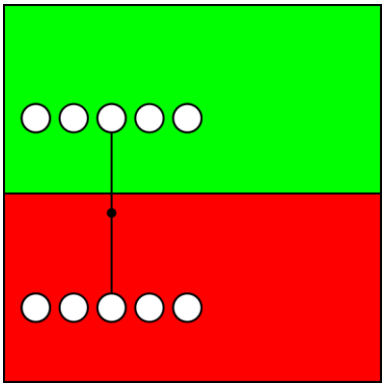


Figure 3.10 The improved landmark detection model for each sample.

The implementation of the landmark detection and tracking algorithm is based on a particle filter framework provided by the open source library MRPT [58], and the algorithm runs with 4000 samples to detect the position of the landmark center at 10 Hz. Initially, the samples are uniformly distributed all over the input images. Then during each iteration, the samples are updated on a new input image, and the same image will be reused if a new one is not yet available, depending on the framerate of the color camera. For each sample, both its position and velocity on the image are updated in order to better capture the landmark on the image. The position and velocity of the samples are implemented as continuous variables, but the positions will be rounded as integers to match corresponding pixels on the image. Large variances are added to the positions and velocities to counteract the varying relative locations among sensor units. When the algorithm loses track of the landmark, the large variances will drive the samples to diverge from each other, making them quickly spread all over the input image again to search for the landmark and attempt to reacquire it. The resampling process is provided by the MRPT library, which is tested to be suitable for the localization sensor. An illustration of the sampling-based landmark detection and tracking algorithm is shown in Figure 3.11, where the samples are shown as cyan dots on the images at corresponding positions. An in-frame boundary can be noticed, especially visible in Figure 3.11(b1) and Figure 3.11(a3), for accommodating the samples to compare colors around their positions. The samples will reverse their moving directions after reaching the boundary. In the sequence represented in Figure

3.11(a1), (a2) and (a3), the initially captured samples in (a1) diverge quickly to cover a larger area with the goal to reacquire the landmark after it is occluded in the image. Reversely, in the sequence shown in Figure 3.11(b1), (b2) and (b3), the initially broad distribution of samples quickly converges to only a few samples on the detected landmark center after the resampling process, when the landmark is revealed to the camera.

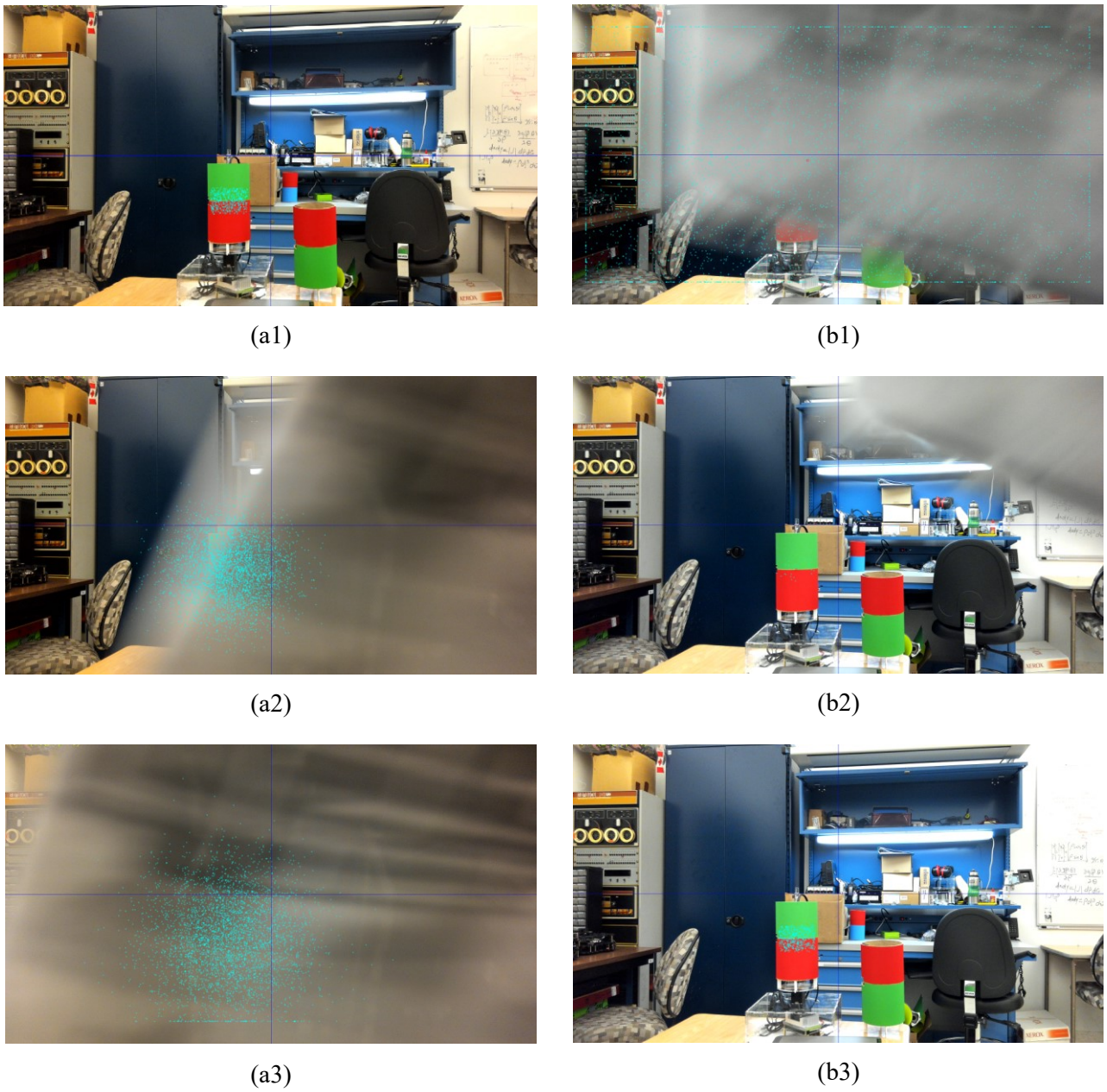


Figure 3.11 Two sequences of observations for detecting a color-coded (green over red) landmark: (a1)(a2)(a3) consecutive frames showing a detected landmark being occluded from the sensor; and (b1)(b2)(b3) consecutive frames showing the landmark being detected while being revealed to the camera. The blue cross in the images represents the center of the image.

The landmark detection and tracking algorithm is developed for continuous execution, and the landmark detection results are generated repetitively. Both the average and the variance of the position of all samples are estimated. The variance helps determine if a landmark is successfully detected, as a small variance indicates that the samples are concentrated on the landmark. If a landmark is detected, the average

position can then be used as the center position of the detected landmark. Additionally, the algorithm does not actively request camera orientation adjustment, although the landmark is detected and tracked within the camera image. The need for adjusting the orientation of the camera to track a moving landmark beyond the camera field of view is determined by the servo controller in parallel, based on the landmark detection and tracking results along with the size of the input image and the timestamp when the image is taken. All the abovementioned information is passed on to the servo controller, which will be discussed in the following subsection.

3.3.3 Servo Controller

The main purpose of the servo controller on the localization sensor is to adjust the orientation of the laser rangefinder and the camera so that the sensor can initiate relative pose measurements. Another important task of the controller is to assist the camera in searching for landmarks mounted on other sensor units whose poses are to be estimated. The two tasks are carried out one at a time depending on whether a landmark is detected. While the variance on the position of the samples in the landmark detection and tracking algorithm is larger than a threshold, the landmark is considered not detected and the controller will perform the search task. Otherwise, the controller adjusts the orientation of the rotary part of the unit based on the detected position of the landmark on the embedded camera image.

Since the camera is installed on each sensor unit with the same horizontal orientation as the laser rangefinder, when the rangefinder on a sensor unit points towards another unit to collect distance measurements, the landmark on the other unit should be horizontally centered on the camera image. Thus, the horizontal position of the landmark center on the image is used as direct feedback for calculating control commands for the servo motor. Let p_C be the horizontal center position on the image when the rangefinder is properly oriented towards the target robotic unit with its own landmark, and p be the current

detected position of the landmark center in the image. The angular error of the current camera orientation to the desired one is given approximately as $(p - p_C)F/R$, where R is the horizontal resolution of the camera image and F is the field of view of the camera. Since the selected servo motor works in position control mode, as discussed in Section 3.2.3, the angular error is directly converted into motor steps and added to the encoder value when the corresponding image is taken to generate a new control command in terms of steps. The update rate of the controller is set to the same as that of the landmark detector, and thus the controller only updates the control command when a new output from the landmark detector is generated.

In the landmark search task, the servo motor rotates the camera so that successive images taken by the camera cover the surrounding environment entirely. Since it is not possible to slow down the speed of the selected servo motor in the position control mode to reduce the effects of motion blurs, the servo motor stops for a while after rotating certain degrees for the detection and tracking algorithm to properly detect a given landmark if it is within the camera's field of view. Once the landmark is detected, the servo controller starts to adjust the orientation of the camera in order to horizontally center the landmark on the camera image.

3.4 Sensor Configuration and Characterization

The purpose of the configuration and characterization processes reported in this section is to ensure that the implemented localization sensor units provide reliable and expected performance. The configuration was carried out while building up the sensor from scratch, as is detailed in Section 3.4.1. In Section 3.4.2, the measurements from each sensing component are characterized, and a probabilistic model is derived for each type of measurement. A calibration procedure based on the iterative least-square

(LS) estimation is presented in Section 3.4.3, which helps to estimate the pose of a localization sensor unit in an offline manner.

3.4.1 Assembly Configuration

The localization sensor employs a laser-cut acrylic enclosure for accurate dimensions, and the stiffness of the material is appropriate for supporting its components. The installed components are aligned as desired, but errors associated with manual assembly still exist, among which the misalignment of the LiDAR and motor with the horizontal plane has the most impact on the performance of the sensor. The assembly errors of the LiDAR and motor are illustrated in Figure 3.12, where S is the LiDAR of one sensor unit taking measurements to another unit, and ϵ_L and ϵ_M are the LiDAR and motor assembly errors, respectively. The motor assembly error, ϵ_M, results from the tilting angle of the plane of rotation of the servo motor away from the horizontal plane, and the LiDAR assembly error, ϵ_L , is the difference angle between the laser beam and the plane of rotation of the servo motor.

Ideally, the laser beam emitted by the LiDARs, $\overline{SM}$, should be perpendicular to the axis of rotation of the landmark, and the axis should be perpendicular to the horizontal plane. Thus, the emitted laser beam should be parallel to the horizontal plane so as to be able to collect measurements towards objects at approximately the same elevation at long distances. However, in reality, both aforementioned assembly errors cumulatively contribute to make the emitted laser beam, $\overline{SM}$, not to perfectly coincide with the horizontal line at the same height as S, $\overline{SN}$.

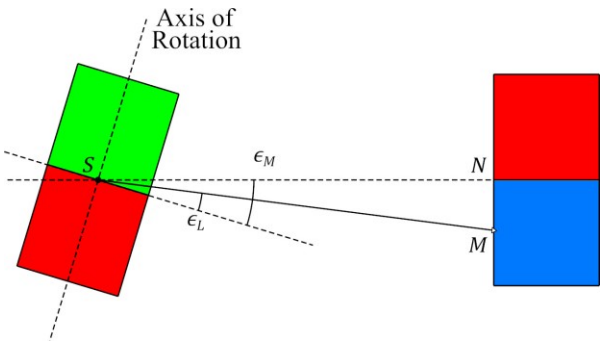


Figure 3.12 A sensor unit located at *S* taking relative measurements to another unit while being affected by both assembly errors on the LiDAR and motor.

During the configuration process and in order to minimize assembly errors, the laser spot projected by the LiDAR is being measured as an important indicator. However, the wavelength of the laser is 905 nm, which is longer than that of visible light. Fortunately, the wavelength can be captured by cameras without infrared filters. A Raspberry Pi NoIR camera is employed to facilitate the process, and video from the camera is streamed by a Raspberry Pi wirelessly. An image of the laser projection on a distant landmark taken from the Raspberry Pi NoIR camera is shown in Figure 3.13.

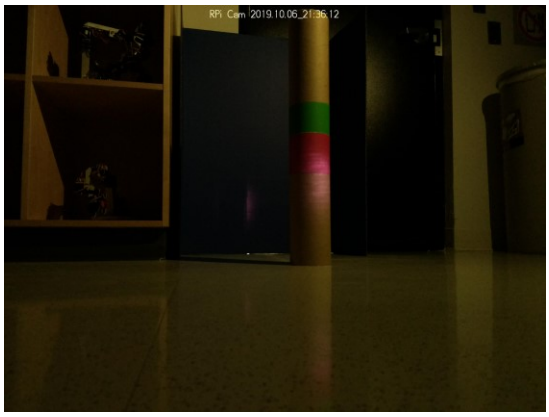


Figure 3.13 An image of the laser projection taken from the Raspberry Pi NoIR camera.

To characterize the assembly errors, each sensor unit is placed at a fixed location while pointing its LiDAR to a vertical panel far away from the unit. The height of the LiDAR in the assembly is measured and marked on the panel for comparison with the laser spot projected by it. As shown in Figure 3.12, both assembly errors on the LiDAR and motor contribute to the difference between the height of the LiDAR

and the height of the laser spot. By observing the position of the laser spot while keeping the LiDAR still and rotating the motor, the effect of the motor installation error can be directly observed. After adjusting the motor to make its shaft as vertical as possible, correction on the LiDAR assembly error can then be achieved by adjusting the LiDAR to minimize the height difference between the projected laser spot and the LiDAR elevation in the assembly.

The assembly error of the embedded cameras can be corrected in a similar way, but since the camera is well fastened on the acrylic board and has a viewing angle of about 60°, such correction on the vertical orientation of the camera is unnecessary. However, the horizontal orientation of the camera should be configured in order to properly control the corresponding servo motor and to collect accurate measurements. The configuration can be easily achieved using the data acquisition software of the localization sensor. Firstly, let an assembled sensor unit successfully detect a stationary landmark placed far away from the units. Then the servo motor is controlled manually to rotate in both directions step-by-step to make sure the laser spot is fully projected on the landmark. The horizontal position of the landmark center on the image of the camera at the time is then deemed as the true horizontal center of the camera.

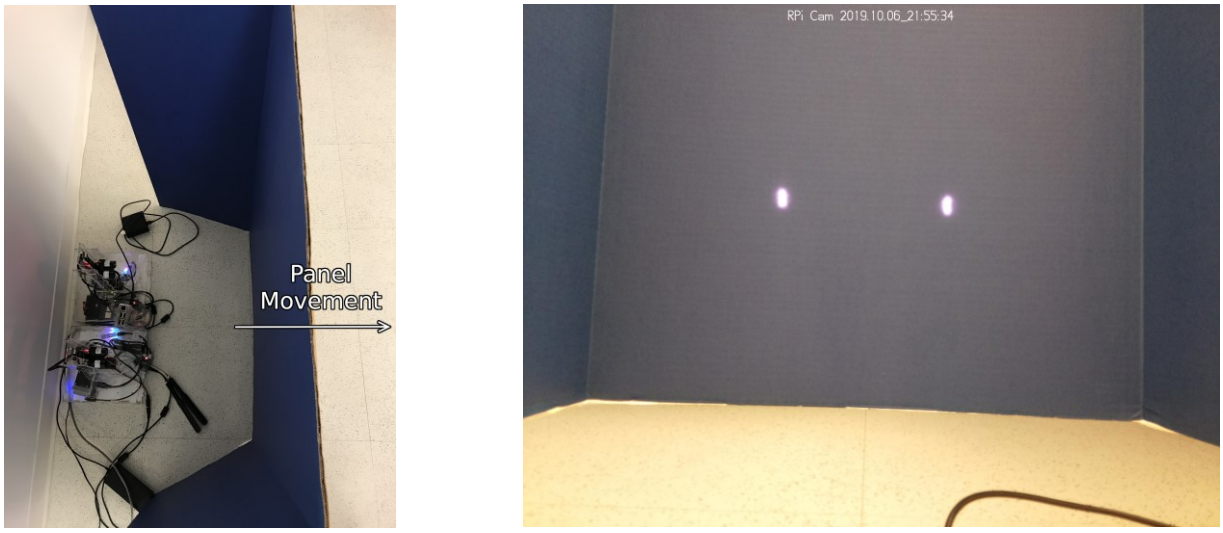
3.4.2 Sensor Characterization

Measurements of distances and angles are inherently corrupted by noise. In order to mitigate its effects, an effective practice is to model the characteristics of the embedded sensing components by comparing their acquired data to the corresponding ground truth. It is also important to characterize every sensing component individually as their characteristics may differ one from another, even though they may be made by the same manufacturer. In this subsection, two implemented instances of the proposed localization sensor are characterized for their distance and angular measurements. The two sensor units are named *Sensor 1* and *Sensor 2*, respectively. As mentioned earlier, the two instances use the same

components except for their processing modules. *Sensor 1* is equipped with a Jetson Nano Kit, while *Sensor 2* is equipped with a Jetson TX2 module on a CTI Orbitty carrier board.

3.4.2.1 LiDAR and Distance Measurements

In the characterization of distance measurements, the LiDAR on each sensor unit collects distance measurements against a flat panel at multiple known distances. Data collection made use of the uniform square floor tiles in the SITE building as a reference, given that the fixed length of the tile sides is 304.8 mm. The panel is placed vertically at the edges of the floor tiles in front of both sensor units and errors are neglected when placing the panel. The acquisition setup is shown in Figure 3.14, where the LiDARs on the two implemented sensor units are characterized at the same time.



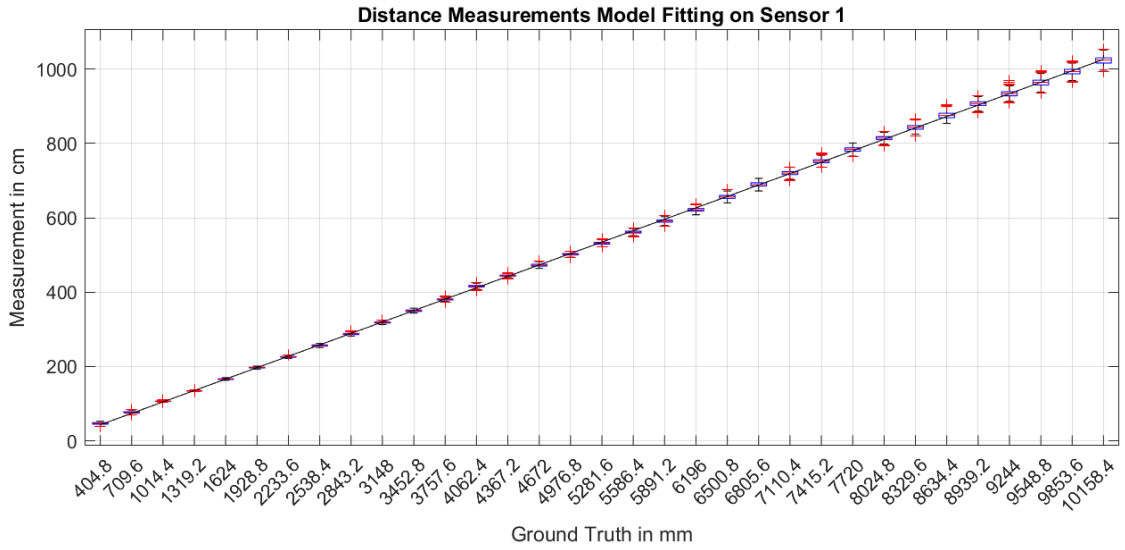
(a) (b)

Figure 3.14 (a) A top view of the data collection setup with two localization sensor units. (b) A picture taken with a Raspberry Pi NoIR camera depicting the two laser projections from the LiDARs.

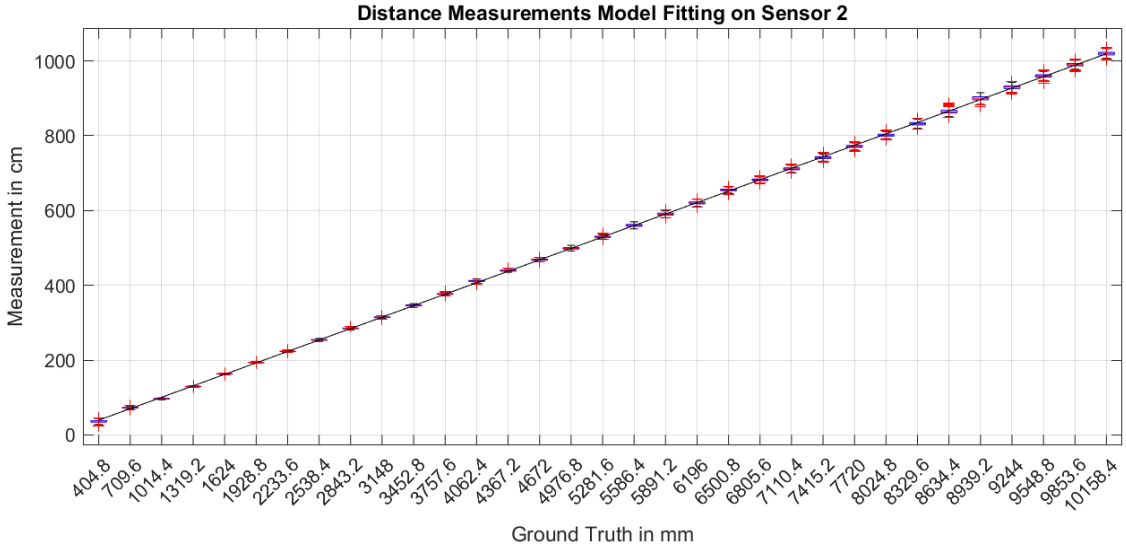
One thousand measurements are collected every 304.8 mm from 1 tile to 33 tiles away. Because of a 100-mm offset resulting from placing the two implemented sensor units beside each other, the actual distance from each LiDAR to the panel ranges from 404.8 mm to 10158.4 mm. The range is selected as

such to cover the nominal minimum working distance of the LiDAR, which is 500 mm, and is limited by the space available in the testing environment. The measurements are shown in Figure 3.15, where they are plotted as boxplots at the measured distances. It can be observed from the plots that the measurements approximately have a linear relationship with the ground truth distances, and therefore the measurements can be fitted into an observation model of a first-order polynomial. The observation model is given as $d_M = f(d) + \epsilon_{d_M} = kd + b + \epsilon_{d_M}$, where k and b are the fitting parameters, d is the actual distance, d_M is the corresponding measurement, and ϵ_{d_M} is the measurement noise. Note that d_M and ϵ_{d_M} are in centimeters while d is in millimetres in this model, as well as on the plots of Figure 3.15.

The analytical forms of the fitted lines of the LiDARs on *Sensor 1* and *Sensor 2* respectively are given as $f_1(x) = 0.1007x + 2.6578$ and $f_2(x) = 0.1004x - 1.1067$, which are plotted as straight lines in Figure 3.15. The two LiDARs have similar characterized polynomials, especially very close k parameters. Parameter k is a scaling factor representing how linearly the measurements are mapped to the ground truths, and it should be approximately 0.1 due to the different units of the measurements (cm) and the actual distances (mm). Parameter b is for compensating the offset of measurements at the distance of 0 mm, which is impossible to be actually measured but can be estimated by the characterized polynomial. Because the offset is affected by many factors such as factory settings and installation errors, the offset of each LiDAR may differ one from another, as is the case in these experimental observation models.



(a)



(b)

Figure 3.15 Characterization of the distance measurements with LiDAR: (a) on *Sensor 1*, (b) on *Sensor 2*.

The standard deviations on the distance measurements for the two LiDARs at each measured distance are plotted in Figure 3.16. For both LiDARs, the standard deviations have similar trends. Within the nominal minimum working distance of 500 mm, both LiDARs manage to collect distance measurements but with a large standard deviation. As the actual distance becomes larger, the standard deviation on both LiDARs first reduces and reaches a minimum at about 1014.4 mm. At larger distances,

the standard deviation on both LiDARs has a positive correlation with the measured distances. Since the distance measurements collected from *Sensor 2* have a smaller standard deviation over most of the characterized range, *Sensor 2* is likely to produce more reliable distance measurements. The experimentally observed standard deviation level is in line with the manufacturer's specifications, in which the reported accuracy is about 0.025 m within five meters and 0.1 m at a longer range, as reported in Section 3.2.2.

Around 4062.4 mm, 6196 mm and 8634.4 mm, the standard deviation on both LiDARs reaches local maxima. Since the two LiDARs are characterized at the same time, these local maxima are considered caused by the setup and the testing environment. The discrepancy in the performance of the LiDARs has been verified to be caused by the off-the-shelf LiDAR devices themselves rather than the different processing modules. By switching the LiDARs to a different localization sensor unit, the performance of the LiDARs remains the same. In addition, for confirming this, a third LiDAR device has been installed on *Sensor 2*, which has the Jetson TX2 module, and the performance was worse than that observed with both characterized LiDAR devices.

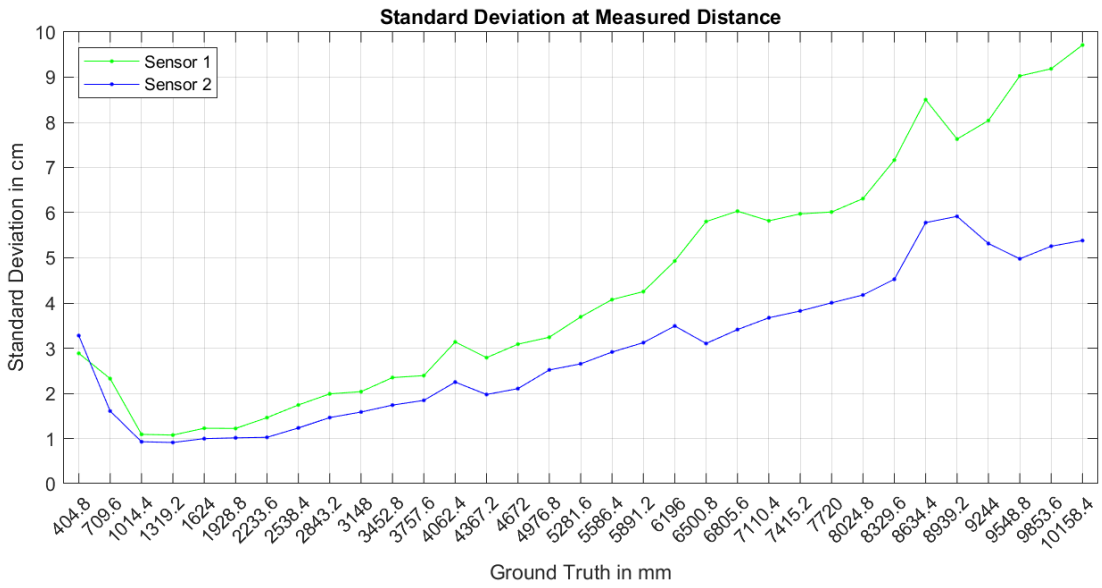
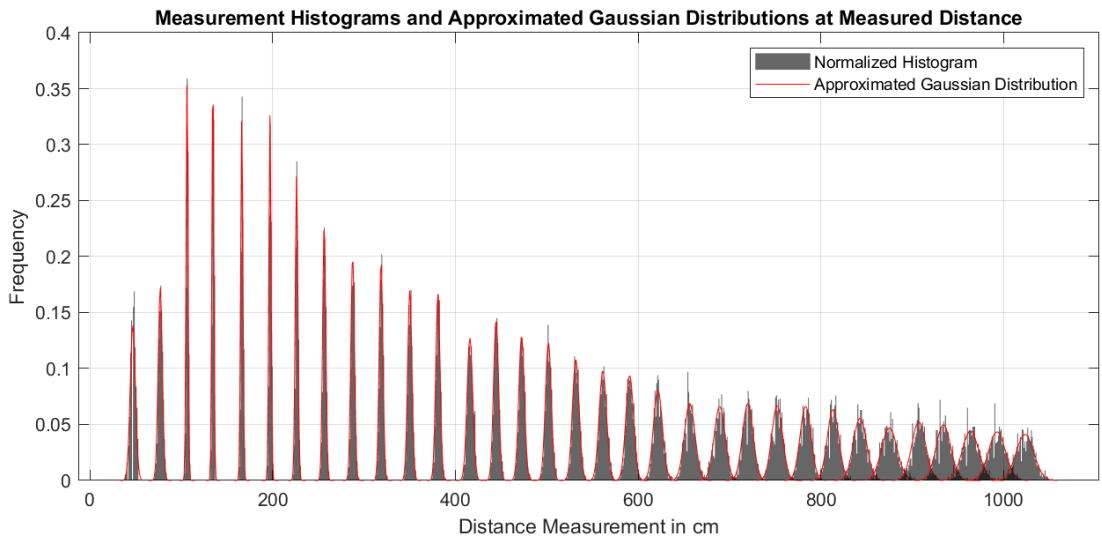
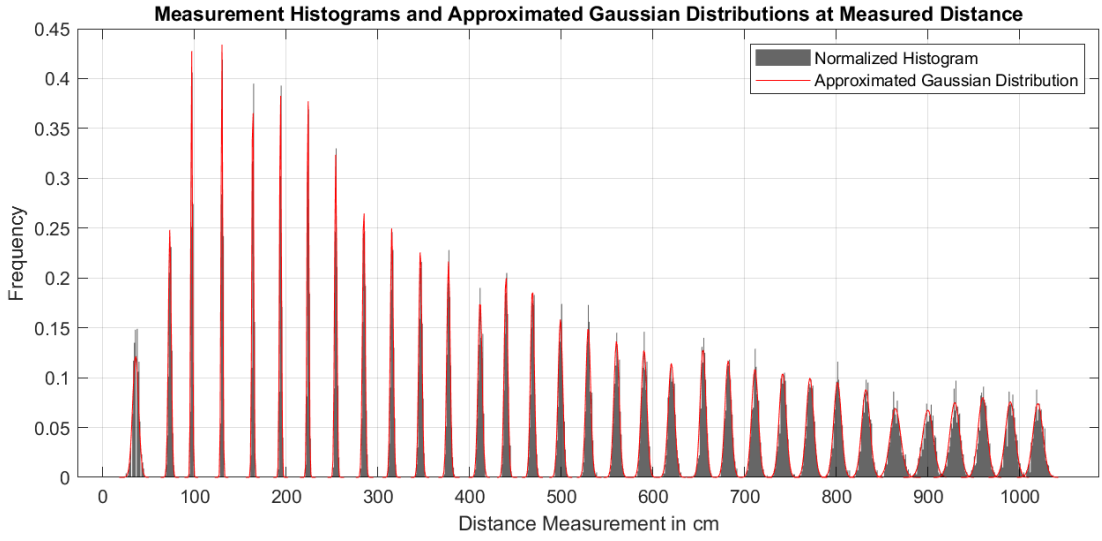


Figure 3.16 Standard deviations of the measurements at each measured distance for both LiDARs.

The measurements of each LiDAR are plotted as histograms at each measured distance in Figure 3.17. At each measured distance, the histogram is normalized in a way such that the width of each bin is 1 and the sum of bin areas is 1. The histograms can be fitted to approximated Gaussian distributions, which are plotted as red lines in the figures. The means and standard deviations of the Gaussian distribution functions are the same as the average and standard deviation of the measurements, respectively at each measured distance.



(a)



(b)

Figure 3.17 Fitting distance measurements at each measured distance to Gaussian distributions: (a) on *Sensor 1*, (b) on *Sensor 2*.

Since the first-order polynomial is properly fitted in Figure 3.15, the distance measurement noise, ϵ_{d_M}, is assumed to be of zero-mean Gaussian distribution. And it is evident that the variance of the measurement error is the same as that of the distance measurement, denoted by $\sigma_{d_M}^2$. Therefore, at a certain distance, d, the corresponding measurement, d_M, has an approximated Gaussian distribution:

$$d_M \sim N(f(d), \sigma_{d_M}^2),$$
$$p(d_M|d) = \frac{1}{\sqrt{2\pi\sigma_{d_M}^2}} \exp\left(-\frac{(d_M - f(d))^2}{2\sigma_{d_M}^2}\right), \tag{3.1}$$

where $f(d)$ is the fitted first order polynomial as shown in Figure 3.15.

As characterized in Figure 3.16, the standard deviation σ_{d_M} is correlated to the true distance d. To simplify the calculation, the value of σ_{d_M} is linearly interpolated from the characterized standard deviation at known distances shown in Figure 3.16 according to a given ground truth d.

Since the measurement noise is representatively mapped as white Gaussian noise, the measurements over the same ground truth are independent and identically distributed. Thus, in the Bayesian context, the likelihood function of certain true distance, d, given N measurements, $\{d_{Mi}\}, i = 1, \dots, N$, is expressed by:

$$\begin{aligned} p(\{d_{Mi}\}|d) &= \prod_{i=1}^{N} p(d_{Mi}|d) \\ &= \frac{1}{\left(2\pi\sigma_{d_M}^2\right)^{\frac{N}{2}}} \exp\left(-\sum_{i=1}^{N} \frac{(d_{Mi} - f(d))^2}{2\sigma_{d_M}^2}\right) \\ &= \frac{1}{\sqrt{N}\left(2\pi\sigma_{d_M}^2\right)^{\frac{N-1}{2}}} \exp\left(\frac{1}{2\sigma_{d_M}^2}\left(\frac{1}{N}\left(\sum_{i=1}^{N} d_{Mi}\right)^2 - \sum_{i=1}^{N} d_{Mi}^2\right)\right) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\mu - f(d)\right)^2}{2\sigma^2}\right) \\ &= \frac{\zeta}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\mu - f(d)\right)^2}{2\sigma^2}\right), \end{aligned} \tag{3.2}$$

where ζ is a scaling factor, $\mu = \frac{1}{N}\sum_{i=1}^{N} d_{Mi}$, $\sigma^2 = \frac{1}{N}\sigma_{d_M}^2$, and $f(d) = kd + b$. The posterior PDF of the true distance given the measurements, $\{d_{Mi}\}$, is given by applying Bayes' theorem:

$$p(d|\{d_{Mi}\}) = \frac{p(\{d_{Mi}\}|d)p(d)}{p(\{d_{Mi}\})}. \tag{3.3}$$

Without loss of generality, the prior PDF, $p(d)$, is assumed to be of a uniform distribution over all possible distances over the characterized range, thus denoted by a constant η. The denominator, $p(\{d_{Mi}\})$, is derived by marginalizing out d in the likelihood function:

$$p(\{d_{Mi}\}) = \int_d p(\{d_{Mi}\}|d)p(d)\mathrm{d}d. \tag{3.4}$$

The above integral is hard to calculate because of the varying standard deviation of the distance measurements, σ_{d_M}. However, the result of the integral can be taken as a normalizing constant and ensures that the integral of the posterior PDF, $p(d|\{d_{Mi}\})$, is equal to 1. Assume the normalizing constant is K, and $K = p(\{d_{Mi}\})$. Therefore, the posterior PDF of d given N measurements, $\{d_{Mi}\}$, has the following form:

$$\begin{aligned} p(d\ |\{d_{Mi}\}) &= \frac{\eta/K}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\mu - f(d)\right)^2}{2\sigma^2}\right) \\ &= \frac{\eta/K}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(kd+b-\mu)^2}{2\sigma^2}\right) \\ &= \frac{\eta/K}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(d-(\mu-b)/k)^2}{2(\sigma/k)^2}\right) \\ &= \frac{1}{\sqrt{2\pi\sigma_d^2}} \exp\left(-\frac{(d-\mu_d)^2}{2\sigma_d^2}\right), \end{aligned} \tag{3.5}$$

where $\mu_d = (\mu - b)/k = \frac{1}{kN}\sum_{i=1}^{N} d_{Mi} - \frac{b}{k}$ and $\sigma_d^2 = \left(\frac{\sigma}{k}\right)^2 = \frac{1}{k^2 N}\sigma_{d_M}^2$. The standard deviation σ_{d_M} will be interpolated from the characterized values as shown in Figure 3.16 according to the average of all measurements. Note that the polynomial function, $f(\cdot)$, can be considered as a linear function that maps the true distances from the domain in millimeters to that in centimeters, and its inverse function, $f^{-1}(\cdot)$, maps the measurements to the domain in millimeters. In addition, according to the results derived in Equation (3.5), using multiple repeated measurements over the same ground truth leads to a more reliable posterior distribution of d, as the variance, σ_d^2, is $1/N$ of the individual variance of the mapped distance measurements.

In order to simplify the calculations, all the distance measurements referred to later in this thesis are mapped to the same domain as the corresponding true distances with the unit of millimeter. Denoting the mapped distance measurements in millimeters as $\{d_{mi}\}$ with lowercase m, the parameters of the posterior

PDF are then rewritten as $\mu_d = \frac{1}{N}\sum_{i=1}^{N} d_{mi}$ and $\sigma_d^2 = \frac{1}{N}\sigma_{d_m}^2$, where $d_{mi} = f^{-1}(d_{Mi})$ and $\sigma_{d_m} = \frac{\mathrm{d}f^{-1}(x)}{\mathrm{d}x}\sigma_{d_M}$.

3.4.2.2 Servo Motor and Angular Measurements

Most angle-based localization methods found in the literature use bearing angles as the angular measurements, because the orientation of the sensors is often assumed to be known. However, the orientation of each localization sensor unit must be estimated most of the time as the unit is designed to be mounted itself on a ground mobile robot. Thus, the angular measurements are chosen as the relative bearing angle with respect to the reference frame of each sensor unit. The range of angle is defined as $[-\pi, \pi)$, and any angle outside of the range will be wrapped as $\theta' = \big((\theta + \pi) \bmod 2\pi\big) - \pi$, where θ is the original angle.

The characterization of the angular measurements is using a slightly different strategy from that of the distance measurements. The servo motor used on each implemented localization sensor unit has an encapsulated contactless absolute encoder, and it is assumed that the feedback from each encoder is accurate for controlling the servo motor. The 4096 steps generated by the encoders for each revolution are thus evenly mapped into the range of $[-\pi, \pi)$ by $\theta_{radian} = \frac{\pi}{2048} \cdot \theta_{steps}$. Therefore, the errors in the angular measurements of the implemented localization sensors are considered to be only related to the landmark detection and tracking algorithm, which is strongly related to the distance between the landmark and the localization sensor units. The angular measurements are then collected against a landmark placed in front of the sensor units at multiple known positions, and the characterized standard deviation of the angular measurements are later plotted against the distance between the sensor units and the landmark in Figure 3.20. As shown in Figure 3.18, the sensor units are symmetrically placed on each side of the

landmark at known positions aligned to the floor tiles, and the landmark is displaced every second floor tiles, or 609.6 mm, along the axis of symmetry of the two units.

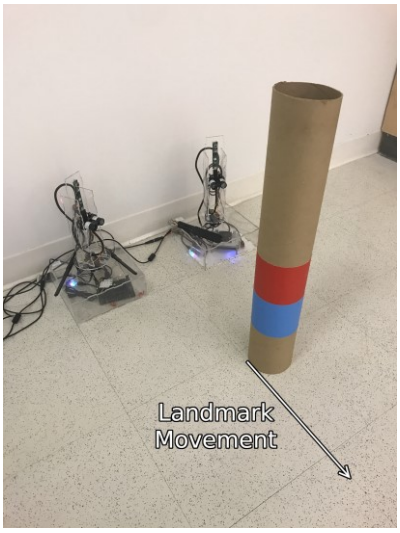
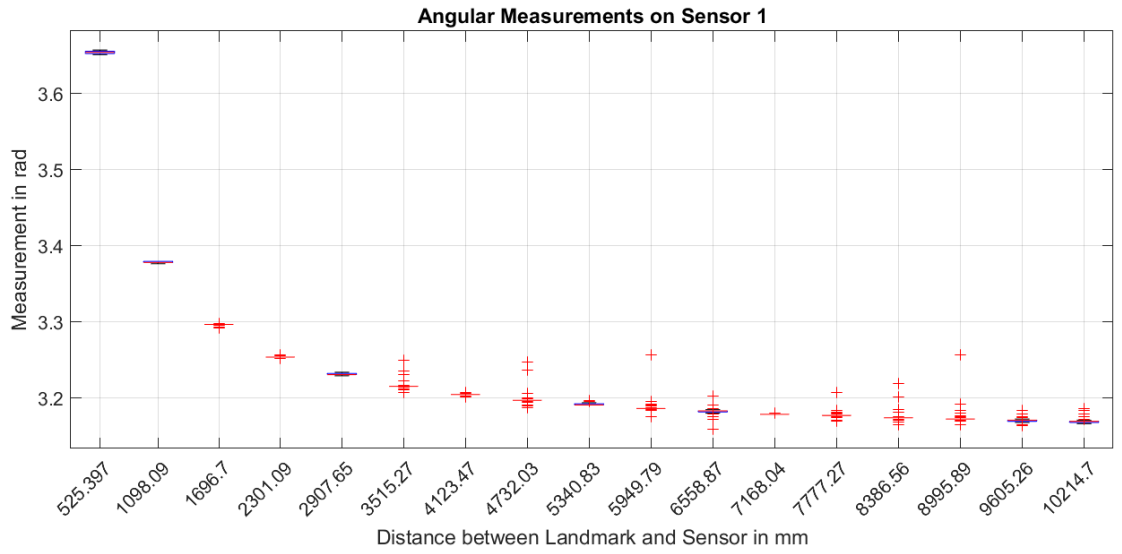
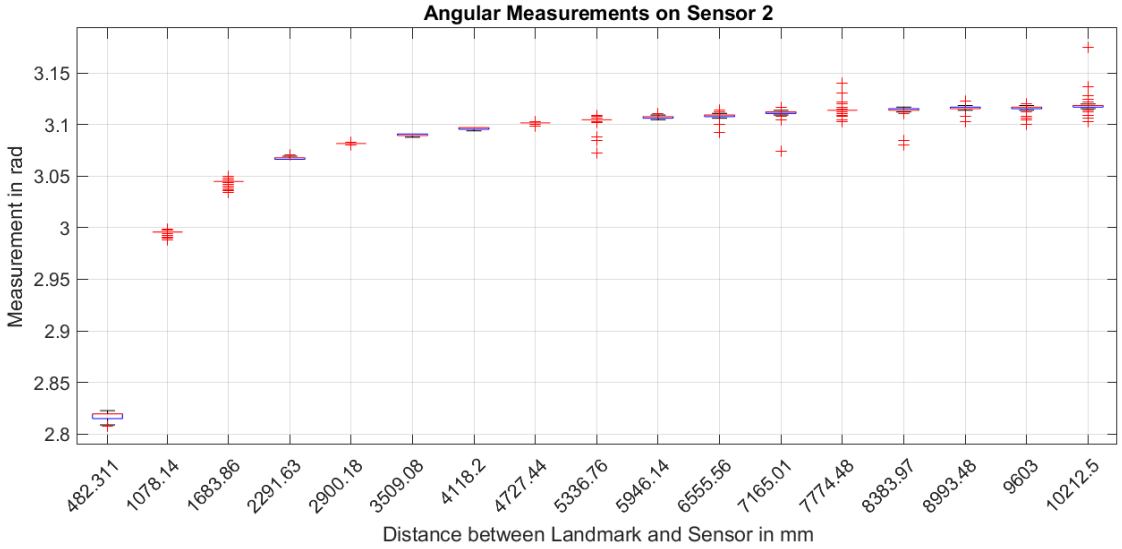


Figure 3.18 Angular measurement characterization setup using one landmark and the two implemented localization sensor units located symmetrically on each side of the landmark.

At each measured position of the landmark, 1000 angular measurements are collected from both sensor units. The collected measurements are shown as boxplots in Figure 3.19 at each measured position, arranged by the Euclidean distance between the landmark and each sensor unit. Since it is assumed that the readings from the servo encoders are reliable, angular measurements and corresponding standard deviations are mapped directly from steps into radians.



(a)



(b)

Figure 3.19 Boxplots of the angular measurements: (a) on *Sensor 1* (b) on *Sensor 2*.

Since the state feedback of the servo motor is considered accurate and the horizontal center of the camera has been calibrated in the previous subsection, the angular measurement error is also considered as Gaussian white noise. Noticeably, an offset term can be characterized by averaging on the angular measurement errors, which directly affects the zero reading of the servo motors. However, because the sensor units are manually placed according to the floors, errors exist in the position of the units and may lead to large angular measurement errors when the distance between the landmark and the sensor units is

small. Therefore, only the angular measurements that are collected when the distance is larger than 4000 mm are taken into consideration for calculating the offset term. With the offset term characterized, the zero reading of the servo motors on *Sensor 1* and *Sensor 2* respectively correspond to the 2037th and 2052nd steps, both of which are raw measurements of steps read directly from the servo encoders.

The standard deviation of the angular measurements at each measured position is given in Figure 3.20 for *Sensor 1* and *Sensor 2.* Beyond the nominal minimal range of operation of 500 mm, the standard deviation on the angular measurements does not show an obvious correlation with the distance between the landmark and the sensor units, but a slight tendency to increase can be observed from the figure.

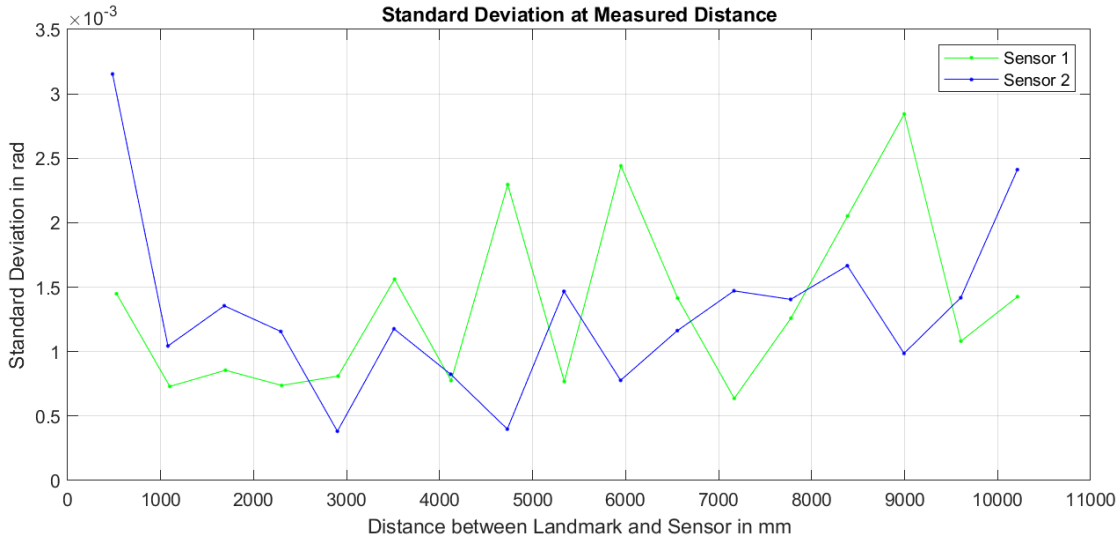


Figure 3.20 Standard deviation of angular measurements for both sensors.

Similar to the characterization of distance measurements, given N repeated angular measurements, $\{\theta_{mi}\}$, over the same ground truth, θ, the posterior PDF of the true angle, θ, is derived as:

$$p(\theta|\{\theta_{mi}\}) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left(-\frac{(\theta-\mu_\theta)^2}{2\sigma_\theta^2}\right), \tag{3.6}$$

where $\mu_\theta = \frac{1}{N}\sum_{i=1}^{N}\theta_{mi}$, $\sigma_\theta^2 = \frac{1}{N}\sigma_{\theta_m}^2$, and σ_{θ_m} is the standard deviation of θ_{mi}. Noticeably, using multiple repeated angular measurements can reduce the variance in the posterior PDF, σ_θ^2, by a ratio of $1/N$,

resulting in a more reliable posterior distribution of θ. Similarly to the characterization of distance measurements, the standard deviation of angular measurements, σ_{θ_m}, is interpolated from the characterized values shown in Figure 3.20 according to the distance between the sensor unit and the landmark.

In the implementation, the standard deviation, σ_{θ_m}, is interpolated by the estimated distance between the sensor unit and a landmark, against which the measurements are taken. When calculating the error term, $(\theta - \mu_\theta)$, the results are always wrapped in the range of $[-\pi, \pi)$ as described earlier in this section. For instance, subtracting $-\frac{3}{4}\pi$ from $\frac{3}{4}\pi$, the result will be $-\frac{1}{2}\pi$ instead of $\frac{3}{2}\pi$. The implemented localization sensor utilizes the aforementioned angle-wrapping equation and wraps the result of the ordinary subtraction within $[-\pi, \pi)$.

The angular standard deviation is important for the implemented localization sensor units and could serve as a metric for the reliability of the landmark detection and tracking algorithm. It should be within a reliable range for the distance measurements to be collected by the LiDARs, which emit laser and measure distances utilizing the reflection of laser on the measured target landmark. If the angular measurement errors are too large, the laser emitted by the LiDARs may completely miss the cylindrical landmark used on the sensor units. The beam divergence of the laser rangefinder is 0.008 rad and the angle of view of the cylindrical landmark when placed at 10 meters is about $2 \cdot \operatorname{atan}\frac{0.053\ \text{m}}{10\ \text{m}} \approx 0.011\ \text{rad}$, where 0.053 m represents the radius of the cylindrical landmark. Since the highest standard deviation shown in Figure 3.20 is about 0.003 rad, the probability of the angular measurements to have an error within −0.003 rad to 0.003 rad is about 68% and the probability is about 95% when the error range doubles. Therefore, it is highly likely that the laser emitted by a LiDAR will be reflected, at least partially, by a cylindrical landmark placed within the working range of 10 meters to support the localization process.

3.4.3 Localization Sensor Calibration

A calibration process for the localization sensor takes place after a localization sensor unit is fully assembled and installed on a working platform, such as a mobile robotic agent. The errors introduced during the procedure of installation often lead to a discrepancy between the actual and desired poses of the unit. This supplementary procedure is intended to overcome such errors and best estimate the pose of the installed localization sensor unit.

The calibration procedure makes use of the iterative LS estimation which is widely used in the field of localization. By fixing the pose of the localization sensor and taking measurements to a landmark which is placed manually at a series of known positions with respect to a given world frame, the pose of the sensor can be estimated with respect to the same frame given an initial close guess. The observation model of the localization sensor is given by:

$$\begin{bmatrix}\rho_i \\ \theta_i\end{bmatrix} = \begin{bmatrix}\sqrt{(x_i-u)^2+(y_i-v)^2} \\ \text{atan2}(y_i-v/x_i-u)-\phi\end{bmatrix}, i = 1,2,\dots,N \tag{3.7}$$

where $[\rho_i, \theta_i]^T$ are the distance and angular measurements to the landmark at the i^{th} position, denoted by $\boldsymbol{x}_i = [x_i, y_i]^T$, and $\boldsymbol{s} = [u, v, \phi]^T$ is the pose of the sensor. Let the observation model to the landmark at the i^{th} position be $[\rho_i, \theta_i]^T = \boldsymbol{f}_i(\boldsymbol{s})$, and the model can be linearized with Taylor-Series expansion at the vicinity of a guess of the pose of the sensor, $\boldsymbol{s}_g = \left[u_g, v_g, \phi_g\right]^T$, with only first-order derivatives:

$$\begin{bmatrix}\rho_i \\ \theta_i\end{bmatrix} \approx \boldsymbol{f}_i(\boldsymbol{s}_g) + \boldsymbol{J}_i(\boldsymbol{s}_g)\begin{bmatrix}u-u_g \\ v-v_g \\ \phi-\phi_g\end{bmatrix}, \tag{3.8}$$

where $\boldsymbol{J}_i(\boldsymbol{s}_g)$ is the Jacobian matrix of $\boldsymbol{f}_i(\boldsymbol{s})$ at $\boldsymbol{s}_g$, which is expressed as:

$$\boldsymbol{J}_i(\boldsymbol{s}_g) = \begin{bmatrix} \frac{u_g - x_i}{\sqrt{(x_i-u_g)^2+(y_i-v_g)^2}} & \frac{v_g - y_i}{\sqrt{(x_i-u_g)^2+(y_i-v_g)^2}} & 0 \\ \frac{y_i - v_g}{(x_i-u_g)^2+(y_i-v_g)^2} & \frac{u_g - x_i}{(x_i-u_g)^2+(y_i-v_g)^2} & -1 \end{bmatrix}. \tag{3.9}$$

Stacking up the equations for all N positions, Equation (3.8) can be rewritten in matrix form:

$$\boldsymbol{r} = \boldsymbol{r}_g + \boldsymbol{A}\boldsymbol{\delta}_{\boldsymbol{s}}, \tag{3.10}$$

where $\boldsymbol{r} = [\rho_1, \theta_1, \dots, \rho_N, \theta_N]^T$, $\boldsymbol{r}_g = \left[\boldsymbol{f}_1^T(\boldsymbol{s}_g), \dots, \boldsymbol{f}_2^T(\boldsymbol{s}_g)\right]^T$, $\boldsymbol{A} = \left[\boldsymbol{J}_1^T(\boldsymbol{s}_g), \dots, \boldsymbol{J}_N^T(\boldsymbol{s}_g)\right]^T$, and $\boldsymbol{\delta}_{\boldsymbol{s}} = \left[u - u_g, v - v_g, \phi - \phi_g\right]^T$. When the measurement errors are taken into consideration, the above equation is rewritten as:

$$\boldsymbol{e} = \boldsymbol{r} - \boldsymbol{r}_g - \boldsymbol{A}\boldsymbol{\delta}_{\boldsymbol{s}}. \tag{3.11}$$

Since the measurement errors are modeled as Gaussian white noise as described in Section 3.4.2, the covariance matrix of $\boldsymbol{e}$ is a diagonal matrix given by $\boldsymbol{\Sigma} = diag(\sigma_{\rho 1}^2, \sigma_{\theta 1}^2, \dots, \sigma_{\rho N}^2, \sigma_{\theta N}^2)$. The iterative LS result is then given by minimizing a cost function, $Q(\boldsymbol{s}_g) = \boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e}$, which yields $\boldsymbol{s}_g' = \boldsymbol{s}_g + (\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{r} - \boldsymbol{r}_g)$. With sufficient measurements, the iterative LS algorithm usually converges after about five iterations, but ten iterations are used in the implementation.

The calibration process estimates the pose of a localization sensor unit by utilizing the measurements made by the unit towards a landmark located in a given, accessible reference frame. The estimation is performed in an offline manner and is aimed at estimating the pose of the sensor unit after installation but before localization operation. It is frequently used in the experiments in Chapter 5 to provide approximated poses of the localization sensor units, because of a lack of alternative means for acquiring the accurate poses of the units with external sensor devices. If a localization sensor is installed on a working platform such as a mobile robot with a known relative pose, the calibration process can also estimate the pose of the platform offline.

3.5 Summary

In this chapter, the design and implementation of an innovative localization sensor are presented in detail. The conceptual design is first discussed, followed by information about the physical implementation and data acquisition process developed for the proposed localization sensor to detect and track a color-coded landmark that supports pose estimation on ground mobile robots. Finally, the sensing components embedded in each localization sensor unit are configured, and the measurements made by each individual component are characterized as a Gaussian distribution. The implemented localization sensor units will be characterized as a whole in Chapter 5, where the performance of the units will also be evaluated collaboratively.

CHAPTER 4 - DERIVATION OF COLLABORATIVE POSE ESTIMATION

The proposed localization sensor is designed for providing pose estimation for ground mobile robots. When using the localization sensor, each robot should have an instance of the sensor installed onboard with a known relative pose to the robot's structure. Given that, since the pose of a robot can be readily estimated by transforming the estimated pose of the sensor unit installed on the robot, only the pose estimation of the proposed localization sensor is discussed in this chapter.

As demonstrated in Chapter 3, complete pose estimation of the localization sensor reaches higher accuracy when collaboration among multiple sensor units is involved. As such, the pose of each localization sensor unit is collaboratively estimated based on the measured relative locations among the sensor units, which are acquired by relative distance and angular measurements. Due to the non-linear relationship of the position estimates and the raw measurements, merging information from multiple sensor units and other sources can be inefficient and time-consuming. Therefore, a linearized model of the collaborative position and pose estimation of the proposed localization sensor is developed in this chapter.

Since measurements are collected against the color-coded landmarks installed on the center of each localization sensor unit, the estimated position of a sensor unit is considered the same as that of its landmark. Therefore, the position estimation of the color-coded landmarks is first obtained as part of the collaborative pose estimation of the localization sensor. The derivations for the landmark position estimation using one and multiple localization sensor units are documented in Sections 4.1 and 4.2, respectively. The pose estimation of the localization sensor is then derived based on the developed models for the landmark position estimation. An inter-calibration process is introduced for the pose estimation in Section 4.3, where two localization sensor units measure against each other. In the inter-calibration process, the estimated position of a sensor unit using the model developed in Section 4.1 can be improved and its orientation can be estimated, given the absolute pose of the other sensor with respect to a chosen

world frame. With more localization taking part in this process, the pose of a localization sensor unit can be estimated collaboratively as described in Section 4.4.

Some naming conventions apply throughout this chapter. Positions are denoted as 2-D vectors, where the position of a landmark is denoted by $\boldsymbol{x} = [x, y]^T$ and the position of a sensor unit is denoted by $\boldsymbol{s} = [u, v]^T$. The orientation of a sensor unit, ϕ, sometimes is used additionally in the square brackets, composing a pose vector, as $[\boldsymbol{s}^T, \phi]^T$. Relative distance and angle from a sensor unit to a landmark are denoted by $\boldsymbol{r} = [\rho, \theta]^T$, and if a subscript m is attached, the vector stands for the measured values, otherwise, it is the actual value. Similarly, the positions, $\boldsymbol{x}$ and $\boldsymbol{s}$, are sometimes attached with a subscript m, denoting corresponding estimated values. Quite often, the variables are attached with an integer in their subscripts, which indicates the variables are related to the sensor unit associated with that number. On some occasions, two integers are present in the subscripts, indicating that the variable is related to both sensor units. The error, mean, variance, and covariance matrix are denoted by ϵ, μ, σ, and $\boldsymbol{\Sigma}$, respectively with the corresponding variable as a subscript. And most importantly, bold letters indicate vectors or matrices, while non-bold letters are scalars, except for operators such as Σ and Π, which are also non-italic.

4.1 Individual Landmark Position Estimation

The individual landmark position estimation is fundamental to the collaborative pose estimation process using the proposed localization sensor. During this process, each sensor unit takes measurements to a landmark and estimates the position of the landmark without any assistance from other units. The distance and angular measurements to a landmark natively form polar coordinates with respect to the local frame of each sensor unit and the estimation results can be obtained by transforming the coordinates to a designated world frame, assuming that the pose of the sensor unit is known. However, the distance between the sensor unit and the landmark may not be measured accurately due to measurement noise and

the shape of the landmark. In order to better estimate the position of the target landmark, the error on distance measurements must be properly modeled.

A top view of a localization sensor unit located at S measuring a cylindrical landmark positioned at point P is shown in Figure 4.1, where R is the radius of the landmark. The distance measurement made by the sensor unit is denoted as d_m.

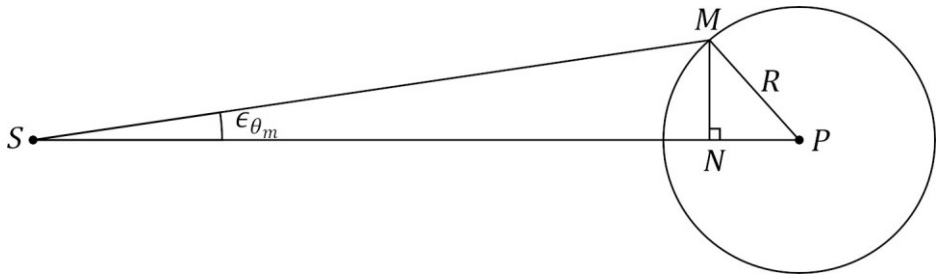


Figure 4.1 A top view of a sensor measuring the distance to a landmark.

The sensor unit is supposed to measure towards the point P, but it is pointing towards the point M due to the existence of noise when controlling its servo motor. Noticeably, the measured distance, $\|\overline{SM}\|$, is independent of the orientation of the landmark because of its circular cross section. Assume that the sensor unit is pointing towards the landmark with an error, ϵ_{θ_m}, and the distance measurement has an error, ϵ_{d_m}, such that:

$$d_m = \|\overline{SM}\| + \epsilon_{d_m}, \tag{4.1}$$

As characterized in Section 3.4.2, both the distance and angular errors are considered additive Gaussian white noise, as expressed below:

$$\epsilon_{d_m} \sim N(0, \sigma_{d_m}^2),$$
$$\epsilon_{\theta_m} \sim N(0, \sigma_{\theta_m}^2). \tag{4.2}$$

In order to derive the actual distance between the sensor unit and the landmark center, $\|\overline{SP}\|$, an auxiliary point N is added, where the line segment $\overline{MN}$ is perpendicular to $\overline{SP}$. The following equations can be derived:

$$\begin{aligned} \|\overline{SN}\| &= \|\overline{SM}\| \cos \epsilon_{\theta_m} = \left(d_m - \epsilon_{d_m}\right) \cos \epsilon_{\theta_m}, \\ \|\overline{MN}\| &= \|\overline{SM}\| \sin \epsilon_{\theta_m} = \left(d_m - \epsilon_{d_m}\right) \sin \epsilon_{\theta_m}, \\ \text{and } \|\overline{NP}\| &= \left\| \sqrt{R^2 - \|\overline{MN}\|^2} \right\| = \left\| \sqrt{R^2 - \left(d_m - \epsilon_{d_m}\right)^2 \sin^2 \epsilon_{\theta_m}} \right\|. \end{aligned} \tag{4.3}$$

The length of $\|\overline{SP}\|$ is then expressed as:

$$\|\overline{SP}\| = \|\overline{SN}\| + \|\overline{NP}\| = \left(d_m - \epsilon_{d_m}\right) \cos \epsilon_{\theta_m} + \left\| \sqrt{R^2 - \left(d_m - \epsilon_{d_m}\right)^2 \sin^2 \epsilon_{\theta_m}} \right\|. \tag{4.4}$$

The above equation is nonlinear with respect to the errors. Since the angular error, ϵ_{θ_m}, is additive Gaussian white noise, $\|\overline{SP}\|$ is approximated by applying the Taylor-Series expansion at the vicinity of the expectation of ϵ_{θ_m}. Let $\|\overline{SP}\| = g(\epsilon_{\theta_m})$, where $g(x) = \left(d_m - \epsilon_{d_m}\right) \cos x + \sqrt{R^2 - \left(d_m - \epsilon_{d_m}\right)^2 \sin^2 x}$, with its derivative being $\frac{dg(x)}{dx} = -\left(d_m - \epsilon_{d_m}\right) \sin x - \left(d_m - \epsilon_{d_m}\right)^2 \sin x \cos x / \sqrt{R^2 - \left(d_m - \epsilon_{d_m}\right)^2 \sin^2 x}$, and thus $g(0) = d_m - \epsilon_{d_m} + R$ and $\frac{\mathrm{d}g(x)}{\mathrm{d}x}\Big|_{x=0} = 0$. The linearized form of Equation (4.4) is given by:

$$\begin{aligned} \|\overline{SP}\| &\approx g(0) + \frac{\mathrm{d}g(x)}{\mathrm{d}x}\Big|_{x=0} \cdot \epsilon_{\theta_m} \\ &= d_m - \epsilon_{d_m} + R. \end{aligned} \tag{4.5}$$

Therefore, the measured distance between the sensor unit and the landmark can be approximated as the sum of the distance measurement, d_m, and the radius of the landmark, R. Let ρ_m be the measured length

of $\overline{SP}$, then $\rho_m = d_m + R$ and its ground truth is given by $\rho = \|\overline{SP}\| = d_m - \epsilon_{d_m} + R = \|\overline{SM}\| + R$, based on Equation (4.5). Since both d_m and ρ_m have the same additive Gaussian white noise, the variances of d_m and ρ_m are the same, namely $\sigma_{\rho_m}^2 = \sigma_{d_m}^2$. ρ_m then has the following Gaussian distribution given the abovementioned information:

$$\rho_m \sim N(\rho, \sigma_{\rho_m}^2). \tag{4.6}$$

The approximated distance measurements to the landmark can then be taken as random variables independent of the angular measurements. In the remaining part of this thesis, distance measurements to a landmark are referred to as the approximated distance measurements, ρ_m.

Let the local polar coordinates of the landmark be $\boldsymbol{r} = [\rho, \theta]^T$ with respect to the local frame of the sensor unit, and $\boldsymbol{r}$ is a 2-D Gaussian variable with mean $\boldsymbol{r}_m = [\rho_m, \theta_m]^T$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}_m} = \begin{bmatrix} \sigma_{\rho_m}^2 & 0 \\ 0 & \sigma_{\theta_m}^2 \end{bmatrix}$, given that ρ and θ are independent variables. Thus, the polar coordinates have the following distribution:

$$\begin{gathered} \boldsymbol{r} \sim N(\boldsymbol{r}_m, \boldsymbol{\Sigma}_{\boldsymbol{r}_m}), \\ p(\boldsymbol{r}|\boldsymbol{r}_m) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{\boldsymbol{r}_m}|}} \exp\left(-\frac{1}{2}(\boldsymbol{r} - \boldsymbol{r}_m)^T \boldsymbol{\Sigma}_{\boldsymbol{r}_m}^{-1}(\boldsymbol{r} - \boldsymbol{r}_m)\right). \end{gathered} \tag{4.7}$$

Representing the position of the landmark in polar coordinates provides a concise and convenient form, however, this form is unusable when merging information from other sensor units. A Cartesian representation is needed, which requires the polar coordinates to be transformed to a Cartesian world frame. An illustration of the transformation is shown in Figure 4.2, where $[x_s, y_s]^T$ is the Cartesian coordinates of the landmark P in the local sensor frame, $X_s O_s Y_s$, and $[x, y]^T$ is the Cartesian coordinates of the landmark in the world frame, XOY.

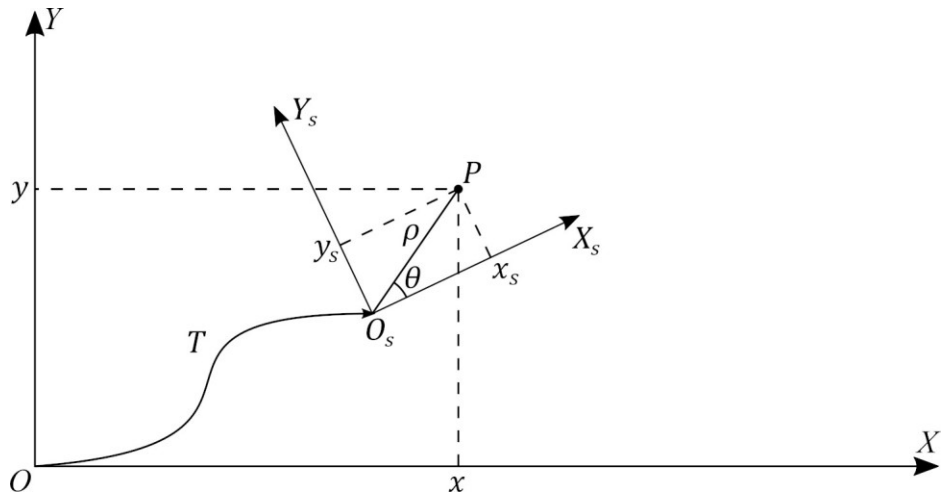


Figure 4.2 Transformation from $X_sO_sY_s$ to XOY.

The transformation from the polar coordinate system to the local Cartesian system is expressed as:

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} \rho\cos\theta \\ \rho\sin\theta \end{bmatrix}. \tag{4.8}$$

Assume that the position of the sensor unit is $\boldsymbol{s} = [u, v]^T$ and its orientation is ϕ, the homogeneous transfer matrix from the sensor frame, $X_sO_sY_s$, to the global frame, XOY, is given by:

$$\boldsymbol{T} = \begin{bmatrix} cos\,\phi & -sin\,\phi & u \\ sin\phi & cos\phi & v \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.9}$$

and the coordinates of the landmark in the global frame is given by the transformation:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \boldsymbol{T}\begin{bmatrix} x_L \\ y_L \\ 1 \end{bmatrix} = \begin{bmatrix} \rho\cos(\theta+\phi)+u \\ \rho\sin(\theta+\phi)+v \\ 1 \end{bmatrix}. \tag{4.10}$$

When written in a more concise form,

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \rho\cos(\theta+\phi)+u \\ \rho\sin(\theta+\phi)+v \end{bmatrix}. \tag{4.11}$$

Noticeably, the transformation from the polar coordinates $\boldsymbol{r}$ to the Cartesian coordinates $\boldsymbol{x}$ is nonlinear.

An approximation of the distribution of $\boldsymbol{x}$ can be acquired by analyzing the probability in the transformation. Let $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{r})$ and $[\boldsymbol{r}, \boldsymbol{r} + \boldsymbol{\delta_r})$ be an infinitesimal interval, where $\boldsymbol{\delta_r} = \left[\delta_\rho, \delta_\theta\right]^T$, the function $\boldsymbol{f}(\boldsymbol{r})$ then transforms the probability on the interval in the polar coordinates system to the corresponding interval in the Cartesian coordinates system, $\left[\boldsymbol{f}(\boldsymbol{r}), \boldsymbol{f}(\boldsymbol{r} + \boldsymbol{\delta_r})\right)$. The latter could be linearized by applying Taylor-Series expansion with only the first-order derivative term at the vicinity of $\boldsymbol{r}$,

$$\boldsymbol{f}(\boldsymbol{r} + \boldsymbol{\delta_r}) \approx \boldsymbol{f}(\boldsymbol{r}) + \boldsymbol{J}(\boldsymbol{r})\boldsymbol{\delta_r}, \tag{4.12}$$

where $\boldsymbol{J}(\boldsymbol{r})$ is the Jacobian matrix of the transformation function $\boldsymbol{f}(\boldsymbol{r})$, and

$$\boldsymbol{J}(\boldsymbol{r}) = \begin{bmatrix} \cos(\theta + \phi) & -\rho \sin(\theta + \phi) \\ \sin(\theta + \phi) & \rho \cos(\theta + \phi) \end{bmatrix}. \tag{4.13}$$

The area of the transformed interval could then be approximated as $|\boldsymbol{J}(\boldsymbol{r})|\delta_\rho\delta_\theta$, and since the probability is the same over the original and transformed intervals, it follows that:

$$p(\boldsymbol{r}|\boldsymbol{r}_m)\delta_\rho\delta_\theta = p(\boldsymbol{x}|\boldsymbol{x}_m)|\boldsymbol{J}(\boldsymbol{r})|\delta_\rho\delta_\theta, \tag{4.14}$$

where $\boldsymbol{x}_m = \boldsymbol{f}(\boldsymbol{r}_m)$ is the estimated position of the landmark, and $|\boldsymbol{J}(\boldsymbol{r})| = \rho$. Therefore, the PDF of $\boldsymbol{x}$ given $\boldsymbol{x}_m$ is expressed as:

$$p(\boldsymbol{x}|\boldsymbol{x}_m) = \frac{1}{\rho} p(\boldsymbol{r}|\boldsymbol{r}_m). \tag{4.15}$$

This approximation continuously utilizes the linearization of the transformation over an infinitesimal area at point $\boldsymbol{r}$, but it is still not a linear form that could be merged with other information. In order to better relate the two variables, the linearization is relaxed by assuming that the measurement errors are small most of the time, indicating that the errors have small variances. The linearization then can be fixed at the vicinity of the measurements, $\boldsymbol{r}_m$. Thus,

$$\boldsymbol{x} \approx \boldsymbol{f}(\boldsymbol{r}_m) + \boldsymbol{J}(\boldsymbol{r}_m)(\boldsymbol{r} - \boldsymbol{r}_m) = \boldsymbol{x}_m + \boldsymbol{J}(\boldsymbol{r}_m)\boldsymbol{\delta}_{\boldsymbol{r}}, \tag{4.16}$$

where $\boldsymbol{x}_m = \begin{bmatrix} \rho_m \cos(\theta_m + \phi) + u \\ \rho_m \sin(\theta_m + \phi) + v \end{bmatrix}$, $\boldsymbol{J}(\boldsymbol{r}_m) = \begin{bmatrix} \cos(\theta_m + \phi) & -\rho_m \sin(\theta_m + \phi) \\ \sin(\theta_m + \phi) & \rho_m \cos(\theta_m + \phi) \end{bmatrix}$, and $\boldsymbol{\delta}_{\boldsymbol{r}} = \left[-\epsilon_\rho, -\epsilon_\theta\right]^T$. Following the above linearization, the position $\boldsymbol{x}$ has also a 2-D Gaussian distribution. The mean and the covariance matrix of $\boldsymbol{x}$ is given by:

$$\begin{gathered} \boldsymbol{\mu}_{\boldsymbol{x}} = E\left[\boldsymbol{x}_m + \boldsymbol{J}(\boldsymbol{r}_m)\boldsymbol{\delta}_{\boldsymbol{r}_m}\right] = \boldsymbol{x}_m, \\ \boldsymbol{\Sigma}_{\boldsymbol{x}_m} = E[(\boldsymbol{x} - \boldsymbol{x}_m)(\boldsymbol{x} - \boldsymbol{x}_m)^T] \\ = E[\boldsymbol{J}(\boldsymbol{r}_m)(\boldsymbol{r} - \boldsymbol{r}_m)(\boldsymbol{r} - \boldsymbol{r}_m)^T \boldsymbol{J}^T(\boldsymbol{r}_m)] \\ = \boldsymbol{J}(\boldsymbol{r}_m)\boldsymbol{\Sigma}_{\boldsymbol{r}_m}\boldsymbol{J}^T(\boldsymbol{r}_m). \end{gathered} \tag{4.17}$$

Therefore,

$$\begin{gathered} \boldsymbol{x} \sim N(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{\Sigma}_{\boldsymbol{x}_m}) \text{ or } \boldsymbol{x} \sim N(\boldsymbol{x}_m, \boldsymbol{J}(\boldsymbol{r}_m)\boldsymbol{\Sigma}_{\boldsymbol{r}_m}\boldsymbol{J}^T(\boldsymbol{r}_m)), \\ p(\boldsymbol{x}|\boldsymbol{x}_m) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{\boldsymbol{x}_m}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_m)^T \boldsymbol{\Sigma}_{\boldsymbol{x}_m}^{-1}(\boldsymbol{x} - \boldsymbol{x}_m)\right), \end{gathered} \tag{4.18}$$

where $\boldsymbol{x}_m = \begin{bmatrix} \rho_m \cos(\theta_m + \phi) + u \\ \rho_m \sin(\theta_m + \phi) + v \end{bmatrix}$, $\boldsymbol{\Sigma}_{\boldsymbol{x}_m} = \boldsymbol{J}(\boldsymbol{r}_m)\boldsymbol{\Sigma}_{\boldsymbol{r}_m}\boldsymbol{J}^T(\boldsymbol{r}_m)$, $\boldsymbol{\Sigma}_{\boldsymbol{r}_m} = \begin{bmatrix} \sigma_{\rho_m}^2 & 0 \\ 0 & \sigma_{\theta_m}^2 \end{bmatrix}$, and $\boldsymbol{J}(\boldsymbol{r}_m) = \begin{bmatrix} \cos(\theta_m + \phi) & -\rho_m \sin(\theta_m + \phi) \\ \sin(\theta_m + \phi) & \rho_m \cos(\theta_m + \phi) \end{bmatrix}$. By maximizing on the conditional PDF presented in the above equation, a deterministic result of the landmark position is given as $\boldsymbol{x}_m$, which is also the measured position of the landmark.

Equation (4.18) is derived based on the assumption that the pose of the sensor unit at S is known, which is unfortunately not the case in general applications. The pose, however, can be estimated by processes such as being measured by another sensor unit whose pose is known, where often the estimated pose comes as a distribution due to the existence of noise. Assume that the position and the orientation of the sensor unit are estimated as $\boldsymbol{s}_m = [u_m, v_m]^T$ and ϕ_m, and without loss of generalization, both have

Gaussian distributions and are independent of the measurements made by the sensor unit, which are given by:

$$\boldsymbol{s} \sim N(\boldsymbol{s}_m, \boldsymbol{\Sigma}_{\boldsymbol{s}_m}),$$
$$\phi \sim N(\phi_m, \sigma_{\phi_m}^2).$$
$$\text{And } p(\boldsymbol{s}|\boldsymbol{s}_m) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{\boldsymbol{s}_m}|}} \exp\left(-\frac{1}{2}(\boldsymbol{s}-\boldsymbol{s}_m)^T \boldsymbol{\Sigma}_{\boldsymbol{s}_m}^{-1}(\boldsymbol{s}-\boldsymbol{s}_m)\right), \tag{4.19}$$
$$p(\phi|\phi_m) = \frac{1}{\sqrt{2\pi}\sigma_{\phi_m}} \exp\left(-\frac{1}{2\sigma_{\phi_m}^2}(\phi-\phi_m)^2\right).$$

To simplify the linearization process, the bearing from the sensor unit to the landmark will be used in the following equations and it is given by $\beta = \phi + \theta$. Since ϕ_m and θ_m are independent and both have Gaussian distributions, the PDF of β given ϕ_m and θ_m can be expressed as:

$$\begin{aligned}
p(\beta|\phi_m,\theta_m) &= \int p(\beta-\phi|\theta_m)p(\phi|\phi_m)\mathrm{d}\phi \\
&= \frac{1}{2\pi\sigma_{\phi_m}\sigma_{\theta_m}} \int \exp\left(-\frac{(\beta-\phi-\theta_m)^2}{2\sigma_{\theta_m}^2} - \frac{(\phi-\phi_m)^2}{2\sigma_{\phi_m}^2}\right)\mathrm{d}\phi \\
&= \frac{1}{2\pi\sigma_{\phi_m}\sigma_{\theta_m}} \int \exp\left(-\frac{\sigma_{\phi_m}^2(\beta-\phi-\theta_m)^2+\sigma_{\theta_m}^2(\phi-\phi_m)^2}{2\sigma_{\theta_m}^2\sigma_{\phi_m}^2}\right)\mathrm{d}\phi \\
&= \frac{\exp\left(-\frac{(\beta-\theta_m-\phi_m)^2}{2\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}\right)}{2\pi\sigma_{\phi_m}\sigma_{\theta_m}} \int \exp\left(-\frac{\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}{2\sigma_{\theta_m}^2\sigma_{\phi_m}^2}\left(\phi - \frac{\sigma_{\phi_m}^2(\beta-\theta_m)+\sigma_{\theta_m}^2\phi_m}{\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}\right)^2\right)\mathrm{d}\phi \\
&= \frac{1}{2\pi\sigma_{\phi_m}\sigma_{\theta_m}} \sqrt{\frac{2\pi\sigma_{\theta_m}^2\sigma_{\phi_m}^2}{\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}} \exp\left(-\frac{(\beta-\theta_m-\phi_m)^2}{2\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}\right) \\
&= \frac{1}{\sqrt{2\pi\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}} \exp\left(-\frac{(\beta-\theta_m-\phi_m)^2}{2\left(\sigma_{\theta_m}^2+\sigma_{\phi_m}^2\right)}\right).
\end{aligned} \tag{4.20}$$

Therefore, the bearing of the landmark, β, has also a Gaussian distribution, $\beta \sim N(\theta_m + \phi_m, \sigma_{\theta_m}^2 + \sigma_{\phi_m}^2)$. The transformation from Equation (4.11) could then be rewritten as:

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \rho\cos\beta + u \\ \rho\sin\beta + v \end{bmatrix}. \tag{4.21}$$

Similar to the previous linearization, let the transformation function be $\boldsymbol{x} = \boldsymbol{g}([\rho, \beta]^T)$, and its Jacobian matrix is $\boldsymbol{J_g}([\rho, \beta]^T) = \begin{bmatrix} \cos\beta & -\rho\sin\beta \\ \sin\beta & \rho\cos\beta \end{bmatrix}$. The above equation can then be linearized at the vicinity of measured values, $[\rho_m, \beta_m]$, by Taylor-Series expansion as:

$$\begin{aligned} \boldsymbol{x} &\approx \boldsymbol{g}([\rho_m, \beta_m]^T) + \boldsymbol{J_g}([\rho_m, \beta_m]^T)\begin{bmatrix} \rho - \rho_m \\ \beta - \beta_m \end{bmatrix} \\ &= \begin{bmatrix} \rho_m\, cos\,\beta_m + u \\ \rho_m\, sin\,\beta_m + v \end{bmatrix} + \boldsymbol{J_g}([\rho_m, \beta_m]^T)\begin{bmatrix} \rho - \rho_m \\ \beta - \beta_m \end{bmatrix}. \end{aligned} \tag{4.22}$$

Given the position of the sensor unit, $\boldsymbol{s}$, the conditional PDF of $\boldsymbol{x}$ is given by:

$$p(\boldsymbol{x}|\boldsymbol{x}_{m|s}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m|s}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_{m|s})^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x} - \boldsymbol{x}_{m|s})\right), \tag{4.23}$$

where $\boldsymbol{x}_{m|s}$ is the estimated position of the landmark given the position of the sensor unit, $\boldsymbol{s}$, and $\boldsymbol{\Sigma}_{x_{m|s}}$ is the corresponding covariance matrix:

$$\begin{aligned} \boldsymbol{x}_{m|s} &= \boldsymbol{g}([\rho_m, \beta_m]^T) = \begin{bmatrix} \rho_m\, cos\,\beta_m + u \\ \rho_m\, sin\,\beta_m + v \end{bmatrix} \\ \boldsymbol{\Sigma}_{x_{m|s}} &= E\left[(\boldsymbol{x} - \boldsymbol{x}_{m|s})(\boldsymbol{x} - \boldsymbol{x}_{m|s})^T\right] \\ &= E\left[\boldsymbol{J_g}([\rho_m, \beta_m]^T)\begin{bmatrix} \rho - \rho_m \\ \beta - \beta_m \end{bmatrix}\begin{bmatrix} \rho - \rho_m \\ \beta - \beta_m \end{bmatrix}^T \boldsymbol{J_g}^T([\rho_m, \beta_m]^T)\right] \\ &= \boldsymbol{J_g}([\rho_m, \beta_m]^T)\begin{bmatrix} \sigma_{\rho_m}^2 & 0 \\ 0 & \sigma_{\beta_m}^2 \end{bmatrix}\boldsymbol{J_g}^T([\rho_m, \beta_m]^T). \end{aligned} \tag{4.24}$$

The subscripts of $\boldsymbol{\Sigma}_{x_{m|s}}$ and $\boldsymbol{x}_{m|s}$ indicate that they are acquired based on the condition of given $\boldsymbol{s}$, which is indeed considered as an unknown random variable. Let $\boldsymbol{x}_\Delta = \boldsymbol{x}_{m|s} - \boldsymbol{s}$, and $\boldsymbol{x}_\Delta$ is not related to the position of the sensor unit, $\boldsymbol{s}$. Then, the conditional PDF of Equation (4.23) can be rewritten as:

$$p(\boldsymbol{x}|\boldsymbol{x}_\Delta, \boldsymbol{s}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m|s}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\Delta - \boldsymbol{s})^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\Delta - \boldsymbol{s})\right). \tag{4.25}$$

The condition on $\boldsymbol{s}$ can be marginalized out by:

$$p(\boldsymbol{x}|\boldsymbol{x}_\Delta) = \int p(\boldsymbol{x}|\boldsymbol{x}_\Delta, \boldsymbol{s}) p(\boldsymbol{s}) \mathrm{d}\boldsymbol{s}. \tag{4.26}$$

Noticeably, the first conditional PDF could be written as a function as $f_1(\boldsymbol{x} - \boldsymbol{s})$, where $f_1(\boldsymbol{x})$ is given by:

$$f_1(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m|s}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\Delta)^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\Delta)\right). \tag{4.27}$$

The second PDF known from Equation (4.19) can be written as:

$$f_2(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{s_m}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{s}_m)^T \boldsymbol{\Sigma}_{s_m}^{-1}(\boldsymbol{x} - \boldsymbol{s}_m)\right). \tag{4.28}$$

Therefore, Equation (4.26) can be expressed by the two functions as:

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{x}_\Delta) &= \int f_1(\boldsymbol{x} - \boldsymbol{s}) f_2(\boldsymbol{s}) \mathrm{d}\boldsymbol{s} \\ &= f_1(\boldsymbol{x}) * f_2(\boldsymbol{x}). \end{aligned} \tag{4.29}$$

The above equation can be easily solved by applying the convolution theorem:

$$p(\boldsymbol{x}|\boldsymbol{x}_\Delta) = \mathcal{F}^{-1}[F_1(\boldsymbol{\omega}) \times F_2(\boldsymbol{\omega})], \tag{4.30}$$

where $\mathcal{F}^{-1}[\cdot]$ is the inverse Fourier transform, and $F_1(\boldsymbol{\omega})$ and $F_2(\boldsymbol{\omega})$ are the Fourier transform of $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$, respectively. Applying the Fourier Transform on $f_1(\boldsymbol{x})$:

$$
\begin{aligned}
F_1(\boldsymbol{\omega}) &= \int f_1(\boldsymbol{x}) \exp(-j2\pi \boldsymbol{x}^T \boldsymbol{\omega})\, \mathrm{d}\boldsymbol{x} \\
&= K \int \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_\Delta)^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x}-\boldsymbol{x}_\Delta) - j2\pi \boldsymbol{x}^T \boldsymbol{\omega}\right) \mathrm{d}\boldsymbol{x} \\
&= K \int \exp\left(-\frac{1}{2}(\boldsymbol{x}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x} - 2\boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x} + \boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta) - j2\pi \boldsymbol{x}^T \boldsymbol{\omega}\right) \mathrm{d}\boldsymbol{x} \\
&= K \int \exp\left(-\frac{1}{2}\left(\boldsymbol{x}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x} - 2\boldsymbol{x}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\left(\boldsymbol{x}_\Delta - j2\pi \boldsymbol{\Sigma}_{x_{m|s}}\boldsymbol{\omega}\right) + \boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta\right)\right) \mathrm{d}\boldsymbol{x} \\
&= K \int \exp\left(-\frac{1}{2}\left((\boldsymbol{x}-\boldsymbol{\Delta})^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x}-\boldsymbol{\Delta}) - \boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{\Delta} + \boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta\right)\right) \mathrm{d}\boldsymbol{x} \\
&= K \exp\left(\frac{1}{2}\boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{\Delta} - \frac{1}{2}\boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta\right) \int \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\Delta})^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x}-\boldsymbol{\Delta})\right) \mathrm{d}\boldsymbol{x} \\
&= K\sqrt{(2\pi)^2 \left|\boldsymbol{\Sigma}_{x_{m|s}}\right|} \exp\left(\frac{1}{2}\boldsymbol{\Delta}^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{\Delta} - \frac{1}{2}\boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta\right) \\
&= \exp\left(\frac{1}{2}\left(\boldsymbol{x}_\Delta - j2\pi \boldsymbol{\Sigma}_{x_{m|s}}\boldsymbol{\omega}\right)^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}(\boldsymbol{x}_\Delta - j2\pi \boldsymbol{\Sigma}_{x_{m|s}}\boldsymbol{\omega}) - \frac{1}{2}\boldsymbol{x}_\Delta^T \boldsymbol{\Sigma}_{x_{m|s}}^{-1}\boldsymbol{x}_\Delta\right) \\
&= \exp\left(-j2\pi \boldsymbol{x}_\Delta^T \boldsymbol{\omega} - 2\pi^2 \boldsymbol{\omega}^T \boldsymbol{\Sigma}_{x_{m|s}}\boldsymbol{\omega}\right),
\end{aligned}
\tag{4.31}
$$

where $K = \frac{1}{2\pi\sqrt{\left|\boldsymbol{\Sigma}_{x_{m|s}}\right|}}$ and $\boldsymbol{\Delta} = \boldsymbol{x}_\Delta - j2\pi \boldsymbol{\Sigma}_{x_{m|s}}\boldsymbol{\omega}$. Similarly, for $f_2(\boldsymbol{x})$,

$$
F_2(\boldsymbol{\omega}) = \exp\left(-j2\pi \boldsymbol{s}_m^T \boldsymbol{\omega} - 2\pi^2 \boldsymbol{\omega}^T \boldsymbol{\Sigma}_{s_m}\boldsymbol{\omega}\right). \tag{4.32}
$$

The Fourier transform of $p(\boldsymbol{x}|\boldsymbol{x}_\Delta)$ is then given by:

$$
\begin{aligned}
\mathcal{F}[p(\boldsymbol{x}|\boldsymbol{x}_\Delta)] &= F_1(\boldsymbol{\omega}) \times F_2(\boldsymbol{\omega}) \\
&= \exp\left(-j2\pi(\boldsymbol{x}_\Delta + \boldsymbol{s}_m)^T \boldsymbol{\omega} - 2\pi^2 \boldsymbol{\omega}^T \left(\boldsymbol{\Sigma}_{x_{m|s}} + \boldsymbol{\Sigma}_{s_m}\right)\boldsymbol{\omega}\right),
\end{aligned}
\tag{4.33}
$$

And the PDF is derived by applying the inverse Fourier transform on the above equation:

$$
\begin{aligned}
p(\boldsymbol{x}|\boldsymbol{x}_\Delta) &= \mathcal{F}^{-1}\left[\exp\left(-j2\pi(\boldsymbol{x}_\Delta + \boldsymbol{s}_m)^T \boldsymbol{\omega} - 2\pi^2 \boldsymbol{\omega}^T \left(\boldsymbol{\Sigma}_{x_{m|s}} + \boldsymbol{\Sigma}_{s_m}\right)\boldsymbol{\omega}\right)\right] \\
&= \frac{1}{2\pi\sqrt{\left|\boldsymbol{\Sigma}_{x_{m|s}} + \boldsymbol{\Sigma}_{s_m}\right|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\Delta - \boldsymbol{s}_m)^T \left(\boldsymbol{\Sigma}_{x_{m|s}} + \boldsymbol{\Sigma}_{s_m}\right)^{-1}(\boldsymbol{x} - \boldsymbol{x}_\Delta - \boldsymbol{s}_m)\right), \\
\text{or } p(\boldsymbol{x}|\boldsymbol{x}_m) &= \frac{1}{2\pi\sqrt{\left|\boldsymbol{\Sigma}_{x_m}\right|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_m)^T \boldsymbol{\Sigma}_{x_m}{}^{-1}(\boldsymbol{x} - \boldsymbol{x}_m)\right),
\end{aligned}
\tag{4.34}
$$

where $\boldsymbol{x}_m = \boldsymbol{x}_\Delta + \boldsymbol{s}_m$ and $\boldsymbol{\Sigma}_{x_m} = \boldsymbol{\Sigma}_{x_{m|s}} + \boldsymbol{\Sigma}_{s_m}$. Substituting in the given information of the pose of the sensor unit,

$$\boldsymbol{x}_m = \begin{bmatrix} \rho_m \cos\beta_m + u_m \\ \rho_m \sin\beta_m + v_m \end{bmatrix} = \begin{bmatrix} \rho_m \cos(\phi_m + \theta_m) + u_m \\ \rho_m \sin(\phi_m + \theta_m) + v_m \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{\boldsymbol{x}_m} = \boldsymbol{J}_{\boldsymbol{g}}\left(\begin{bmatrix} \rho_m \\ \beta_m \end{bmatrix}\right) \begin{bmatrix} \sigma_{\rho_m}^2 & 0 \\ 0 & \sigma_{\beta_m}^2 \end{bmatrix} \boldsymbol{J}_{\boldsymbol{g}}^T\left(\begin{bmatrix} \rho_m \\ \beta_m \end{bmatrix}\right) + \boldsymbol{\Sigma}_{\boldsymbol{s}_m} \tag{4.35}$$

$$= \boldsymbol{J}_{\boldsymbol{g}}\left(\begin{bmatrix} \rho_m \\ \phi_m + \theta_m \end{bmatrix}\right) \begin{bmatrix} \sigma_{\rho}^2 & 0 \\ 0 & \sigma_{\theta}^2 + \sigma_{\phi}^2 \end{bmatrix} \boldsymbol{J}_{\boldsymbol{g}}^T\left(\begin{bmatrix} \rho_m \\ \phi_m + \theta_m \end{bmatrix}\right) + \boldsymbol{\Sigma}_{\boldsymbol{s}_m},$$

where $\boldsymbol{J}_{\boldsymbol{g}}\left(\begin{bmatrix} \rho_m \\ \phi_m + \theta_m \end{bmatrix}\right) = \begin{bmatrix} \cos(\phi_m + \theta_m) & -\rho_m \sin(\phi_m + \theta_m) \\ \sin(\phi_m + \theta_m) & \rho_m \cos(\phi_m + \theta_m) \end{bmatrix}$. It is noticeable that in the case of Equation (4.35), once the pose of the sensor unit, $[\boldsymbol{s}_m, \phi_m]^T$, becomes more certain, $\left|\boldsymbol{\Sigma}_{\boldsymbol{s}_m}\right|$ and $\sigma_{\phi_m}^2$ become smaller, and the result approximates that derived in Equation (4.18).

4.2 Collaborative Landmark Position Estimation

By applying the Taylor-Series expansion on the coordinates transformation detailed in the previous section, an approximate linear model of the individual position estimation result, $\boldsymbol{x}$, is derived in Equation (4.18) or Equation (4.35) depending on the prior knowledge on the sensor unit's pose. For each localization sensor unit, a deterministic estimate of the landmark position can be obtained by maximizing on the PDF of the vector $\boldsymbol{x}$. However, better results can be achieved taking advantage of the Gaussian form of the linearized distribution, where information from other sources can be efficiently merged.

As shown in Figure 4.3, a landmark placed at *P* is being located by *Sensor 1* and *Sensor 2* simultaneously, whose local frames are $X_1O_1Y_1$ and $X_2O_2Y_2$, respectively.

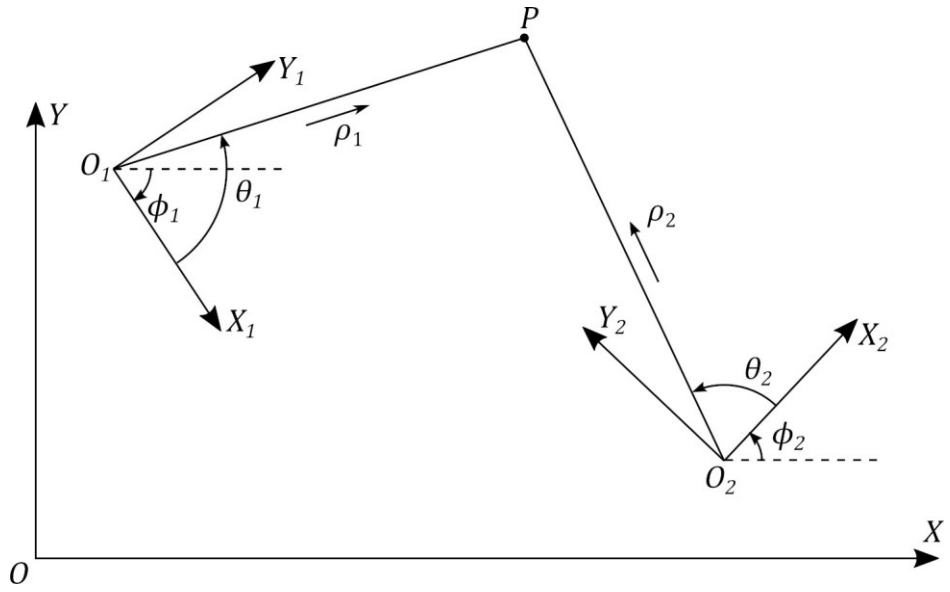


Figure 4.3 Collaborative position estimation of a landmark with two localization sensor units.

In the example, assume *Sensor 1* is located at $\boldsymbol{s}_1 = [u_1, v_1]^T$ with orientation ϕ_1, *Sensor 2* is located at $\boldsymbol{s}_2 = [u_2, v_2]^T$ with orientation ϕ_2. Furthermore, assume that $[\boldsymbol{s}_1^T, \phi_1]^T$ and $[\boldsymbol{s}_2^T, \phi_2]^T$ are independent of each other. The position and orientation of the two units are given by the following distributions:

$$\begin{aligned} \boldsymbol{s}_1 &\sim N(\boldsymbol{s}_{m1}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{m1}}), \\ \phi_1 &\sim N(\phi_{m1}, \sigma_{\phi_{m1}}^2), \\ \boldsymbol{s}_2 &\sim N(\boldsymbol{s}_{m2}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{m2}}), \\ \phi_2 &\sim N(\phi_{m2}, \sigma_{\phi_{m2}}^2). \end{aligned} \tag{4.36}$$

Let the measurements of *Sensor 1* be $\boldsymbol{r}_{m1} = [\rho_{m1}, \theta_{m1}]^T$ and the measurements of *Sensor 2* be $\boldsymbol{r}_{m2} = [\rho_{m2}, \theta_{m2}]^T$. The individual position estimation results of the two sensor units, $\boldsymbol{x}_{m1}$ and $\boldsymbol{x}_{m2}$, have the following distribution functions, given Equation (4.34) and (4.35):

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{x}_{m1}) &= \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m1}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_{m1})^T \boldsymbol{\Sigma}_{\boldsymbol{x}_{m1}}^{-1} (\boldsymbol{x} - \boldsymbol{x}_{m1})\right), \\ p(\boldsymbol{x}|\boldsymbol{x}_{m2}) &= \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m2}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_{m2})^T \boldsymbol{\Sigma}_{\boldsymbol{x}_{m2}}^{-1} (\boldsymbol{x} - \boldsymbol{x}_{m2})\right), \end{aligned} \tag{4.37}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{x}_{m1}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{x}_{m2}}$ are the covariance matrices of the individual position estimates and

$$\boldsymbol{x}_{m1} = \begin{bmatrix} \rho_{m1}\cos(\phi_{m1}+\theta_{m1}) + u_1 \\ \rho_{m1}\sin(\phi_{m1}+\theta_{m1}) + v_1 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{x_{m1}} = \boldsymbol{J_g}\left(\begin{bmatrix} \rho_{m1} \\ \phi_{m1}+\theta_{m1} \end{bmatrix}\right)\begin{bmatrix} \sigma^2_{\rho_{m1}} & 0 \\ 0 & \sigma^2_{\theta_{m1}}+\sigma^2_{\phi_{m1}} \end{bmatrix}\boldsymbol{J_g}^T\left(\begin{bmatrix} \rho_{m1} \\ \phi_{m1}+\theta_{m1} \end{bmatrix}\right) + \boldsymbol{\Sigma_{s_{m1}}},$$

$$\boldsymbol{x}_{m2} = \begin{bmatrix} \rho_{m2}\cos(\phi_{m2}+\theta_{m2}) + u_2 \\ \rho_{m2}\sin(\phi_{m2}+\theta_{m2}) + v_2 \end{bmatrix}, \tag{4.38}$$

$$\boldsymbol{\Sigma}_{x_{m2}} = \boldsymbol{J_g}\left(\begin{bmatrix} \rho_{m2} \\ \phi_{m2}+\theta_{m2} \end{bmatrix}\right)\begin{bmatrix} \sigma^2_{\rho_{m2}} & 0 \\ 0 & \sigma^2_{\theta_{m2}}+\sigma^2_{\phi_{m2}} \end{bmatrix}\boldsymbol{J_g}^T\left(\begin{bmatrix} \rho_{m2} \\ \phi_{m2}+\theta_{m2} \end{bmatrix}\right) + \boldsymbol{\Sigma_{s_{m2}}},$$

$$\boldsymbol{J_g}\left(\begin{bmatrix} \rho \\ \theta \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{bmatrix}.$$

The collaborative position estimation result can be obtained by applying Bayes' theorem on the above distributions:

$$\begin{aligned} p(\boldsymbol{x}|\,\boldsymbol{x}_{m1},\boldsymbol{x}_{m2}) &= \eta p(\boldsymbol{x}|\,\boldsymbol{x}_{m1})p(\boldsymbol{x}|\boldsymbol{x}_{m2}) \\ &= \frac{\eta\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_{m1})^T\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}(\boldsymbol{x}-\boldsymbol{x}_{m1})-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_{m2})^T\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}(\boldsymbol{x}-\boldsymbol{x}_{m2})\right)}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m1}}||\boldsymbol{\Sigma}_{x_{m2}}|}} \\ &= \frac{\eta\exp\left(-\frac{1}{2}(\boldsymbol{x}^T(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}+\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}})\boldsymbol{x}-2\boldsymbol{x}^T(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}\boldsymbol{x}_{m1}+\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}\boldsymbol{x}_{m2})+\boldsymbol{x}^T_{m1}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}\boldsymbol{x}_{m1}+\boldsymbol{x}^T_{m2}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}\boldsymbol{x}_{m2})\right)}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_{m1}}||\boldsymbol{\Sigma}_{x_{m2}}|}} \\ &= \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_m}|}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_m)^T\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_m}(\boldsymbol{x}-\boldsymbol{x}_m)\right), \end{aligned} \tag{4.39}$$

where η is a normalization factor, $\boldsymbol{x}_m = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}+\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}\right)^{-1}(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}\boldsymbol{x}_{m1}+\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}\boldsymbol{x}_{m2})$, and $\boldsymbol{\Sigma}_{\boldsymbol{x}_m} = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m1}}+\boldsymbol{\Sigma}^{-1}_{\boldsymbol{x}_{m2}}\right)^{-1}$.

The collaborative probability density function can be easily extended to the case where information from more sensor units is available. Assuming that N units collaboratively estimate the position of the landmark placed at P, the position and the orientation of the i^{th} unit are provided as independent Gaussian distributions:

$$\begin{aligned} \boldsymbol{s}_i &\sim N\left(\boldsymbol{s}_{mi}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{mi}}\right), \\ \phi_i &\sim N\left(\phi_{mi}, \sigma^2_{\phi_{mi}}\right). \end{aligned} \tag{4.40}$$

The PDF of the landmark position based on all the given information can be derived as:

$$\begin{aligned} p(\boldsymbol{x}|\{\boldsymbol{x}_{mi}\}) &= \eta \prod_{i=1}^{N} p(\boldsymbol{x}|\, \boldsymbol{x}_{mi}) \\ &= \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}_{x_m}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_m)^T \boldsymbol{\Sigma}_{x_m}^{-1}(\boldsymbol{x}-\boldsymbol{x}_m)\right), \end{aligned} \tag{4.41}$$

where η is a normalization factor, $\boldsymbol{x}_{mi} = \boldsymbol{f}(\boldsymbol{r}_{mi}) = \begin{bmatrix} \rho_{mi}\cos(\phi_{mi}+\theta_{mi}) + u_{mi} \\ \rho_{mi}\sin(\phi_{mi}+\theta_{mi}) + v_{mi} \end{bmatrix}$ is the individual position estimate made by the i^{th} sensor unit, its corresponding covariance matrix is $\boldsymbol{\Sigma}_{x_{mi}} = \boldsymbol{J}_{\boldsymbol{g}}\left(\begin{bmatrix} \rho_{mi} \\ \phi_{mi}+\theta_{mi} \end{bmatrix}\right) \begin{bmatrix} \sigma^2_{\rho_{mi}} & 0 \\ 0 & \sigma^2_{\theta_{mi}} + \sigma^2_{\phi_{mi}} \end{bmatrix} \boldsymbol{J}_{\boldsymbol{g}}^T\left(\begin{bmatrix} \rho_{mi} \\ \phi_{mi}+\theta_{mi} \end{bmatrix}\right) + \boldsymbol{\Sigma}_{\boldsymbol{s}_{mi}}$, $\boldsymbol{\Sigma}_{\boldsymbol{x}_m} = \left(\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\boldsymbol{x}_{mi}}^{-1}\right)^{-1}$, and $\boldsymbol{x}_m = \left(\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\boldsymbol{x}_{mi}}^{-1}\right)^{-1} \sum_{i=1}^{N}\left(\boldsymbol{\Sigma}_{\boldsymbol{x}_{mi}}^{-1}\boldsymbol{x}_{mi}\right)$. Noticeably, the form of $\boldsymbol{x}_m$ can be considered as the average of the individual estimates, $\{\boldsymbol{x}_{mi}\}$, weighted by their corresponding inversed covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{x}_{mi}}^{-1}$.

4.3 Localization Sensor Inter-Calibration

Inter-calibration is the process where two localization sensor units take relative distance and angular measurements towards each other, which is an essential part of the pose estimation of the localization sensor. In this process, the measured relative location between the sensor units is improved by merging the distance measurements from both sensor units. If the pose of one sensor unit is given, the pose of the other sensor unit could be estimated, both of which are within the same reference frame. For the sensor unit whose pose is unknown, its position is estimated based on the previously developed model for individual landmark position estimation, but with a merged estimated distance. The orientation of the sensor unit can be obtained by applying geometric constraints found in the inter-calibration process,

however, the uncertainties in the angular measurements will not be reduced. An example of inter-calibration is shown in Figure 4.4.

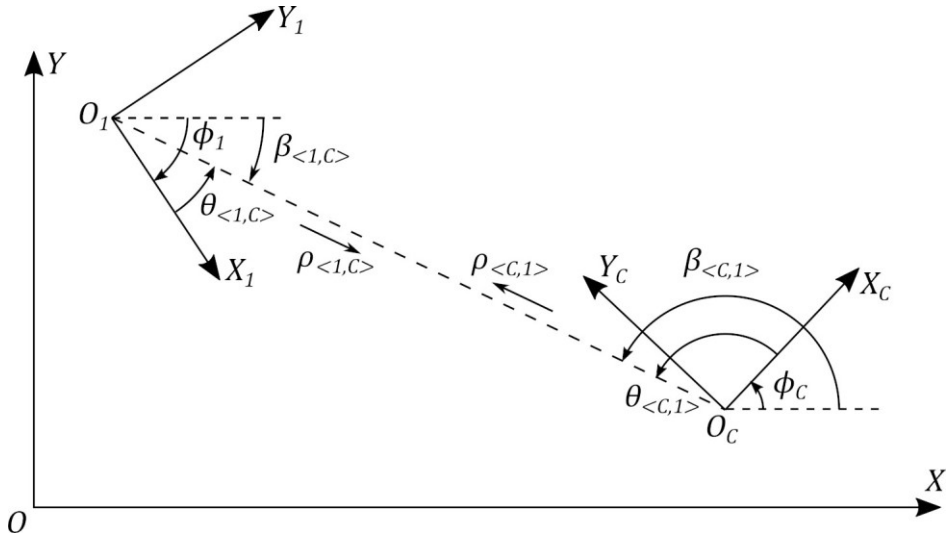


Figure 4.4 An example of inter-calibration between two sensor units.

Sensor 1 on the left is placed at the position $\boldsymbol{s}_1 = [u_1, v_1]^T$ with orientation ϕ_1, and both are defined as Gaussian distributions as follows:

$$\begin{aligned} \boldsymbol{s}_1 &\sim N(\boldsymbol{s}_{m1}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{m1}}), \\ \phi_1 &\sim N(\phi_{m1}, \sigma^2_{\phi_{m1}}). \end{aligned} \tag{4.42}$$

The position and orientation of *Sensor C*, $\boldsymbol{s}_C = [u_C, v_C]^T$ and ϕ_C, are unknown and are to be inter-calibrated. It is assumed here that both sensor units are built the same way and possess the same capabilities to measure distance and bearing angle towards landmarks carried by other sensor units, but each sensor unit has its own noise characteristics. Let the distance and angular measurements made by *Sensor 1* to *Sensor C* be denoted by $\rho_{m<1,C>}$ and $\theta_{m<1,C>}$, and the measurements made by *Sensor C* to *Sensor 1* be $\rho_{m<C,1>}$ and $\theta_{m<C,1>}$. The measurements have the following distributions:

$$
\begin{aligned}
p\left(\rho_{<1,C>}\middle|\rho_{m<1,C>}\right) &= \frac{1}{\sqrt{2\pi}\sigma_{\rho_{m<1,C>}}}\exp\left(-\frac{\left(\rho_{<1,C>}-\rho_{m<1,C>}\right)^2}{2\sigma^2_{\rho_{m<1,C>}}}\right),\\
p\left(\theta_{<1,C>}\middle|\theta_{m<1,C>}\right) &= \frac{1}{\sqrt{2\pi}\sigma_{\theta_{m<1,C>}}}\exp\left(-\frac{\left(\theta_{<1,C>}-\theta_{m<1,C>}\right)^2}{2\sigma^2_{\theta_{m<1,C>}}}\right),\\
p\left(\rho_{<C,1>}\middle|\rho_{m<C,1>}\right) &= \frac{1}{\sqrt{2\pi}\sigma_{\rho_{m<C,1>}}}\exp\left(-\frac{\left(\rho_{<C,1>}-\rho_{m<C,1>}\right)^2}{2\sigma^2_{\rho_{m<C,1>}}}\right),\\
p\left(\theta_{<C,1>}\middle|\theta_{m<C,1>}\right) &= \frac{1}{\sqrt{2\pi}\sigma_{\theta_{m<C,1>}}}\exp\left(-\frac{\left(\theta_{<C,1>}-\theta_{m<C,1>}\right)^2}{2\sigma^2_{\theta_{m<C,1>}}}\right),
\end{aligned}
\tag{4.43}
$$

where $\rho_{<1,C>}$, $\theta_{<1,C>}$, $\rho_{<C,1>}$ and $\theta_{<C,1>}$ are the corresponding true values of the measurements, and $\sigma_{\rho_{m<1,C>}}$, $\sigma_{\theta_{m<1,C>}}$, $\sigma_{\rho_{m<C,1>}}$, and $\sigma_{\theta_{m<C,1>}}$ are the corresponding standard deviations of the measurements. Some geometric constraints are imposed in Figure 4.4 in the context of inter-calibration:

$$
\begin{aligned}
\rho_{<1,C>} &= \rho_{<C,1>},\\
\text{and } \beta_{<C,1>} &= \pi + \beta_{<1,C>},
\end{aligned}
\tag{4.44}
$$

where $\beta_{<1,C>} = \phi_1 + \theta_{<1,C>}$ and $\beta_{<C,1>} = \phi_C + \theta_{<C,1>}$ are the bearing angles of the line segment $\overline{O_1O_C}$ observed by the two sensor units, respectively.

According to Equation (4.35), the position estimate of *Sensor C*, represented by the landmark that it carries, with only measurements of *Sensor 1* is given as:

$$
\begin{aligned}
\boldsymbol{s}_C &\sim N\left(\boldsymbol{s}_{mC<1,C>}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{mC<1,C>}}\right),\\
\boldsymbol{s}_{mC<1,C>} &= \begin{bmatrix}\rho_{m<1,C>}\, cos\left(\phi_{m1}+\theta_{m<1,C>}\right)+u_{m1}\\ \rho_{m<1,C>}\, sin\left(\phi_{m1}+\theta_{m<1,C>}\right)+v_{m1}\end{bmatrix},\\
\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC<1,C>}} &= \boldsymbol{J_g}\left(\begin{bmatrix}\rho_{m<1,C>}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right)\begin{bmatrix}\sigma^2_{\rho_{m<1,C>}} & 0\\ 0 & \sigma^2_{\theta_{m<1,C>}}+\sigma^2_{\phi_{m1}}\end{bmatrix}\boldsymbol{J_g}^T\left(\begin{bmatrix}\rho_{m<1,C>}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right)+\boldsymbol{\Sigma}_{\boldsymbol{s}_{m1}},\\
\boldsymbol{J_g}\left(\begin{bmatrix}\rho_{m<1,C>}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right) &= \begin{bmatrix}\cos\left(\phi_{m1}+\theta_{m<1,C>}\right) & -\rho_{m<1,C>}\sin\left(\phi_{m1}+\theta_{m<1,C>}\right)\\ \sin\left(\phi_{m1}+\theta_{m<1,C>}\right) & \rho_{m<1,C>}\cos\left(\phi_{m1}+\theta_{m<1,C>}\right)\end{bmatrix},
\end{aligned}
\tag{4.45}
$$

where $\boldsymbol{s}_{mC<1,C>}$ is the estimated position of *Sensor C* made by *Sensor 1*, and $\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC<1,C>}}$ is its covariance matrix. Since the distance measurements are made by both sensor units measuring the same line segment,

$\overline{O_1O_C}$, the distance measurement, $\rho_{m<1,C>}$, with standard deviation, $\sigma_{\rho_{m<1,C>}}$, can be merged with that from *Sensor C* by applying Bayes' theorem:

$$p\left(\rho_{\{1,C\}}\middle|\rho_{m<1,C>},\rho_{m<C,1>}\right)=\eta p\left(\rho_{<1,C>}\middle|\rho_{m<1,C>}\right)p\left(\rho_{<C,1>}\middle|\rho_{m<C,1>}\right), \tag{4.46}$$

where η is a normalization factor. The subscript using curly brackets indicates that the value is obtained engaging measurements between *Sensor 1* and *Sensor C*, and the numbers inside the curly brackets are interchangeable. The merged distribution is given as:

$$p\left(\rho_{\{1,C\}}\middle|\rho_{m<1,C>},\rho_{m<C,1>}\right)=\frac{1}{\sqrt{2\pi}\rho_{m\{1,C\}}}\exp\left(-\frac{\left(\rho_{<1,C>}-\rho_{m\{1,C\}}\right)^2}{2\sigma^2_{\rho_{m\{1,C\}}}}\right), \tag{4.47}$$

where $\rho_{m\{1,C\}}=\frac{\sigma^2_{\rho_{m<C,1>}}\rho_{m<1,C>}+\sigma^2_{\rho_{m<1,C>}}\rho_{m<C,1>}}{\sigma^2_{\rho m<1,C>}+\sigma^2_{\rho_{m<C,1>}}}=\frac{\frac{1}{\sigma^2_{\rho m<1,C>}}\rho_{m<1,C>}+\frac{1}{\sigma^2_{\rho_{m<C,1>}}}\rho_{m<C,1>}}{\frac{1}{\sigma^2_{\rho m<1,C>}}+\frac{1}{\sigma^2_{\rho_{m<C,1>}}}}$ is the merged distance measurement by both sensor units, and its variance is given by $\sigma^2_{\rho_{m\{1,C\}}}=\frac{\sigma^2_{\rho_{m<1,C>}}\sigma^2_{\rho_{m<C,1>}}}{\sigma^2_{\rho_{m<1,C>}}+\sigma^2_{\rho_{m<C,1>}}}=\frac{1}{\frac{1}{\sigma^2_{\rho_{m<1,C>}}}+\frac{1}{\sigma^2_{\rho_{m<C,1>}}}}$. Noticeably, $\rho_{m\{1,C\}}$ is also the average of $\rho_{m<1,C>}$ and $\rho_{m<C,1>}$ weighted by their inversed variances, and the standard deviation of $\rho_{m\{1,C\}}$ is smaller than those of both $\rho_{m<1,C>}$ and $\rho_{m<C,1>}$, forming a more certain posterior probability distribution of the ground truth, $\rho_{\{1,C\}}$. The position estimation of *Sensor C* is then improved as:

$$\begin{aligned}
\boldsymbol{s}_C &\sim N\left(\boldsymbol{s}_{mC\{1,C\}},\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{1,C\}}}\right),\\
\boldsymbol{s}_{mC\{1,C\}} &=\begin{bmatrix}\rho_{m\{1,C\}}\,cos\left(\phi_{m1}+\theta_{m<1,C>}\right)+u_{m1}\\ \rho_{m\{1,C\}}\,sin\left(\phi_{m1}+\theta_{m<1,C>}\right)+v_{m1}\end{bmatrix},\\
\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{1,C\}}} &=\boldsymbol{J_g}\left(\begin{bmatrix}\rho_{m\{1,C\}}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right)\begin{bmatrix}\sigma^2_{\rho_{m\{1,C\}}} & 0\\ 0 & \sigma^2_{\theta_{m<1,C>}}+\sigma^2_{\phi_{m1}}\end{bmatrix}\boldsymbol{J_g^T}\left(\begin{bmatrix}\rho_{m\{1,C\}}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right)+\boldsymbol{\Sigma}_{\boldsymbol{s}_{m1}},\\
\boldsymbol{J_g}\left(\begin{bmatrix}\rho_{m\{1,C\}}\\ \phi_{m1}+\theta_{m<1,C>}\end{bmatrix}\right) &=\begin{bmatrix}\cos\left(\phi_{m1}+\theta_{m<1,C>}\right) & -\rho_{m\{1,C\}}\sin\left(\phi_{m1}+\theta_{m<1,C>}\right)\\ \sin\left(\phi_{m1}+\theta_{m<1,C>}\right) & \rho_{m\{1,C\}}\cos\left(\phi_{m1}+\theta_{m<1,C>}\right)\end{bmatrix},
\end{aligned} \tag{4.48}$$

where $\boldsymbol{s}_{mC\{1,C\}}$ is the inter-calibrated position of *Sensor C* with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{1,C\}}}$.

The second constraint in Equation (4.44) is useful for estimating the orientation of *Sensor C*, as it can be rewritten as:

$$\phi_C = \pi + \phi_1 + \theta_{<1,C>} - \theta_{<C,1>} = \pi + \beta_{<1,C>} - \theta_{<C,1>}, \tag{4.49}$$

where $\beta_{<1,C>}$ is the bearing of *Sensor C* observed by *Sensor 1*, and according to Equation (4.20), its distribution is given by $\beta_{<1,C>} \sim N(\beta_{m<1,C>}, \sigma^2_{\beta_{m<1,C>}})$, where $\beta_{m<1,C>} = \theta_{m<1,C>} + \phi_{m1}$ and $\sigma^2_{\beta_{m<1,C>}} = \sigma^2_{\theta_{m<1,C>}} + \sigma^2_{\phi_{m1}}$. Since $\beta_{m<1,C>}$ and $\theta_{m<C,1>}$ are independent measurements, a similar process to that in Equation (4.20) could be applied to obtain the conditional PDF of ϕ_C given these measurements. Thus,

$$\begin{aligned} p(\phi_C|\beta_{m<1,C>}, \theta_{m<C,1>}) &= \int p(\phi_C - \pi + \theta_{<C,1>}|\beta_{m<1,C>}) p(\theta_{<C,1>}|\theta_{m<C,1>}) \mathrm{d}\theta_{<C,1>} \\ &= \frac{1}{\sqrt{2\pi\left(\sigma^2_{\beta_{m<1,C>}} + \sigma^2_{\theta_{m<C,1>}}\right)}} \exp\left(-\frac{(\phi_C - \pi - \beta_{m<1,C>} + \theta_{m<C,1>})^2}{2\left(\sigma^2_{\beta_{m<1,C>}} + \sigma^2_{\theta_{m<C,1>}}\right)}\right). \end{aligned} \tag{4.50}$$

According to the above equation, the orientation of *Sensor C*, ϕ_C, has the following distribution given ϕ_{m1}, $\theta_{m<1,C>}$ and $\theta_{m<C,1>}$:

$$\begin{aligned} \phi_C &\sim N\left(\pi + \beta_{m<1,C>} - \theta_{m<C,1>}, \sigma^2_{\beta_{m<1,C>}} + \sigma^2_{\theta_{m<C,1>}}\right), \\ \text{or } \phi_C &\sim N\left(\pi + \phi_{m1} + \theta_{m<1,C>} - \theta_{m<C,1>}, \sigma^2_{\theta_{m<1,C>}} + \sigma^2_{\phi_{m1}} + \sigma^2_{\theta_{m<C,1>}}\right), \end{aligned} \tag{4.51}$$

and the PDF of ϕ_C is given as:

$$p(\phi_C|\phi_{m1}, \theta_{m<1,C>}, \theta_{m<C,1>}) = \frac{1}{\sqrt{2\pi}\sigma_{\phi_{C\{1,C\}}}} \exp\left(-\frac{(\phi_C - \phi_{mC\{1,C\}})^2}{2\sigma^2_{\phi_{mC\{1,C\}}}}\right), \tag{4.52}$$

where $\phi_{mC\{1,C\}} = \pi + \phi_{m1} + \theta_{m<1,C>} - \theta_{m<C,1>}$ and $\sigma^2_{\phi_{mC\{1,C\}}} = \sigma^2_{\theta_{m<1,C>}} + \sigma^2_{\phi_{m1}} + \sigma^2_{\theta_{m<C,1>}}$. Note that all the angles are defined within the range $[-\pi, \pi)$, and thus the results of the addition and subtraction operations are always wrapped to the particular range by $\theta' = \big((\theta + \pi) \bmod 2\pi\big) - \pi$, where θ denotes the original angles.

It is noticeable that in the inter-calibration process, the errors in the angular measurements are combined in the estimated orientation of *Sensor C*, and its corresponding variance is the sum of the variance on the orientation of *Sensor 1* and those on the angular measurements from both *Sensor 1* and *Sensor C*. If the inter-calibrated sensor unit is directly used for inter-calibrating another sensor unit without any improvement, the uncertainties in the estimated pose of that extra sensor unit will accumulate and grow even larger in both position and orientation. Therefore, the estimated orientation usually should be improved to reduce the uncertainty in the orientation. As mentioned earlier, one simple way is to take repeated angular measurements over the same ground truth to reduce the uncertainties in $\theta_{m<1,C>}$ and $\theta_{m<C,1>}$, but the uncertainty in ϕ_{m1} remains untouched. Another way to reduce the uncertainties on orientation is through collaboration with multiple localization sensor units, as described in the following subsection.

4.4 Collaborative Sensor Pose Estimation

An example of multiple localization sensor units working collaboratively to estimate the pose of *Sensor C* is shown in Figure 4.5, where the position and orientation of *Sensor C* is denoted by $\boldsymbol{s}_C = [u_C, v_C]^T$ and ϕ_C respectively. In the example, the pose of *Sensor C* is collaboratively estimated with another three sensor units with known poses, *Sensor 1*, *Sensor 2* and *Sensor 3* respectively located at O_1, O_2, and O_3. Note that as mentioned before, angles are defined in the range of $[-\pi, \pi)$, and the bearing angle measured by *Sensor 3* shown in Figure 4.5 is thus changed accordingly.

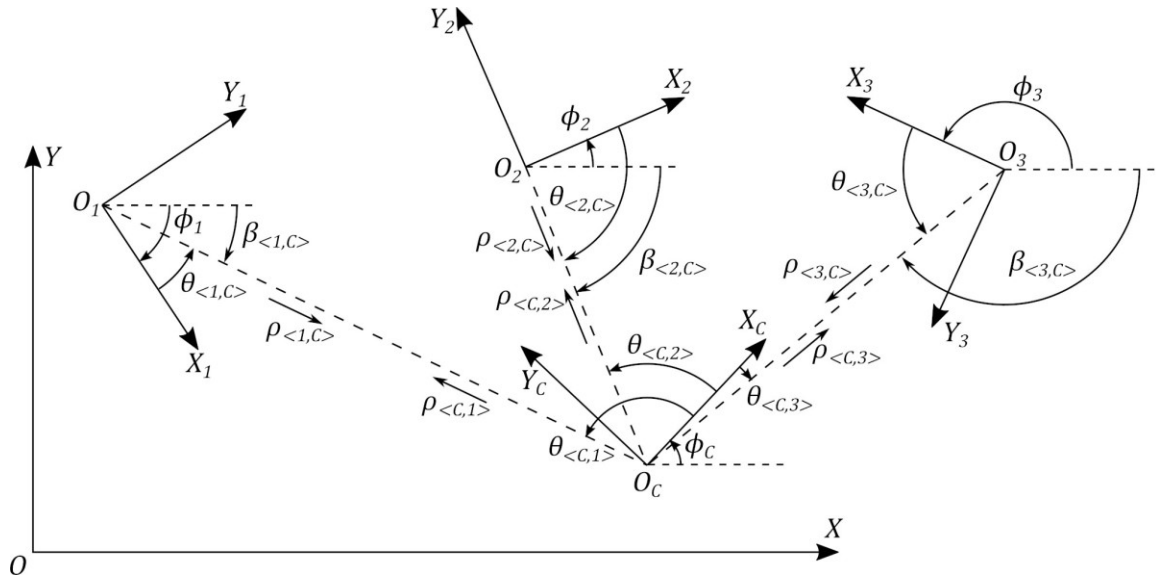


Figure 4.5 An example of collaborative pose estimation.

Assume that N available sensor units participate in the process, each with a known pose, and the N poses are independent of each other. The position and orientation of the i^{th} sensor unit are given as Gaussian distributions as follows:

$$\begin{aligned} \boldsymbol{s}_i &\sim N\left(\boldsymbol{s}_{mi}, \boldsymbol{\Sigma}_{\boldsymbol{s}_{mi}}\right), \\ \phi_i &\sim N\left(\phi_{mi}, \sigma^2_{\phi mi}\right), \\ i &= 1,2,\dots,N. \end{aligned} \tag{4.53}$$

The relative measurements made by the i^{th} sensor unit to *Sensor C* is denoted by $\boldsymbol{r}_{m<i,C>} = \left[\rho_{m<i,C>}, \theta_{m<i,C>}\right]^T$, while the corresponding measurements made by *Sensor C* towards the i^{th} sensor unit is denoted by $\boldsymbol{r}_{m<C,i>} = \left[\rho_{m<C,i>}, \theta_{m<C,i>}\right]^T$. Since the measurements are collected by different units with individually characterized sensing components, $\boldsymbol{r}_{m<i,C>}$ and $\boldsymbol{r}_{m<C,i>}$ are considered independent of each other.

In the process of collaborative pose estimation of *Sensor C*, the position of *Sensor C* is first estimated collaboratively by the N sensor units with known poses, similar to that in the collaborative landmark position estimation described in Section 4.2. Secondly, *Sensor C* inter-calibrates with each of the N sensor

units, resulting in N individually calibrated orientation estimates. The individually calibrated orientation estimates from the inter-calibration process then merge to formulate a collaboratively estimated orientation of *Sensor C*, and the uncertainties reside in ϕ_i and relative angular measurements are reduced on the merged orientation estimate. The previously estimated position of *Sensor C* also benefits from the inter-calibration process and is improved by merging distance measurements collected by *Sensor C*.

The collaborative position estimation of *Sensor C* using measurements from the N sensor units whose poses are known is given by Equation (4.41). With the distance measurements merged with the corresponding ones made from *Sensor C*, as described in Equation (4.48), the improved collaborative position estimation of *Sensor C* is given by:

$$
\begin{aligned}
p\left(\boldsymbol{s}_C \middle| \{\boldsymbol{s}_{mC\{i,C\}}\}\right) &= \frac{1}{2\pi\sqrt{\left|\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC}}\right|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{s}_{mC})^T \boldsymbol{\Sigma}_{\boldsymbol{s}_{mC}}^{-1}(\boldsymbol{x}-\boldsymbol{s}_{mC})\right), \\
\boldsymbol{s}_{mC\{i,C\}} &= \begin{bmatrix} \rho_{m\{i,C\}}\, cos\left(\phi_{mi}+\theta_{m<i,C>}\right)+u_{mi} \\ \rho_{m\{i,C\}}\, sin\left(\phi_{mi}+\theta_{m<i,C>}\right)+v_{mi} \end{bmatrix}, \\
\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{i,C\}}} &= \boldsymbol{J_g}\left(\begin{bmatrix} \rho_{m\{i,C\}} \\ \phi_{mi}+\theta_{m<i,C>} \end{bmatrix}\right) \begin{bmatrix} \sigma^2_{\rho_{m\{i,C\}}} & 0 \\ 0 & \sigma^2_{\theta_{m<i,C>}}+\sigma^2_{\phi_{mi}} \end{bmatrix} \boldsymbol{J_g}^T\left(\begin{bmatrix} \rho_{m\{i,C\}} \\ \phi_{mi}+\theta_{m<i,C>} \end{bmatrix}\right)+\boldsymbol{\Sigma}_{\boldsymbol{s}_{mi}}, \qquad (4.54) \\
\boldsymbol{J_g}\left(\begin{bmatrix} \rho_{m\{i,C\}} \\ \phi_{mi}+\theta_{m<i,C>} \end{bmatrix}\right) &= \begin{bmatrix} \cos\left(\phi_{mi}+\theta_{m<i,C>}\right) & -\rho_{m\{i,C\}}\sin\left(\phi_{mi}+\theta_{m<i,C>}\right) \\ \sin\left(\phi_{mi}+\theta_{m<i,C>}\right) & \rho_{m\{i,C\}}\cos\left(\phi_{mi}+\theta_{m<i,C>}\right) \end{bmatrix}, \\
\boldsymbol{s}_{mC} &= \left(\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{i,C\}}}^{-1}\right)^{-1} \sum_{i=1}^{N}\left(\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{i,C\}}}^{-1}\boldsymbol{s}_{mC\{i,C\}}\right), \text{ and } \boldsymbol{\Sigma}_{\boldsymbol{s}_{mC}} = \left(\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{i,C\}}}^{-1}\right)^{-1},
\end{aligned}
$$

where $\boldsymbol{s}_{mC\{i,C\}}$ is the improved individual position estimation of *Sensor C* made by the i^{th} sensor unit with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC\{i,C\}}}$, and $\boldsymbol{s}_{mC}$ is the improved collaborative position estimate with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{s}_{mC}}$.

The orientation estimation of *Sensor C* using multiple sensor units is based on Equation (4.52) of the inter-calibration process for two localization sensor units. The conditional PDF of the orientation, ϕ_C, given the angular measurements in between the i^{th} sensor unit and *Sensor C* can be expressed as:

$$p(\phi_C|\phi_{mi},\theta_{m<i,C>},\theta_{m<C,i>}) = \frac{1}{\sqrt{2\pi}\sigma_{\phi_{mC\{i,C\}}}}\exp\left(-\frac{(\phi_C-\phi_{mC\{i,C\}})^2}{2\sigma^2_{\phi_{mC\{i,C\}}}}\right), \tag{4.55}$$

where $\sigma^2_{\phi_{mC\{i,C\}}} = \sigma^2_{\theta_{m<i,C>}} + \sigma^2_{\phi_{mi}} + \sigma^2_{\theta_{m<C,i>}}$ and $\phi_{mC\{i,C\}} = \pi + \phi_{mi} + \theta_{m<i,C>} - \theta_{m<C,i>}$. By applying Bayes' theorem on the individual estimates, $\phi_{mC\{i,C\}}$, the orientation of *Sensor C* is collaboratively estimated as:

$$\begin{aligned} p(\phi_C|\{\phi_{mC\{i,C\}}\}) &= \eta\prod_{i=1}^{N} p(\phi_C|\phi_{mC\{i,C\}}) \\ &= \frac{1}{\sqrt{2\pi}\sigma_{\phi_{mC}}}\exp\left(-\frac{(\phi_C-\phi_{mC})^2}{2\sigma^2_{\phi_{mC}}}\right), \end{aligned} \tag{4.56}$$

where η is a normalization factor, $\sigma^2_{\phi_{mC}} = 1\Big/\sum_{i=1}^{N}\frac{1}{\sigma^2_{\phi_{mC\{i,C\}}}}$, and $\phi_{mC} = \sum_{i=1}^{N}\frac{\phi_{mC\{i,C\}}}{\sigma^2_{\phi_{mC\{i,C\}}}}\Big/\sum_{i=1}^{N}\frac{1}{\sigma^2_{\phi_{mC\{i,C\}}}}$.

The collaborative pose estimation model of the proposed localizations is derived in this section by performing inter-calibration between the sensor unit to be estimated and the sensor units with known poses. The model can also be divided as collaborative sensor position estimation with merged distance measurements and collaborative sensor orientation estimation as given in Equation (4.54) and (4.56). As demonstrated in these equations, the uncertainties achieved in the individual position and orientation estimation models are reduced in the collaborative pose estimation model by merging the individual position and orientation estimates, leading to a more reliable pose estimation result. It is worth noting that the measurements taken by the sensor unit to be calibrated are necessary for estimating its orientation but not its position, since its position can be estimated solely by other sensor units, while its orientation can only be estimated when the unit itself participates in the process.

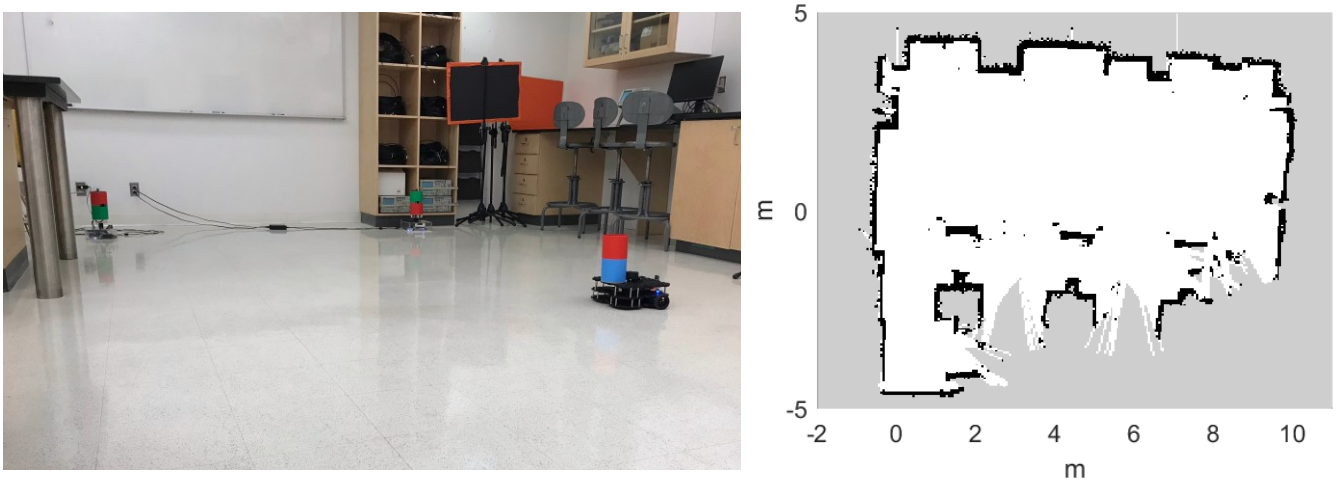
4.5 Summary

In this chapter, a collaborative pose estimation framework using the proposed localization sensor is derived. The position estimation of the localization sensor, represented by the landmark it carries, is first

developed using either an individual sensor unit or multiple sensor units that work collaboratively. An inter-calibration process is then presented, where the pose of a sensor unit can be calibrated in the same reference frame as that of the other sensor unit engaged in the process. Finally, based on the developed methods, a scheme for collaborative pose estimation of a localization sensor unit is formulated using the inter-calibration process with each of the other sensor units. By introducing the collaboration among multiple localization sensor units, the collaborative pose estimation scheme has demonstrated to have the advantage of reducing uncertainties over individual landmark position using one localization sensor unit and inter-calibration for estimating orientation using a pair of localization sensor units. The estimation methods developed in this chapter will be experimentally validated and evaluated in Chapter 5.

CHAPTER 5 - EXPERIMENTAL VALIDATION

Two instances of the proposed localization sensor are fully implemented in this work, which are named *Sensor 1* and *Sensor 2* respectively and correspond to the characterized sensor units in Section 3.4.2. The components on the two sensor units are the same except that *Sensor 1* uses a Jetson Nano Kit as its embedded processing module, while *Sensor 2* uses a Jetson TX2 module with a CTI Orbitty carrier board. The details about the experimental validation of the proposed localization sensor design are documented in this chapter. Tests have taken place in an indoor lab environment over an area of approximately 3×10 m as shown in Figure 5.1(a), and a 2-D occupancy grid map of the testing area is shown in Figure 5.1(b), the details of which will be discussed in Section 5.3. The wall with a whiteboard in the picture corresponds to that on the left on the map. The indoor lab environment is paved with uniform square tiles with a side of 304.8 mm, which are taken advantage of as reference locations when placing the implemented sensor units and stationary landmarks.



(a) (b)

Figure 5.1 (a) Picture of the testing environment; (b) 2-D occupancy grid map of the testing area.

In order to validate the proposed localization sensor for its desired applications, the two implemented sensor units are tested and validated in several use cases. The implemented sensor units are

first tested for individual and collaborative position estimation of a stationary landmark in Section 5.1. Then, the implemented sensor units are tested against a moving landmark installed on a mobile robot, which is detailed in Section 5.2. In Section 5.3, the inter-calibration process, detailed in Section 4.3, between the two implemented sensor units is tested, where the pose of one unit is assumed unknown and to be inter-calibrated. The ground truth values of the pose of both sensor units in the above experiments are obtained with the localization sensor calibration process presented in Section 3.4.3. This is motivated by the fact that manually positioning the sensor units is prone to large errors and there is currently no alternative high accuracy sensing equipment accessible to us for obtaining absolute poses. Due to the reason that only two sensor units are implemented, the collaborative pose estimation of the localization sensor where multiple localization sensor units can be fully calibrated in terms of position and orientation, as proposed in Section 4.4, has not been tested in this work, as it would have required the availability of at least a third unit of the proposed localization sensor.

5.1 Individual and Collaborative Landmark Position Estimation

Each localization sensor unit has a uniquely color-coded cylindrical landmark installed and centered on top, and thus the position of the sensor unit is considered the same as the position of the corresponding landmark. Therefore, instead of using a third localization sensor unit, a standalone landmark is tested against the two implemented units, and the position of the landmark is estimated both individually and collaboratively in this experiment. The purpose of this experiment is to validate the proposed position estimation methods provided in Sections 4.1 and 4.2 using the proposed localization sensor.

In the analysis of the position estimation results, several error terms are defined. The position estimation error is the Euclidean distance between the estimated position and the true position. The

distance and angular measurement errors are defined as the difference between the distance or angular measurements and their corresponding true values. The error equations are given as follows:

$$\begin{aligned} \epsilon_{position} &= \|\boldsymbol{x}_{estimated} - \boldsymbol{x}_{actual}\|, \\ \epsilon_{distance} &= \rho_{measurement} - \rho_{actual}, \\ \epsilon_{angle} &= \theta_{measurement} - \theta_{actual}. \end{aligned} \tag{5.1}$$

A frequently used evaluation measure, the root-mean-square error (RMSE), is introduced for comparing the errors acquired during the experiment. Assuming N errors of either position estimates, distance measurements or angular measurements, the RMSE of the N errors is defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\epsilon_i^2}, \tag{5.2}$$

where ϵ_i is the i^{th} error.

5.1.1 Experimental Setup

The experimental setup for the position estimation of a stationary landmark is shown in Figure 5.2. The x and y axes are selected as shown in the figure parallel to the sides of the floor tiles, and the localization sensor units are manually placed on the y axis. The landmark is placed at the corners of the floor tiles with offsets along the selected x and y axes, and the tile sides at the corners are always tangent to the outside of the landmark. For instance, the coordinates of the landmark are the same as the coordinates of the corresponding tile corners plus the corresponding offsets on both axes with an absolute value of 53 mm, which is equivalent to the radius of the landmark. These are considered as the ground truth position coordinates of the landmark in the experimental analysis of Section 5.1.2. The slight discrepancy between the actual positions of the landmark and the positions given according to the tiles is neglected during the whole experimental process.

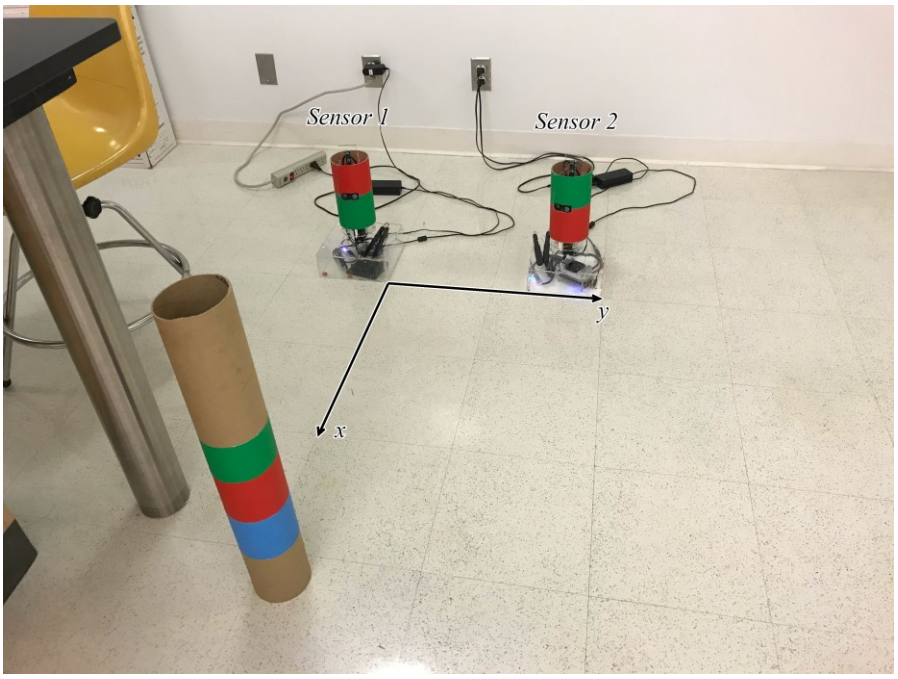


Figure 5.2 Setup for position estimation of a stationary landmark. The additional green section on the landmark is for testing sensor software, but only red and blue sections are used for this test.

Four trials are carried out during the experiment with different distances between the two sensor units. The distance between the two units is increased by an interval of about 609.6 mm along the y axis after each trial. Before starting each trial, the standalone landmark is placed at several positions in front of the sensor units to collect necessary data for calibrating the poses of the units, using the localization sensor calibration process provided in Section 3.4.3. For each calibration process, both angular and distance measurements are recorded for 100 times at each landmark position. The approximated poses of the two units are obtained according to the floor tiles, and both the approximated and calibrated poses of the units are given in Table 5.1. It is worth mentioning that during the experiment, *Sensor 1* is kept unmoved in all four trials, but its calibrated pose varies throughout the trials. The errors in the calibrated poses will lead to measurement biases and eventually result in unexpected estimation errors.

	Sensor 1		*Sensor 2*	
	Approximated Pose (mm, mm, rad)	Calibrated Pose (mm, mm, rad)	Approximated Pose (mm, mm, rad)	Calibrated Pose (mm, mm, rad)
Trial 1	$[-100, -100, 0]^T$	$[-91.6, -104.2, -0.0074]^T$	$[-100, 509.6, 0]^T$	$[-91.0, 498.2, 0.0190]^T$
Trial 2	$[-100, -100, 0]^T$	$[-92.5, -96.4, -0.0100]^T$	$[-100, 1119.2, 0]^T$	$[-98.3, 1112.0, 0.0171]^T$
Trial 3	$[-100, -100, 0]^T$	$[-91.2, -97.7, -0.0089]^T$	$[-100, 1728.8, 0]^T$	$[-100.4, 1719.4, 0.0182]^T$
Trial 4	$[-100, -100, 0]^T$	$[-89.2, -98.8, -0.0077]^T$	$[-100, 2338.4, 0]^T$	$[-98.4, 2334.9, 0.0141]^T$

Table 5.1 The approximated and calibrated poses of the sensor units.

After the sensor calibration process is completed, the experimental procedure involves placing the landmark in the same way as described above with an interval of two tiles, or 609.6 mm, on both x and y axes, totaling 90 positions over the testing area. The positions used during the experiment may overlap with the positions used for the sensor calibration process, but the measurements in the two processes are taken separately. Since the measurements are taken while the landmark is stationary, the requirement on the exact timestamps of the measurements is dropped. At each position of the landmark, both angular and distance measurements are collected continuously for 10 times, and the effects of reducing uncertainties by using multiple measurements are also analyzed in this test.

5.1.2 Results and Analysis

Both individual and collaborative position estimates of a stationary landmark of all four trials are shown in Figure 5.3 to Figure 5.6, respectively. The poses of the two sensor units are also plotted on the left of the figures. For each location of the landmark, the ten individual position estimates are plotted as dots using the same color (green or blue) as the corresponding sensor units, while the collaborative position estimates, which leverage raw measurements from the two sensor units, are plotted as red dots. Due to some unexpected measurements, there are a few position estimates far away from the

corresponding ground truth positions. Position estimates with a distance to corresponding ground truth larger than 304.8 mm are considered as outliers, and the outliers are connected to their ground truth locations by straight dashed lines, in order to keep track of their actual association with a specific location of the landmark.

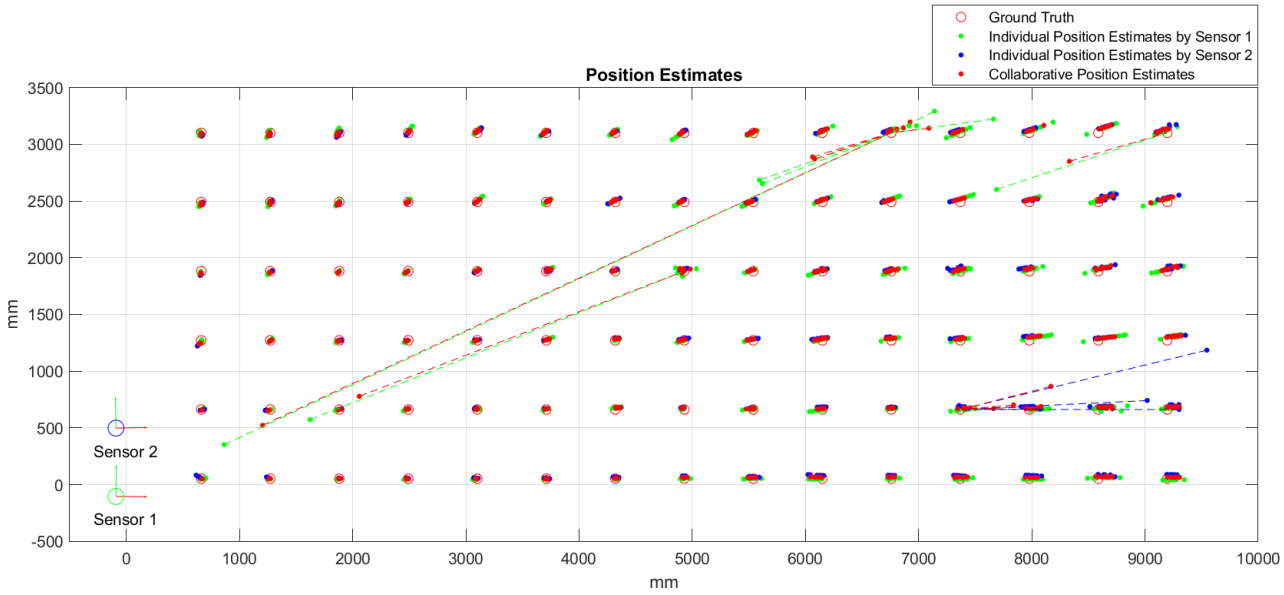


Figure 5.3 Position estimates of a stationary landmark in *Trial 1*.

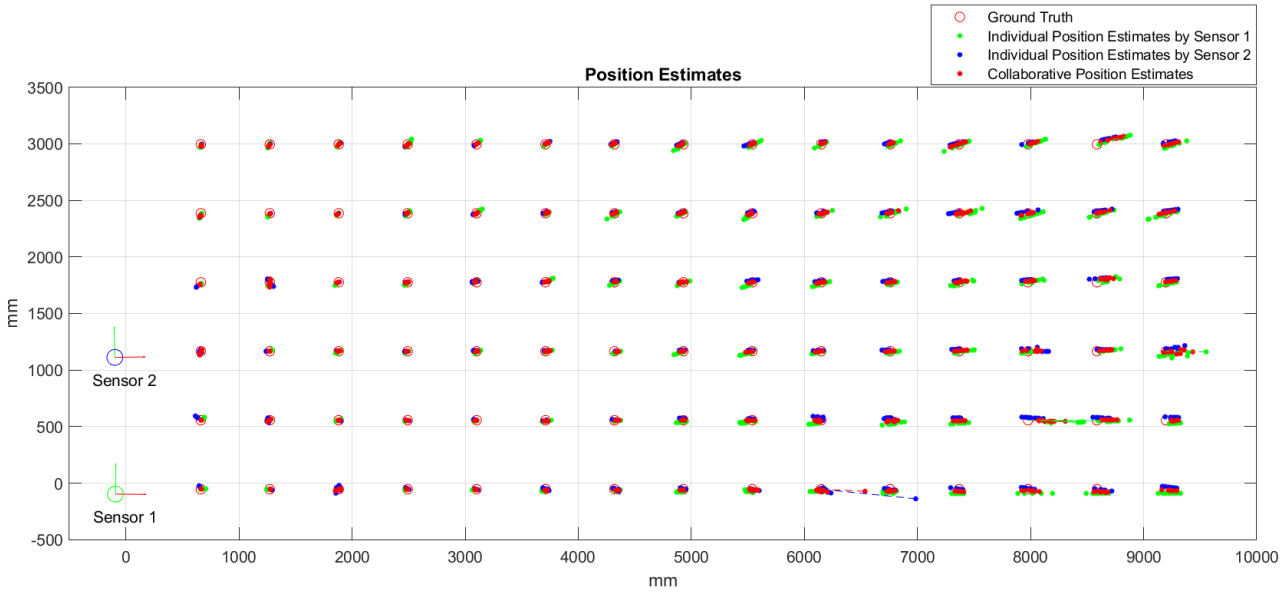


Figure 5.4 Position estimates of a stationary landmark in *Trial 2*.

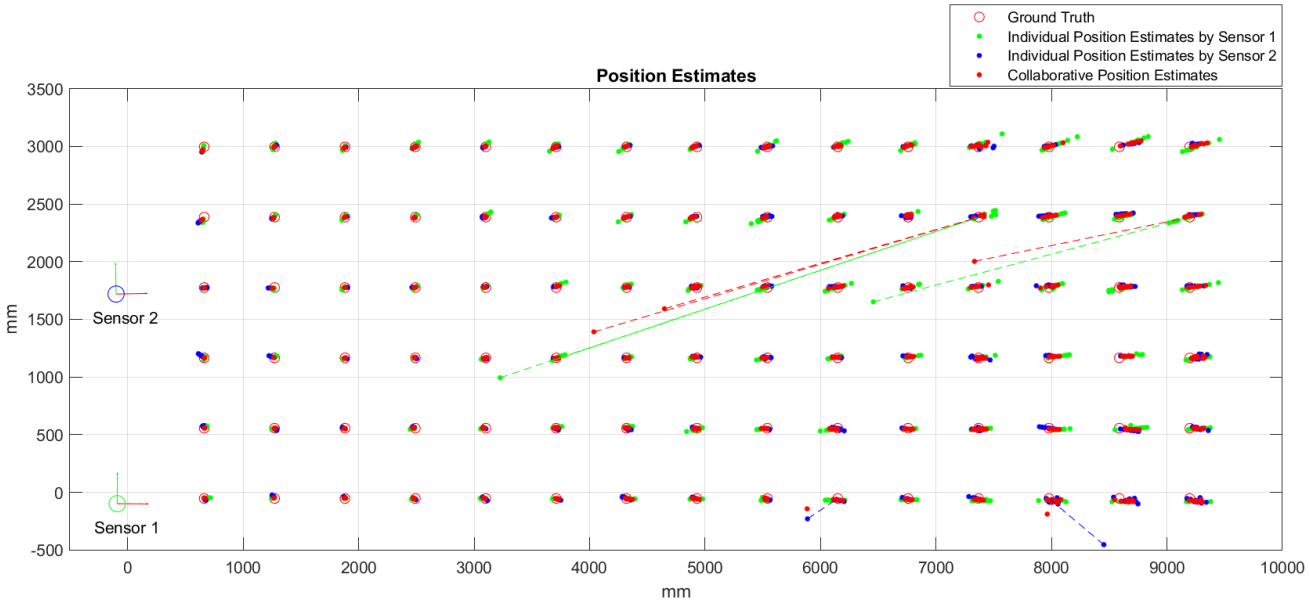


Figure 5.5 Position estimates of a stationary landmark in *Trial 3*.

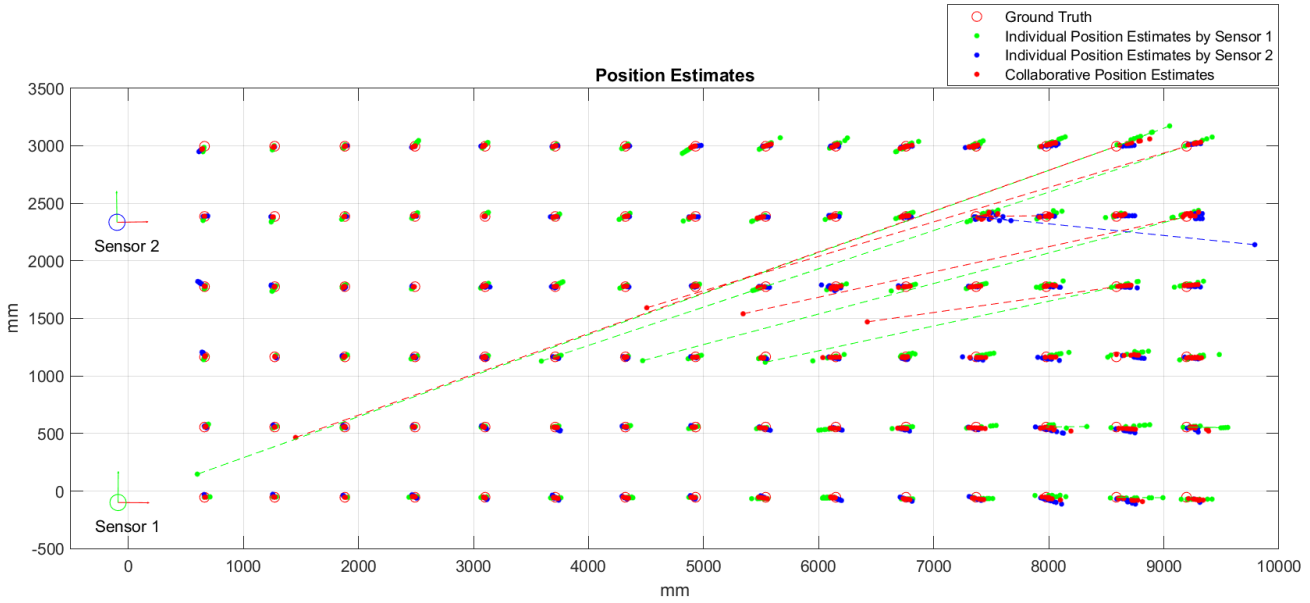
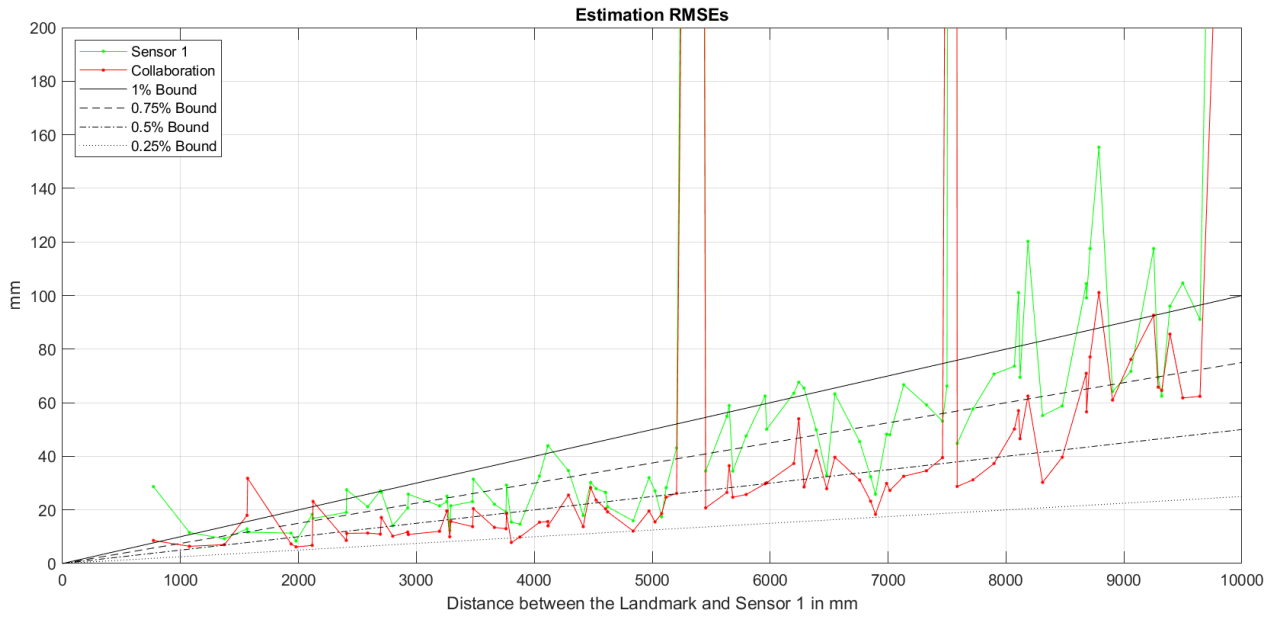


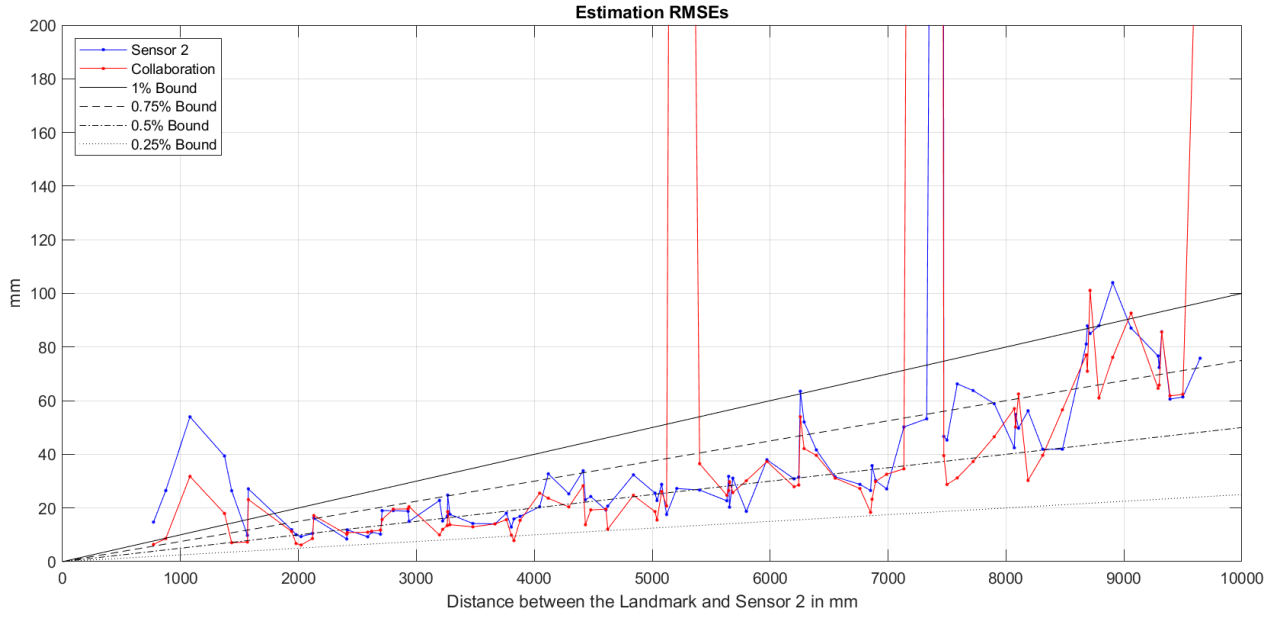
Figure 5.6 Position estimates of a stationary landmark in *Trial 4*.

The RMSEs of all 10 estimates at each position from both the individual and collaborative landmark position estimations in the four trials are plotted in Figure 5.7 to Figure 5.10, respectively. In the figures, the RMSEs are rearranged according to the distance from the actual position of the landmark to each sensor unit, and the RMSEs of collaborative position estimates are plotted with respect to each sensor unit

respectively in two subfigures. Four percentage error bounds, from 0.25% to 1% with an interval of 0.25%, are plotted as straight lines, $y = px$, where p is a particular percentage and x is the distance separating the landmark from the sensor. The figures are zoomed in on the range of the vertical axis from 0 to 200 mm, neglecting RMSEs larger than 200 mm, which are likely contributed by outliers.

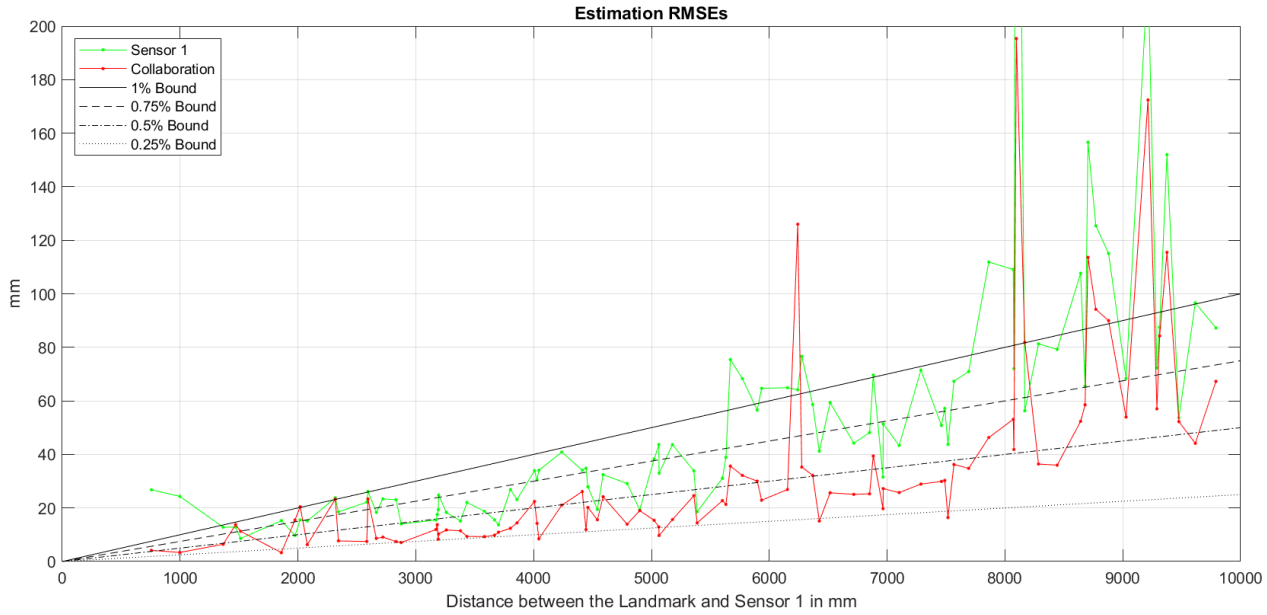


(a)

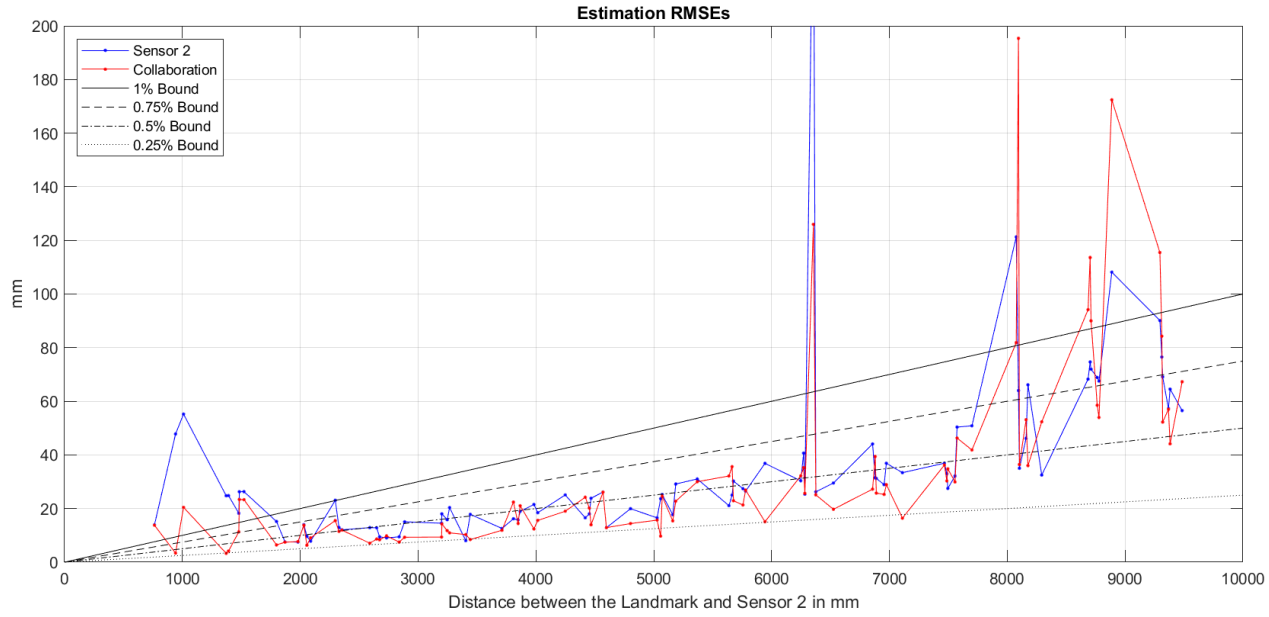


(b)

Figure 5.7 Position estimation RMSEs comparison between individual and collaborative landmark position estimations in *Trial 1*. (a) *Sensor 1;* (b) *Sensor 2*.

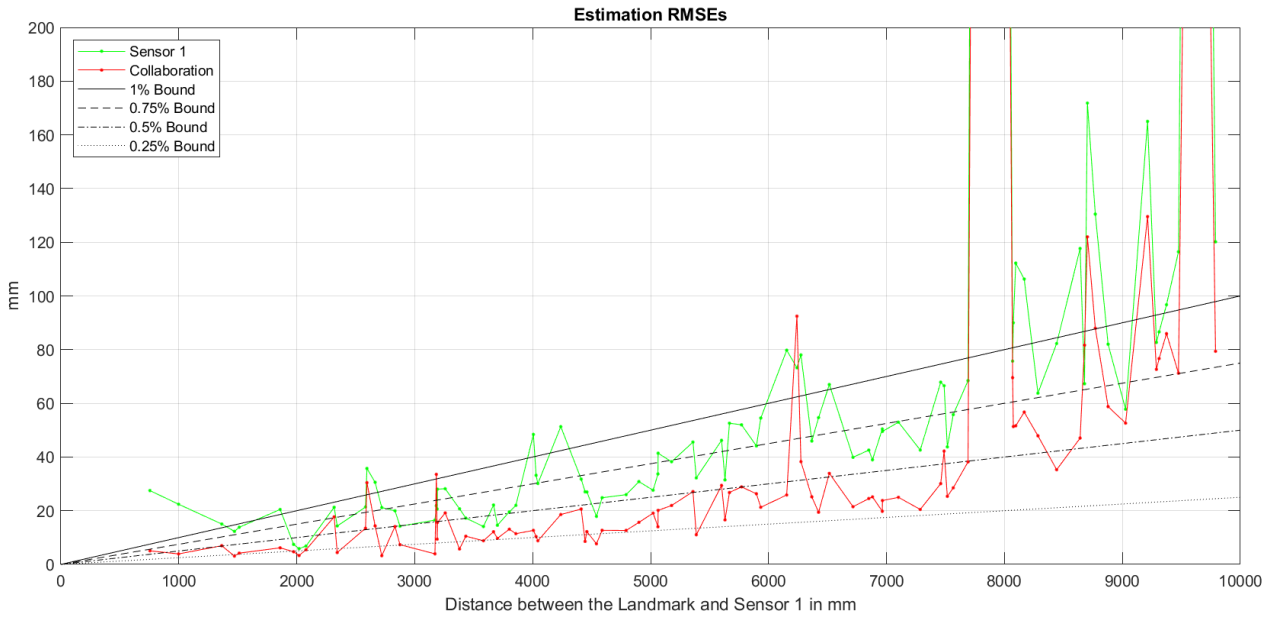


(a)

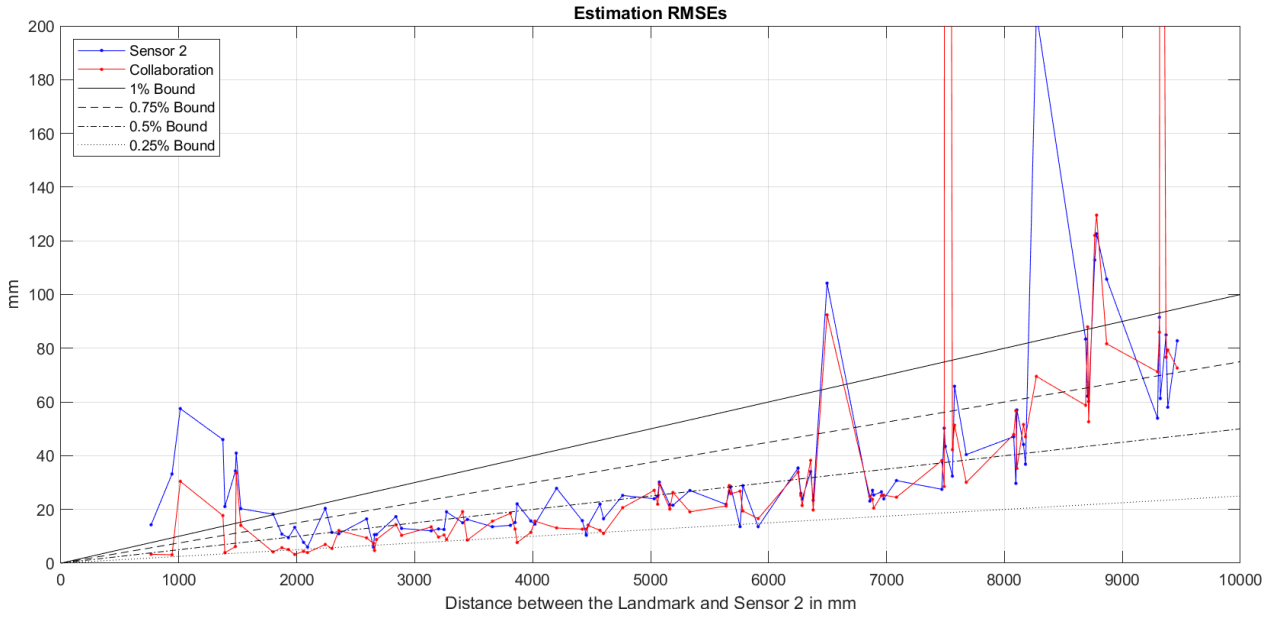


(b)

Figure 5.8 Position estimation RMSEs comparison between individual and collaborative landmark position estimations in *Trial 2*. (a) *Sensor 1;* (b) *Sensor 2*.

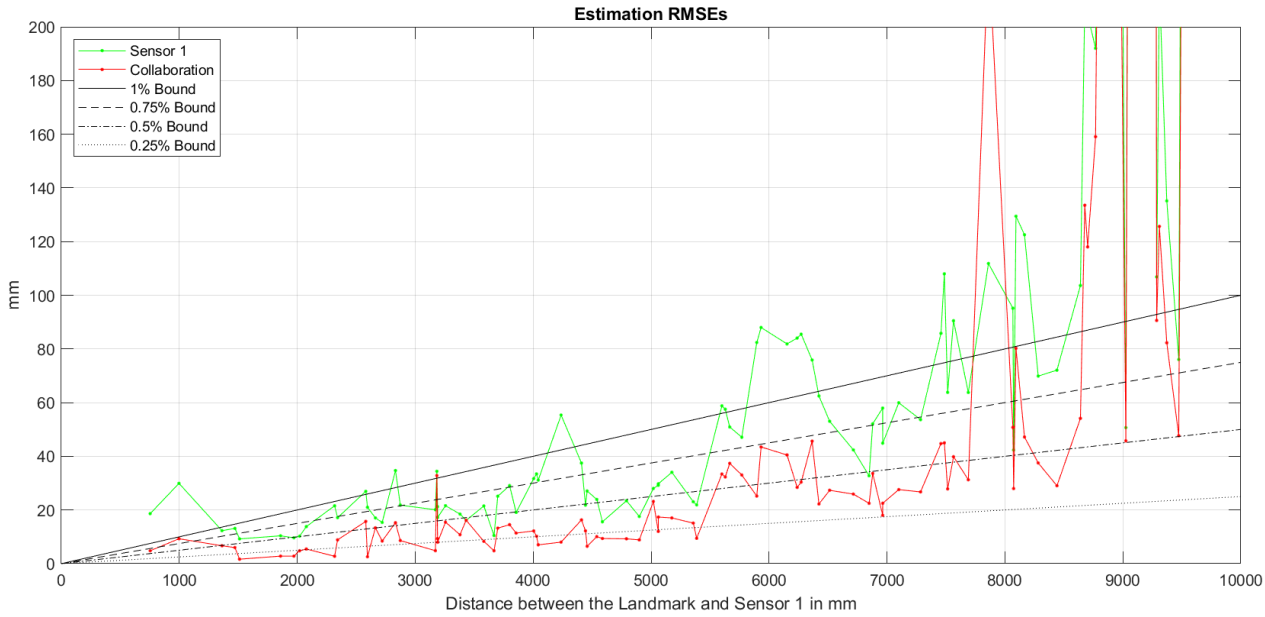


(a)

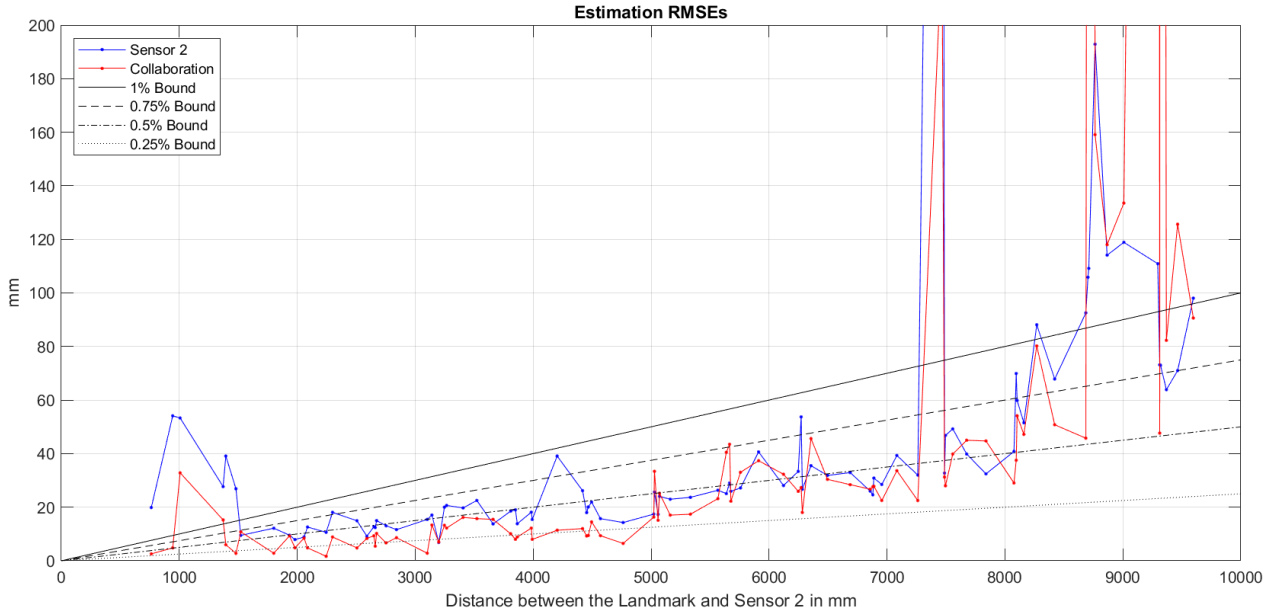


(b)

Figure 5.9 Position estimation RMSEs comparison between individual and collaborative landmark position estimations in *Trial 3.* (a) *Sensor 1;* (b) *Sensor 2*.



(a)



(b)

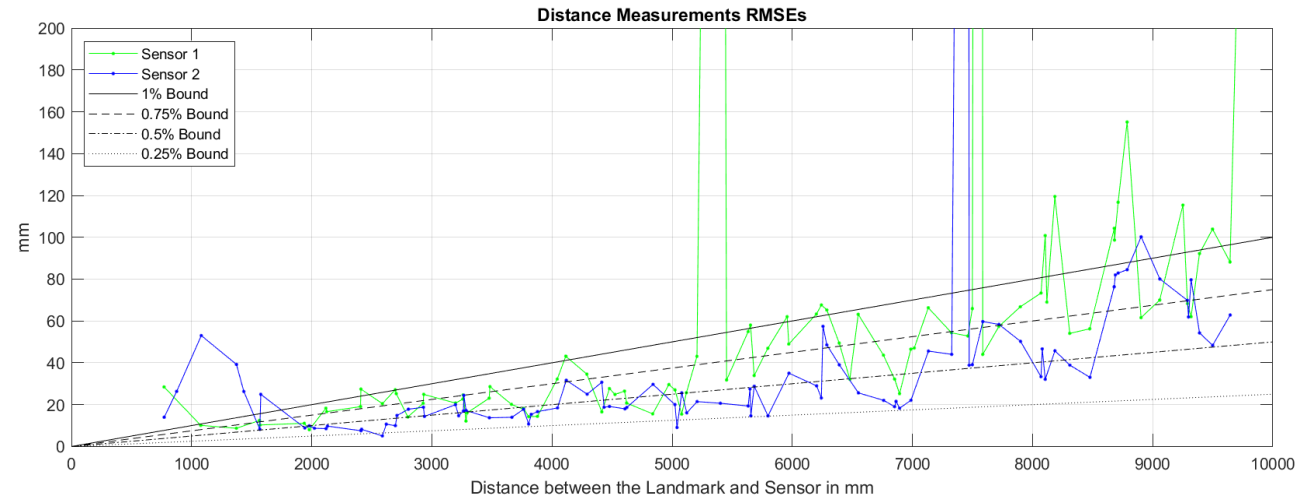
Figure 5.10 Position estimation RMSEs comparison between individual and collaborative landmark position estimations in *Trial 4*. (a) *Sensor 1;* (b) *Sensor 2*.

As observed on the above figures, the position estimates made by *Sensor 1* exhibit a larger variance than those by *Sensor 2*, and the position estimation RMSEs from *Sensor 1* mostly range from 0.5% to 1%, while the position estimation RMSEs from *Sensor 2* range from 0.25% to 0.75% with most RMSEs around

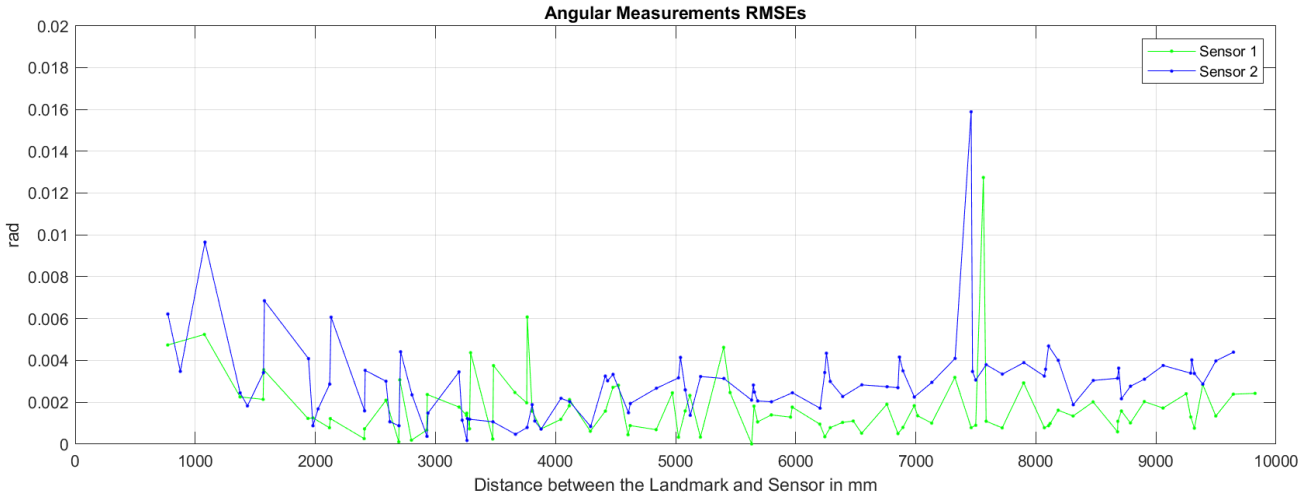
0.5%. The collaborative position estimates, as the weighted average of individual position estimates, clearly show a more concentrated distribution and appear closer to the ground truth values at many positions than the individual position estimates by either sensor units, except for the occasions caused by outliers. The RMSEs of the collaborative position estimates generally range from 0.25% to 0.5%, and a slight decrease in the RMSEs can be noticed as the distance between the two sensor units increases.

At about 1000 mm, both sensor units exhibit consistently uncommon position estimation errors. The standard deviations of the distance measurements from both sensor units are around their minimum values, according to their characteristics provided in Section 3.4.2, and the corresponding estimates in Figure 5.3 to Figure 5.6 show clear biases on the distances between a sensor unit and its estimates. Similar biases can be observed on *Sensor 1* at around 4000 mm and 6000 mm. At distances larger than 8000 mm, both sensor units show large estimation RMSEs, but the appearance of these RMSEs are less consistent than those within 8000 mm. Notable positive biases on distances between position estimates and corresponding sensor units can be observed in Figure 5.3 to Figure 5.6 at long distances, especially those from *Sensor 1* in *Trial 4*.

The RMSEs of the ten raw distance and angular measurements collected at each position in each trial are plotted in Figure 5.11 to Figure 5.14 in order to further analyze the position estimation errors. The distance measurements RMSEs are also plotted against the same percentage error bounds as those in the position estimation RMSEs.

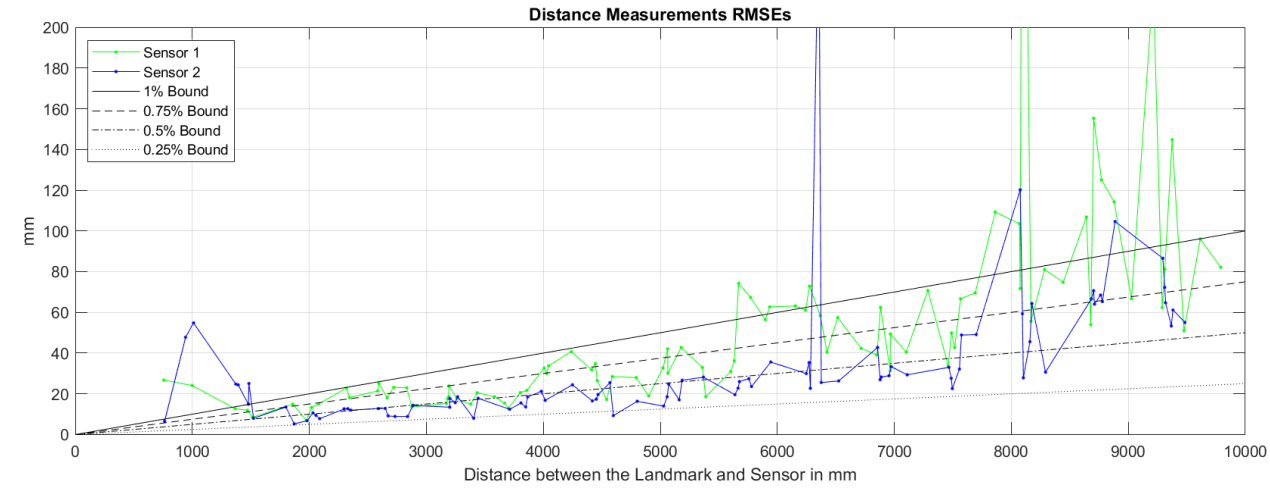


(a)

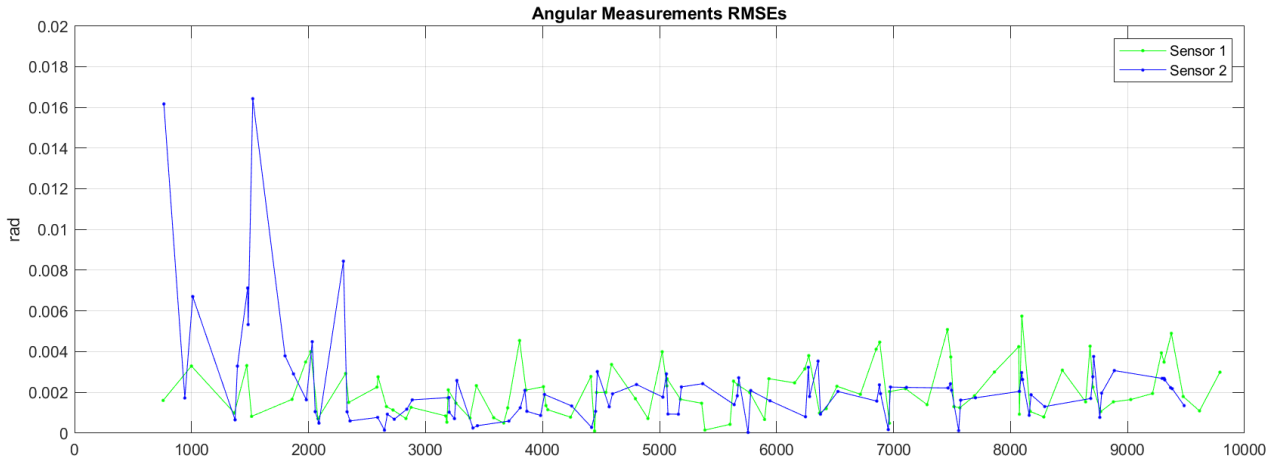


(b)

Figure 5.11 Raw measurements RMSEs in *Trial 1*. (a) Distance measurements RMSEs; (b) Angular measurements RMSEs.

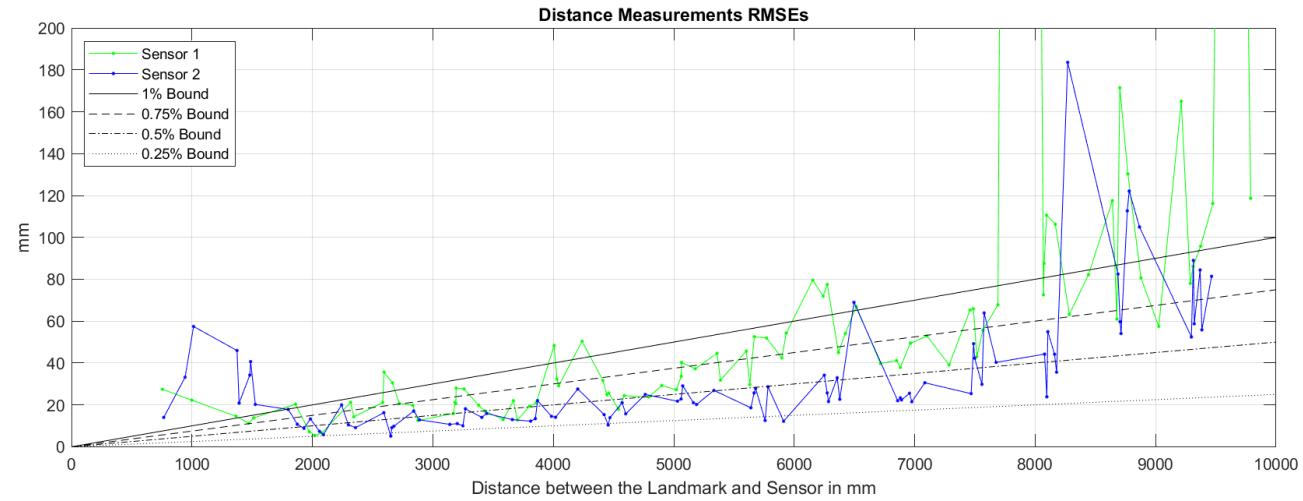


(a)

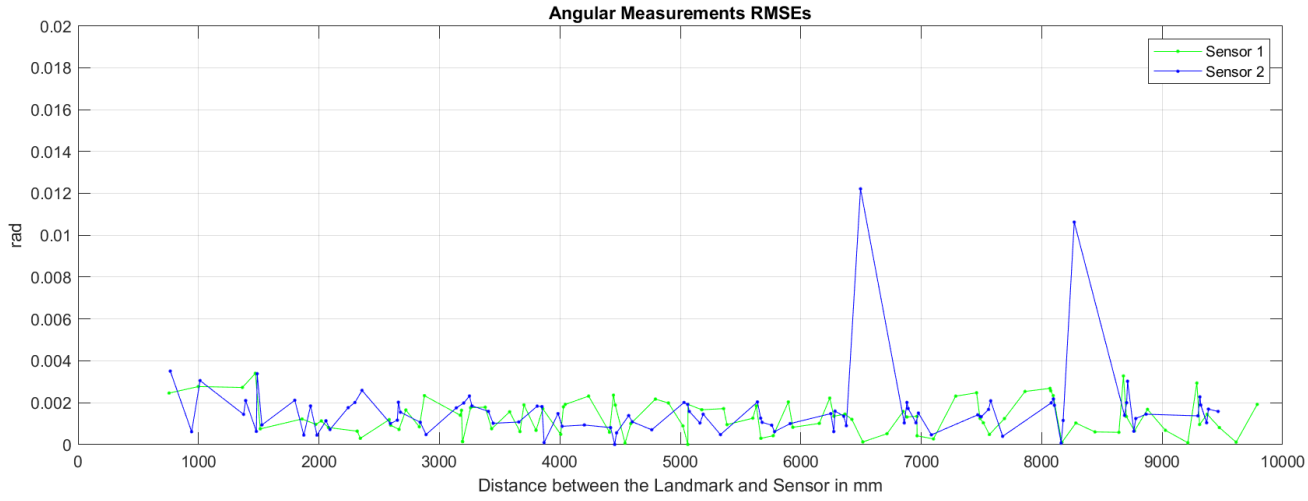


(b)

Figure 5.12 Raw measurements RMSEs in *Trial 2*. (a) Distance measurements RMSEs; (b) Angular measurements RMSEs.

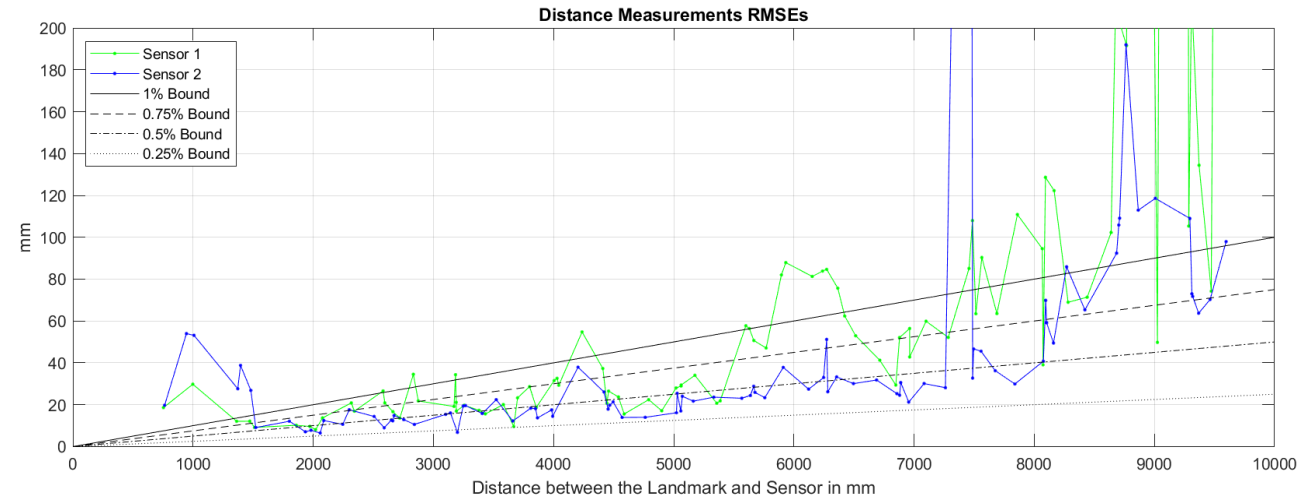


(a)

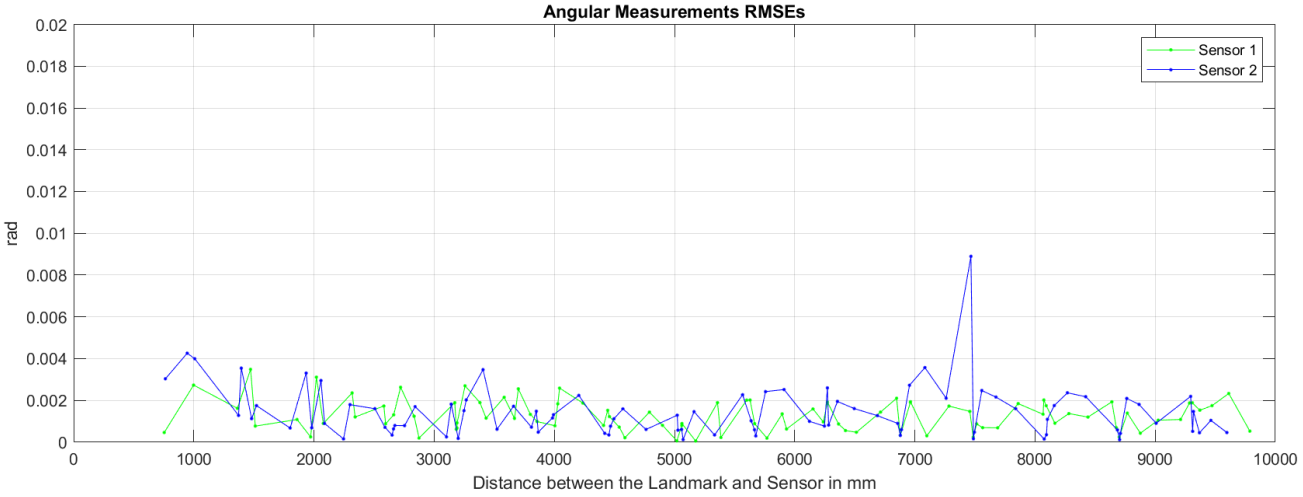


(b)

Figure 5.13 Raw measurements RMSEs in *Trial 3*. (a) Distance measurements RMSEs; (b) Angular measurements RMSEs.



(a)



(b)

Figure 5.14 Raw measurements RMSEs in *Trial 4*. (a) Distance measurements RMSEs; (b) Angular measurements RMSEs.

The angular measurements RMSEs from the two sensor units are approximately within the same range of 0 to 0.004 rad except for occasional large values. Noticeably, the angular measurements RMSEs are two times larger than the resolution of the servo motor, $\pi/2048 \approx 0.0015$ rad. But since the poses of the sensor units are not perfectly accurate but calibrated as described in Section 3.4.3, the errors resulting from the sensor calibration process also contribute to the angular measurements RMSEs as biases along with the angular measurement error, which is characterized as Gaussian white noise with a varying

standard deviation within 0.003 rad in Section 3.4.2. Therefore, the angular measurements RMSEs are considered to fall within an acceptable range.

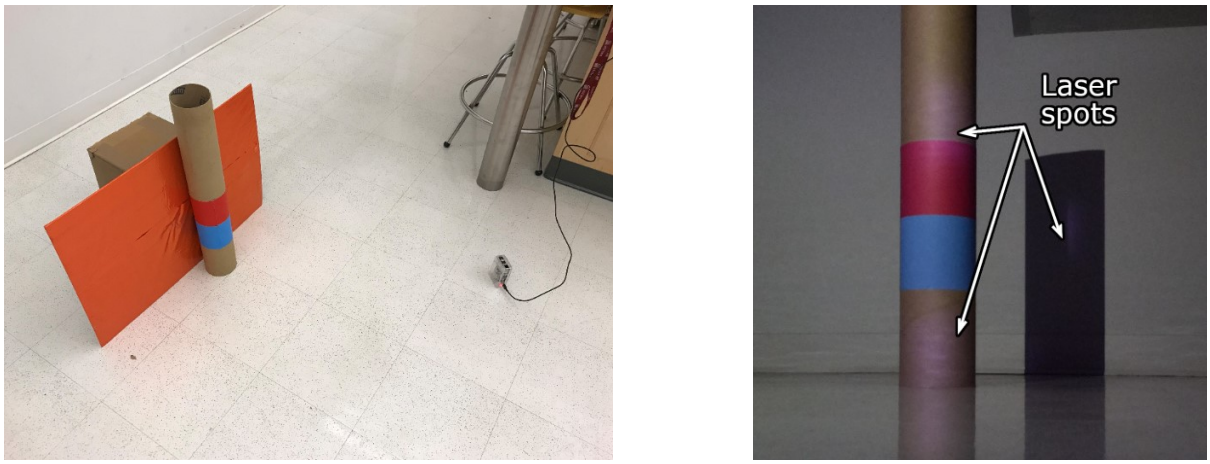
The plots of the distance measurements RMSEs for each sensor unit resemble those of the individual position estimation RMSEs and their ranges are almost the same. Clearly, the position estimation errors are majorly contributed by the raw distance measurement errors, which explains the biases observed earlier on the position estimates and the corresponding RMSEs. It is also revealed that the different performance of the two sensor units on position estimation is mainly caused by distance measurement errors. As discussed in Section 3.4.2 during the sensor characterization, the LiDAR on *Sensor 1* has higher standard deviations and thus possibly larger errors than the LiDAR on the other unit.

The estimation biases at distances such as 1000 mm are caused by the selected linear LiDAR model extracted from the sensor characterization provided in Section 3.4.2, where the distance measurement error is modeled as Gaussian white noise. However, the expected distance measurement error may not always be exactly zero, resulting in estimation biases at several distances. The biases can be corrected by further leveraging the collected measurements during the characterization process, but the linear model is considered here as sufficient for experimentally validating the concept of the proposed localization sensor. Besides, a more accurate characterization of the LiDARs used on the implemented sensor units requires a more thorough survey of LiDAR measurements, which can be time-consuming and inefficient.

The inconsistent estimation biases at distances longer than 8000 mm are due to the nature of the LiDARs and the testing environment. As the laser beam emitted by the LiDAR devices has a divergence of 0.008 rad, its dispersion increases over distance. Also due to the existence of angular measurement errors, the cylindrical landmark used in this experiment occasionally fails to fully reflect the laser beam at longer distances, where the laser beam reaches beyond the landmark and casts on background surfaces, which contributes to the positive biases. The effects of the laser beam casting only partially on the

landmark are examined in a setup shown in Figure 5.15(a). The landmark is placed at about 9 meters away from both sensor units, and the orange flat panel is placed behind the landmark during the test, supported by a cardboard box, to block the laser beam from casting on the back wall. The raw distance measurements recorded directly from the two LiDARs during the test are shown in Figure 5.16, where the measurements are in centimeters.

Initially, the landmark is placed in position without the panel and the cardboard box, and an image showing the laser beam cast on both the landmark and the wall behind it is taken by a Raspberry Pi NoIR camera, as shown in Figure 5.15(b), where the lights in the lab environment are turned off for capturing the laser spots, while the sensor units are running as usual with the landmark lightened up by a flashlight from a cell phone. At about 1990 s in Figure 5.16, the flat panel is placed behind the landmark, and it is removed at about 2020 s. Over the period, the level of the raw distance measurements from each sensor unit is slightly reduced as shown in Figure 5.16. This demonstrates that when the orange panel is present, it blocks the laser beam emitted by the sensor units and acts as an object at approximately the same position as the landmark to the sensor units. The orange panel is covered by a reflective material, and thus the laser beam cast on the panel will either be reflected back to the LiDARs, providing more faithful measurements to the landmark, or be reflected in other directions without affecting the distance measurements at all. Therefore, since the large biases caused by the landmark partially reflecting the laser beam appear primarily at distances longer than 8000 mm, where the landmark is closer to the wall behind it, the implemented localization sensor units are deemed reliable with a working distance up to 8 meters. Using landmarks with a larger diameter could help extend the working distance of the implemented sensor units but the localization device would become bulkier and more cumbersome. Another solution is to use LiDAR devices with lower divergence on their projected beam, which would better preserve the usability of the localization sensor on small mobile robots.



(a) (b)

Figure 5.15 (a) Setup for examing the effects of laser beam only partially casting on the landmark; (b) laser beam cast on both the landmark and the wall behind it, captured by a Raspberry Pi NoIR camera.

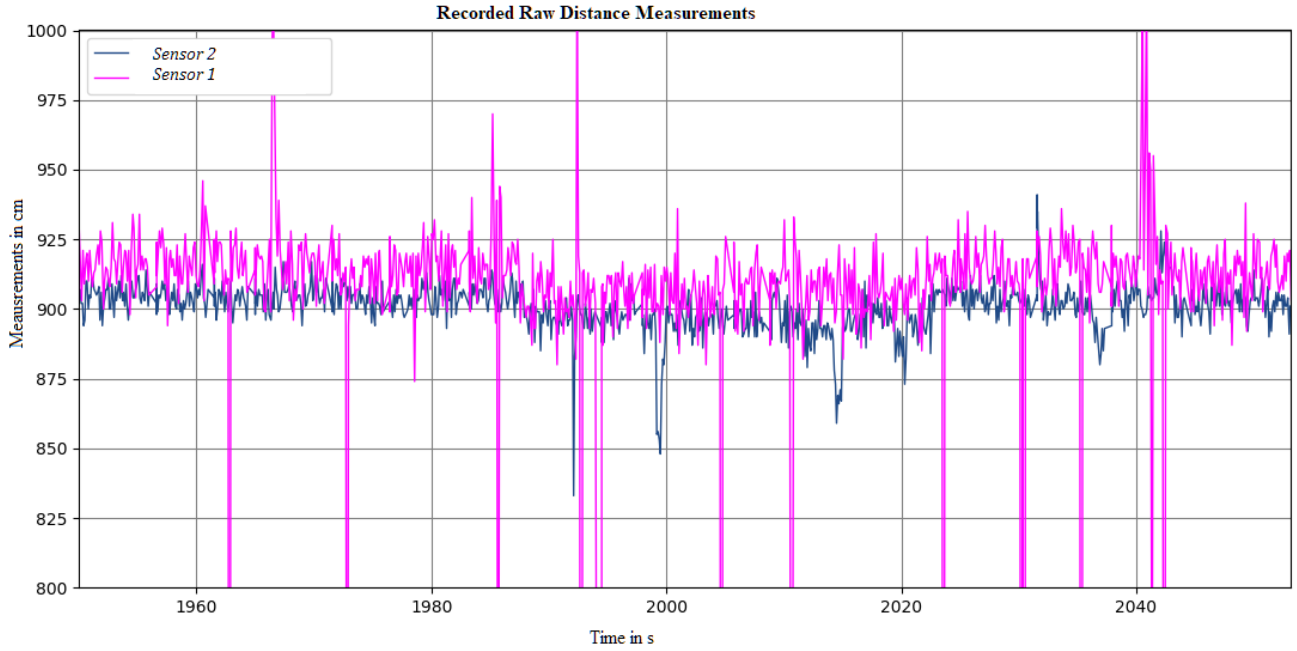
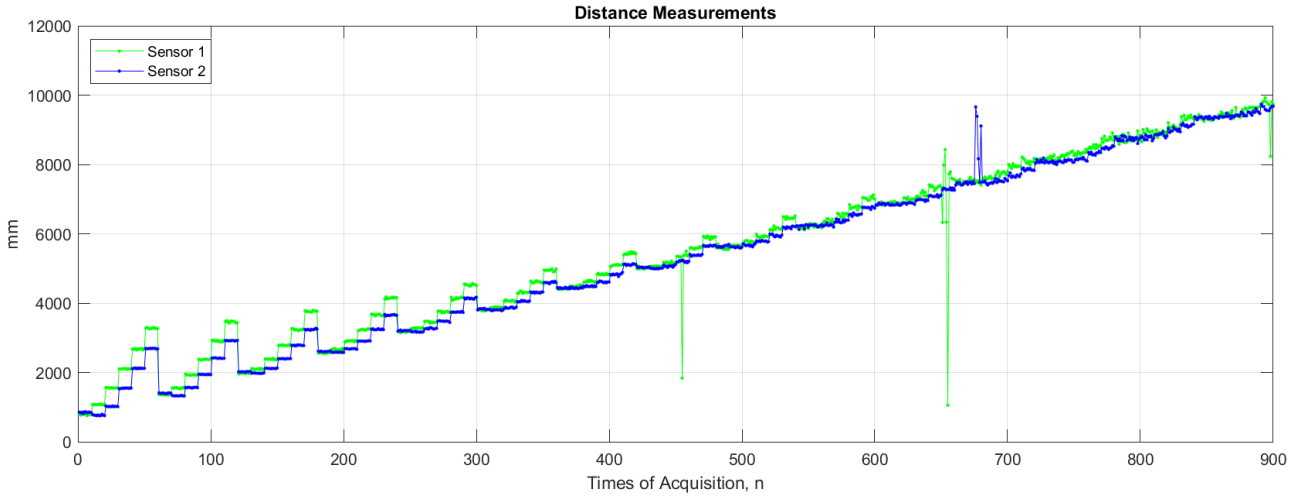


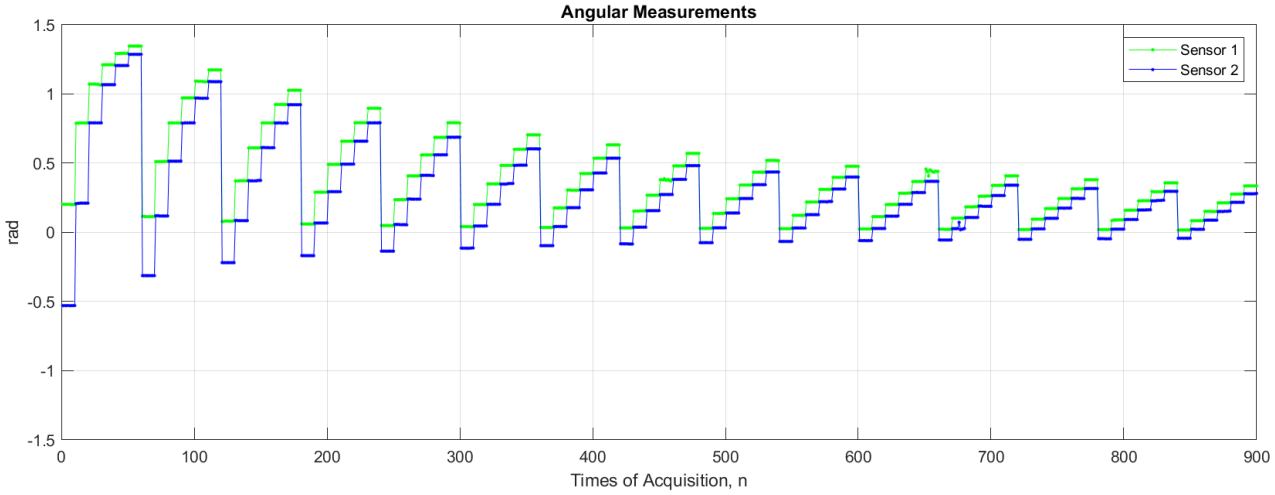
Figure 5.16 Recorded raw distance measurements while examining the effects of laser beam only partially casting on the landmark.

It is noticeable that *Sensor 1* tends to generate more outliers than *Sensor 2*, and in Figure 5.11 to Figure 5.14 depicting the measurements RMSEs, large distance measurements RMSEs are often accompanied by large angular measurements RMSEs. The raw distance and angular measurements are therefore plotted in Figure 5.17 to Figure 5.20 for each trial, respectively. The landmark is manually placed

over 90 positions in the testing environment and measurements are taken repeatedly for 10 times at each position, a total of 900 distance measurements and 900 angular measurements are recorded. Instead of temporal or spatial order, the measurements are arranged in acquisition order, where the times of acquisition, n, is used for the horizontal axis in Figure 5.17 to Figure 5.20. Furthermore, as shown in Figure 5.3 to Figure 5.6, the landmark is first placed incrementally along the vertical axis for six times with the same horizontal coordinate, and the procedure is then repeated with an increased horizontal coordinate for a total of 15 times along the horizontal axis, resulting in the steps-like plots.

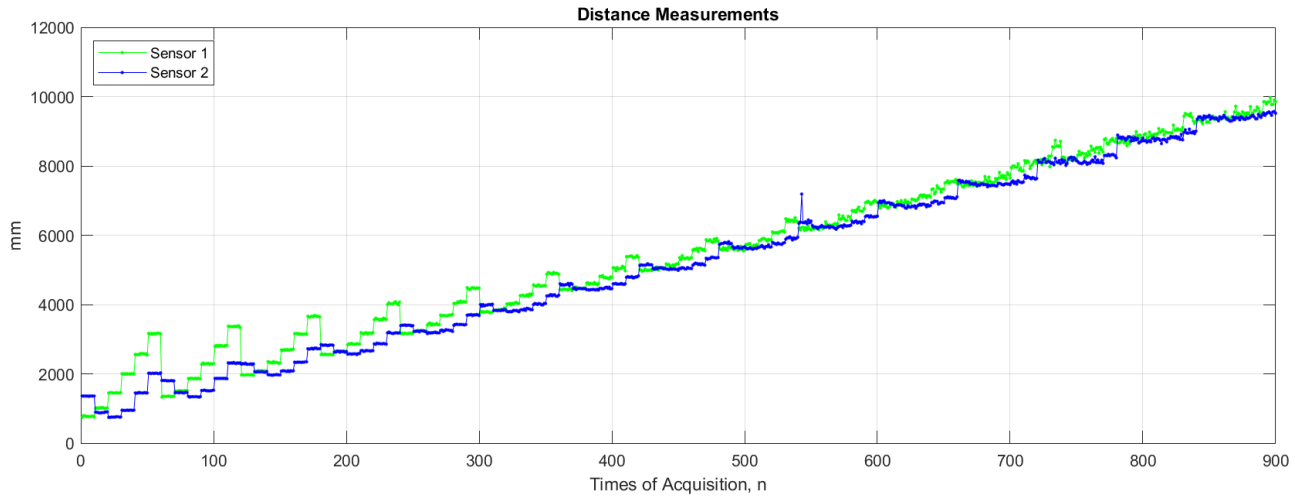


(a)

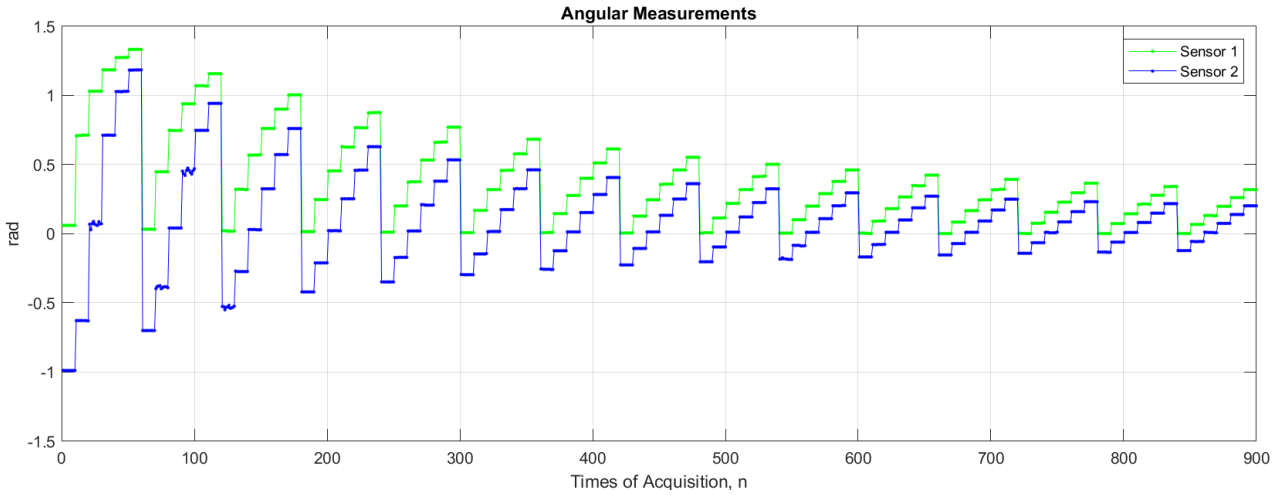


(b)

Figure 5.17 Measurements arranged in acquisition order in *Trial 1*. (a) Distance measurements; (b) angular measurements.

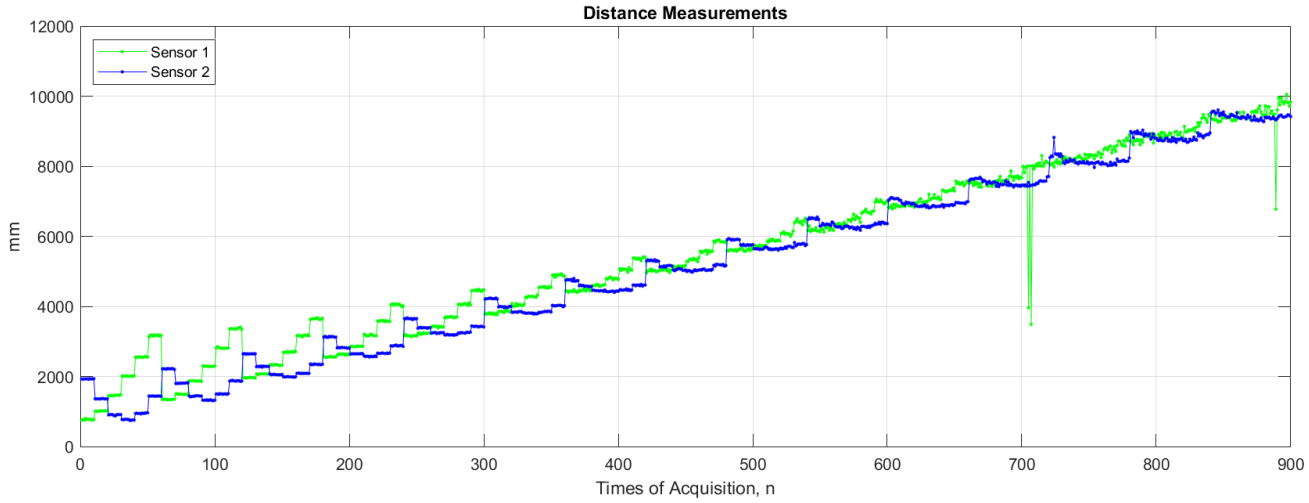


(a)

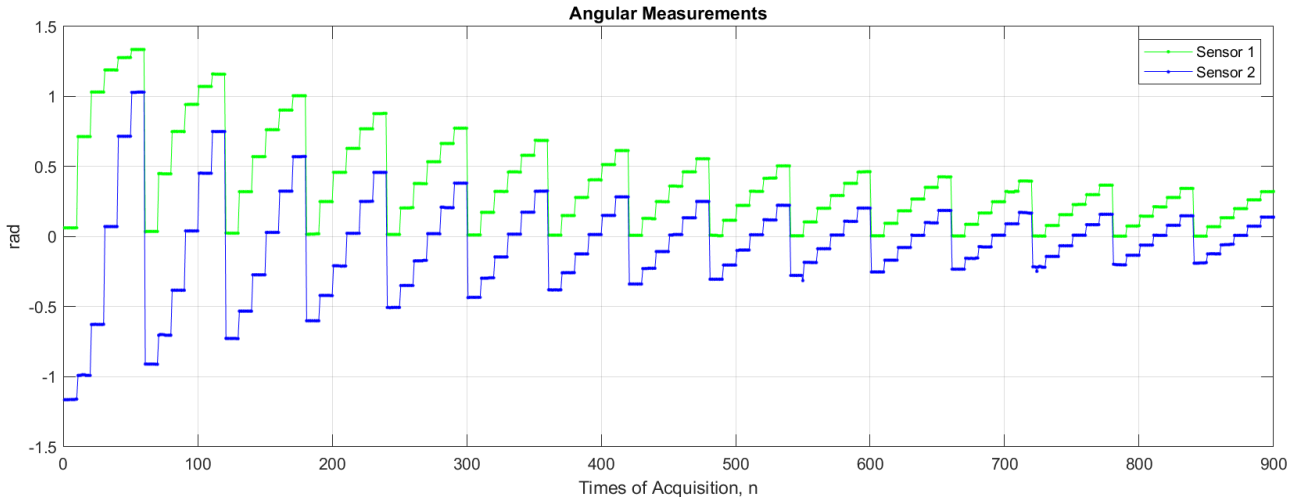


(b)

Figure 5.18 Measurements arranged in acquisition order in *Trial 2*. (a) Distance measurements; (b) angular measurements.

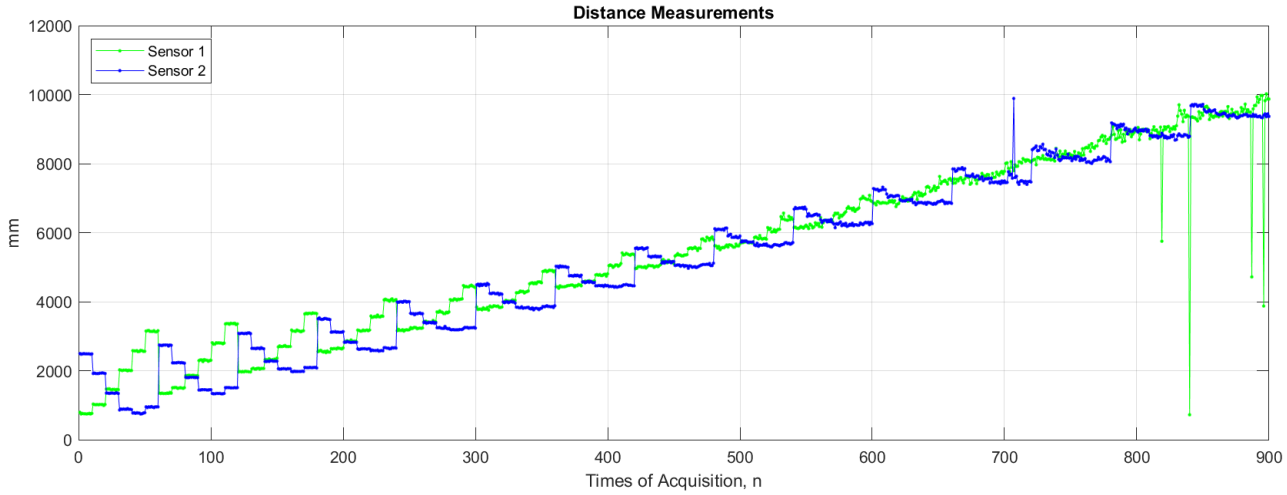


(a)

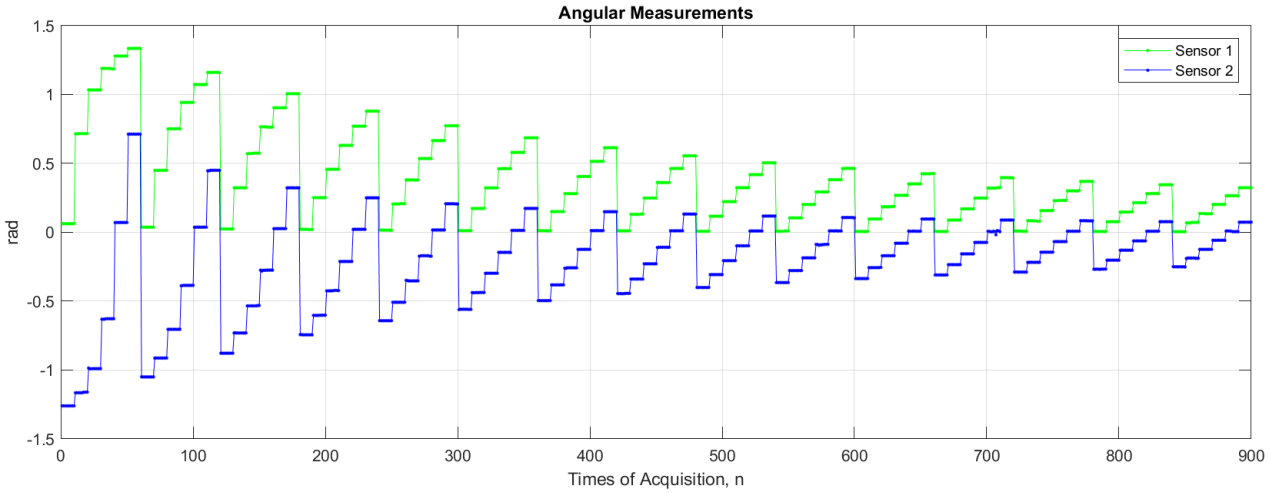


(b)

Figure 5.19 Measurements arranged in acquisition order in *Trial 3*. (a) Distance measurements; (b) angular measurements.



(a)



(b)

Figure 5.20 Measurements arranged in acquisition order in *Trial 4*. (a) Distance measurements; (b) angular measurements.

In all trials, the angular measurements are consistent with only a few disturbances on both sensor units. The distance measurements, however, have shown some unexpected large errors at several distances as spikes in the figures which directly led to the occurrence of outliers. As shown in the above figures, outliers with positive distance measurement errors are almost always accompanied by disturbances in the angular measurements, while outliers with negative distance measurement errors are only limited to *Sensor 1* and without any relation to unusual angular measurements. Based on our observations, positive

distance outliers are generated when the measurements were collected while the motor was still in motion. As a result, the laser beams emitted by the LiDARs missed the landmark, partially or completely, and the objects behind the landmark contributed to the distance measurements. In future refinement of the technology, filtering techniques such as Kalman filter could be applied to the implemented sensor units to avoid such type of outliers, and the servo controller should be further improved to provide a faster and more stable response. The negative distance outliers are mainly due to the nature of the LiDAR on *Sensor 1*. The characterization of the LiDARs in Section 3.4.2 uses a flat panel for collecting measurements, but in the experiments, the LiDARs collect measurements against a cylindrical landmark with a radius of 53 mm. While the cylindrical shape diffuses the laser beam emitted by the LiDARs to a wide range of angles, a lower fraction of the energy in the laser beam reaches back to the LiDAR device than that during the sensor characterization process presented in Section 3.4.2. When the strength of the received laser beam is lower than a pre-set threshold, a raw measurement of 1 cm is automatically generated, which becomes a smaller than usual value after the interpolation of measurements described in Section 3.3.1. To reduce this type of outliers, larger landmarks or LiDAR devices with lower beam divergence could be beneficial.

5.2 Individual Position Tracking of a Moving Landmark

The implemented localization sensor units are tested in this experiment with a landmark rigidly mounted on a mobile robot, where the landmark simulates a third localization sensor unit on a robot whose position is tracked individually by each of the two static sensor units. This experiment is to validate the capability to estimate the position of a moving sensor unit, which puts higher requirements on the response of the landmark detection and tracking algorithm and the servo motor controller. During the experiment, the ground truth of neither the position of the landmark nor the pose of the mobile robot is accessible due to the lack of precise location estimation equipment in our facilities. Therefore, in this experiment, the

landmark position estimation errors are evaluated only qualitatively. However, the usability of the implemented sensors for tracking a moving landmark and the primary sources of error are discussed.

To partially make up for not knowing ground truth location of the robot, an alternative localization method is utilized in this experiment since the mobile robot, a Turtlebot3 Waffle Pi [59], comes with a 360° 2-D laser scanner capable of sensing over a range of 0.12 m to 3.5 m and an accuracy of about ±5% at most distances. A 2-D occupancy grid map is first built by the robot using the laser scanner as shown in Figure 5.1(b) using GMapping [60] with the default settings of the robot, and then the robot is localized according to the map using an implementation of Monte-Carlo localization method [61]. The processes of map-building and robot localization use laser scan data from the 2-D laser scanner as well as the odometry data generated by the robot.

5.2.1 Experimental Setup

As shown in Figure 5.21, the setup of the localization sensor units is similar to that in the previous experiment which was conducted with a stationary landmark, and the same procedure has been carried out for calibrating the pose of the sensor units. The landmark is fixed on top of the mobile robot behind the 2-D laser scanner. As such it blocks the laser scanner from measuring distances to objects behind the robot for an angle of about 60°, where the measurements are automatically ignored since the scanner is unable to detect within 0.12 meters.

In the testing area, the robot is manually controlled to follow a zig-zag path along the floor tile lines, as depicted in Figure 5.21. The zig-zag path has multiple turns to ensure a fair level of complexity in the position estimation process.

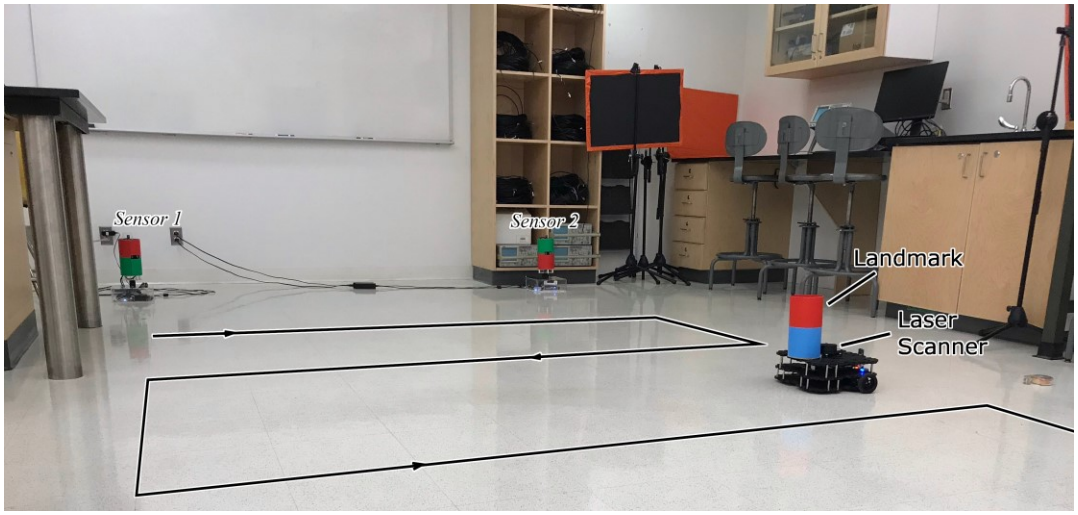


Figure 5.21 Turtlebot3 Waffle Pi robot with a landmark installed onboard.

The pose of the mobile robot, which carries the landmark to be located, is first estimated by the Monte-Carlo localization method. Then, the position estimate of the landmark is acquired by transforming the estimated position of the mobile robot that carries it with respect to the relative position of the landmark to the robot, measuring $[-220.8, 0]^T$, where both units are millimeters. In order to compare the position estimates of the landmark by the Monte-Carlo localization method and the implemented localization sensor units, the pose of the two sensor units is transformed into the reference frame of the occupancy grid map by manually measuring the geometries of the testing environment. It is worth noting that the occupancy grid map is built using all 360° measurements acquired by the 2-D laser scanner without the presence of the landmark.

In this experiment, two trials are carried out with the same setup, where the robot follows the same path but with different speeds. The robot moves along the path at a maximum speed of 0.1 m/s in the first trial, while in the second trial, the maximum speed is increased to 0.15 m/s. The speed of the mobile robot calculated from the recorded odometry data in the two trials is shown in Figure 5.22. At the corners of the zig-zag path, the robot first stops and turns around in the same position. Both localization sensor units are kept unmoved during the two trials.

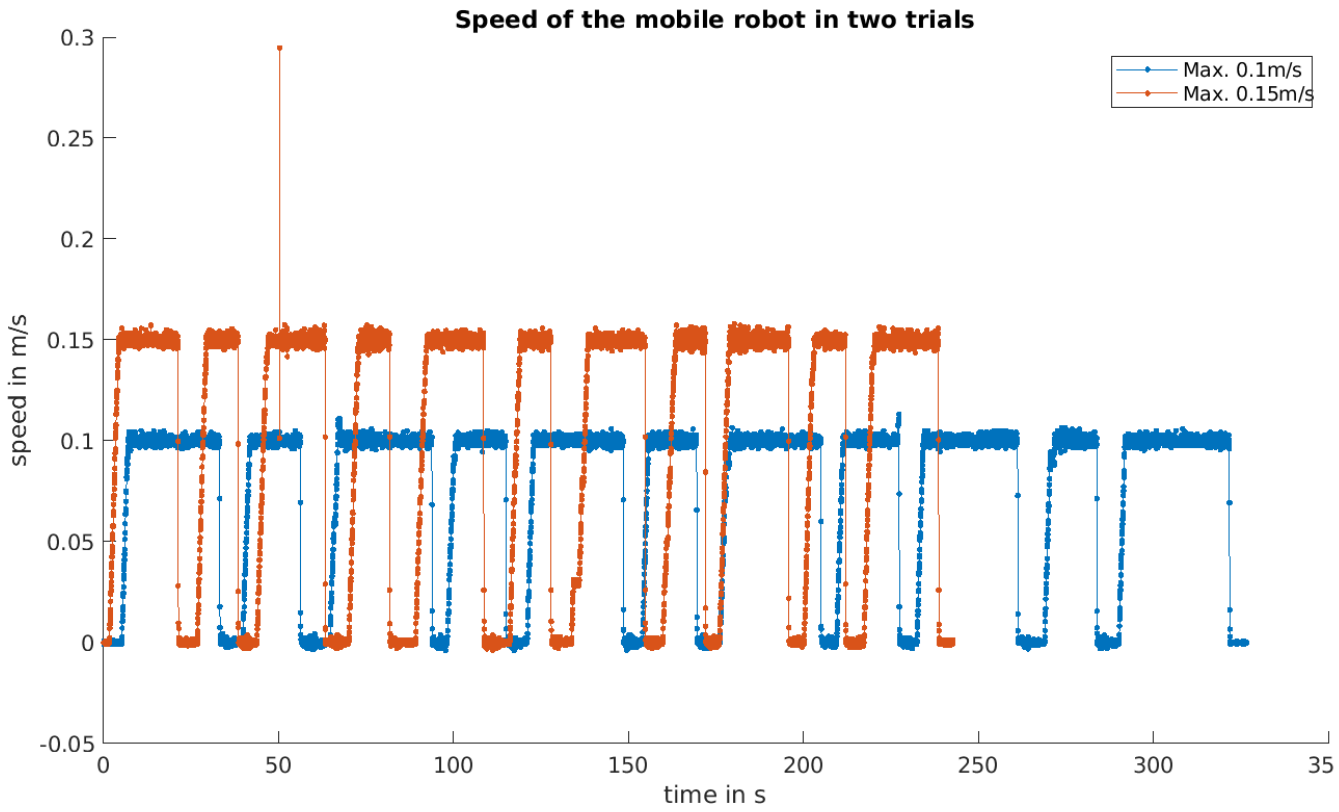
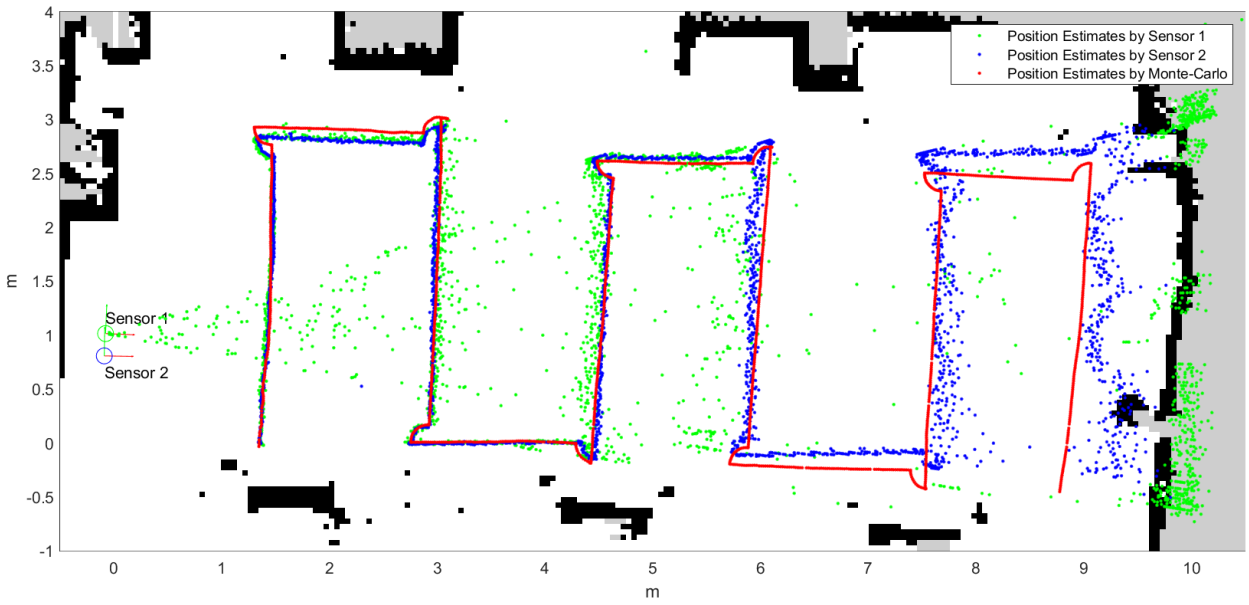


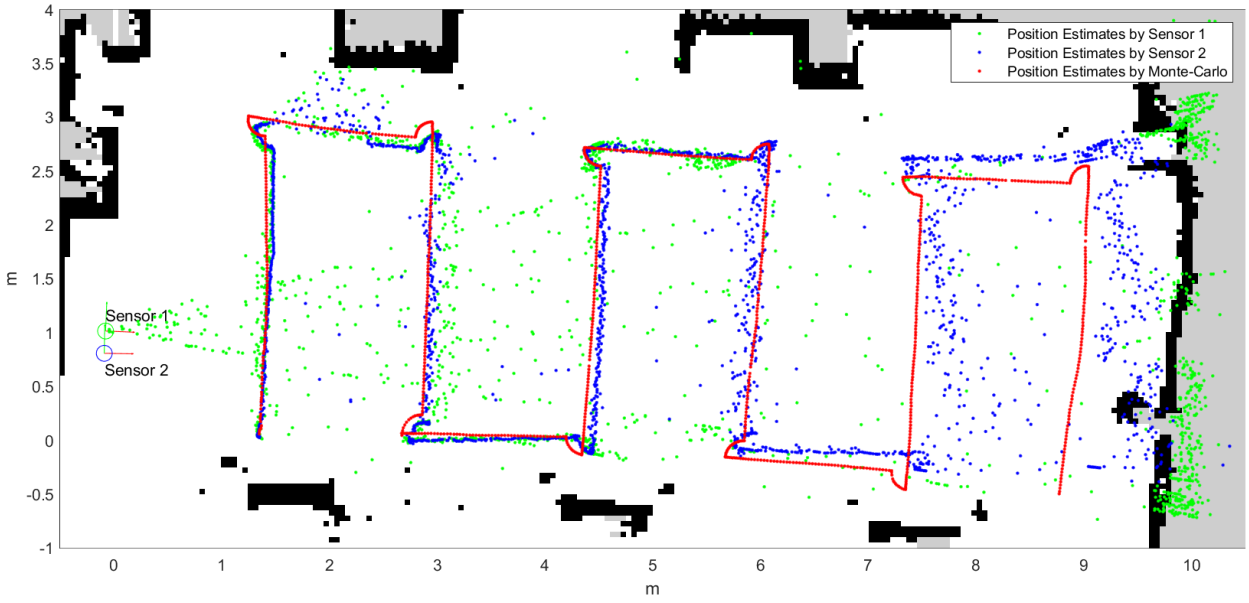
Figure 5.22 The recorded speed of the mobile robot in two trials.

5.2.2 Results and Analysis

The landmark position estimates given by the localization sensor units and the Monte-Carlo localization method are shown in Figure 5.23 along the trajectory for both trials. The localization sensor units are plotted on the left of the map, and the position estimates made by the sensor units are shown as dots with the same color as the corresponding unit. The position estimates by the Monte-Carlo localization method, based on real-time laser scan data and odometry data, are shown as red dots. In the experiment, the mobile robot starts from the lower-left corner and ends at the lower right corner on the 2-D occupancy grid map.



(a)



(b)

Figure 5.23 Landmark position estimates in two trials. (a) The first trial with the robot moving at a maximum speed of 0.1 m/s; and (b) the second trial with the robot moving at a maximum speed of 0.15 m/s.

In the above figures, the position estimates by the sensor units match with the position estimates made by the Monte-Carlo localization method at closer distances with some notable deviations. At distances longer than about 6 meters to the sensor units, the position estimates by the Monte-Carlo localization method are evidently different from the position estimates by the sensor units. Because of the

manual navigation of the mobile robot, at each turning point of the zig-zag path, the robot stops and turns accordingly. While the robot is stationary, the landmark position estimates have the same characteristics as those in the previous experiment discussed in Section 5.1.2, where the angular measurement RMSEs are mostly within the range from 0 to 0.004 rad. Therefore, the landmark position estimates by the sensor units are considered more reliable than the estimates by the Monte-Carlo localization method, within proper distance ranges under different speeds of the robot. In addition, the Monte-Carlo localization method is limited by the resolution and accuracy of the occupancy grid map, which requires an exhaustive environment survey before the localization operation, and it cannot be used in feature-less environments, such as wide open spaces without walls or objects to create shape profiles, since the method requires laser scans of the working environment to match the prebuilt map. On the contrary, the proposed localization sensor does not require any prior knowledge of, or acquisition on, the working environment, and it therefore does not rely on any existing visual or shape features in the environment. In scenarios such as search and rescue or first responders support, where the working environment may not have abundant features and is not accessible before the localization operation, methods like the Monte-Carlo localization method may fail, while the proposed localization sensor will still succeed in providing pose estimation for mobile robots.

The position estimates by the two sensor units have shown similar performance to that in the previous experiment with a stationary landmark in Section 5.1. Based on the position estimates depicted in Figure 5.23, *Sensor 1* generates more outliers than *Sensor 2*, and the position estimates of *Sensor 1* have larger variance. At long distances, positive bias can be observed on the distance measurements by both sensor units, and many position estimates by *Sensor 1* are indeed made against the far wall on the right. With the increase in the speed of the mobile robot, the variance of the position estimates by both sensor units increases as well, and the positive bias on distance measurements from *Sensor 2* at long distances

becomes larger. The raw measurements from both sensor units collected in the two trials are shown in Figure 5.24 and Figure 5.25.

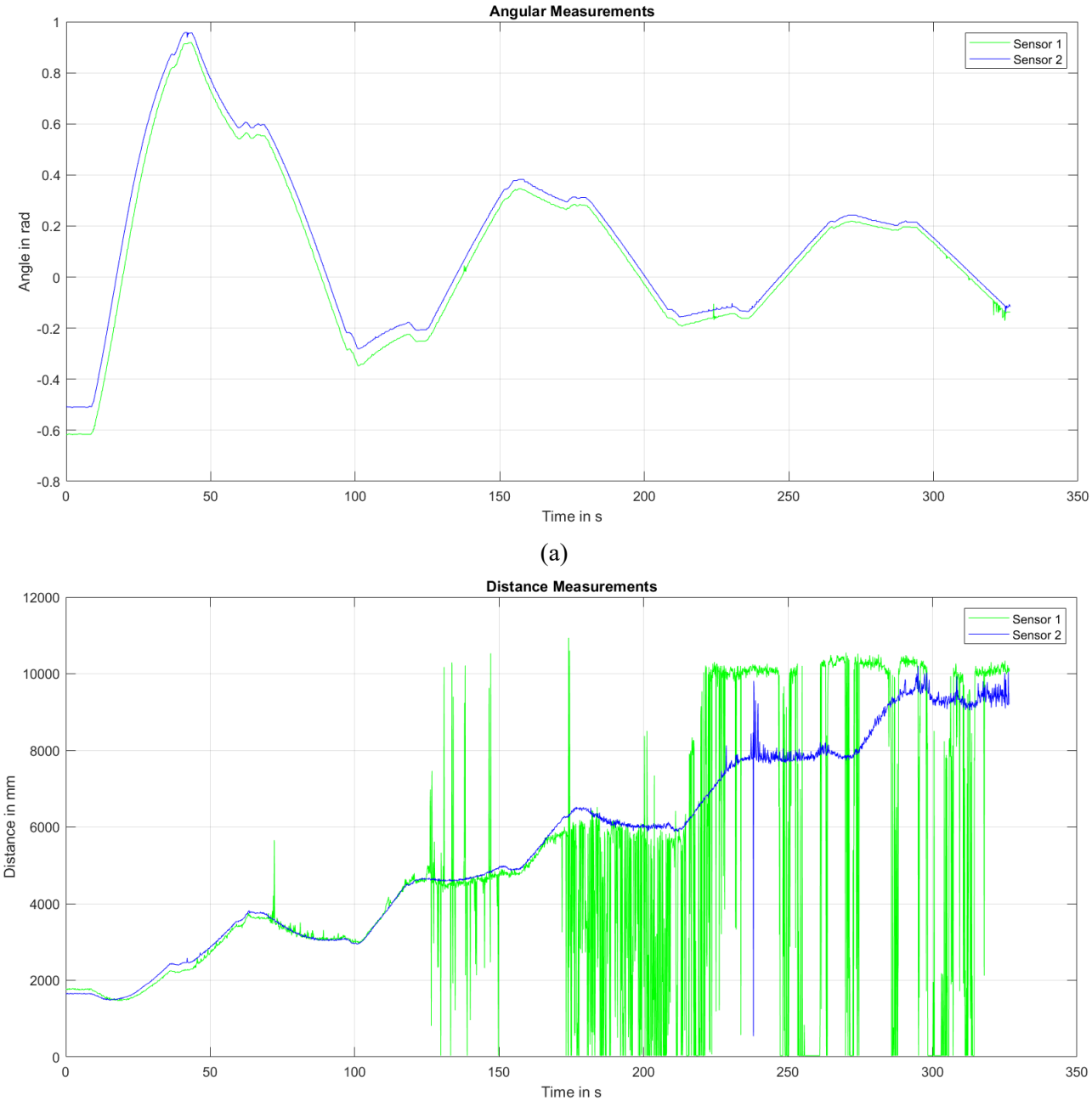
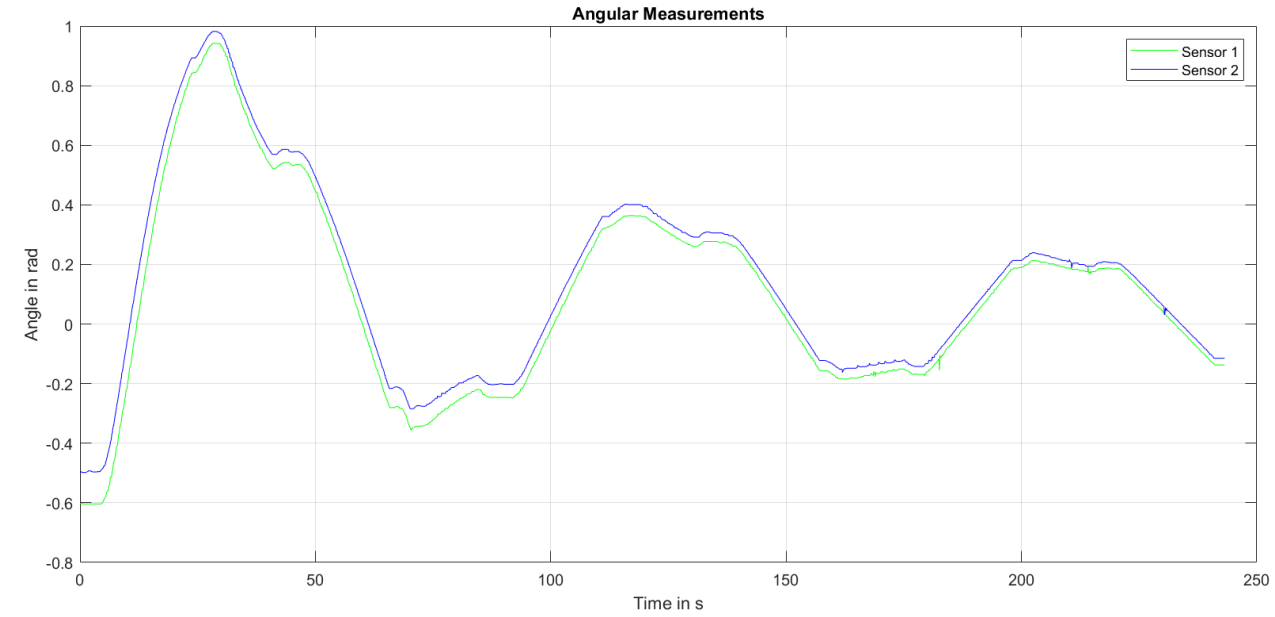
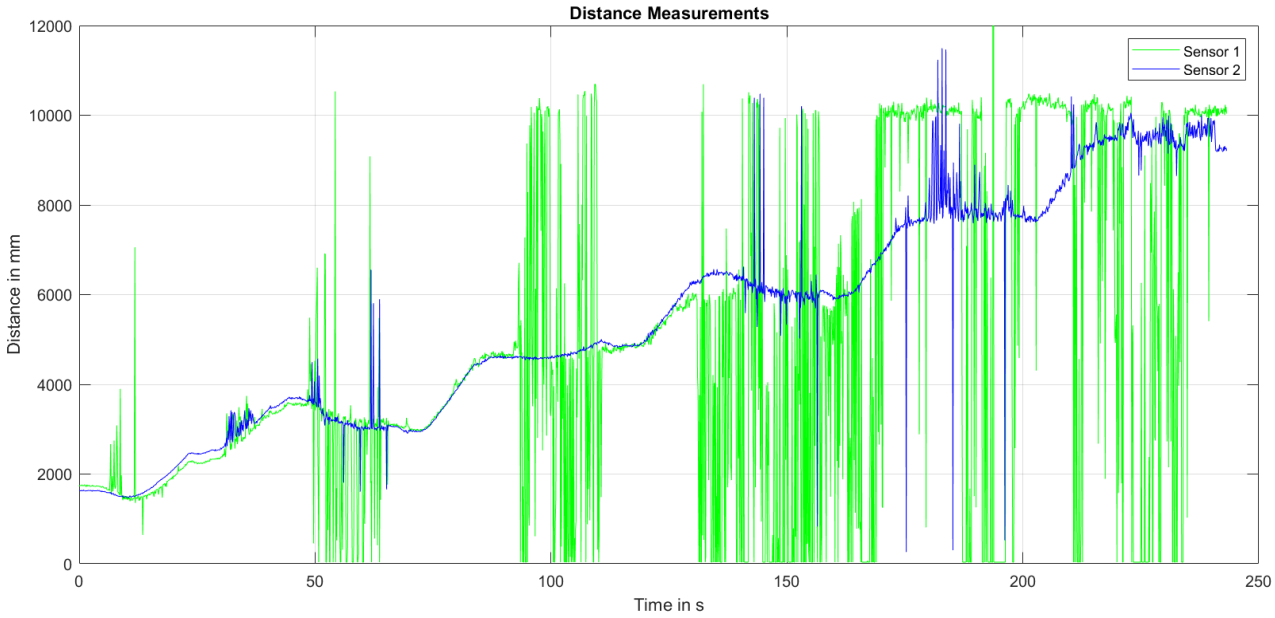


Figure 5.24 Measurements from both sensor units in the first trial with the robot moving at a maximum speed of 0.1 m/s: (a) angular measurements; and (b) distance measurements.



(a)



(b)

Figure 5.25 Measurements from both sensor units in the second trial with the robot moving at a maximum speed of 0.15 m/s: (a) angular measurements; and (b) distance measurements.

At the speed of 0.1 m/s, *Sensor 2* provides significantly more stable distance measurements than *Sensor 1*. Distance measurements from *Sensor 2* are consistent until about 8000 mm, while distance measurements from *Sensor 1* start to degrade at about 4000 mm. With the speed increased to 0.15 m/s,

distance measurements from *Sensor 1* exhibit many disturbances at shorter distances, while distance measurements from *Sensor 2* remain stable and consistent within 8000 mm with only a few disturbances at about 3500 mm. At the higher speed, the positive bias on the distance measurements, caused by the laser beam emitted by the LiDARs partially or completely missing the landmark surface, appears more frequently at long distances on both sensor units. Contrary to the distance measurements, the angular measurements from both sensor units are very stable and mostly consistent in the two trials. A few disturbances shown as small spikes can be observed in Figure 5.24(a) and Figure 5.25(a), but the disturbances can be reduced by applying filtering techniques.

The more frequent appearance of the disturbances on the distance measurements at the higher speed indicates that the response of the servo motor is slightly delayed. The time spent on image processing of the visual landmark detection and tracking algorithm is a dominant reason. As the speed of the mobile robot increases, the displacement of the robot during such processing time also increases, and the implemented servo controller fails to keep up with the moving landmark, resulting in more disturbances on distance measurements. The processing time of the visual landmark detection and tracking algorithm could be reduced if the computational power of the Jetson modules could be fully leveraged, as the current implementation does not make use of their CUDA cores and relies completely on the CPU. On the other hand, better designs of the servo controller could be applied to make the servo response more robust against potential delays. Other means, including rectifying the camera images for more accurate servo control commands, may also be helpful to improve the servo response.

It is worth noting that the difference in the processing modules of the sensor units, respectively the Jetson TX2 module and Jetson Nano kit, does not have significant effects on the performance of the two sensor units with the developed software. As shown in Figure 5.24(b) and Figure 5.25(b), no evident delay on the angular measurements could be observed by *Sensor 2*, which employs the less powerful Jetson

Nano kit, compared to its more powerful counterpart. Furthermore, the angular measurements from both sensor units in *Trial 1* are plotted in a close-up view in Figure 5.26 around 180 s, at which the distance measurements on *Sensor 1* are noisier than *Sensor 2*, and the time of the local minima and maxima, indicating that the mobile robot stops and turns around, are approximately the same. Though the Jetson Nano kit used on one of the sensor units is a less powerful processing module, it is as competent as the more powerful Jetson TX2 module for the implementation of the localization sensor. Because of the lower cost of the Jetson Nano kit, the full cost of future implementations of the localization sensor will be greatly reduced.

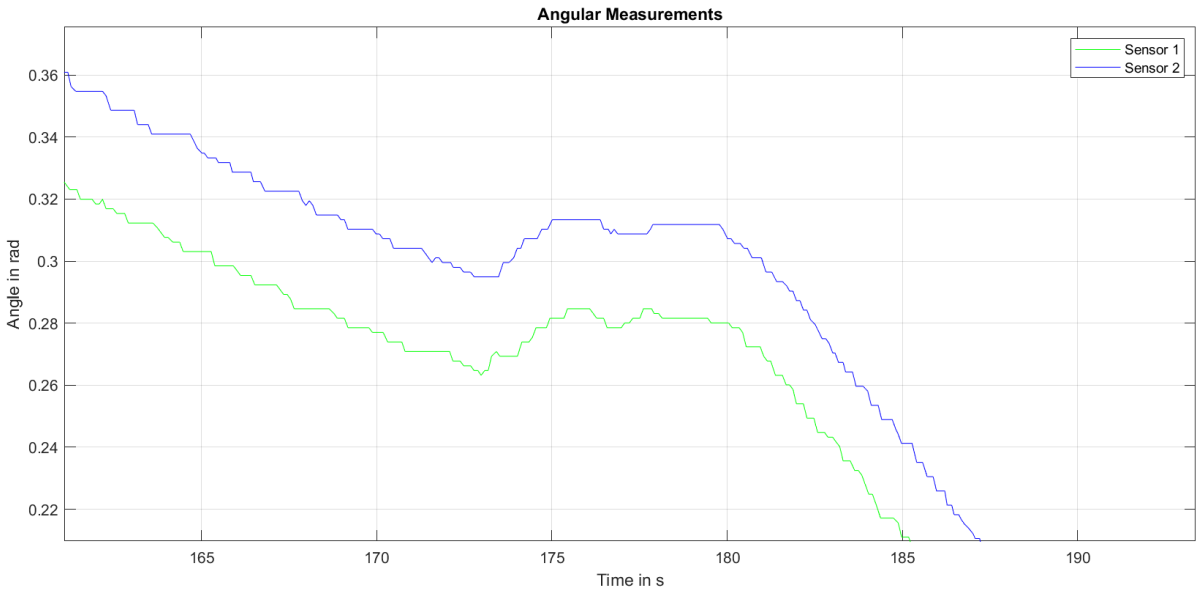


Figure 5.26 Angular measurements in *Trial 1* zoomed in around 180 s.

5.3 Inter-Calibration of Localization Sensor

In this experiment, the two implemented localization sensor units are tested to validate the proposed inter-calibration process described in Section 4.3. During the inter-calibration process, the two localization sensor units collect relative distance and angular measurements against each other, and the pose of one sensor unit can thus be estimated with respect to the same reference frame as that of the other unit.

The errors introduced in Equation (5.1) are used for the evaluation of the inter-calibration between the two implemented sensor units, and the errors on the inter-calibrated pose also serve as an important metrics. Therefore, all the error equations used in this experiment are defined as follows:

$$\begin{aligned} \epsilon_{position} &= \|\boldsymbol{s}_{ic} - \boldsymbol{s}_{actual}\|, \\ \epsilon_{\rho} &= \rho_{measurement} - \rho_{actual}, \\ \epsilon_{\theta} &= \theta_{measurement} - \theta_{actual}, \\ \epsilon_{u} &= u_{ic} - u_{actual}, \\ \epsilon_{v} &= v_{ic} - v_{actual}, \\ \epsilon_{\phi} &= \phi_{ic} - \phi_{actual}, \end{aligned} \tag{5.3}$$

where $\boldsymbol{s} = [u, v]^T$ denotes the position of a localization sensor unit and ϕ denotes its orientation, and $[\rho, \theta]^T$ denotes the relative distance and angle. The subscript ic stands for inter-calibration. The actual values of the position and orientation of a localization sensor unit are given by the sensor calibration process detailed in Section 3.4.3, and the actual values of the measurements are deducted based on the calibrated position of the two sensor units involved in the inter-calibration process.

As presented in Section 4.3, the inter-calibrated position of a sensor unit has reduced uncertainties over the estimate provided by the individual landmark position estimation introduced in Section 4.1. In addition, the uncertainties in the angular measurements are summed up and propagated into the inter-calibrated orientation. To analyze the propagation of uncertainties in the inter-calibration process, multiple relative measurements are taken repeatedly by both implemented sensor units, and the standard deviations of the raw relative measurements, inter-calibrated distance, and inter-calibrated orientation are also reported. When using N relative measurements collected by each sensor unit, the inter-calibration process is executed repeatedly for N times. During each execution, only one relative distance measurement and one relative angular measurement from both sensor units are used. Therefore, after all the executions with N measurements, N individual inter-calibrated poses will be generated, and the average and standard

deviation of all N poses will be provided and analyzed. During the inter-calibration process, the estimated standard deviations of the raw relative measurements will be interpolated using the average of all N measurements as described in Section 3.4.2. For the repeated N executions of the inter-calibration process, the RMSEs of the inter-calibrated position, inter-calibrated orientation, and relative measurements are also used in the analysis as an auxiliary metric. The RMSE is defined the same as that in Equation (5.2), given N errors of corresponding variable:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\epsilon_i^2}, \tag{5.4}$$

where ϵ_i is the i^{th} error of the corresponding variable.

5.3.1 Experimental Setup

The setup is shown in Figure 5.27, where the implemented localization sensor units are randomly placed in the testing area. The extra landmark with red and blue sections in the figure is used for calibrating the pose of the sensor units as described in Section 3.4.3. The sensor calibration process is similar to those in the previous experiments in both Section 5.1 and Section 5.2, however, since the LiDARs on the units can only be considered as reliable within certain ranges, the landmark is placed near each sensor unit and the sensor calibration process is performed respectively for each unit.

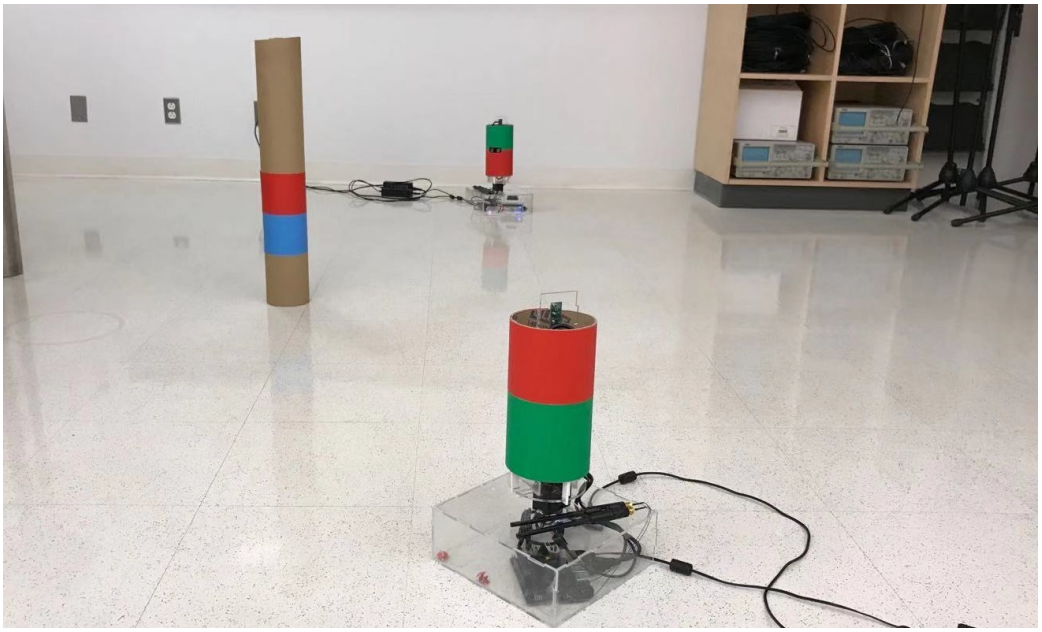


Figure 5.27 Experimental setup for inter-calibration between two localization sensors.

Five trials are carried out in this experiment with a different distance between the two sensor units in each trial, and 100 relative measurements of distance and angle are collected repeatedly by each sensor unit. At the beginning of each trial, the pose of each sensor unit is calibrated as described in Section 3.4.3 using measurements collected against the extra landmark with red and blue sections, manually placed at a series of known positions. The calibrated poses of the sensor units in all five trials are shown in Table 5.2. Though the pose of each sensor unit is calibrated in each trial as given conditions, during the experiment, the pose of *Sensor 2* is assumed unknown and to be inter-calibrated with *Sensor 1*, and the inter-calibrated pose of *Sensor 2* is in the same chosen world frame as the pose of *Sensor 1*.

	Calibrated Pose of *Sensor 1* (mm, mm, rad)	Calibrated Pose of *Sensor 2* (mm, mm, rad)
Trial 1	$[2571.6, 1079.4, -0.8113]^T$	$[11.2, 1151.5, 0.3435]^T$
Trial 2	$[3480.8, 2540.9, -1.8989]^T$	$[6.4, 1160.6, 0.3371]^T$
Trial 3	$[4621.5, 1539.0, -2.4650]^T$	$[131.8, 95.8, -1.0644]^T$
Trial 4	$[6752.7, 889.9, -2.1762]^T$	$[124.5, 102.8, -1.0685]^T$
Trial 5	$[9420.6, 2919.9, -1.8347]^T$	$[120.7, 104.6, -1.070]^T$

Table 5.2 Calibrated poses of the two sensor units in all trials.

The ground truth of the relative distance and angular measurements are shown in Table 5.3, where $\rho_{\{1,2\}}$ denotes the actual distance between the two sensor units, $\theta_{<1,2>}$ denotes the actual relative angle of *Sensor 2* with respect to *Sensor 1*, and $\theta_{<2,1>}$ denotes the actual relative angle of *Sensor 1* with respect to *Sensor 2*. Note that the actual values are derived based on the calibrated poses of the sensor units provided in Table 5.2 and thus are subject to errors residing within the sensor calibration process.

	$\rho_{\{1,2\}}$ (mm)	$\theta_{<1,2>}$ (rad)	$\theta_{<2,1>}$ (rad)
Trial 1	2561.5	-2.3584	-0.3717
Trial 2	3738.6	-0.8659	0.0410
Trial 3	4663.9	-0.4323	1.3387
Trial 4	6674.7	-0.8472	1.1867
Trial 5	9716.7	-1.0129	1.3635

Table 5.3 Ground truth of relative distance and angular measurements.

During this experiment, it is observed that the LiDAR rangefinders used on the implemented sensor units may interfere with each other when both LiDAR rangefinders are working and directly facing towards each other. In order to avoid such interference, measurements are taken by only one sensor unit at a time while the LiDAR rangefinder on the other sensor is not working.

5.3.2 Results and Analysis

The inter-calibrated position and orientation of *Sensor 2* given the pose of *Sensor 1* in all trials are shown in Table 5.4. The inter-calibrated pose of *Sensor 2* in the same reference frame as that of *Sensor 1* is denoted by $\left[u_{m2\{1,2\}}, v_{m2\{1,2\}}, \phi_{m2\{1,2\}}\right]^T$. Since the position of *Sensor 2* is inter-calibrated repeatedly, the RMSE of its inter-calibrated positions and orientations in each trial is also provided in Table 5.4. For position RMSEs, the percentage of the RMSE over the distance between the sensor units, $\rho_{\{1,2\}}$, is given in corresponding parenthesis.

In all following tables, *N* represents the number of measurements individually used in the inter-calibration process, and thus the inter-calibration process will correspondingly execute *N* times. μ is the average of its corresponding variable over *N* values, and ϵ is the error calculated according to Equation (5.3) using the average, μ, as the inter-calibrated value or measurements. The standard deviation and RMSE of a variable are calculated on *N* values, and cells are marked with dark gray when $N = 1$.

Trial	N	$u_{m2\{1,2\}}$ (mm)		$v_{m2\{1,2\}}$ (mm)		Position RMSE (mm)	$\phi_{m2\{1,2\}}$ (rad)			
		μ	ϵ	μ	ϵ		μ	ϵ	σ	RMSE
1	1	11.5	0.3	1161.5	9.9	9.9(0.4%)	0.3330	−0.0105	0.0000	0.0105
	50	15.0	3.7	1163.2	11.7	14.1(0.6%)	0.3323	−0.0112	0.0011	0.0113
	100	18.2	7.0	1163.2	11.7	15.7(0.6%)	0.3323	−0.0112	0.0010	0.0113
2	1	6.5	0.2	1163.3	2.7	2.7(0.1%)	0.3361	−0.0010	0.0000	0.0010
	50	11.4	5.0	1165.2	4.7	13.5(0.4%)	0.3361	−0.0010	0.0000	0.0010
	100	13.5	7.2	1166.1	5.5	14.3(0.4%)	0.3361	−0.0010	0.0000	0.0010
3	1	134.3	2.4	102.1	6.2	6.6(0.1%)	−1.0645	−0.0001	0.0000	0.0001
	50	126.8	−5.1	98.6	2.8	15.8(0.3%)	−1.0633	0.0011	0.0011	0.0016
	100	125.8	−6.1	99.0	3.2	16(0.3%)	−1.0632	0.0013	0.0008	0.0015
4	1	112.1	−12.5	108.9	6.1	13.9(0.2%)	−1.0672	0.0013	0.0000	0.0013
	50	94.0	−30.6	111.1	8.3	40.2(0.6%)	−1.0686	−0.0001	0.0012	0.0012
	100	95.7	−28.9	111.2	8.4	38.3(0.6%)	−1.0685	0.0000	0.0012	0.0012
5	1	−42.4	−163.1	50.1	−54.5	172.0(1.8%)	−1.0662	0.0034	0.0000	0.0034
	50	28.5	−92.2	68.1	−36.5	105.9(1.1%)	−1.0644	0.0053	0.0008	0.0053
	100	47.9	−72.9	74.3	−30.3	197.4(2.0%)	−1.0644	0.0052	0.0007	0.0053

Table 5.4 Experimental results of inter-calibration between two sensors in all five trials using different numbers of measurements.

The percentages following the position RMSEs are calculated with respect to the actual distance between the two sensor units, which are provided in Table 5.3 as $\rho_{\{1,2\}}$. The Position RMSEs from Table 5.4 are further plotted in Figure 5.28. As previously validated, the individual position estimation of *Sensor 1* has a RMSE ranging between 0.5% to 1%. Therefore, the two bounds are also plotted Figure 5.28. The position RMSEs are mostly within 0.5% of the corresponding distance between the sensor units in the first four trials. In the last trial, the distance between the sensor units is longer than 8000 mm and the RMSEs are much larger than 1% of the distance. Such observation coincides with the RMSEs from the individual and collaborative landmark position estimation described in Section 5.1.2. However, with more

measurements taking part in the inter-calibration process and the uncertainties in the measurements reduced, the position RMSE does not decrease but increases most of the time.

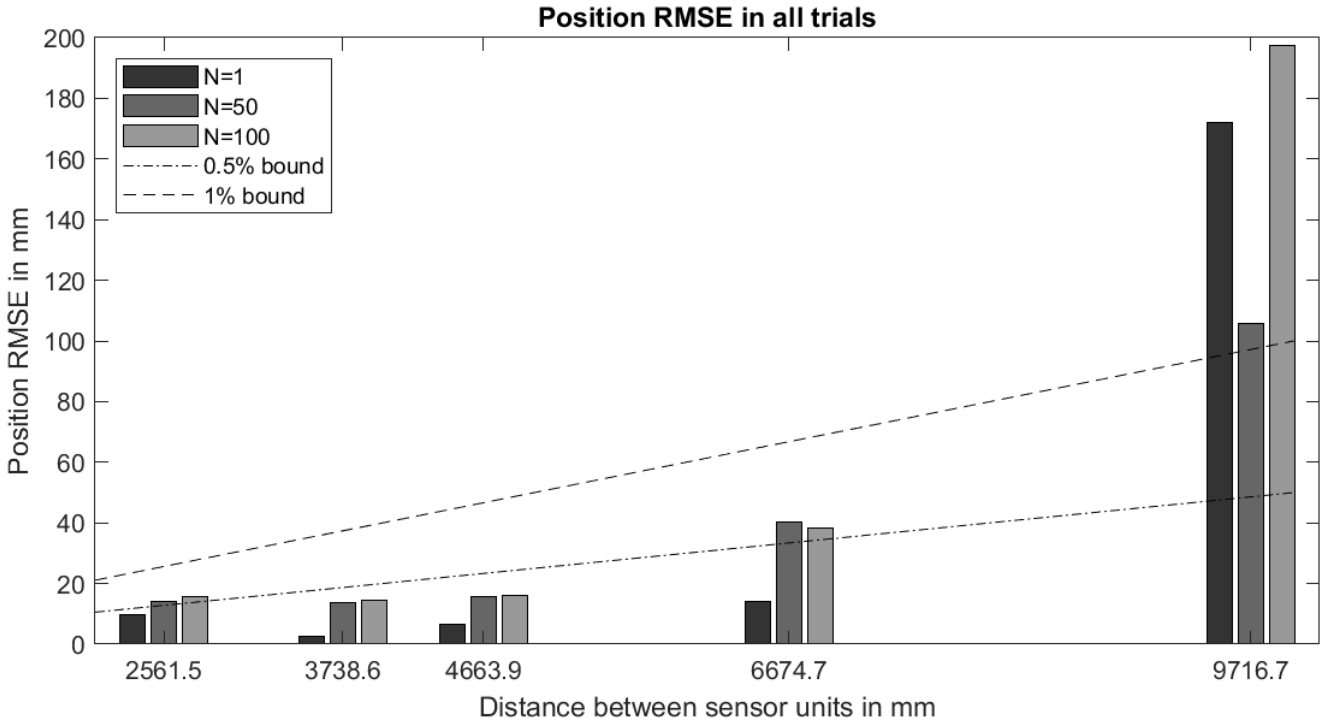


Figure 5.28 Position RMSE in all trials. The distance between sensor units increase from *Trial 1* to *Trial 5*.

The inter-calibrated orientation of *Sensor 2*, $\phi_{m2\{1,2\}}$, are very consistent in the trials with standard deviations smaller than the resolution of the servo motor, which is about 0.0015 rad, but the orientation RMSEs are often larger than the standard deviation of $\phi_{m2\{1,2\}}$ and are approximately the same as the absolute value of the error of the average inter-calibrated orientation of *Sensor 2*.

In order to further analyze the unexpected RMSEs, the raw relative measurements are shown in Table 5.5 and Table 5.6, along with the inter-calibrated distance between the two sensor units and the inter-calibrated orientation of *Sensor 2*. The measurements made by *Sensor 1* towards *Sensor 2* are denoted by $\left[\rho_{m<1,2>}, \theta_{m<1,2>}\right]^T$ and the measurements made by *Sensor 2* towards *Sensor 1* are denoted by $\left[\rho_{m<2,1>}, \theta_{m<2,1>}\right]^T$. The ground truth of the measurements is given in Table 5.3, and $\rho_{\{1,2\}}$ is the ground

truth of both $\rho_{m<1,2>}$ and $\rho_{m<2,1>}$. The inter-calibrated distance using $\rho_{m<1,2>}$ and $\rho_{m<2,1>}$ in the inter-calibration process, as defined in Equation (4.47), is denoted by $\rho_{m\{1,2\}}$, and its ground truth is also given by $\rho_{\{1,2\}}$.

Trial	N	$\rho_{m<1,2>}$ (mm)				$\rho_{m<2,1>}$ (mm)				$\rho_{m\{1,2\}}$ (mm)			
		μ	ϵ	σ	RMSE	μ	ϵ	σ	RMSE	μ	ϵ	σ	RMSE
1	1	2558.4	−3.0	0.0	3.0	2563.0	1.6	0.0	1.6	2561.5	0.0	0.0	0.0
	50	2545.7	−15.7	11.0	19.1	2564.4	3.0	6.7	7.2	2558.0	−3.4	6.4	7.2
	100	2544.5	−16.9	11.2	20.2	2560.0	−1.4	8.8	8.8	2554.8	−6.7	7.4	9.9
2	1	3720.1	−18.5	0.0	18.5	3747.8	9.2	0.0	9.2	3737.4	−1.2	0.0	1.2
	50	3719.1	−19.5	18.1	26.5	3740.0	1.4	17.1	17.0	3732.2	−6.4	11.7	13.2
	100	3719.0	−19.6	16.4	25.5	3736.4	−2.2	15.8	15.8	3729.9	−8.7	11.2	14.1
3	1	4673.2	9.3	0.0	9.3	4653.8	−10.2	0.0	10.2	4659.9	−4.0	0.0	4.0
	50	4656.1	−7.8	25.9	26.8	4673.7	9.8	17.3	19.7	4668.1	4.1	14.6	15.1
	100	4659.4	−4.5	26.0	26.2	4673.4	9.5	17.1	19.4	4668.9	5.0	14.3	15.1
4	1	6758.2	83.5	0.0	83.5	6664.9	−9.8	0.0	9.8	6686.4	11.7	0.0	11.7
	50	6806.5	131.8	44.7	139.0	6673.7	−1.0	31.0	30.7	6704.1	29.4	24.5	38.1
	100	6801.3	126.6	43.3	133.8	6673.0	−1.7	28.3	28.2	6702.4	27.7	23.2	36.1
5	1	10044.6	327.9	0.0	327.9	9840.9	124.2	0.0	124.2	9888.6	171.9	0.0	171.9
	50	9967.4	250.7	71.6	260.5	9769.0	52.3	47.7	70.5	9815.6	98.8	36.7	105.3
	100	9899.5	182.8	733.2	752.1	9762.6	45.9	49.3	67.2	9795.2	78.5	181.7	197.1

Table 5.5 Raw relative distance measurements collected by both sensor units and inter-calibrated relative distance in all five trials using different numbers of measurements.

Similar to the errors observed in the experiment of collaborative landmark position estimation in Section 5.1.2, the position RMSE is majorly caused by the errors from the relative distance measurements. The RMSE of the raw relative distance measurements collected by both sensor units does not reduce evidently with more repeated measurements, but the RMSE and the standard deviation on the inter-calibrated distance are reduced when compared to those on the raw relative distance measurements.

Therefore, the uncertainties within the raw distance measurements are reduced in the inter-calibration process, and the inter-calibrated distance can be more reliable with fewer errors at times when compared to the raw distance measurements.

Trial	N	$\theta_{m<1,2>}$ (rad)				$\theta_{m<2,1>}$ (rad)				$\phi_{m2\{1,2\}}$ (rad)			
		μ	ϵ	σ	RMSE	μ	ϵ	σ	RMSE	μ	ϵ	σ	RMSE
1	1	−2.3623	−0.0039	0.0000	0.0039	−0.3651	0.0066	0.0000	0.0066	0.3330	−0.0105	0.0000	0.0105
	50	−2.3631	−0.0046	0.0011	0.0048	−0.3651	0.0066	0.0000	0.0066	0.3323	−0.0112	0.0011	0.0113
	100	−2.3631	−0.0046	0.0010	0.0047	−0.3651	0.0066	0.0000	0.0066	0.3323	−0.0112	0.0010	0.0113
2	1	−0.8652	−0.0007	0.0000	0.0007	0.0414	0.0004	0.0000	0.0004	0.3361	−0.0010	0.0000	0.0010
	50	−0.8652	−0.0007	0.0000	0.0007	0.0414	0.0004	0.0000	0.0004	0.3361	−0.0010	0.0000	0.0010
	100	−0.8652	−0.0007	0.0000	0.0007	0.0414	0.0004	0.0000	0.0004	0.3361	−0.0010	0.0000	0.0010
3	1	−0.4034	−0.0011	0.0000	0.0011	1.3376	−0.0011	0.0000	0.0011	−1.0645	−0.0001	0.0000	0.0001
	50	−0.4032	−0.0009	0.0006	0.0011	1.3367	−0.0020	0.0008	0.0021	−1.0633	0.0011	0.0011	0.0016
	100	−0.4033	−0.0010	0.0004	0.0011	1.3364	−0.0023	0.0006	0.0024	−1.0632	0.0013	0.0008	0.0015
4	1	−0.8483	−0.0011	0.0000	0.0011	1.1842	−0.0025	0.0000	0.0025	−1.0672	0.0013	0.0000	0.0013
	50	−0.8489	−0.0018	0.0008	0.0019	1.1850	−0.0017	0.0008	0.0019	−1.0686	−0.0001	0.0012	0.0012
	100	−0.8489	−0.0018	0.0008	0.0019	1.1849	−0.0018	0.0008	0.0019	−1.0685	0.0000	0.0012	0.0012
5	1	−1.0124	0.0005	0.0000	0.0005	1.3606	−0.0029	0.0000	0.0029	−1.0662	0.0034	0.0000	0.0034
	50	−1.0121	0.0008	0.0008	0.0011	1.3591	−0.0044	0.0002	0.0044	−1.0644	0.0053	0.0008	0.0053
	100	−1.0121	0.0008	0.0007	0.0011	1.3591	−0.0044	0.0002	0.0044	−1.0644	0.0052	0.0007	0.0053

Table 5.6 Raw relative angular measurements collected by both sensor units with the inter-calibrated orientation of *Sensor 2* in all five trials using different numbers of measurements.

The raw relative angular measurements are stable with rather small standard deviations in the trials. The RMSEs of the angular measurements by both sensor units are slightly larger than the corresponding standard deviations and majorly result from the errors of the average relative angular measurements. As defined in Section 4.3, the variances of the relative angular measurements by both sensor units directly sum up into the corresponding variance of the inter-calibrated orientation. Therefore, the standard

deviation of the inter-calibrated orientation is larger than those of the raw relative angular measurements most of the time. In addition to the aforementioned variances, the errors of the raw angular measurements also sum up into the error of $\phi_{m2\{1,2\}}$. However, as the errors of the raw relative angular measurements can be either positive or negative, the summed error of $\phi_{m2\{1,2\}}$ may become smaller, such as in *Trial 4*, or much larger, such as in *Trial 5*. As such, the uncertainties in the raw relative measurements can not be reliably reduced and require other processes such as collaboration with more localization sensor units.

The distance and angular measurement errors are representatively modelled as Gaussian white noise in Section 3.4.2, but the characteristics of the errors may not be perfectly modelled and the unmodelled components of the errors may have adverse effects on the performance of the implemented localization sensor units. At the beginning of each trial, the sensor calibration process, presented in Section 3.4.3, is carried out to estimate the pose of the sensor units and use the estimates as ground truth. However, the calibration process is subject to the unmodelled components of the measurement errors, which in turn affects the analysis of the performance in the experiments, leading to the unexpected RMSEs when using more repeated measurements. Though methods providing more accurate prior knowledge about the pose of the localization sensor units should be investigated in order to improve the accuracy in the analysis, the encouraging results have showcased the potential of the inter-calibration process. The latter ensures that the position estimate of a localization sensor unit can be more reliably acquired than using individual landmark position estimation and that the orientation estimation of the sensor unit is viable. The feasibility and usability of the proposed collaborative sensor pose estimation scheme are therefore demonstrated, with the inter-calibration process as its foundation.

5.4 Summary

In this chapter, the two implemented localization sensor units are experimentally validated in an indoor lab environment with three experiments. First, the sensor units are validated against a stationary landmark. The landmark position estimation using one of the implemented sensor units has reached about 0.25% to 0.75% RMSE with respect to the distance between the landmark and the sensor unit, within a working range of 8 meters. With collaboration from another localization sensor unit, the landmark position estimation RMSE is further reduced to mostly 0.25% to 0.5% in the working range of 8 meters, though the second unit has slightly worse performance. Second, the implemented sensor units are validated with a moving landmark, and the performance of the sensor units is compared to a Monte-Carlo localization method. The implemented localization sensor units successfully track and estimate the position of the moving landmark at the speeds of 0.1 m/s and 0.15 m/s while performing individually within a working distance of up to 6 meters. The localization sensor units offer similar performance to the Monte-Carlo localization method within their working range, and quite often, the two units have exhibited more reliable position estimates. In the last experiment, a proposed inter-calibration process is validated using the two sensor units, and one of the units is successfully inter-calibrated given the pose of the other unit.

In summary, the proposed localization sensor should have the following specifications:

Working condition:	Tested indoors with proper lighting
Working range:	0.5 to 8 meters
Distance measurement accuracy:	About 0.75% of the distance between the landmark and a localization sensor
Angular measurement accuracy:	Up to about 0.004 rad and mostly within 0.002 rad
Individual position estimation accuracy:	About 0.75% of the distance between the landmark and a localization sensor
Collaborative position estimation accuracy:	About 0.5% of the distance between the landmark and a localization sensor
Tracking ability of a moving landmark:	Up to 6 meters with 0.1 m/s; up to 4 meters with 0.15 m/s

Table 5.7 Specifications for the proposed localization sensor.

Limited by the LiDAR devices integrated into the implemented localization sensor units, the minimum working distance is 0.5 meters and the position estimation errors, majorly contributed by the distance measurement errors, are positively correlated to the distance between a sensor unit and the landmark being measured. However, unlike motion tracking systems such as HTC VIVE, the proposed localization sensor does not require any extra system to be installed in the working environment, and the compact design of the sensor can be easily installed on board of a mobile robot, ensuring that the proposed sensor has the potential to cover a rather large area. The performance of the proposed localization sensor is also better than that of ordinary GPS solutions and comparable to differential GPS solutions in outdoor environments. In addition, by leveraging collaboration among multiple sensor units, the proposed localization method can provide pose estimation for mobile robots without relying on any knowledge or information of the environment both before and during the localization operation. Thus, the localization sensor prevails over solutions such as SLAM and visual odometry in scenarios where the working environment is not accessible before the localization operation or where the environment lacks features or textures to support localization.

CHAPTER 6 - CONCLUSION

In this chapter, a summary is provided in Section 6.1. The contributions of the thesis are underlined in Section 6.2. In Section 6.3, potential future work is discussed to extend the work presented in the thesis.

6.1 Summary

The thesis starts with an extensive review of localization solutions. As identified in the literature, localization solutions are mostly non-collaborative and rely heavily on observable features or textures readily available in the working environment, and the predominant solutions are usually computationally expensive for mobile robotics. The thesis then proposes an innovative collaborative localization sensor for mobile robots aiming at alleviating the requirements on the working environment. When installed on multiple mobile robots, implemented instances of the proposed sensor measure their relative location with respect to each other and provide pose estimation for the robots collaboratively. Because of the collaboration among the sensor units, the proposed sensor requires neither prior knowledge nor existing information of the environment. Thus, the proposed localization sensor is suitable for applications within feature-less working environments or contexts that do not provide access to the working environment before the localization operation takes place. The details of configuring and characterizing the proposed localization sensor are documented, where the distance and angular measurement errors are representatively mapped as Gaussian white noise. A sensor calibration process is provided along with the sensor configuration and calibration. The process estimates the pose of a sensor unit in an offline manner, and the process is repeatedly used in the later experiments.

A collaborative pose estimation scheme is developed in the thesis using the proposed localization sensor. A landmark position estimation using one or multiple localization sensor units is first presented, where the collaborative position estimation with multiple sensor units is proved theoretically to generate

more reliable position estimates with smaller RMSE than individual position estimation using a single sensor unit. Then, an inter-calibration process is introduced between two localization sensor units, where the two sensor units collect measurements against each other. The inter-calibration process enables the capability of the proposed sensor to estimate the orientation of a sensor unit, and it is demonstrated that the inter-calibrated position is more reliable than the proposed individual landmark position estimation. Finally, based on the developed collaborative landmark position estimation and sensor inter-calibration processes, a collaborative sensor pose estimation approach is derived. The approach is more reliable with reduced uncertainties and errors in estimations for both the orientation and position of a sensor unit than previous models.

Two instances of the proposed localization sensor are implemented, and the proposed estimation methods are experimentally validated in an indoor environment with proper lighting conditions. The implemented localization sensor units accurately estimated the position of a stationary landmark both individually and collaboratively, and the collaboratively estimated pose of the landmark has shown the lower RMSEs most of the time. Moreover, the sensor units managed to track a moving landmark while estimating its position. The inter-calibration process is also successfully validated using the implemented sensor units. Although only two sensor units are implemented and the proposed collaborative pose estimation scheme was not fully validated, the encouraging results of the experiments carried out in this work demonstrate the feasibility and usability of the proposed localization sensor on collaborative pose estimation when multiple sensor units are available.

6.2 Contributions

The first and main contribution of this thesis consists of the design, implementation, characterization and extensive experimental evaluation of an innovative localization sensor device that can be easily

mounted on various models of ground mobile robots. The pose of the robots that carry such a device, each with an onboard cylindrical landmark with a unique visible pattern, can be estimated within a certain range from the localization device. The main benefit of this solution is that it is self-contained, independent from existing features in the environment, and does not require the installation of any devices in the environment.

A sampling-based landmark detection and tracking algorithm from a standard color camera is developed and embedded on the proposed localization sensor to recognize and take measurements on other implemented sensor instances. This algorithm forms the fundamental capability of the localization sensor to automatically detect and track landmarks, each associated with a mobile robot, as they move in the environment. This represents a valuable feature of the localization solution for applications in the context of autonomous robotics.

A rigorous extension to the pose estimation process is proposed that supports collaborative pose estimation of mobile agents when working with multiple implemented instances of the proposed localization sensor. A linearized model of the collaborative pose estimation is presented in this work, which provides a simple way to merge information from other sensors. The original methodology significantly improves the accuracy of pose estimation by merging raw distance and angular measurements from all available localization sensor units. This concept finds important applications in swarm robotics.

A protocol for automated inter-calibration in between two localization sensor units is also introduced to assist with the initial distribution of localization sensors in the environment, especially when the sensor units are meant to remain static in position while tracking other units mounted on mobile robots. The inter-calibration procedure facilitates operation and contributes to improving the overall pose estimation and tracking performance.

Two instances of the proposed sensor are implemented and validated on several use cases in this work, but the proposed sensor design can be readily replicated because of its detailed design, construction from off-the-shelf components, and reusable software packages.

This work has led to the publication of [62].

6.3 Future work

Throughout the experiments, the ground truths of the sensor pose are given by the sensor calibration process presented in Section 3.4.3. Though the calibration process was the best available method to acquire the sensor poses in the context of this work, the accuracy of the calibration process has affected the validation of the proposed localization sensor. The sensor characterization should be improved with more accurate models of the measurements to reduce the errors resulting from the calibration process. Alternatively, a better solution for defining the actual sensor poses can be employed for validating the proposed localization sensor.

On the implemented instances of the proposed sensor, the selected LiDAR rangefinder devices have shown variable performance and both rangefinders have limitations on measuring the landmark used in the work at long distances even when it is not in motion. Without reducing the usability of the proposed sensor by increasing the size of the landmark, a more accurate and reliable solution for measuring distances between sensor units is to be investigated.

The response of the servo controller and the landmark detection and tracking algorithm should be improved to counteract the time dedicated to image processing, which introduces delays that can significantly degrade the performance of the localization sensor when landmark tracking is involved. The landmark detection and tracking algorithm should harness the full image processing power provided by

the Jetson processing modules and provide more responsive and reliable performances. And the servo controller should be improved to be able to counteract the effects of potential delays.

The current localization method only makes use of the distance and angular measurements collected by the localization sensor, and the performance of the localization sensor is greatly affected by measurement noise. While the sensor tracks a landmark, useful information such as the previous position estimates of the landmark could be obtained and contribute to better performance of the localization sensor. Filtering techniques, for instance, take advantage of such information to construct a state transition model. And by using the state transition model, a predicted state of the landmark is made available to mitigate the effects of noisy measurements.

Finally, the collaborative pose estimation strategy with the proposed sensor was partially validated with two instances of the proposed sensor in Chapter 5. More instances of the proposed sensor should be implemented, and the collaborative pose estimation with multiple instances should be further validated.

BIBLIOGRAPHY

[1] P. Misra, B. P. Burke, and M. M. Pratt, "GPS Performance in Navigation," in *Proceedings of the IEEE*, vol. 87, no. 1, pp. 65–85, 1999.

[2] M. B. Kjærgaard, H. Blunck, T. Godsk, T. Toftkjær, D. L. Christensen, and K. Grønbæk, "Indoor Positioning Using GPS Revisited," *International Conference on Pervasive Computing*, vol. 6030, pp. 38–56, 2010.

[3] M. Windolf, N. Götzen, and M. Morlock, "Systematic Accuracy and Precision Analysis of Video Motion Capturing Systems—Exemplified on the Vicon-460 System," *Journal of Biomechanics*, vol. 41, no. 12, pp. 2776–2780, 2008.

[4] M. Borges, A. Symington, B. Coltin, T. Smith, and R. Ventura, "HTC Vive: Analysis and Accuracy Improvement," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2610–2615, 2018.

[5] J. J. Caffery, "New Approach to the Geometry of TOA Location," in *Proceedings of* the *IEEE Vehicular Technology Conference*, vol. 4, no. 52, pp. 1943–1949, 2020.

[6] D. E. Manolakis, "Efficient Solution and Performance Analysis of 3-D Position Estimation by Trilateration," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1239–1248, 1996.

[7] F. Thomas and L. Ros, "Revisiting Trilateration for Robot Localization," *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 93–101, Feb. 2005.

[8] W. H. Foy, "Position-Location Solutions by Taylor-Series Estimation," *IEEE Transactions on Aerospace and Electronic Systtems*, vol. AES-12, no. 2, pp. 187–194, Mar. 1976.

[9] Z. Yang and Y. Liu, "Quality of Trilateration: Confidence-Based Iterative Localization," in *Proceedings of the IEEE International Conference on Distributed Computing Systems*, pp. 446–453, 2008.

[10] A. E. Waadt, C. Kocks, S. Wang, G. H. Bruck, and P. Jung, "Maximum Likelihood Localization Estimation based on Received Signal Strength," in *Proceedings of the IEEE 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, pp. 1–5, 2010.

[11] D. J. Torrieri, "Statistical Theory of Passive Location Systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, pp. 183–198, March 1984.

[12] B. T. Fang, "Simple Solutions For Hyperbolic And Related Position Fixes," *IEEE Transactions on Aerospace Electronic Systems*, vol. 26, no. 5, pp. 748–753, Sept. 1990.

[13] R. G. Stansfield, "Statistical Theory of D.F. Fixing," *Journal of the Institution of Electrical Engineers - Part IIIA Radiocommunication*, vol. 94, no. 15, pp. 762–770, March-April 1947.

[14] M. Gavish and A. J. Weiss, "Performance Analysis of Bearing-Only Target Location Algorithms," *IEEE Transactions on Aerospace Electronic Systems*, vol. 28, no. 3, pp. 817–828, July 1992.

[15] K. Spingarn, "Passive Position Location Estimation Using the Extended Kalman Filter," *IEEE Transactions on Aerospace Electronic Systems*, vol. AES-23, no. 4, pp. 558–567, July 1987.

[16] Y. Wang and K. C. Ho, "An Asymptotically Efficient Estimator in Closed-Form for 3-D AOA

Localization Using a Sensor Network," *IEEE Transactions on Wireless Communication*, vol. 14, no. 12, pp. 6524–6535, Dec. 2015.

[17] V. Y. Zhang, A. K. S. Wong, K. T. Woo, and R. W. Ouyang, "Hybrid TOA/AOA-based Mobile Localization with and without Tracking in CDMA Cellular Networks," in *Proceedings of the IEEE Wireless Communication and Netwworking Conference (WCNC)*, 2010.

[18] J. Yin, Q. Wan, S. Yang, and K. C. Ho, "A Simple and Accurate TDOA-AOA Localization Method Using Two Stations," *IEEE Signal Processing Lettetters*, vol. 23, no. 1, pp. 144–148, Jan. 2016.

[19] C. Drocourt, L. Delahoche, C. Pegard, and A. Clerentin, "Mobile Robot Localization Based on an Omnidirectional Stereoscopic Vision Perception System," in *Proceedings of the IEEE International Conference on Robotices and Automation*, vol. 2, pp. 1329–1334, 1999.

[20] Z. Cao, S. Liu, and J. Röning, "Omni-directional Vision Localization Based on Particle Filter," in *Proceedings of the 4th IEEE International Conference on Image and Graphics (ICIG)*, pp. 478–483, 2007.

[21] S. He and S. H. G. Chan, "Wi-Fi Fingerprint-based Indoor Positioning: Recent Advances and Comparisons," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1. pp. 466–490, 2016.

[22] S. He, S. H. G. Chan, L. Yu, and N. Liu, "SLAC: Calibration-Free Pedometer-Fingerprint Fusion for Indoor Localization," *IEEE Transactions on Mobible Computing*, vol. 17, no. 5, pp. 1176–1189, May 2018.

[23] J. Borenstein and L. Feng, "UMBmark: A Method for Measuring, Comparing, and Correcting Dead-reckoning Errors in Mobile Robots," *Univ. of Michigan*, 1994. [Online]. Available: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/3753/bac6477.0001.001.pdf?sequence=5&isAllowed=y. [Accessed: 18-Feb-2020].

[24] J. Borenstein and L. Feng, "Measurement and Correction of Systematic Odometry Errors in Mobile Robots," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 6, pp. 845–857, 1996.

[25] J. Borenstein and L. Feng, "Gyrodometry: a New Method for Combining Data from Gyros and Odometry in Mobile Robots," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 423–428, 1999.

[26] B. Barshan and H. F. Durrant-whyte, "Inertial Navigation Systems for Mobile Robots," *IEEE Transactions on Robotics and Automation*, vol. 11, no. 3, pp. 328–342, 1995.

[27] Y. Fuke and E. Krotkov, "Dead Reckoning for a Lunar Rover on Uneven Terrain," in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, pp. 411–416, 1996.

[28] K. Lingemann, H. Surmann, A. Nüchter, and J. Hertzberg, "Indoor and Outdoor Localization for Fast Mobile Robots," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, pp. 2185–2190, 2004.

[29] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611–625, Jul. 2016.

[30] D. Scaramuzza and F. Fraundorfer, "Visual Odometry [Tutorial]," *IEEE Robotics & Automation Magazine*, vol. 18, no. 4, pp. 80–92, Dec. 2011.

[31] F. Fraundorfer and D. Scaramuzza, "Visual Odometry: Part II: Matching, Robustness, Optimization, and Applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, IEEE, pp. 78–90, June

2012.

[32] D. G. Lowe, “Distinctive Image Features from Scale-invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[33] H. Bay, T. Tuytelaars, and L. Gool, “SURF: Speeded Up Robust Features,” *European Conference on Computer Vision*, vol. 3951, pp. 404–417, 2006.

[34] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: an Efficient Alternative to SIFT or SURF,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2564–2571, 2011.

[35] H. J. Chien, C. C. Chuang, C. Y. Chen, and R. Klette, “When to use what feature? SIFT, SURF, ORB, or A-KAZE Features for Monocular Visual Odometry,” in *Proceedings of the International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 1–6, 2016.

[36] E. Olson, J. Leonard, and S. Teller, “Fast Iterative Alignment of Pose Graphs with Poor Initial Estimates,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2262–2269, 2006.

[37] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, “Bundle Adjustment – a Modern Synthesis,” *International Workshop on Vision Algorithms*, vol. 1883, pp. 298–372, 2000.

[38] J. Engel, V. Usenko, and D. Cremers, “A Photometrically Calibrated Benchmark For Monocular Visual Odometry,” 2016, *arXiv:1607.02555*.

[39] J. Engel, T. Schöps, and D. Cremers, “LSD-SLAM: Large-Scale Direct Monocular SLAM,” *European Conference on Computer Vision*, pp. 834–849, 2014.

[40] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, “ORB-SLAM: A Versatile and Accurate Monocular SLAM System,” *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

[41] R. Mur-Artal and J. D. Tardos, “ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras,” *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.

[42] W. Hess, D. Kohler, H. Rapp, and D. Andor, “Real-time Loop Closure in 2D LIDAR SLAM,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1271–1278, 2016.

[43] D. Dardari and A. Conti, “A Sub-optimal Hierarchical Maximum Likelihood Algorithm for Collaborative Localization in Ad-hoc Networks,” in *Proceedings of the 1st Annual IEEE Communications Society Conference on Sensor Ad Hoc Communications and Networks (IEEE SECON)*, pp. 425–429, 2004.

[44] X. Li, “Collaborative Localization With Received-Signal Strength in Wireless Sensor Networks,” *IEEE Transactions Vehicular Technology*, vol. 56, no. 6, pp. 3807–3817, Nov. 2007.

[45] T. Jia and R. M. Buehrer, “On the Optimal Performance of Collaborative Position Location,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 374–383, Jan. 2010.

[46] L. W. Chan, J. R. Chiang, Y. C. Chen, C. N. Ke, J. Hsu, and H. H. Chu, “Collaborative Localization: Enhancing WiFi-based Position Estimation with Neighborhood Links in Clusters,” *International Conference on Pervasive Computing*, vol. 3968, pp. 50–66, 2006.

[47] D. Taniuchi, X. Liu, D. Nakai, and T. Maekawa, “Spring Model Based Collaborative Indoor Position Estimation with Neighbor Mobile Devices,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 268–277, Mar. 2015.

[48] T. Higuchi, H. Yamaguchi, and T. Higashino, “Context-supported Local Crowd Mapping via Collaborative Sensing with Mobile Phones,” *Pervasive and Mobile Computing*, vol. 13, pp. 26–51, 2014.

[49] e-con Systems, “See3CAM_130 - Technical Documents,” 2020. [Online]. Available: https://www.e-consystems.com/doc_13MP_autofocus_USB3_Camera.asp. [Accessed: 18-Feb-2020].

[50] Garmin, “Lidar Lite v3 Operation Manual and Technical Specifications,” 2016. [Online]. Available: http://static.garmin.com/pumac/LIDAR_Lite_v3_Operation_Manual_and_Technical_Specifications.pdf. [Accessed: 18-Feb-2020].

[51] ROBOTIS, “ROBOTIS e-Manual: XL430-W250-T,” 2020. [Online]. Available: http://emanual.robotis.com/docs/en/dxl/x/xl430-w250/. [Accessed: 18-Feb-2020].

[52] NVIDIA, “Jetson TX2 Series System-on-Module Data Sheet,” 2018. [Online]. Available: http://developer.nvidia.com/embedded/dlc/jetson-tx2-series-modules-data-sheet. [Accessed: 18-Feb-2020].

[53] Connect Tech Inc., “Orbitty Carrier for NVIDIA Jetson TX2/TX2i/TX1 User Guide,” 2019. [Online]. Available: http://connecttech.com/pdf/CTIM-ASG003_Manual.pdf. [Accessed: 18-Feb-2020].

[54] NVIDIA, “NVIDIA Jetson Nano System-on-Module Data Sheet,” 2020. [Online]. Available: https://developer.nvidia.com/embedded/dlc/jetson-nano-system-module-datasheet. [Accessed: 18-Feb-2020].

[55] M. Quigley *et al.*, “ROS: an Open-source Robot Operating System,” in *ICRA Workshop on Open Source Software*, 2009.

[56] RabbitMQ, “RabbitMQ Documentation,” 2020. [Online]. Available: https://www.rabbitmq.com/documentation.html. [Accessed: 18-Feb-2020].

[57] G. Jang, S. Kim, W. Lee, and I. Kweon, “Color Landmark Based Self-Localization for Indoor Mobile Robots,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1037–1042, 2002.

[58] J. Blanco, “Contributions to Localization, Mapping and Navigation in Mobile Robotics,” Ph.D. dissertation, Dept. de Ingenier´ıa de Sistemas y Automática, University of Málaga, Málaga, Spain, 2009.

[59] ROBOTIS, “ROBOTIS e-Manual: TurtleBot3,” 2020. [Online]. Available: http://emanual.robotis.com/docs/en/platform/turtlebot3/overview/. [Accessed: 18-Feb-2020].

[60] G. Grisetti, C. Stachniss, and W. Burgard, “Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters,” *IEEE Transaction on Robotics*, vol. 23, no. 1, pp. 34–46, Feb. 2007.

[61] D. Fox, W. Burgard, F. Dellaert, and S. Thrun, “Monte Carlo Localization: Efficient Position Estimation for Mobile Robots,” in *Proceedings of the National Conference on Artificial Intelligence*, pp. 343–349, 1999.

[62] S. Yang and P. Payeur, “Collaborative Localization Sensor for Mobile Robots in Feature-Free Environments,” in *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, May 25-28 2020.