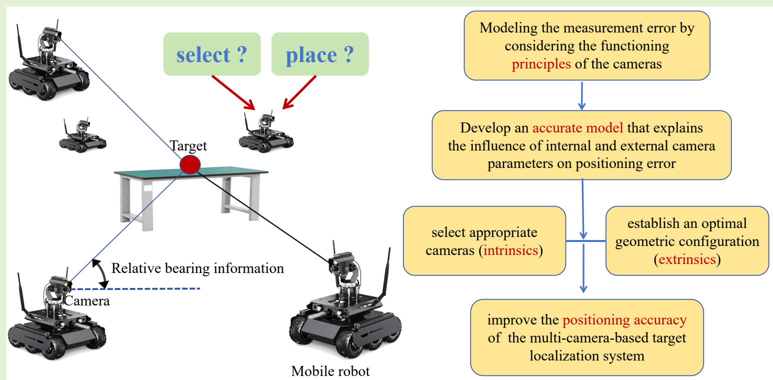
Optimal placement and selection of camera-based target localization system using relative bearing measurements

Kangwen Lin, Liang Zhang, Zhaochen Liu, Yefei Huang, Zexu Zhang

***Abstract*—As passive optical sensors, cameras provide low cost, low power consumption, and rich texture information, providing a more advanced positioning method than wireless sensor networks. However, in recent works, the direct influence of geometric placement (extrinsics) and camera selection (intrinsics) on target localization accuracy has received limited attention. This work investigates the optimal placement and selection problem for localizing a single target, using a multi-robot system with cameras as relative bearing sensors. We initiate our analysis by examining the mechanisms underlying measurement errors in camera-based relative bearing measurements. Specifically, we model two primary sources of these errors: geometric errors resulting from camera optical center offsets and imaging plane distortions, and random errors that occur during the measurement process. We then apply the Fisher Information Matrix (FIM) and the D-optimal criterion to characterize the localization accuracy of the target, which uniformly incorporates all the extrinsic and intrinsic parameters of cameras, including the positions, focal length, resolution, etc. Two typical scenarios of target localization are analyzed to show its potential in practical deployment by maximizing the D-optimality criterion of the FIM. Finally, numerical simulations validated the correctness and effectiveness of the proposed methods.**



***Index Terms*—Target localization, Relative bearing measurement, Camera-based bearing measurement, Fisher Information Matrix, D-optimality criterion.**

I. Introduction

Target localization, which involves obtaining the target's position information through observation and signal processing, is one of the core functions of multi-sensor systems. It has already played a significant role in the lunar exploration [1], surveillance and rescue mission [2], and so on. Among many well-developed target localization systems, cameras, utilizing passive optical imaging technology, outperform radio-based UWB [3], LiDAR [4], and ultrasonic devices on embedded platforms due to their low cost, low power consumption, and rich texture information [5]. Therefore, the problem of target localization that uses a camera to obtain the relative bearing between the target and the optical center of the camera has attracted increasing attention in recent years.

It is well known that the observation geometry and internal parameters of sensors significantly impact the localization accuracy of the target being measured. Current research on optimizing sensor placement and selection is mostly oriented toward general or radio-propagation-based measurement models, with little work dedicated to camera-based relative bearing measurement models. [6]–[9] used general models for the optimal placement of sensors under the target localization problem, where the measurement process is modeled by directly corrupting the Gaussian noise into the true value of the quantity being measured. For example, the general model of Arrival of Angle (AOA) measurement used in [8] is derived by computing the arctan of the target-sensor relative position, then adding Gaussian noise to simulate real-world imperfections. The referenced Time Difference of Arrival (TDOA) measurement model is established by computing target-sensor distance and incorporating constant and Gaussian noise to simulate real-world errors in [6]. However, sensor-

Correspondence: Liang Zhang (liangzhang@ahu.edu.cn)

This work was supported in part by the State Key Laboratory of Micro-Spacecraft Rapid Design and Intelligent Cluster under grant No. MS01240110, the National Natural Science Foundation of China under grant No. 62303002, the Key Program of Natural Science Foundation of Anhui Higher Education Institutions of China under grant No. 2022AH050068

Kangwen Lin, Liang Zhang, and Zhaochen Liu are with the School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China. (e-mail: z23301057@stu.ahu.edu.cn; liangzhang@ahu.edu.cn; z23301083@stu.ahu.edu.cn). Yefei Huang and Zexu Zhang are with the Deep Space Exploration and Research Center, School of Astronautics, Harbin Institute of Technology, Harbin 150001, China. (e-mail: huangyefei@hit.edu.cn; zexuzhang@hit.edu.cn).

specific sensing principles and internal architectures introduce complexities in bearing and ranging measurement errors beyond Gaussian noise descriptions. Efforts have been made to take into account the sensing principle of radio propagation to the target localization cases in [10], [11], where the waveforms of received signals are examined as the model of distance measurement, enabling the inference of the positions and internal characteristics of radio devices (e,g. clock offset, multi-path effect, node priority et al.).

Despite those achievements, the sensing principle of cameras for AOA measurement remains overlooked in the optimization problem of target localization in recent works. This paper focuses on the optimization of sensor layout and parameter selection for a target localization system using multiple cameras for relative bearing measurement. After analyzing two kinds of measurement errors, geometric errors stemming from camera optical center offsets and imaging plane distortions, and random errors occurring during the measurement process, respectively, a correlation model between the accuracy of the target localization and the internal and external parameters of the cameras is established. Specifically, the random and geometric errors of the camera during the relative bearing measurement process are characterized using a mathematical modeling approach based on normal distribution. Furthermore, through the FIM, we established the relationship between camera hardware parameters and positioning accuracy, thereby providing a method for selecting cameras in target localization scenarios.

The main contributions are threefold:

1) We introduce an error model that incorporates the functioning principles of the camera-based relative bearing measurements, integrating both the random and geometric components.
2) We mathematically formulate the positioning accuracy of the multi-camera-based cooperative localization system, exploiting all the intrinsic and extrinsic parameters, which can provide insights for optimizing the layout and parameters of cameras.
3) Finally, we present two typical scenarios for the optimal placement and selection of cameras: the single-target localization problem when the target is within the camera placement area, and when it is outside.

The rest of this article is organized as follows. The related works of this paper are introduced in Section II. The problem formulation is presented in Section III. Section IV is the method of system model, where we also mathematically formulate the positioning accuracy of the multi-camera-based target localization system's intrinsic and extrinsic parameters. In section V, the D-optimality criterion of the FIM is used to discuss the optimal placement of the target to be measured within and outside the camera area. Finally, the simulation results are presented in Section VI, and Section VII concludes this article.

II. Related Works

Target localization: Estimating the position of a target involves the use of various types of sensors in combination with different localization algorithms and environmental features. Sensors can be broadly categorized based on their working principles into wireless sensors [12] and optical sensors [13]. Sensors are classified as either bearing measurement sensors or ranging measurement sensors.

Ranging-based sensors usually obtain the distance information from the sensor to the target, ultimately estimating the position of the target based on the sensor's location. For instance, Time of Arrival (TOA) [14] utilizes the time taken for a signal to travel from the transmitter to the receiver, allowing for distance calculation using known signal propagation speeds (such as the speed of light or sound). TDOA [15] measures the time differences at which a common signal reaches different receivers, enabling the determination of the signal source's location by analyzing these time discrepancies. An unmanned aerial vehicle (UAV) equipped with a stereo camera can also perform distance measurement tasks based on pinhole imaging principles. Bearing-based sensors estimate the target's location through the relative angle between the sensor and the target. For example, AOA [16] employs multiple antennas arranged in an array at its receiving end. The sensor calculates relative azimuth angles based on time differences in signals arriving at each antenna. Studies have shown that, compared to ranging measurement sensors, bearing measurement sensors are more cost-effective and exhibit stronger resistance to interference [17].

Optimal placement of sensors: With the advancement of autonomous perception technology in robotics, wireless sensors (e.g., LiDAR and WLAN), which rely on wireless communication technology and can be applied to positioning and monitoring within small ranges, are being employed for target localization [18]. Consequently, extensive research has been conducted on optimizing the placement of sensors to achieve better precision in positioning. Researchers have considered factors such as the observation models relevant to practical application scenarios, signal propagation paths, and signal delays. They extracted sensors' location information and communication errors into a joint state estimation based on a normal distribution. Utilizing the FIM, [19] proposed various optimization criteria that serve as a foundation for investigating optimal network topologies for wireless sensors. Specifically, the Cramér-Rao Lower Bound (CRLB) [20] represents the minimum variance to be minimized in the positioning estimation process. The Squared Position Error Bound (SPEB) [10] denotes the sum of squares to reduce the position error. Both optimization objective functions are based on FIM and can be utilized to optimize the system configuration.

However, we have to acknowledge that in complex environments like outer space [21], these wireless sensor systems encounter challenges, including limited power supply and insufficient measurement range. In such cases, optical sensors (such as cameras) are usually employed by researchers usually employ the technology of object recognition and feature extraction to precisely analyze the relative bearing and ranging information that optical sensors obtain, which can replace wireless sensors for large-scale localization and navigation due to their advantages of being information-rich, highly

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3657144

accurate, and energy-efficient. In addition, optical sensor facilitates non-contact measurement, allowing adaptation to complex environments and target localization, which eliminates the drawback of wireless sensors needing to communicate with users for accurate positioning. Finally, in unknown and complex environments, optical sensors are less susceptible to interference from electromagnetic waves or other noise [22], [23]. However, the research on optimizing the deployment of optical sensors for target localization has not attracted widespread attention.

Lastly, we summarize that many studies focus on optimizing sensor placement but often overlook the significant impact that the internal parameters of sensors can have on the overall positioning accuracy. This paper explicitly examines the sensing principles of cameras for relative bearing measurement and analyzes the sources of measurement error during the positioning process. By employing FIM modeling, we investigate the optimal arrangement of multiple cameras to maximize cooperative localization accuracy. Additionally, this study applies FIM to evaluate how camera hardware parameters, such as resolution, focal length, the number of cameras, and their distance from targets, affect positioning precision.

III. Problem Formulation

We consider Multi-Robot System (MRS) where each robot equipped with a camera can be utilized to determine the location of a target using a relative bearing-measurement method, which is often described by relative angle information between the target and the camera. In the following manuscript, we use a simplified method to represent the i-th camera as the camera mounted on the i-th mobile robot. The principle of relative bearing measurement using the camera is shown in Fig. 1a; the relative angle between the target and the camera can be calculated by measuring the pitch angle (β) and yaw angle (α). Given the intrinsic parameters of cameras, we can calculate the relative bearing of the target with respect to the camera's optical center by the pixel coordinates of the image centroid of the target on the image plane. Note that we can simultaneously derive the relative yaw and pitch angle from such a measurement process. In this work, we concentrate on revealing how the intrinsics and extrinsics of cameras can contribute to the accuracy of target localization using the multi-camera-based target localization system by FIM and D-optimality criteria. Therefore, without loss of generality, this work mainly analyzes the target localization problem with relative bearing measurements by cameras in the 2-D space, i.e., only the yaw angle measurement is considered. This can be realized in practice by setting the height of the cameras' optical centers to be identical, ensuring that the imaging planes of the cameras are perpendicular to the ground. As shown in Fig. 1b, let denote the coordinates of the i-th mobile robot and the target in the world frame of reference $\mathcal{O}_W$ as (x_{ci}, y_{ci}) and (x_t, y_t), respectively. The true value of the relative angle between the target and the i-th camera can be expressed as:

$$\theta_i = \arctan \frac{y_{ci} - y_t}{x_{ci} - x_t} \tag{1}$$

Furthermore, based on the principle of camera-based relative bearing measurement, the relative angle between the i-th camera and the target, which can also be possessed by

$$\theta_i = \psi_i - \alpha_i \tag{2}$$

where α_i represents the relative bearing between the optical axis of the camera and the target, which can be measured by the i-th camera to the target in the camera coordinate system, as shown in Fig. 1c. ψ_i is the angle of the i-th mobile robot's heading direction, which can be obtained from the odometry of the mobile robot. In this manuscript, the camera is rigidly mounted, and the optical axis direction of the camera is consistent with the heading direction of the robot. Practically, the mobile robots need to control their yaw angle ψ to keep the visibility of the target to the carried cameras due to the limited field of vision (FOV). Therefore, once the exact location of a given target is provided, if two or more robots equipped with bearing-measurement sensors can compute the relative angles between the target and themselves, the position of the target can be estimated. The calculation of the target position is based on the following

$$\begin{bmatrix} y_{c1} - x_{c1}\tan\theta_N \\ y_{c2} - x_{c2}\tan\theta_N \\ \vdots \\ y_{cN} - x_{cN}\tan\theta_N \end{bmatrix} = \begin{bmatrix} -\tan\theta_1 & 1 \\ -\tan\theta_2 & 1 \\ \vdots & \vdots \\ -\tan\theta_N & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{3}$$

where $(x_{ci}, y_{ci}), i = 1, 2..., N$ is the coordinates of cameras in world coordinates, (x_t, y_t) represents the target. $\theta_i, i = 1, 2, ...N$ stands for the relative angle between the target and the cameras.

Let us introduce $\boldsymbol{X}$ as the vector of unknown parameters in the camera-based target localization system, and we use $\boldsymbol{\theta}$ as the vector to represent the relative angle of the observation by cameras. The FIM of the camera-based target localization system based on bearing measurement can be expressed as:

$$\text{FIM} = \mathbb{E}_{\boldsymbol{X}} \left\{ -\frac{\partial^2}{\partial \boldsymbol{X} \partial \boldsymbol{X}^T} \ln g(\boldsymbol{\theta}|\boldsymbol{p}) \right\} \tag{4}$$

where $g(\boldsymbol{\theta}|\boldsymbol{p})$ is the joint probability density function of the parameter vector $\boldsymbol{X}$. Without loss of generality, we consider the Gaussian model for the camera-based target localization system, the conditional probability density function of $\boldsymbol{\theta}$ can be given by

$$g(\boldsymbol{\theta}|\boldsymbol{p}) = \frac{1}{\sqrt{(2\pi)^N \det(\boldsymbol{\Sigma})}} \times \exp\left\{ -\frac{1}{2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\boldsymbol{p}))^T \boldsymbol{\Sigma}^{-1} ((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\boldsymbol{p})) \right\} \tag{5}$$

Where $\boldsymbol{p} = [x_t, y_t]^T$ is the target's position, and $\boldsymbol{\Sigma}$ is the covariance matrix of relative bearing measurement based on the camera, which is computed by the measurement error.

Previous efforts [6]–[9] focused on employing a general error model that the measurement error assumes to be independent, Gaussian, zero-mean, with a random variance. However, the sensing principle of measurement remains overlooked in the optimization problem of target localization. In this paper, we denote the coordinates of the i-th mobile robot's camera imaging plane center point (x_{0i}, y_{0i}, z_{0i}), the centroid pixel

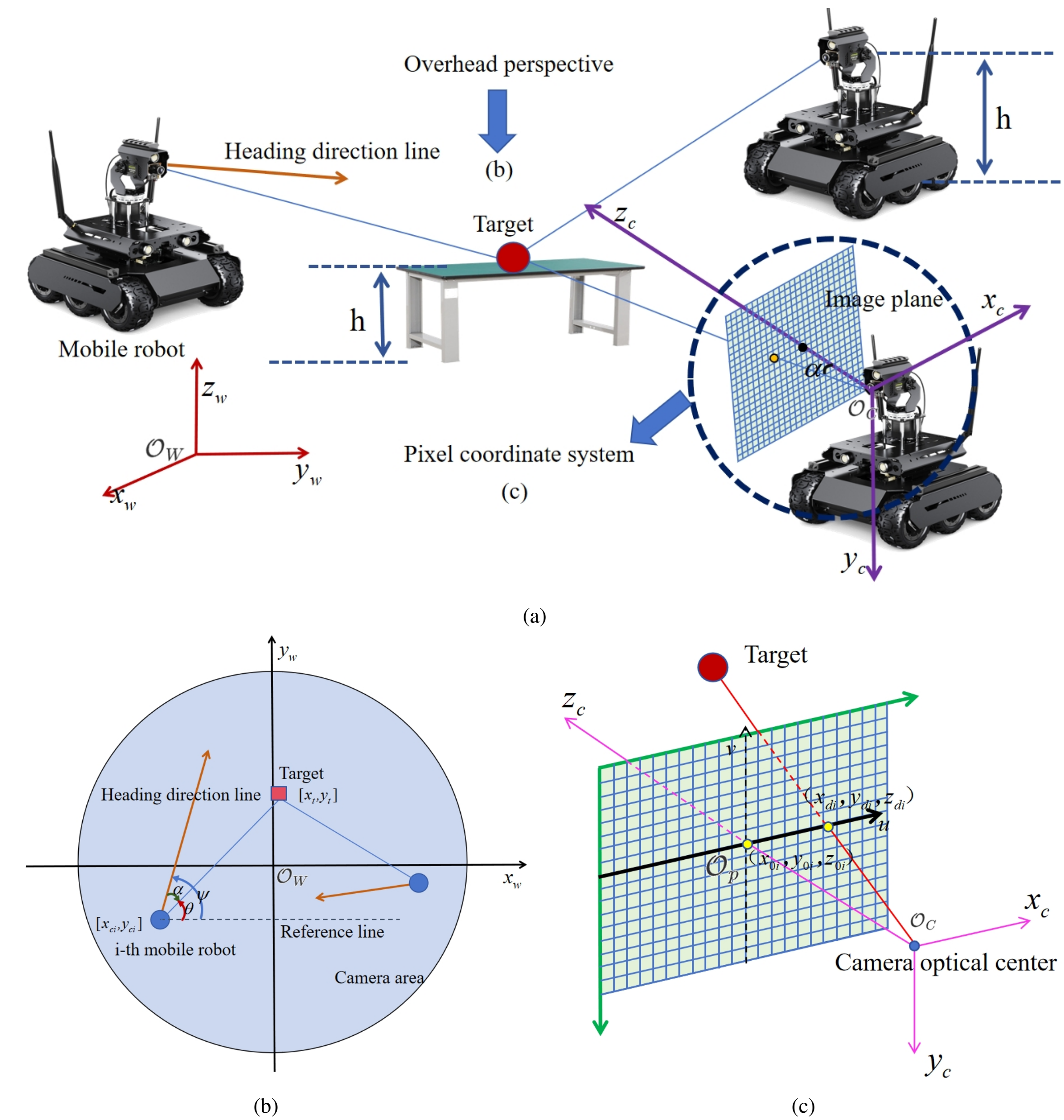


Fig. 1: The principle of relative bearing measurement using cameras

coordinate of the target extracted on the imaging plane is (x_{di}, y_{di}, z_{di}) in camera coordinate. The absolute value of the measurement angle of the i-th mobile robot can be obtained as follows:

$$|\alpha_i| = \arctan \frac{\sqrt{(y_{di} - y_{0i})^2 + (x_{di} - x_{0i})^2}}{f_i} \tag{6}$$

where f_i is the focal length of the i-th camera, and α_i is positive in a counterclockwise rotating direction and negative otherwise. It is noteworthy that the uncertainty in the centroid coordinates on the imaging plane and the optical center position of the camera can affect the relative bearing measurement.

Specifically, fluctuations in the edge pixel points of the target on the imaging plane can lead to variations in the extracted centroid coordinates. This error is influenced by the camera's resolution and focal length. Additionally, optical center shifts and imaging plane distortions caused by manufacturing processes can significantly impact relative bearing calculations, the error of which is closely related to the focal length. We provide a detailed analysis of the specific relationship in Section IV. Consequently, the influence of the intrinsic parameters of the camera-based target localization system should not be overlooked. Then we use focal length and resolution vector $\boldsymbol{C}_i = [f_i, F_i]$ to stand for the intrinsic parameters of the i-th camera, and $\boldsymbol{C} = [\boldsymbol{C}_i^T, \boldsymbol{C}_2^T, \cdots, \boldsymbol{C}_N^T]$.

Based on the analysis in (2) (5) (6), the position and the intrinsic parameters of all cameras can influence the accuracy of camera-based target localization. Then we could describe the unknown parameter vector $\boldsymbol{X}$ as

$$\boldsymbol{X} = [\boldsymbol{\theta}, \boldsymbol{C}] \tag{7}$$

where $\boldsymbol{\theta}$ is the vector of all the relative angles between the user and the cameras that are equipped on the mobile robots.

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \cdots, \boldsymbol{\theta}_N^T] \tag{8}$$

where $\boldsymbol{\theta}_i = [\theta_i, 0], i = 1, 2, \cdots, N$, $[x_{ci}, y_{ci}]$ is the position of the i-th camera and $[x_t, y_t]$ is the position of the target.

Let us introduce an estimator $\hat{\boldsymbol{X}}$ of the unknown vector $\boldsymbol{X}$ based on the observation. The mean squared error (MSE) matrix of $\hat{\boldsymbol{X}}$ follow the inequality

$$\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{X}} \left\{(\hat{\boldsymbol{X}} - \boldsymbol{X})(\hat{\boldsymbol{X}} - \boldsymbol{X})^T\right\} \succeq \text{FIM}_{\boldsymbol{X}}^{-1} \tag{9}$$

where $\text{FIM}_{\boldsymbol{X}}$ is the FIM for the parameter vector $\boldsymbol{X}$, which contain the information of θ. Inspired by [10], we note that parameter θ includes the target's position. Consequently, let us denote the $\hat{\boldsymbol{p}}$ as the estimated vector of the target's position, which follows from (9).

$$\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{X}} \left\{(\hat{\boldsymbol{p}} - \boldsymbol{p})(\hat{\boldsymbol{p}} - \boldsymbol{p})^T\right\} \succeq [\text{FIM}_{\boldsymbol{X}}^{-1}]_{2\times 2} \tag{10}$$

and

$$\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{X}} \left\{||(\hat{\boldsymbol{p}} - \boldsymbol{p})||^2\right\} \geq tr\left\{[\text{FIM}_{\boldsymbol{X}}^{-1}]_{2\times 2}\right\} \tag{11}$$

If in 3-D localization, we should consider the $tr\left\{[\text{FIM}_{\boldsymbol{X}}^{-1}]_{3\times 3}\right\}$. Lastly, we concentrate on finding the minimum of $tr\left\{[\text{FIM}_{\boldsymbol{X}}^{-1}]\right\}$ to reduce the lower bound of the MSE in target localization, which is the CRLB. However, for a non-singular 2×2 matrix FIM, the following relations are held:

$$tr(\text{FIM}^{-1}) = \frac{tr(\text{FIM})}{\det(\text{FIM})} \tag{12}$$

The (12) shown that $\text{FIM}^{-1} \propto \frac{1}{\det(\text{FIM})}$. We focus on seeking the maximum $\det(\text{FIM})$ that can also reduce the lower bound of MSE, which is the D-optimality criterion.

Remark 1: The FIM serves as a metric for assessing the accuracy and effectiveness of parameter estimation. It can be utilized in conjunction with various optimization criteria to evaluate the impact of the geometric configuration between sensors and the target on localization accuracy. Specifically, the A-optimality aims to minimize the trace of the inverse FIM, thereby minimizing the sum of the diameters of the error ellipse. The D-optimality seeks to minimize the determinant of the FIM, thus minimizing the area of the error ellipse. The E-optimality focuses on minimizing the maximum eigenvalue of the inverse FIM, achieving the shortest possible major axis of the error ellipse. These criteria can all be employed to optimize sensor placement to enhance localization accuracy. In this study, we consider selecting the D-optimality.

In conclusion, in this paper, we acknowledge that both extrinsic and intrinsic parameters of the cameras in use significantly impact the precision of target localization. More specifically, extrinsic parameters establish the geometric configuration of the cooperative measurement. Concurrently, intrinsic parameters, including focal length and resolution, contribute to the propagation of target pixel uncertainty on the image plane to the yaw angle, particularly in the presence of measurement noise. Therefore, the strategic selection and placement of cameras emerge as vital factors in minimizing the localization uncertainty of the target. This study aims to construct a mathematical model that elucidates this relationship, thereby facilitating the optimization of camera selection and placement. Considering the (11) and (12), the problem of optimal placement and selection of the cameras can be stated as the following problem:

$$\boldsymbol{X} = \arg \max_{\boldsymbol{X}=[\boldsymbol{\theta},\boldsymbol{C}]} \det(\text{FIM}_{\boldsymbol{X}}) \tag{13}$$

IV. System Modeling Method

In this paper, we consider the target localization scenario based on relative bearing measurements from robots equipped with cameras. The relative bearing measurement of the i-th robot can be represented as follows:

$$\hat{\theta}_i = \hat{\psi}_i - \hat{\alpha}_i \tag{14}$$

where

$$\begin{aligned} \hat{\psi}_i &= \psi_i + \omega_i' \\ \hat{\alpha}_i &= \alpha_i + \omega_i'' \end{aligned} \tag{15}$$

where ψ_i is the true value of the i-th robot's heading direction angle. We consider that ω_i' is a Gaussian noise with zero-mean and stochastic variance. Besides, the ω_i'' is the measurement error from the i-th camera. In this paper, we assume that all mobile robots are equipped with identical cameras, i.e., $\omega_1', \omega_2', \cdots, \omega_N'$ are independent and identically distributed, similarly, $\omega_1'', \omega_2'', \cdots, \omega_N''$ are independent and identically distributed, and we define that $\omega_i' \backsim \mathcal{N}(0, \sigma_0^2)$.

Remark 2: The robot's position and heading information are derived from its odometer and inertial measurement unit (IMU). While external factors are disregarded, the primary source of error stems from the sensor's internal noise, typically modeled as Gaussian white noise. This noise can lead to varying degrees of offset in the camera's heading angle.

This paper aims to develop an accurate model that explains the influence of internal and external camera parameters on positioning error, rather than model the error as previously known but parameter-independent general Gaussian noise. To achieve this, we need to study the relationship between the relative bearing measurement error and the camera's intrinsic parameters. By analyzing the pinhole imaging model, we know that the error in camera relative bearing measurement primarily stems from the imaging error of the target on the image plane and the intrinsic bias of the camera. Mathematically, ω'' consists of random error ω_a and geometric error ω_b. In addition, each camera in our effort is installed such that its optical axis height aligns with the target height, and the imaging plane is perpendicular to the ground. Without loss of generality, we consider only the yaw angle while disregarding the effects of the pitch angle on actual bearing measurements. Consequently, before analyzing the errors in the target localization of camera-based bearing measurement, the following assumptions are made.

Assumption 1: Regardless of whether the deviation is caused by random errors or geometric error, the impact on the centroid position of the target on the imaging plane, when observed within the pixel coordinate system $\mathcal{O}_P$, manifests as a fluctuation of the centroid position under error-free conditions along the u-axis.

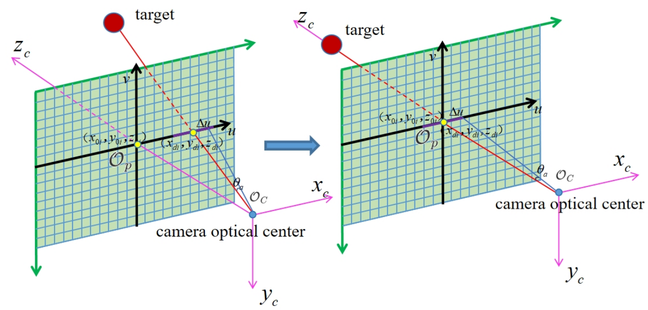


Fig. 2: Relative bearing measurement model change

Random error arises from stochastic noise, specifically manifested as illumination noise due to unstable external light sources and random noise in camera imaging. These factors cause fluctuations at the edges of the target's image on the image plane, ultimately leading to variations in the centroid coordinates around their original position, with the magnitude of these fluctuations Δu can be modeled as a Gaussian model with a mean of 0 and a variance of σ_r^2, i.e. $\Delta u \backsim \mathcal{N}(0, \sigma_r^2)$. In the camera coordinate system, the random error indicating the intrinsic bias ω_a of the i-th mobile robot is represented as

$$\begin{aligned}\omega_a =& \arctan \frac{\sqrt{(x_{d_i} + \Delta u - x_{0_i})^2 + (y_{d_i} - y_{0_i})^2}}{f} - \\ & \arctan \frac{\sqrt{(y_{d_i} - y_{0_i})^2 + (x_{d_i} - x_{0_i})^2}}{f}\end{aligned} \tag{16}$$

However, the expression of the intrinsic bias is complicated and it is difficult to describe it by a mathematical model, the method of relative bearing measurements can be modified such that the orientation line and the line of the camera's optical center and the target are in coincident i.e. ignore the measurement angle, The process of change is shown in the Fig. 2, the purple line in which is where the centroid coordinates might appear, the random error can be expressed as:

$$\omega_a = \arctan \frac{\Delta u}{f} \tag{17}$$

Given that the value of Δu is small, the intrinsic bias can be approximated to $\arctan \frac{\Delta u}{f} \approx \frac{\Delta u}{f} + o(\frac{\Delta u}{f})$ based on the Taylor expansion, the random error can be considered as a normal distribution with a mean of 0 and a variance of $\frac{\sigma_r^2}{f^2}$ i.e. $w_a \backsim \mathcal{N}(0, \frac{\sigma_r^2}{f^2})$ by the omission of the higher-order term.

Remark 3: After utilizing the Taylor expansion while neglecting higher-order terms, ω_a is approximately equal to $\frac{\Delta u}{f}$; the use of such approximation methods does not affect the actual results of error analysis. In practice, random error values are typically very minor, and the impact of neglecting higher-order terms can be completely disregarded. Regarding the relationship between camera resolution and random error, although numerous studies have been conducted on this topic, no specific relationship has been established. In this paper, a direct proportional relationship is employed for approximation, which can be considered as an open problem for future research.

We note that the random error arises from fluctuations in the centroid coordinates on the imaging plane, closely linked to the pixel size of the camera's sensor. To achieve a more precise calculation of the centroid coordinates of all the imaging points of a target on the imaging plane, it is often desirable for the target's image to cover a larger number of pixels. A greater amount of pixel data can lead to a more accurate computation of the centroid coordinates. The relationship between pixel size and resolution is as follows:

$$s = \sqrt{\frac{1}{F}} \tag{18}$$

where s is the pixel size and F is the camera resolution. According to previous efforts ' experience [24], [25], with an increase in pixel size, the measurement error associated with the optical sensor also escalates. then we consider that ω_a is approximately proportional to s, which can be expressed as $\omega_a \propto \sqrt{\frac{1}{F}}$, on account of $w_a \backsim \mathcal{N}(0, \frac{\sigma_r^2}{f^2})$, $\frac{\sigma_r}{f} \propto \sqrt{\frac{1}{F}}$,i.e.,

$$\frac{\sigma_r}{f} = k\sqrt{\frac{1}{F}} \tag{19}$$

where k is a positive constant.

In addition, under ideal conditions, the projection of the i-th camera's optical center onto the image plane (x_{oi}, y_{oi}, z_{oi}) should coincide with the geometric center of the image plane. However, the projection position may also shift due to optical center offset and imaging plane distortion, which is called geometric error ω_b, resulting from the camera installation itself and long-term use. We note that the existence of the optical center could be the projection position offset along one direction; besides, the imaging plane distortion leads to the projection position along the other direction. Inspired by the paper about probability statistics, the combined influence of these two factors can cause the projection position offset from the center of the imaging center, and the amplitude of the offset is Δv. In this study, the height of the camera's optical center is aligned with the target's height, so the geometric errors only affect the camera-based bearing measurements in the yaw direction. Consequently, we can assume the projection coordinate after the shift is $(x_{oi} + \Delta v, y_{oi}, z_{oi})$, which can be expressed as:

$$\omega_b = \arctan \frac{\Delta v}{f} \tag{20}$$

In our work, we consider that the offset of the projection coordinate in the imaging plane derives from optical offset and imaging plane distortion is independent and identically distributed, which is the normal distribution with a mean of 0 and a variance of σ_g^2. Inspired by the [26], the magnitude of this resultant displacement follows a Rayleigh distribution. The probability density of the length of Δv satisfies the following relation.

$$f(\Delta v) = \frac{\Delta v}{\sigma_g^2} e^{-\frac{\Delta v^2}{2\sigma_g^2}} \ , \Delta v > 0 \tag{21}$$

Through calculation, $\mathbb{E}(\Delta v) = 1.25\sigma_g$ and $var(\Delta v) = 0.43\sigma_g^2$, the $\mathbb{E}(*)$ stands for mathematical expectation, the $var(*)$ stands for variance. Employing the identical analytical approach as that for the random error, the value of Δv is extremely small. ω_b can be regarded as equivalent to $\frac{\Delta v}{f} + o(\frac{\Delta v}{f})$ by utilizing the Taylor expansion, then the $\omega_b \approx \frac{\Delta v}{f}$ after omitting the higher-order terms. According to the law of large numbers [27] and the central limit theorem [28], Δv can be approximated by the normal distribution i.e. $\omega_b \backsim \mathcal{N}(\frac{1.25\sigma_g}{f}, \frac{0.43\sigma_g^2}{f^2})$.

Since random error ω_a and geometric error ω_b are independent of each other, ω_i'' can also be considered to follow a normal distribution

$$\omega_i'' \backsim \mathcal{N}(\frac{1.25\sigma_g}{f}, \frac{\sigma_r^2}{f^2} + \frac{0.43\sigma_g^2}{f^2}) \tag{22}$$

Besides, we consider that the error from the i-th mobile robot's heading direction ω_i' and the error from the i-th camera ω_i'' are independent. Ultimately, we can derive the error model for the camera-based target localization system using bearing measurements.

$$\omega_i \backsim \mathcal{N}(\frac{1.25\sigma_g}{f}, \frac{\sigma_r^2}{f^2} + \frac{0.43\sigma_g^2}{f^2} + \sigma_0^2) \tag{23}$$

Remark 4: In our work, the errors of camera-based bearing measurement come from three parts, one of which is the error from the robot's odometer and IMU. This result of the robot shaking, and has no connection with the camera. Random error arises from stochastic noise, specifically manifested as illumination noise due to unstable external light sources and noise in camera imaging. Lastly, the geometric error comes from the optical center offset and the image plane distortion. In fact, the mechanisms underlying these three types of errors differ significantly, resulting in minimal shared information among them. Therefore, we can consider them as completely independent errors.

Combining (14) the bearing measurement between the target an i-th camera is

$$\hat{\theta}_i = \theta_i + \omega_i \tag{24}$$

Consequently, the camera-based bearing measurements can be expressed in a vector form, as

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \boldsymbol{\omega} \tag{25}$$

Where $\boldsymbol{\theta} = \left[\theta_1, \theta_2, \cdots, \theta_N\right]^T \boldsymbol{\omega} = \left[\omega_1, \omega_2, \cdots, \omega_N\right]^T$, the covariance matrix of camera measurements is denoted by

$$\boldsymbol{\Sigma} = \mathbb{E}((\boldsymbol{\omega} - \mathbb{E}(\boldsymbol{\omega}))(\boldsymbol{\omega} - \mathbb{E}(\boldsymbol{\omega}))^T) \tag{26}$$

This covariance matrix can be obtained as

$$\boldsymbol{\Sigma} = \delta^2 \mathbf{I} \tag{27}$$

Where $\mathbf{I}$ is an identity matrix of order N and $\delta^2 = \frac{\sigma_r^2}{f^2} + \frac{0.43\sigma_g^2}{f^2} + \sigma_0^2$.

We have modeled the measurement error by considering the functioning principles of the cameras, by which all the intrinsic parameters are included in the model, then the well-known FIM in (4) is applied to evaluate the relationship between the geometrical configuration of multiple cameras and the localization accuracy of the target. Given that the error ω, the FIM derived from bearing measurements can be directly expressed as follows:

$$\text{FIM} = \nabla_{\boldsymbol{p}} \boldsymbol{\theta}(\boldsymbol{p})^T \boldsymbol{\Sigma}^{-1} \nabla_{\boldsymbol{p}} \boldsymbol{\theta}(\boldsymbol{p}) \tag{28}$$

where

$$\nabla_{\boldsymbol{p}} \boldsymbol{\theta}(\boldsymbol{p}) = \begin{bmatrix} \frac{\partial \theta_1}{\partial x_{t1}} & \frac{\partial \theta_1}{\partial y_{t1}} \\ \vdots & \vdots \\ \frac{\partial \theta_N}{\partial x_{tN}} & \frac{\partial \theta_N}{\partial y_{tN}} \end{bmatrix} = \begin{bmatrix} -\frac{\sin\theta_1}{d_1} & \frac{\cos\theta_1}{d_1} \\ \vdots & \vdots \\ -\frac{\sin\theta_N}{d_N} & \frac{\cos\theta_N}{d_N} \end{bmatrix} \tag{29}$$

By substituting (29) and (27), the FIM can be expressed as

$$\text{FIM} = \frac{1}{\delta^2} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{30}$$

where

$$\begin{aligned} A_{11} &= \sum_{i=1}^{N} \frac{\sin^2\theta_i}{d_i^2} \quad A_{22} = \sum_{i=1}^{N} \frac{\cos^2\theta_i}{d_i^2} \\ A_{12} &= A_{21} = \sum_{i=1}^{N} \frac{\sin\theta_i \cos\theta_i}{d_i^2} \end{aligned} \tag{31}$$

and $\delta^2 = \sigma_0^2 + \frac{k^2}{F} + \frac{0.43\sigma_g^2}{f^2}$, which shows the functional relationship between camera resolution, focal length, number, observation distance, and positioning accuracy.

We then propose the evaluation metric for optimizing the geometrical configuration and intrinsic parameters of the multi-camera-based target localization system using the D-optimality as

$$\det(\text{FIM}) = A_{11}A_{22} - A_{12}A_{21} \tag{32}$$

In order to refine the problem of optimal placement and selection of the cameras in camera-based target localization, we put the calculation result in (32) into (13).

Remark 5: The CRLB represents the theoretical lower limit on the variance of any unbiased parameter estimator. However, the estimator of θ in this work is biased, i.e., $E(\hat{\theta}) = \theta + \frac{1.25\sigma_g}{f}$, which seems to prove that the FIM and D-optimality criteria we derived are invalid. On the contrary, we notice that if $\hat{\theta}$ is a biased estimate, then $E(||\hat{\theta} - \theta||^2) \geq tr(\text{FIM}^{-1})(1 + \frac{db(\theta)}{d\theta})$ where $E(\hat{\theta}) = \theta + b(\theta)$ and $b(\theta) = \frac{1.25\sigma_g}{f}$ which have no relationship with the θ, i.e., $\frac{db(\theta)}{d\theta} = 0$. Consequently, we can conclude that although θ is a biased estimation in our work, it has no influence on the result of $E(||\hat{\theta} - \theta||^2)$. The accurate model we proposed using FIM and D-optimality is reasonable.

V. Optimal Deployment Scenarios

In this section, we examine the proposed D-optimality of the FIM as the performance metric for optimizing the placement of multiple cameras. This issue is explored in two separate scenarios: the single-target localization problem when the target is within the camera placement area, and when it is outside.

A. Scenario 1: Target being inside the camera placement area

When the target is inside the camera area, no constraint is exerted on the selectable angles θ_i. Accordingly, the optimal solution in this situation is the same as that of the unconstrained placement matter.

In the following theorem, the maximum value of the determinant of the FIM is calculated and derived, and the conditions for achieving the optimal camera placement are provided.

Theorem 1: The maximum value of the determinant of the FIM for the angle measurement localization with N sensors is equal to

$$\det(\mathbf{FIM}) = \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{d_i^2} \tag{33}$$

and this maximum value is attained when

$$\sum_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \tag{34}$$

Generally, the bearing measurement geometry of all cameras fulfilling the condition of equation (33) cannot be acquired in a closed-form format. Consider the special case where all cameras possess the same measuring distance, i.e., $d_1 = d_2 = \cdots = d_N = d$. One of the evident optimal geometric solutions satisfying the condition of equation (33) is

$$\theta_i = \frac{2\pi i}{N} + \theta_0, i = 0, 1...N \tag{35}$$

Where θ_0 is any angle, then the camera is divided to meet the uniform angle array (UAA).

Proof:

$$\det(\mathbf{FIM}) = \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{d_i^2} - \left(\frac{1}{\delta^2}\sum_{i=1}^{N}\frac{\sin\theta_i\cos\theta_i}{d_i^2}\right)^2 \tag{36}$$

The following inequalities can be represented as the $\det(\mathbf{FIM})$

$$\det(\mathbf{FIM}) \overset{a}{\leq} \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{d_i^2} \tag{37}$$

and the equality in (a) is held when the following conditions are satisfied:

$$\sum_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \tag{38}$$

Finally, the conclusion can be reached

$$\underset{\boldsymbol{X}}{max}\,\det(\mathbf{FIM}) = \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{d_i^2} \tag{39}$$

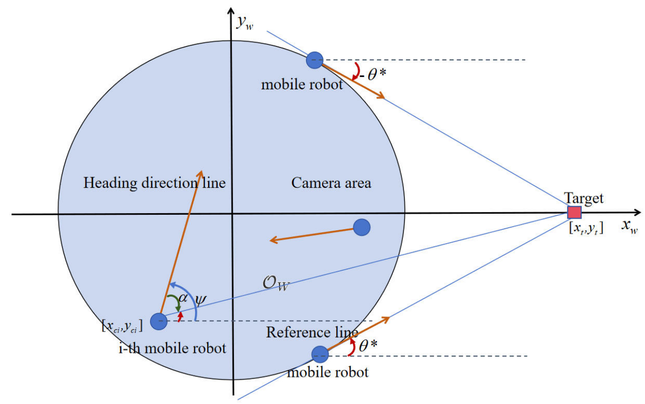


Fig. 3: Target outside the camera area

B. Scenario 2: Target being outside camera placement area

When the target is outside the camera area, let us consider the center of the camera area as the origin of the world coordinate system, and the target is located on the positive half-axis of the x-axis of the world coordinate system. As shown in Fig. 3 under this condition, the absolute value of the maximum relative bearing (θ^*) between the target and the camera is always less than $90°$, i.e., $|\theta_i| \leq \theta^* < 90°$ $i = 1, 2, ..., N$. Indeed, when the target is outside the camera area, certain constraints are added to the optimization problem such that

$$\begin{aligned} &\max_{\boldsymbol{X}}\ \det(\mathbf{FIM}) \\ &s.t.\ \ |\theta_i| \leq \theta^* < 90°\ , i = 1, 2, \cdots, N \end{aligned} \tag{40}$$

The following theorem states the optimal camera placement under communication constraints.

Theorem 2: Depending on whether the number of sensors N is even or odd, the solution of the optimization about the source outside the camera area can be divided into the following two cases:

1)N is even

If $\theta^* \geq 45°$, then the optimal angular relative bearing $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\sum_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \quad \sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2} = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{d_i^2} \tag{41}$$

If $\theta^* < 45°$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\sum_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \quad \sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2} = \sum_{i=1}^{N}\frac{\sin^2\theta^*}{d_i^2} \tag{42}$$

proof:

$$\det(\mathbf{FIM}) \leq \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{di^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{di^2} \tag{43}$$

and that equality holds when

$$\sum_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \tag{44}$$

Then $\det(\mathbf{FIM})$ can be expressed as

$$\det(\mathbf{FIM}) = \frac{1}{\delta^4}\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}(\sum_{i=1}^{N}\frac{1}{d_i^2} - \sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}) \tag{45}$$

Under the unconstrained condition, the optimal $\sum\limits_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}$ is

$$(\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2})_{opt} = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{d_i^2} \tag{46}$$

However, we must note that $|\theta_i| \leq \theta^* < 90°$, so the value of

$$\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2} \leq \sum_{i=1}^{N}\frac{\sin^2\theta^*}{d_i^2} \tag{47}$$

Then we could calculate the extreme situation when

$$(\sum_{i=1}^{N}\frac{\sin^2\theta^*}{d_i^2}) = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{d_i^2} \tag{48}$$

that means $\sin^2\theta^* = \frac{1}{2}$ and $\theta^* = 45°$

If $\theta^* \geq 45°$, which means $(\sum\limits_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2})_{opt}$ is in the feasible region, the optimal placement of the camera is

$$(\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2})_{opt} = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{d_i^2} \tag{49}$$

On the other hand, if $\theta^* < 45°$, which means no matter how to change the relative bearing between the target and the camera, the $(\sum\limits_{i=1}^{N}\frac{\sin^2\theta^*}{d_i^2})_{opt}$ is not satisfied. Consequently, when $|\theta_i| \in (0, \theta^*)$, the value of $\det(\mathbf{FIM})$ is gradually increasing in size, and the optimal placement of the camera is

$$(\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2})_{opt} = \sum_{i=1}^{N}\frac{\sin^2\theta^*}{d_i^2} \tag{50}$$

discussion: Consider the special case with the same distance variance for all cameras,i.e., $d_1 = d_2 = \cdots = d_N = d$, for the N is even, the equality $\sum\limits_{i=1}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0$ is held when $\theta_i = -\theta_{\frac{N}{2}+i},\ i = 1, 2, \cdots, \frac{N}{2}$.

If $\theta^* \geq 45°$, the optimal camera placement is obtained as follows

$$\begin{aligned} &\sum_{i=1}^{N}\sin^2\theta_i = \frac{N}{2} \\ &\theta_i = -\theta_{\frac{N}{2}+i} \quad i = 1, 2, \cdots, \frac{N}{2} \end{aligned} \tag{51}$$

If $\theta^* < 45°$, the optimal camera placement is obtained as follows

$$\begin{aligned} &\theta_i = \theta^* \\ &\theta_{\frac{N}{2}+i} = -\theta^* \qquad i = 1, 2, \cdots, \frac{N}{2} \end{aligned} \tag{52}$$

■

2)N is odd

If $\theta^* \geq \arcsin(\sqrt{\frac{\sum_{i=1}^{N}\frac{1}{d_i^2}}{2\sum_{i=2}^{N}\frac{1}{d_i^2}}})$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &\theta_1 = 0° \\ &\sum_{i=2}^{N}\frac{\sin^2\theta_i}{d_i^2} = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{d_i^2} - \frac{\sin^2\theta_1}{d_1} \\ &\sum_{i=2}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \end{aligned} \tag{53}$$

If $\theta^* < \arcsin(\sqrt{\frac{\sum_{i=1}^{N}\frac{1}{d_i^2}}{2\sum_{i=2}^{N}\frac{1}{d_i^2}}})$ and $\theta^* \geq 45°$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &\theta_1 = 0° \\ &\sum_{i=2}^{N}\frac{\sin^2\theta_i}{d_i^2} = \sum_{i=2}^{N}\frac{\sin^2\theta^*}{d_i^2} \\ &\sum_{i=2}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \end{aligned} \tag{54}$$

If $\theta^* < 45°$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &|\theta_1| = \theta^* \\ &\sum_{i=2}^{N}\frac{\sin^2\theta_i}{d_i^2} = \sum_{i=2}^{N}\frac{\sin^2\theta^*}{d_i^2} \\ &\sum_{i=2}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0 \end{aligned} \tag{55}$$

Proof:

$$\begin{aligned} &\det(\mathbf{FIM}) = \\ &= \frac{1}{\delta^4}\left(\sum_{i=1}^{N}\frac{\sin^2\theta_i}{d_i^2}\sum_{i=1}^{N}\frac{\cos^2\theta_i}{d_i^2} - \left(\sum_{i=1}^{N}\frac{\sin\theta_i\cos\theta_i}{d_i^2}\right)^2\right) \\ &= \frac{1}{\delta^4}\left(\left(\frac{\sin^2\theta_1}{d_1^2} + A\right)\left(\sum_{i=1}^{N}\frac{1}{d_i^2} - \left(\frac{\sin^2\theta_1}{d_1^2} + A\right)\right) - \right. \\ &\left.\left(\frac{\sin\theta_1\cos\theta_1}{d_1} + \sum_{i=2}^{N}\frac{\sin\theta_i\cos\theta_i}{d_i}\right)^2\right) \\ &\overset{b}{\leq} \frac{1}{\delta^4}\left(\left(\frac{\sin\theta_1}{d_1} + A\right)\left(\sum_{i=1}^{N}\frac{1}{d_i^2} - \left(\frac{\sin\theta_1}{d_1} + A\right)\right) - \right. \\ &\left.\left(\frac{\sin\theta_1\cos\theta_1}{d_1}\right)^2\right) \end{aligned} \tag{56}$$

where $A = \sum\limits_{i=2}^{N}\frac{\sin^2\theta_i}{d_i^2}$, and that equality in (b) is held when $\sum\limits_{i=2}^{N}\frac{\sin 2\theta_i}{d_i^2} = 0$. Based on the derivative of $\det(\mathbf{FIM})$

with respect to A, the optimal $\sum_{i=2}^{N} \frac{\sin^2 \theta_i}{d_i^2}$ is obtained as

$$(\sum_{i=2}^{N} \frac{\sin^2 \theta_i}{d_i^2})_{opt} = \frac{1}{2} \sum_{i=1}^{N} \frac{1}{d_i^2} - \frac{\sin^2 \theta_1}{d_1} \tag{57}$$

As noted in (47), we compute the extreme condition as follows:

$$(\sum_{i=2}^{N} \frac{\sin^2 \theta^*}{d_i^2}) = \frac{1}{2} \sum_{i=1}^{N} \frac{1}{d_i^2} - \frac{\sin^2 \theta_1}{d_1} \tag{58}$$

that means

$$\theta^* = \arcsin \sqrt{\frac{\frac{1}{2} \sum\limits_{i=1}^{N} \frac{1}{d_i^2} - \frac{\sin^2 \theta_1}{d_1}}{\sum\limits_{i=2}^{N} \frac{1}{d_i^2}}} \tag{59}$$

If $\theta^* \geq \arcsin(\sqrt{\frac{\sum_{i=1}^{N} \frac{1}{d_i^2} - \frac{2\sin^2 \theta_1}{d_1^2}}{2\sum_{i=2}^{N} \frac{1}{d_i^2}}})$, which means $(\sum_{i=2}^{N} \frac{\sin^2 \theta_i}{d_i^2})_{opt}$ is in the feasible region, Through putting the $(\sum_{i=2}^{N} \frac{\sin^2 \theta_i}{d_i^2})_{opt}$ in (56)

$$\det(\mathrm{FIM}) = \frac{1}{4\delta^4} (\sum_{i=1}^{N} \frac{1}{d_i^2})^2 - \frac{1}{\delta^4} \frac{\sin^2 \theta_1 \sin^2 \theta_1}{d_1^2} \tag{60}$$

In order to maximize $\det(\mathrm{FIM})$, we should set θ_1 to zero. If $\theta^* < \arcsin(\sqrt{\frac{\sum_{i=1}^{N} \frac{1}{d_i^2} - \frac{2\sin^2 \theta_1}{d_1^2}}{2\sum_{i=2}^{N} \frac{1}{d_i^2}}})$, which means $(\sum\limits_{i=2}^{N} \frac{\sin^2 \theta_i}{d_i^2})_{opt}$ is not in the feasible region, When $\theta_i = \theta^*$, the value of $\det(\mathrm{FIM})$ is maximum. Through putting the $\sum\limits_{i=2}^{N} \frac{\sin^2 \theta^*}{d_i^2}$ in (56), the subsequent task lies in discussing the optimum layout of the camera when θ^* and θ_1 keep varying. Let us use $B = \sum\limits_{i=2}^{N} \frac{\sin^2 \theta^*}{d_i^2}$ to redescribe the $\det(\mathrm{FIM})$

$$\begin{aligned} &\det(\mathrm{FIM}) = \\ &\frac{1}{\delta^4}\left(\left(B\sum_{i=1}^{N}\frac{1}{d_i^2} - B^2\right) + \frac{\sin^2\theta_1}{d_1^2}\left(-2B + \sum_{i=2}^{N}\frac{1}{d_i^2}\right)\right) \\ &= \frac{1}{\delta^4}\left(\sum_{i=1}^{N}\frac{1}{d_i^2}B - B^2 + \frac{\sin^2\theta_1}{d_1^2}\left(1 - 2\sin^2\theta^*\right)\sum_{i=2}^{N}\frac{1}{d_i^2}\right) \end{aligned} \tag{61}$$

In this condition, If $\theta^* \geq 45°$, then the $1 - 2\sin^2\theta^* \leq 0$ and in order to maximize the $\det(\mathrm{FIM})$, the θ_1 must be set to 0,i.e.,$\theta_1 = 0$.

In addition, If $\theta^* < 45°$, then the $1 - 2\sin^2\theta^* > 0$ and in order to maximize the $\det(\mathrm{FIM})$, We need to maximize the $\frac{\sin^2\theta_1}{d_1^2}$. thus,the angle $|\theta_1|$ must be set to θ^*,i.e.,$|\theta_1| = \theta^*$.

discussion: Consider the special case with the same distance variance for all cameras, i.e., $d_1 = d_2 =, \cdots, = d_N = d$. for odd N and the distance variance for all cameras are same, the equality $\sum\limits_{i=2}^{N} \frac{\sin 2\theta_i}{d_i^2} = 0$ is held when $\theta_i = -\theta_{\frac{N-1}{2}+i}$, $i = 2, 3, \cdots, \frac{N+1}{2}$.

If $\theta^* \geq \arcsin(\sqrt{\frac{\sum_{i=1}^{N} \frac{1}{d_i^2}}{2\sum_{i=2}^{N} \frac{1}{d_i^2}}})$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &\theta_1 = 0° \\ &\sum_{i=2}^{N} \sin^2\theta_i = \frac{N}{2} \\ &\theta_i = -\theta_{\frac{N-1}{2}+i},\ i = 2, 3, \cdots, \frac{N+1}{2} \end{aligned} \tag{62}$$

If $\theta^* < \arcsin(\sqrt{\frac{\sum_{i=1}^{N} \frac{1}{d_i^2}}{2\sum_{i=2}^{N} \frac{1}{d_i^2}}})$ and $\theta^* \geq 45°$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &\theta_1 = 0° \\ &\theta_i = \theta^*,\ i = 2, 3, \cdots, \frac{N+1}{2} \\ &\theta_i = -\theta^*,\ i = \frac{N+3}{2}, \frac{N+5}{2}, \cdots, N \end{aligned} \tag{63}$$

If $\theta^* < 45°$, then the optimal angular geometry $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$ satisfy as follows

$$\begin{aligned} &|\theta_1| = \theta^* \\ &\theta_i = \theta^*,\ i = 2, 3, \cdots, \frac{N+1}{2} \\ &\theta_i = -\theta^*,\ i = \frac{N+3}{2}, \frac{N+5}{2}, \cdots, N \end{aligned} \tag{64}$$

■

VI. Simulation

In this section, we aim to design two types of simulation experiments. First, we aim to verify the validity of the optimized deployment methodology for a multi-camera-based target localization system across various scenarios, demonstrating that our proposed model, which concerns the localization accuracy of intrinsic and extrinsic camera parameters, can enhance target localization precision through optimized layout strategies. Additionally, we design simulation experiments to investigate the influence of intrinsic and extrinsic camera parameters on target localization accuracy in relation to the number of cameras, distance to the target, resolution, and focal length. Finally, we utilize our proposed mathematical model to provide suitable methods for camera selection.

A. Verification of the derivation of the optimal placement theory

In this subsection, we aim to verify the derivation of the optimal deployment of the multi-camera target localization system in Section V. We assume that the target is located at $(x_t, y_t) = [0, 0]$, and all cameras' focal length and resolution are $6mm$ and 1980×1024, each camera-to-target distance

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3657144

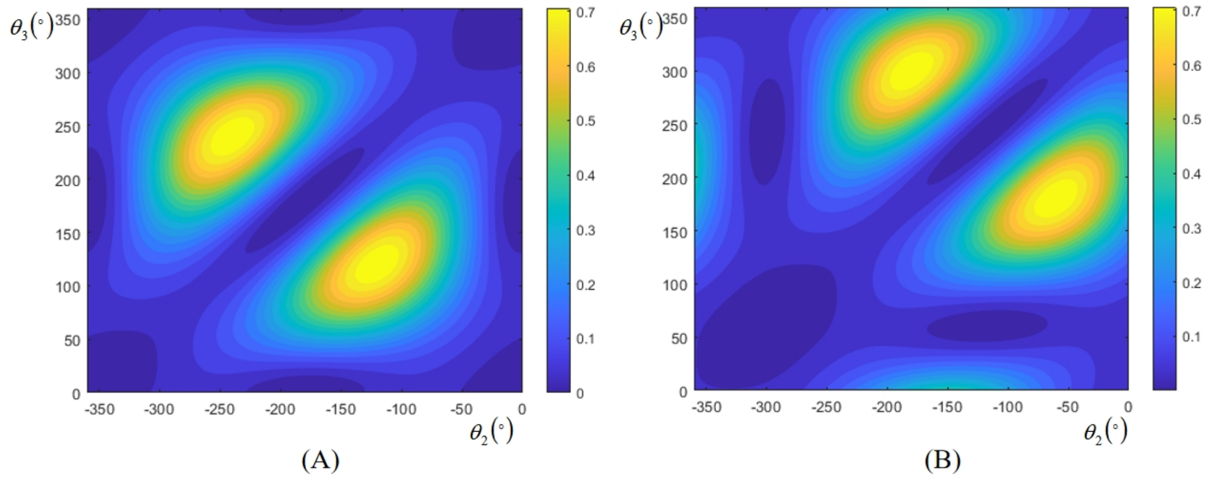


Fig. 4: Two-dimensional contour plot of $\det(\mathrm{FIM})$ when the target is within the camera area, and the camera is at different positions.

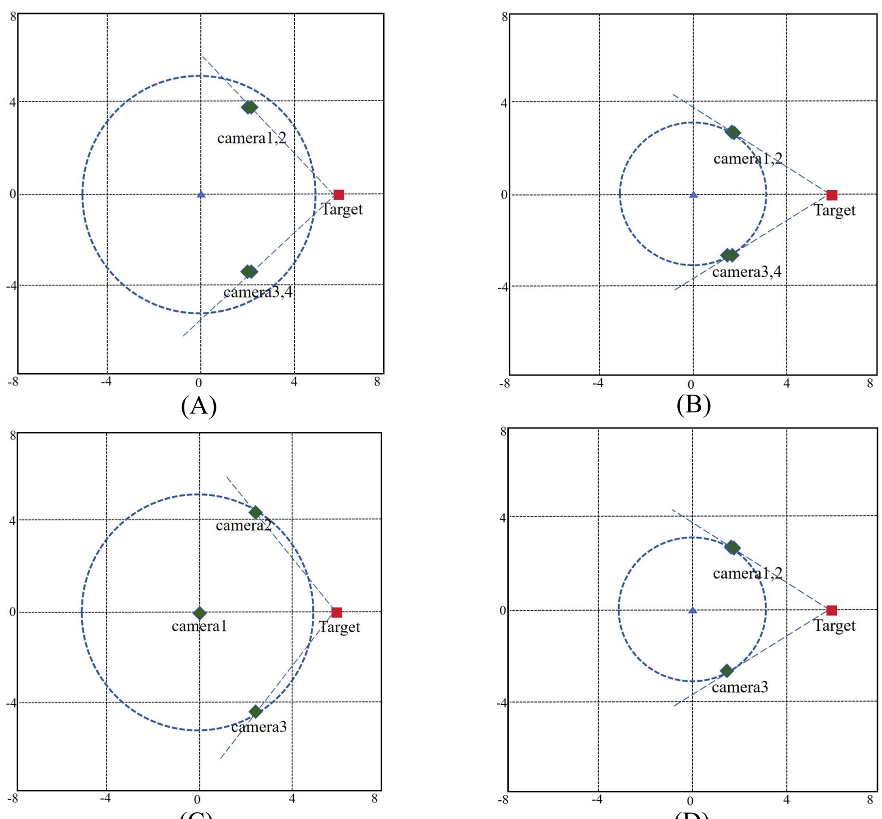


Fig. 5: The simulation about *Scenario 3,4*, which shows the optimal placement of cameras in different initial conditions, (A) $\theta^* = 51.32°, N = 4$, (B) $\theta^* = 30°, N = 4$, (C) $\theta^* = 51.32°, N = 3$, (D) $\theta^* = 30°, N = 3$

is maintained at 3, then we can easily calculate the bearing-measurement error approximately follows a normal distribution. Next, we design the two simulations to examine the correctness of *Theorem 1* in *Scenario 1* and *Scenario 2*. In the following simulation scenarios, we use θ_N to represent the relative angle between the i-th camera and the target.

Scenario 1: In this scenario, the location of the first camera is $[-3, 0]$, which shows the $\theta_1 = 0°$. However, the positions of the second and third cameras are stochastic; in order to determine the optimal geometric configuration of cameras, the positions of the second and third cameras have been varied constantly. Specifically, we constantly change the θ_2 from $0°$ to $360°$ and the θ_3 from $-360°$ to $0°$. Finally, the $\det(\mathrm{FIM})$ in (32) of cameras in different placement conditions can be computed using exhaustive search, and the heatmap of $\det(\mathrm{FIM})$ under different camera geometric configurations is shown in Fig. 4 (A).

Scenario 2: In this scenario, the initial position of the first camera has been changed into $[-\frac{3}{2}, -\frac{3\sqrt{3}}{2}]$, then the $\theta_1 = 60°$. Similar to *Scenario 1*, we constantly vary the relative angle θ_1 and θ_2. Finally, the heatmap can be acquired in Fig. 4 (B) to stand for the different geometric configuration in order to compare with the outcome in *Scenario 1*.

This paper employs a scaling method in Section V to determine the optimal camera placement when the target is within or outside the camera area through detailed and rigorous derivations. From the previous two scenarios, it is obviously shown that the optimal placement of all cameras in the target localization system is UAA when the target is in the camera area based on exhaustive search, which confirms our derivation in *Theorem 1*. In addition, we focus on the optimal deployment method when the target is outside the camera area. If we continue to use the exhaustive method to verify Theorem 2, it will result in an extremely large amount of computation. However, we note that many optimization algorithms, such as gradient ascent [29], Particle Swarm Optimization (PSO) [30], and genetic algorithms [31] are helpful. The purpose of employing these algorithms is to rapidly and correctly compute the optimal solution that satisfies maximization (32) although the calculation process is extremely complex. In this manuscript, PSO is a simple and efficient algorithm capable of rapidly computing optimal values in complex environments and is selected to perform auxiliary calculations to analyze the correctness of *Theorem 2* in *Scenario 3* and *Scenario 4*.

Scenario 3: In this scenario, the position of the target is located on $[6, 0]$ and 4 cameras are used to complete the positioning task, and the bearing measurement distance from all cameras to the target is 6. In addition, we firstly assume that the camera area is a circle whose center is $[0, 0]$ and radius is $3\sqrt{3}$, and then the $\theta^* = 51.32°$ can be easily computed by geometrical relationship. We search for the maximum value of (32) that represents optimal target localization accuracy using PSO to obtain the optimal geometric configuration of all cameras. Similarly, the same method is used to calculate the optimal solution when the radius of the camera area is 3 and the value of $\theta^* = 30°$. Finally, the outcome of optimal placement is shown in Fig. 5 (A)-(B).

Scenario 4: In this scenario, the initial condition and the method of simulation experiment are similar to *Scenario 3*; the only difference is that the number of cameras we used in this scenario is three. Lastly, the optimal deployment result is shown in Fig. 5 (C)-(D).

We carefully compare the results of the terminal placement obtained from PSO in Fig. 5 with our theoretical analysis *Theorem 2* where the target is outside the camera area, we can easily find that our theoretical derivations align with the situation that conform to the maximum $\det(\text{FIM})$, which also indicate the optimal positioning accuracy in practical localization. All the simulations in four scenarios demonstrate that the mathematical metrics for extrinsic parameters (the position of the cameras) of the multi-camera-based target localization system, as formulated in (32) can be used to derive methods for optimizing camera placement for enhancing localization accuracy.

B. Analysis of camera selection methods with different intrinsic parameters

In this subsection, we focus on the relationship between the accuracy of the camera-based bearing measurement localization system and the cameras' intrinsic parameters. Without loss of generality, we deploy five different MRS using bearing measurement for target localization; each MRS is equipped with four cameras with the same parameters whose geometric configuration conforms to the UAA. However, cameras in different MRS have different intrinsic parameters, which are shown in Table I. Through simulation experiments with different camera intrinsic parameters, a comparative analysis can be conducted to determine the optimal camera parameters that achieve the highest localization accuracy. Then, we conduct the simulation under the circumstances of $k = 10^3$, $\omega' \backsim \mathcal{N}(0, 1)$ and $\sigma_g = 0.01$.

Specifically, in the 2-D x–y mission space, the target is located at $[0, 0]$, and the position of four robots are $[3, 3]$, $[3, -3]$, $[-3, 3]$, and $[-3, -3]$. Using the parameter optimization metric introduced in (32), we compute the $\det(\text{FIM})$ of the multi-camera-based target localization system in different sets of MRS. This enables us to select optimal cameras' intrinsic parameters from the five candidate sets, the one that yields the highest target localization accuracy.

In addition, we implement an Extended Kalman Filter (EKF) based on (3) to emulate a realistic localization process, where camera-based bearing measurements are used to estimate the target position. We then compute the MSE of the localization estimates to empirically validate the proposed accurate model. Each MRS performs m experiments using camera-based measurements, with the i-th experiment yielding an estimated target position $\hat{\boldsymbol{T}}_i$. The MSE can be represented as

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(\hat{\boldsymbol{T}}_i - \boldsymbol{T})^2 \tag{65}$$

where $\boldsymbol{T}$ is the true position of the target.

By calculating $\det(\text{FIM})$ and MSE of the target localization systems using camera-based bearing measurement, the result of which is shown in TABLE I, where m is 100. We can easily find that the 8-th set of MRS equipped with the cameras whose intrinsic parameter is that the Resolution is 1980×1080 and the focal length is $12mm$. This simulation shows that the estimation metric in (32) gives us a convenient method to select the most suitable cameras in target localization.

In conclusion, from the analysis in the previous analysis, we can ascertain that as the camera resolution and focal length increase, the mean squared error (MSE) of target position estimation in a multi-camera localization system decreases, indicating higher localization accuracy. Therefore, when selecting camera intrinsic parameters, priority should be given to those that are optimal in both resolution and focal length. However, if the intrinsic parameters are randomly paired, such as having the maximum focal length but suboptimal resolution, it is advisable to utilize the accuracy model proposed in (32) to prioritize the selection of cameras that maximize localization accuracy.

TABLE I: Intrinsic parameters for different kinds of cameras

Camera class	1	2	3	4	5	6	7	8	9	10
Resolution	800×960	800×960	800×960	640×480	640×480	640×480	1920×1080	1920×1080	1280×1024	1280×1024
Focal length	9mm	10mm	11mm	9mm	10mm	11mm	8mm	12mm	8mm	12mm
det(FIM)	0.525	1.166	2.383	0.035	0.079	0.167	2.421	44.552	0.883	17.417
MSE($\times 10^{-3}$)	3.21	3.09	2.71	5.21	4.62	3.78	2.27	1.06	3.13	1.84

VII. CONCLUSIONS

In this paper, the problem of selecting appropriate cameras (intrinsics) and establishing an optimal geometric configuration (extrinsics) of a multi-camera-based target localization system to improve the positioning accuracy is addressed. We model the positioning system using a Gaussian model by analyzing the error from camera relative bearing measurements. In this way, we transfer the noise generated by the mobile robot and the error within the camera to the error in computing the relative bearing measurements. Subsequently, the FIM and the D-optimization criteria are employed to meticulously analyze how to position multiple cameras both within and outside the camera area to attain the maximum positioning accuracy and derive the relationship between the positioning accuracy and the intrinsics of the camera (including the focal length, resolution).

In our current work, there are many areas that need to be improved. Firstly, in this work, we consider that the position of the target is fixed. However, utilizing imprecise prior target position information has significant practical value and broad research potential. Besides, we note that the value of bearing-measurement θ in our manuscript is not unbiased; we will find a more suitable, accurate model rather than CRLB to evaluate the localization performance. Lastly, we hope to extend our work to three-dimensional space and obtain a more practical conclusion of the deployment and selection method of the camera.

REFERENCES

[1] P. Arm, G. Waibel, J. Preisig, T. Tuna, R. Zhou, V. Bickel, G. Ligeza, T. Miki, F. Kehl, H. Kolvenbach, *et al.*, "Scientific exploration of challenging planetary analog environments with a team of legged robots," *Science robotics*, vol. 8, no. 80, p. eade9548, 2023.

[2] Z. Fang and A. V. Savkin, "Strategies for optimized uav surveillance in various tasks and scenarios: A review," *Drones*, vol. 8, no. 5, p. 193, 2024.

[3] Y. Liu and Y. Bao, "Real-time remote measurement of distance using ultra-wideband (uwb) sensors," *Automation in Construction*, vol. 150, p. 104849, 2023.

[4] X. Zhang, K. Kwon, J. Henriksson, J. Luo, and M. C. Wu, "A large-scale microelectromechanical-systems-based silicon photonics lidar," *Nature*, vol. 603, no. 7900, pp. 253–258, 2022.

[5] J. M. Rothberg, T. S. Ralston, A. G. Rothberg, J. Martin, J. S. Zahorian, S. A. Alie, N. J. Sanchez, K. Chen, C. Chen, K. Thiele, *et al.*, "Ultrasound-on-chip platform for medical imaging, analysis, and collective intelligence," *Proceedings of the National Academy of Sciences*, vol. 118, no. 27, p. e2019339118, 2021.

[6] M. Sadeghi, F. Behnia, and R. Amiri, "Optimal geometry analysis for tdoa-based localization under communication constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3096–3106, 2021.

[7] M. Sadeghi, F. Behnia, and R. Amiri, "Optimal sensor placement for 2-d range-only target localization in constrained sensor geometry," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2316–2327, 2020.

[8] S. Xu, "Optimal sensor placement for target localization using hybrid rss, aoa and toa measurements," *IEEE Communications Letters*, vol. 24, no. 9, pp. 1966–1970, 2020.

[9] S. Xu, L. Wu, K. Doğançay, and M. Alaee-Kerahroodi, "A hybrid approach to optimal toa-sensor placement with fixed shared sensors for simultaneous multi-target localization," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1197–1212, 2022.

[10] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—part i: A general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, 2010.

[11] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband localization—part ii: Cooperative networks," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4981–5000, 2010.

[12] K. Romer and F. Mattern, "The design space of wireless sensor networks," *IEEE wireless communications*, vol. 11, no. 6, pp. 54–61, 2004.

[13] M. F. Ferreira, E. Castro-Camus, D. J. Ottaway, J. M. López-Higuera, X. Feng, W. Jin, Y. Jeong, N. Picqué, L. Tong, B. M. Reinhard, *et al.*, "Roadmap on optical sensors," *Journal of Optics*, vol. 19, no. 8, p. 083001, 2017.

[14] J. Shen, A. F. Molisch, and J. Salmi, "Accurate passive location estimation using toa measurements," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2182–2192, 2012.

[15] S. He and X. Dong, "High-accuracy localization platform using asynchronous time difference of arrival technology," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1728–1742, 2017.

[16] N. Rogel, D. Raphaeli, and O. Bialer, "Time of arrival and angle of arrival estimation algorithm in dense multipath," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5907–5919, 2021.

[17] S. Pant, S. Umesh, and S. Asokan, "Knee angle measurement device using fiber bragg grating sensor," *IEEE Sensors Journal*, vol. 18, no. 24, pp. 10034–10040, 2018.

[18] H. Chen, Q. Shi, R. Tan, H. V. Poor, and K. Sezaki, "Mobile element assisted cooperative localization for wireless sensor networks with obstacles," *IEEE transactions on wireless communications*, vol. 9, no. 3, pp. 956–963, 2010.

[19] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, "Quantum fisher information matrix and multiparameter estimation," *Journal of Physics A: Mathematical and Theoretical*, vol. 53, no. 2, p. 023001, 2020.

[20] F. Nielsen, "Cramér-rao lower bound and information geometry," *Connected at Infinity II: A Selection of Mathematics by Indians*, pp. 18–37, 2013.

[21] B. Yang, J. Li, and H. Zhang, "Resilient indoor localization system based on uwb and visual–inertial sensors for complex environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021.

[22] M. A. Jamshed, K. Ali, Q. H. Abbasi, M. A. Imran, and M. Ur-Rehman, "Challenges, applications, and future of wireless sensors in internet of things: A review," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 5482–5494, 2022.

[23] G. K. Ijemaru, K. L.-M. Ang, and J. K. Seng, "Wireless power transfer and energy harvesting in distributed sensor networks: Survey, opportunities, and challenges," *International journal of distributed sensor networks*, vol. 18, no. 3, p. 15501477211067740, 2022.

[24] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics letters*, vol. 33, no. 2, pp. 156–158, 2008.

[25] M. J. Mlodzianoski, M. F. Juette, G. L. Beane, and J. Bewersdorf, "Experimental characterization of 3d localization techniques for particle-tracking and super-resolution microscopy," *Optics express*, vol. 17, no. 10, pp. 8264–8277, 2009.

[26] D. Roy, "Discrete rayleigh distribution," *IEEE transactions on reliability*, vol. 53, no. 2, pp. 255–260, 2004.

[27] L. E. Baum and M. Katz, "Convergence rates in the law of large numbers," *Transactions of the American Mathematical Society*, vol. 120, no. 1, pp. 108–123, 1965.
[28] M. Peligrad and S. Utev, "Central limit theorem for linear processes," *The Annals of Probability*, vol. 25, no. 1, pp. 443–456, 1997.
[29] K. Lin, L. Zhang, K. Wu, and Z. Liu, "Optimal deployment of multi-robot based regional positioning system using relative bearing measurement," *IEEE Robotics and Automation Letters*, vol. 10, no. 10, pp. 10602–10609, 2025.
[30] M. A. M. De Oca, T. Stutzle, M. Birattari, and M. Dorigo, "Frankenstein's pso: a composite particle swarm optimization algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 1120–1132, 2009.
[31] J. Vasconcelos, J. A. Ramirez, R. Takahashi, and R. Saldanha, "Improvements in genetic algorithms," *IEEE Transactions on magnetics*, vol. 37, no. 5, pp. 3414–3417, 2002.

Kangwen Lin is graduated from the School of Mechanical and Electrical Engineering, Anhui JianZhu University in 2023 with a Bachelor of Engineering degree in Automation. He is currently pursuing his master's degree in control engineering at the School of Electrical Engineering and Automation, Anhui University. His research interests center around cooperative localization and coverage control.

Liang Zhang is currently a lecturer at the School of Engineering and Automation, Anhui University, China. His research interests are mainly around the multi-robot or multi-agent system (MAS/MRS), including distributed estimation, cooperative control, planning under uncertainty. He received his B.Eng. degree in Automation from Shandong University, China in 2015 and Ph.D. degree in the School of Astronautics, Harbin Institute of Technology, China in 2021. He was also a visiting Ph.D. student in the Autonomous System Lab, ETH Zurich, Switzerland from 2018 to 2019 under the foundation from China Scholarship Council.

Zhaochen Liu is graduated from the School of Electronic Information, Zhongyuan Institute of Technology in 2023 with a Bachelor of Engineering degree in Aircraft control and information engineering. He is currently pursuing his master's degree in control engineering at the School of Electrical Engineering and Automation, Anhui University. His research interests center around Cluster dynamics bionics.

Yefei Huang received his B.S. and M.S. degrees in aerospace engineering from Harbin Institute of Technology, Harbin, China, in 2016 and 2019, respectively. He is currently working toward the Ph.D in the Deep Space Exploration Research Center at Harbin Institute of Technology. His research interest is the application of vision-based methods in astronautics.

Zexu Zhang Zexu Zhang is a full professor at the Harbin Institute of Technology, China, where he is also the director of the Institute of Aircraft Dynamics and Control. His research interests include aircraft autonomous navigation and control, intelligent cooperative perception and autonomous decision-making in drone swarm, data visualization. He was elected director of the Committee of Space Intelligence of the Chinese Society of Space Research(2021-2025).