

Bearing-Only Active Sensing Under Merged Measurements

Luke Calkins^{1,2}, Phil Baldoni¹, James McMahon¹, Corbin Wilhelmi¹, and Michael M. Zavlanos²

Abstract—In this paper we propose an algorithm to actively track multiple moving targets using a bearing-only sensor in the presence of merged measurements. Merged measurements arise from sensor resolution constraints and therefore targets that are close in relative bearing to the sensor get reported as a single group measurement. We employ a merged measurement model in a nonlinear joint probabilistic data association filter for tracking multiple targets through merging events. We also propose an online adaptive planning algorithm that maneuvers the sensor in order to increase tracking performance. We introduce a novel method based on forward value iteration that incorporates the merged measurement information into the planning strategy. The resulting trajectory is biased away from situations where merged measurements occur, as this leads to more uncertainty in the target state estimates. We demonstrate our algorithm both in simulation as well as onboard real unmanned ground vehicles. This is the first time bearing-only tracking with merged measurements has been accomplished with a mobile sensor in practice. Furthermore, to the best of our knowledge, this is the first time merged measurement data association information has been utilized to effectively plan for such situations.

Index Terms—Reactive and Sensor-Based Planning, Visual Tracking

I. INTRODUCTION

IN order to remain discrete and avoid detection, it is advantageous to employ passively collected data for detection and tracking. For example, systems have used camera footage to track pedestrians [1], [2] and acoustics to detect and track underwater phenomena [3], [4]. In contrast to active tracking, where the sensing system emits some signal into the environment (i.e. LIDAR, RADAR, SONAR), a passive tracking system only collects information generated by the target itself. Multiple challenges arise in the passive tracking problem: (i) overcoming environmental noise, (ii) bearing only measurements (a lack of relative range information), (iii) and the merging of measurements from multiple targets. Merged measurements arise in scenarios where one target occludes others, or when targets are closely spaced relative to the



Fig. 1: Test area for the UGV experiments in the Naval Research Laboratory’s “Laboratory for Autonomous Systems Research” (LASR) High Bay.

sensor resolution. When targets are merged, the sensor reports the group as a single detected target as opposed to multiple distinct detections. In the case where measurements merge, a single measurement must be associated with multiple targets to correctly interpret the passive data. When data is not associated correctly this often leads to degraded tracker performance in the form of track coalescence (where two or more tracks merge onto a single target) and the premature deletion of tracks [5]–[7].

In this work, we focus on the difficulties associated with merged measurements for a passive, bearing-only sensor. In this particular scenario, bearing-only measurements are collected on-board a mobile ground vehicle using two monocular cameras, one facing port and the other starboard. Using the real time RGB-D camera feed, data is passed through an off the shelf, pre-trained neural network [8] to detect targets (other ground vehicles) and their relative bearing. A sensor resolution model modeling the probability that two targets will merge is incorporated into a joint probabilistic data association filter (JPDAF) [9] which performs the data association and tracking. We also introduce a novel planning algorithm that maneuvers the sensing agent in order to increase the tracking performance.

A. Related Work

Autonomous tracking using bearing-only sensors has received the most attention in underwater environments onboard autonomous underwater vehicles (AUVs). In our previous work [4], we provide field results for a passive sonar multi-target tracker based on a particle filter and a simple keep broadside behavior. Other works using sonar optimize an information-metric over some horizon [10], [11].

When tracking multiple targets simultaneously in the presence of clutter, the process consists of two steps. First, measurements are associated to targets, and then single target filtering is applied. Two broad categories encompass this

Manuscript received: October, 15, 2020; Revised January, 12, 2021; Accepted February, 24, 2021.

This paper was recommended for publication by Editor Pauline Pounds upon evaluation of the Associate Editor and Reviewers’ comments.

This work is supported in part by the Office of Naval Research (ONR) through the NRL Base Program and by ONR under agreement #N00014-18-1-2374

¹Luke Calkins, Phil Baldoni, James McMahon, and Corbin Wilhelmi are with the Naval Research Laboratory, Washington, D.C. 20375, USA {luke.calkins, philip.baldoni, james.mcmahon, corbin.wilhelmi}@nrl.navy.mil

²Luke Calkins and Michael M. Zavlanos are with the department of Mechanical Engineering and Material Science, Duke University, Durham, NC 27708, USA {william.calkins, michael.zavlanos}@duke.edu

Digital Object Identifier (DOI): see top of this page.

approach, namely the JPDAF [9] and multiple hypothesis tracking (MHT) [12]. Within the last decade, the probability hypothesis density (PHD) filter [13] has gained a lot of traction as it does not perform data association. However, it does not produce labeled tracks over time but rather an unlabeled distribution of target states at every time step. A review of the state-of-the-art for multi-target tracking is given in [14].

The importance of considering sensor resolution limitations in tracking algorithms is described in [15], however they point out that 99% of the tracking literature does not consider it. The first method for incorporating merged measurements to an arbitrary number of targets in a JPDAF context was given in [5]. Methods also exist to incorporate merged measurements into MHT [6] and PHD filters [7].

In the robotics and controls community, researchers have been developing efficient methods to compute control policies to aid in active information acquisition and tracking [16], [17]. Forward value iteration [18] provides a way to optimize a sensor's trajectory by considering the dynamics of the sensor itself by solving an optimal control problem. The reduced value iteration method [19] provides a more computationally efficient approach to FVI with a sacrifice in optimality. The method proposed in this paper is based off the FVI algorithm, but considers the effect of merged measurements in the planning process. Most similar to the desired objective in this work is the work on data association aware planning in SLAM [20], and active target tracking using a team of robots utilizing a PHD filter [21]. None of the described works in the active sensing literature consider the effects of merged measurements into the tracking or planning problems.

B. Contributions

To the best of our knowledge, this is the first exploration of a merged measurement JPDAF in a bearing-only tracking scenario with an autonomous mobile sensor. We also incorporate the merged measurement information and data association into our planning strategy, which has not been explored before. This has the benefit of driving the sensor to configurations where merging will not occur when possible, resulting in better tracking performance. A further contribution of this work is the demonstration of the proposed algorithm in a real practical tracking application with autonomous ground vehicles equipped with cameras. Transitioning to experiments presents many challenges including noisy sensors, accurate state estimation, and the asynchronous running of different algorithms. We utilize a pre-trained neural network for our target detection thus resulting in many missed detections and false alarms. We are able to demonstrate robustness of our algorithm to these challenges.

The rest of the paper is organized as follows. In Section II, the governing motion and measurement models are defined. The tracking algorithm is presented in Section III and the planning strategy in Section IV. Experimental and numerical results are presented in Section V before concluding remarks are given in Section VI.

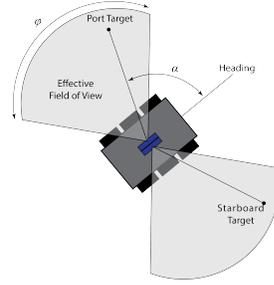


Fig. 2: Sensor field of view. Blind spots directly ahead and behind the vehicle.

II. PRELIMINARIES

Here we present the governing equations utilized in the development of our proposed algorithm.

A. Constant Velocity Motion Model

We represent the state of the target by $\mathbf{x}_k \triangleq [x_k \ y_k \ \dot{x}_k \ \dot{y}_k]^T$, where (x_k, y_k) is the target's position and (\dot{x}_k, \dot{y}_k) its velocity. We assume targets of interest are moving with approximately constant velocity according to the following second-order kinematic model with acceleration disturbances

$$\begin{aligned} \mathbf{x}_k &= f_k(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ &= \underbrace{\begin{bmatrix} 1 & 0 & T_k & 0 \\ 0 & 1 & 0 & T_k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{F}_k} \underbrace{\begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \dot{x}_{k-1} \\ \dot{y}_{k-1} \end{bmatrix}}_{\mathbf{x}_{k-1}} + \underbrace{\begin{bmatrix} (w_x)_k \\ (w_y)_k \\ (w_{\dot{x}})_k \\ (w_{\dot{y}})_k \end{bmatrix}}_{\mathbf{w}_k}, \end{aligned} \quad (1)$$

where \mathbf{F}_k is the state transition matrix, and $T_k \triangleq t_k - t_{k-1}$ is the sampling interval. The zero-mean Gaussian process noise vector, $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ has covariance matrix \mathbf{Q}_k given by

$$\mathbf{Q} = \sigma_a \begin{bmatrix} \frac{T_k^4}{4} & 0 & \frac{T_k^3}{2} & 0 \\ 0 & \frac{T_k^4}{4} & 0 & \frac{T_k^3}{2} \\ \frac{T_k^3}{2} & 0 & T_k^2 & 0 \\ 0 & \frac{T_k^3}{2} & 0 & T_k^2 \end{bmatrix} \quad (2)$$

where σ_a^2 is the variance of the zero-mean acceleration disturbance assumed equal in both coordinate directions.

B. Passive Bearing-Only Sensor Model

The observing vehicle, also referred to as the ownship, has state $\mathbf{p}_k = [x_k^0, y_k^0, \psi_k]^T$ where (x_k^0, y_k^0) is its position and ψ_k is its heading measured counterclockwise from the positive x axis. We employ a limited field-of-view (FOV) bearing only camera sensor. The bearing of the target relative to the ownship's position is

$$\beta_k(\mathbf{x}_k, \mathbf{p}_k) = \text{atan}(\Delta y_k / \Delta x_k) \quad (3)$$

where $\Delta x_k(\mathbf{x}_k, \mathbf{p}_k) = x_k - x_k^0$ and $\Delta y_k(\mathbf{x}_k, \mathbf{p}_k) = y_k - y_k^0$ are relative distances. The bearing measured in the reference frame of the sensor is $\alpha_k = \beta_k - \psi_k$. Thus, α_k is measured counterclockwise from the ownship's heading. The measure-

ment reported by the sensor is

$$z_k = h_k(\mathbf{x}_k, v_k; \mathbf{P}_k) = \begin{cases} \alpha_k(\mathbf{x}_k, \mathbf{P}_k) + v_k & \alpha_k \in \text{FOV} \\ \emptyset & \text{otherwise} \end{cases} \quad (4)$$

where $v_k \sim \mathcal{N}(0, \sigma_b^2)$ is zero-mean Gaussian noise with variance σ_b^2 and \emptyset is the empty set. The effective FOV φ is centered around the 90 and 270 degree α angles on both the port and starboard sides respectively. See Figure 2 for reference. Therefore, the sensor reports a bearing measurement if the relative bearing is within its field of view, otherwise, no measurement is reported. In terms of Bayesian filtering, (4) is referred to as the likelihood function, $\mathcal{L}(z_k|\mathbf{x}_k)$. The likelihood function gives the relative probability of observing a bearing measurement for all target states $\mathbf{x} \in \mathbb{R}^4$. However, it is not necessarily a probability distribution and it need not integrate to unity across the target state space. Bearing measurements resulting from targets outside the field of view have zero likelihood.

With the motion model (1) and likelihood function (4), the standard extended Kalman Filter (EKF) [22] can be applied to recursively track the time varying target state \mathbf{x}_k . Note that when tracking a target, if the estimated position is outside the sensor's FOV, only the prediction step in the EKF will be performed and not a measurement update. An EKF is needed due to the nonlinear measurement model (4). Therefore, linearization of (4) around the predicted target state $\mathbf{x}_{k|k-1}$ is needed at each time step to perform the measurement update. This process works well for single target tracking in the absence of clutter or merged measurements.

III. MERGED MEASUREMENT MODELING AND TRACKING

In this paper, we seek to track multiple targets in the presence of merged measurements and random clutter. When tracking multiple targets, there is a measurement origin uncertainty problem, as the sensor does not know which measurement came from which target. We employ a merged-measurement variant of the joint probabilistic data association filter (JPDAF). The JPDAF is a "soft association" technique where every measurement within the validation gate of each target it assigned to the target with an appropriate probability [9].

A. JPDAF with Merged Measurements

We follow the methodology in [5] for modeling the merging of measurements and subsequent incorporation into the estimator. We summarize this method here as it will be utilized in the development of the planning strategy in Section IV. We assume that the number of targets is arbitrary but known. A *group* is defined as a set of targets that gives rise to a single, joint measurement due to the resolution characteristics of the sensor. The merging events are modeled as graphs, where nodes represent targets and an edge between nodes represents the event that two targets are pairwise unresolved. Therefore, a group of targets is considered unresolved if there exists a path in the graph connecting the targets. A three target graph is



Fig. 3: Representation of merging events. Target one is resolved while targets two and three are unresolved, producing a merged measurement

shown in Figure 3 in which two targets are unresolved and thus produce a group measurement, and one target is resolved. A probabilistic representation of the event that targets are merged is needed, conditioned on the target states and the ownship state. This probability depends on the sensor characteristics as well as the type of processing used, but is given here as a simple intuitive model that is mathematically tractable. Given a merging graph \mathcal{V} , the probability of the merging graph given the target state, \mathbf{x}_k , where $\mathbf{x}_k = [\mathbf{x}_k^{(1)}, \mathbf{x}_k^{(2)}, \dots, \mathbf{x}_k^{(N)}]$ is the combined state for N total targets is

$$p(\mathcal{V}|\mathbf{x}_k) = \prod_{\mathcal{S}_e} P_u(\mathcal{S}_e) \prod_{\mathcal{S}_0} (1 - P_u(\mathcal{S}_0)), \quad (5)$$

where \mathcal{S}_e is the set of unresolved pairs in the graph \mathcal{V} and \mathcal{S}_0 is the set of pairwise targets that are resolved, and therefore do not have an edge connecting them. P_u is the probability that two targets are pairwise unresolved and is dependent on the joint state \mathbf{x}_k . Consider the set \mathcal{S} of all pairwise targets, such that $\mathcal{S} = \mathcal{S}_e \cup \mathcal{S}_0$. For a certain pair of targets $\mathbf{x}_k^{(i)}, \mathbf{x}_k^{(j)}$, the probability that they are unresolved is given by

$$P_u(\mathbf{x}_k^{(i)}, \mathbf{x}_k^{(j)}) = e^{-\frac{1}{2}(\Delta_\beta^{i,j}/\alpha_\beta)^2} \quad (6)$$

where $\Delta_\beta^{i,j}$ is the difference in bearing between targets i and j , and α_β describes the resolution capabilities of the sensor. In (6), two targets that are at the same relative bearing will merge with probability one, with decreasing probability as the difference in bearing grows, with decay rate controlled by parameter α_β . The merging probability can also be written as a scaled multi-variate Gaussian as

$$P_u(\mathbf{x}_k^{(i)}, \mathbf{x}_k^{(j)}) = |2\pi\mathbf{R}_u|^{1/2} \mathcal{N}(0; \Delta_\beta^{i,j}, \mathbf{R}_u), \quad (7)$$

where $\mathbf{R}_u = \alpha_\beta^2$. Rewriting (6) as a normal distribution simplifies the calculation of the posterior density described in Section III-C.

B. Group Measurement Model

The set of measurements at time step k , denoted \mathbf{Z}_k is described by

$$\mathbf{Z}_k = \Pi_p \begin{bmatrix} \mathbf{Z}_k^c \\ \mathbf{Z}_k^t \end{bmatrix} \quad (8)$$

consisting of a set of clutter generated measurements \mathbf{Z}_k^c and target generated measurements \mathbf{Z}_k^t . The matrix Π_p is a M_k -dimensional random permutation matrix that describes the data association uncertainty as the sensor cannot discriminate between clutter and target generated measurements and the order in which they appear in \mathbf{Z}_k is random.

We seek a likelihood function $p(\mathbf{Z}_k|\mathcal{V}, \mathbf{d}, \mathbf{x}_k)$, describing the measurements conditioned on a merging event \mathcal{V} , data association vector \mathbf{d} , and joint target states \mathbf{x} . The data association vector \mathbf{d} of length M_k has entries $d(j) = i$ if measurement

j is assigned to target i , or $d(j) = 0$ if measurement j is assigned to clutter. The group measurement model used in this paper, like [5], assumes that a group measurement comes from the group center, i.e., the average bearing of the target group. However, we do not assume a linear measurement model in this work as in [5], but rather develop an extended Kalman filter (EKF)-like approach based on the linearization of the nonlinear bearing-only measurement model. For a group of n_g targets (possibly of size one), whose states are stacked in the group state \mathbf{x}_k^g , their group measurement, $z_k^{t,(j)}$ is given by

$$\begin{aligned} z_k^{t,(j)} &= h_k^{n_g}(\mathbf{x}_k^g) + u_k^{g,n_g} \\ &= \frac{1}{n_g} (h_k(\mathbf{x}_k^{g,1}) + \dots + h_k(\mathbf{x}_k^{g,n_g})) + u_k^{g,n_g} \end{aligned} \quad (9)$$

where $u_k^{g,n_g} \sim \mathcal{N}(0, \mathbf{R}_k^{n_g})$ is the group measurement noise which depends on the number of targets in the group. As assumed in most merged measurement tracking literature, the group measurement noise is greater for larger target groups.

For the clutter measurements, \mathbf{Z}_k^c , it is common to assume that these measurements are independent and distributed uniformly throughout the sensor's field of view (FOV) with the number of measurements having a Poisson distribution with rate parameter λ . If the data association \mathbf{d} is known, then the clutter measurements at time k have the probability density function

$$p(\mathbf{Z}_k^c | \mathbf{d}, \mathcal{V}) = \frac{1}{|\text{FOV}|^{M_{k,c}}} \quad (10)$$

and the target measurements, in which measurements from different groups are assumed independent, have the probability density function

$$p(\mathbf{Z}_k^t | \mathcal{V}, \mathbf{d}, \mathbf{x}_k) = \prod_{i=1}^{M_{k,t}} p(z_k^{t,(i)} | \mathbf{d}, \mathcal{V}, \mathbf{x}_k) \quad (11)$$

where $p(z_k^t | \mathbf{d}, \mathcal{V}, \mathbf{x})$ is the group measurement model (9). The total number of measurements at time step k is given by $M_k = M_{k,t} + M_{k,c}$, where $M_{k,t}$ and $M_{k,c}$ are the number of target measurements and clutter measurements respectively. In reality, since data association is unknown, the measurement update for each feasible association is performed resulting in a Gaussian mixture with appropriate association weight.

C. Posterior Density

Now, the posterior density can be calculated when the resolution events and data association are unknown. The unknown merging events are considered by marginalizing over all possible merging graphs \mathcal{V} .

$$p(\mathbf{x}_k | \mathbf{Z}^k) = \sum_{\mathcal{V}} P(\mathbf{x}_k, \mathcal{V} | \mathbf{Z}^k) \quad (12)$$

Since the data association \mathbf{d} is also unknown, a marginalization over all possible data associations is also performed resulting in the full posterior density

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{Z}^k) &\propto \sum_{\mathcal{V}} p(\mathcal{V} | \mathbf{x}_k) \sum_{\mathbf{d} \in \mathcal{D}(\mathcal{V})} p(\mathbf{d} | \mathcal{V}, \mathbf{x}_k) \\ &\quad \times p(\mathbf{Z}_k | \mathcal{V}, \mathbf{d}, \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{x}_{k-1}). \end{aligned} \quad (13)$$

In (13), $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ is the predicted target density, equivalent to the prediction step for an EKF [22]. The density $p(\mathbf{Z}_k | \mathcal{V}, \mathbf{d}, \mathbf{x}_k) = p(\mathbf{Z}_k^t | \mathcal{V}, \mathbf{d}, \mathbf{x}_k) \times p(\mathbf{Z}_k^c | \mathbf{d}, \mathcal{V})$. In order to calculate (13), an expression for the data association probability $p(\mathbf{d} | \mathcal{V}, \mathbf{x}_k)$ is needed (alg. 1:6). This is given as [5]

$$\begin{aligned} p(\mathbf{d} | \mathcal{V}, \mathbf{x}_k) &= P_c(M_{k,c}) \frac{(M_k - M_t)!}{M_k!} \\ &\quad \times \prod_{j:d_j=0} (1 - P_D^j) \prod_{j:d_j=1} P_D^j, \end{aligned} \quad (14)$$

where P_D^j is the probability of detection for measurement index j , $\{j : d(j) = 0\}$ is the set of clutter detections, and $\{j : d(j) > 0\}$ is the set of target generated detections. $P_c(M_{k,c})$ is the probability of receiving $M_{k,c}$ clutter measurements which is given by the Poisson pmf with parameter λ . Intuitively, for a given merging graph \mathcal{V} , the target updates are calculated using an EKF for each data association hypothesis, resulting in a Gaussian posterior belief. This result is weighted by $p(\mathbf{d} | \mathbf{x}_k, \mathcal{V}) \times p(\mathbf{Z}_k^t | \mathbf{x}_k, \mathcal{V}, \mathbf{d}) \times p(\mathbf{Z}_k^c | \mathbf{d}, \mathcal{V})$ (alg. 1:7). The result is a Gaussian mixture with more likely data associations carrying more weight. The JPDAF approach is to approximate this Gaussian mixture with a single Gaussian using first and second order moment matching which therefore outputs a single Gaussian belief for each target (alg. 1:11).

With merged measurements, the final step is the marginalization over all possible merging graphs \mathcal{V} . This involves multiplying the Gaussian mixture over all data associations $\mathbf{d} \in \mathcal{D}(\mathcal{V})$ by the probability of the graph $p(\mathcal{V} | \mathbf{x}_k)$ (alg. 1:8). By defining $\pi^{(i,j)}(k) = \delta_{k,i} - \delta_{k,j}$ $k = 1, \dots, n_g$, the probability of the merging graph $p(\mathcal{V} | \mathbf{x}_k)$ can be redefined to be in terms of the measurement model (4) as opposed to $\Delta_{\beta}^{i,j}$. Since $\Delta_{\beta}^{i,j} = \pi^{(i,j)} [h(\mathbf{x}^{(i)}), h(\mathbf{x}^{(j)})]^T$, we can write the multiplication of $P_u(\mathbf{x}_k^{(i)}, \mathbf{x}_k^{(j)})$ with a measurement-updated Gaussian mixture component $\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \Sigma_{k|k})$ as an EKF-like measurement update

$$\begin{aligned} P_u(\mathbf{x}_k^{(i)}, \mathbf{x}_k^{(j)}) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_k, \Sigma_k) &= |2\pi \mathbf{R}_{u, N_{res}}|^{1/2} \times \\ &\mathcal{N}(\mathbf{0}; \pi^{(i,j)} [h(\mathbf{x}^{(i)}), h(\mathbf{x}^{(j)})]^T, \mathbf{R}_{u, N_{res}}) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_k, \Sigma_k). \end{aligned} \quad (15)$$

This multiplication can be rewritten as

$$\begin{aligned} \mathcal{N}(\mathbf{0}; \pi^{(i,j)} [h(\mathbf{x}^{(i)}), h(\mathbf{x}^{(j)})]^T, \mathbf{R}_{u, N_{res}}) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_k, \Sigma_{k|k}) &= \\ \mathcal{N}(\mathbf{0}; \pi^{(i,j)} [h(\mathbf{x}^{(i)}), h(\mathbf{x}^{(j)})]^T, \mathbf{S}^{i,j}) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_k^{i,j}, \Sigma_{k|k}^{i,j}), \end{aligned} \quad (16)$$

where $\mathbf{S}^{i,j}$, $\hat{\mathbf{x}}_{k|k}^{i,j}$, and $\Sigma_{k|k}^{i,j}$ come from the standard EKF equations. This multiplication is the equivalent of an EKF measurement update. The resolution update can be interpreted as a fictitious measurement of zero of the separation of the target bearings. Since $p(\mathcal{V} | \mathbf{x}_k)$ is the product of P_u terms, each multiplication with a P_u term is a measurement-like update as given above in (16). There are also $(1 - P_u)$ terms in (5), and therefore a measurement updated density is split into two components for each $(1 - P_u)$ term defined by the graph \mathcal{V} . The set of components generated by a resolution update for a graph \mathcal{V} and data association hypothesis \mathbf{d} are collected in the set $\mathcal{U}(\mathcal{V})$. The merged measurement JPDAF algorithm is

Algorithm 1 Merged Measurement JDPDF Tracking

Require: $\hat{\mathbf{x}}_k = [\hat{\mathbf{x}}_k^{(1)}, \dots, \hat{\mathbf{x}}_k^{(N)}]$, $\Sigma_k = \text{diag}\{\Sigma_k^{(1)}, \dots, \Sigma_k^{(N)}\}$
Require: $\mathbf{Z}_k = \{\mathbf{z}_k^{(1)}, \dots, \mathbf{z}_k^{(M_k)}\}$

- 1: Perform prediction step $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}$, $\Sigma_{k|k-1} = \mathbf{F}\Sigma_{k-1|k-1}\mathbf{F}^T + \mathbf{Q}$
- 2: Generate all feasible graphs \mathcal{V}
- 3: **for** each graph **do**
- 4: formulate all data association hypotheses $\mathcal{D}(\mathcal{V})$
- 5: **for** each data association hypotheses $\mathbf{d} \in \mathcal{D}(\mathcal{V})$ **do**
- 6: Calculate prior probability $\Pr\{\mathbf{d}|\mathcal{V}, \mathbf{x}_k\}$
- 7: Perform measurement update, which yields a scaled Gaussian $c^{\mathcal{V}, \mathbf{d}} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{\mathcal{V}, \mathbf{d}}, \Sigma_{k|k}^{\mathcal{V}, \mathbf{d}})$
- 8: Update with resolution model. The result is a sum of scaled Gaussians $\sum_{u \in \mathcal{U}(\mathcal{V})} c^{\mathcal{V}, u, \mathbf{d}} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{\mathcal{V}, u, \mathbf{d}}, \Sigma_{k|k}^{\mathcal{V}, u, \mathbf{d}})$
- 9: **end for**
- 10: **end for**
- 11: Approximate Gaussian mixture with single Gaussian using moment matching

$$\hat{\mathbf{x}}_{k|k} = \sum_{\mathcal{V}} \sum_{u \in \mathcal{U}(\mathcal{V})} \sum_{\mathbf{d} \in \mathcal{D}(\mathcal{V})} c^{\mathcal{V}, u, \mathbf{d}} \hat{\mathbf{x}}_{k|k}^{\mathcal{V}, u, \mathbf{d}}$$

$$\Sigma_{k|k} = \sum_{\mathcal{V}} \sum_{u \in \mathcal{U}(\mathcal{V})} \sum_{\mathbf{d} \in \mathcal{D}(\mathcal{V})} c^{\mathcal{V}, u, \mathbf{d}} \{\Sigma_{k|k}^{\mathcal{V}, u, \mathbf{d}} + (\hat{\mathbf{x}}_{k|k}^{\mathcal{V}, u, \mathbf{d}} - \hat{\mathbf{x}}_{k|k})(\hat{\mathbf{x}}_{k|k}^{\mathcal{V}, u, \mathbf{d}} - \hat{\mathbf{x}}_{k|k})^T\}$$

summarized in Algorithm 1. A complexity analysis for this type of tracking algorithm is out of the scope of this paper, refer to [23] for details.

IV. DATA ASSOCIATION AWARE PLANNING

Here we introduce a planning approach in order to maneuver the ownship vehicle while simultaneously tracking the targets in order to achieve optimal tracking performance. Our method is similar to the approach presented in [19] using forward value iteration (FVI). FVI progressively grows a search tree where nodes represent reachable states $(\mathbf{p}_t, \hat{\mathbf{x}}_t, \Sigma_t)$ from the current initial state $(\mathbf{p}_0, \hat{\mathbf{x}}_0, \Sigma_0)$. For each node \mathbf{p}_t is the ownship state, $\hat{\mathbf{x}}_t$ is the stacked target state estimates, and $\Sigma_t = \text{diag}\{\Sigma_{t,1}, \dots, \Sigma_{t,N}\}$ is the covariance of all N targets.

Assume the ownship vehicle is governed by the following dynamical equation

$$\mathbf{p}_{t+1} = f(\mathbf{p}_t, u_t) \quad (17)$$

where $u_t \in \mathcal{U}$ is the control from a set of finite controls \mathcal{U} . It is assumed that the state of the ownship vehicle is known. Given an initial pose and target distribution, we seek an optimal control sequence u_1, \dots, u_T over horizon T that minimizes the mutual information of the final target state \mathbf{x}_T and the measurement set $\mathbf{z}_{1:T}$. This is a stochastic optimal control problem for which closed-loop policies are preferable. However, under the assumption of linear dynamics and a linear measurement model with respect to the target state, this becomes a deterministic optimal control problem [18]

$$\begin{aligned} \min_{u \in \mathcal{U}^T} \quad & \log \det(\Sigma_T) \\ \text{s.t.} \quad & \mathbf{p}_{t+1} = f(\mathbf{p}_t, u), \quad t = 0, \dots, T-1 \\ & \Sigma_{t+1} = \rho_{\mathbf{p}_{t+1}}(\Sigma_t) \quad t = 0, \dots, T-1, \end{aligned} \quad (18)$$

where $\rho_{\mathbf{p}_{t+1}}(\cdot)$ is the Kalman filter Riccati map. The control problem (18) is solved using FVI by propagating the covariance matrix from each node to all children nodes by enumerating through each action in \mathcal{U} . By expanding the initial node using all possible actions in \mathcal{U} , the set of nodes S_1 reachable at time $t = 1$ are calculated. From any node in the tree, there exists one edge for every action in \mathcal{U} connecting to one time step ahead in the tree. If the motion model and measurement models are linear, then the covariance updates do not depend on the actual target trajectory. However, since we have a nonlinear measurement model (4), linearization is needed and model predictive control can be applied creating an adaptive policy.

In this paper, we propose an alternative method for updating the covariance matrices in (18) that incorporates the information from the merged measurement model. Since the measurement noise is assumed larger for larger target groups, updates from merged measurements will result in larger target covariances, thus creating a larger cost in (18). By performing a covariance update similar to that of the merged measurement JPDF tracker described in Section III, the planning algorithm will drive the vehicle to configurations where targets will not merge in the measurement space, thus leading to better tracking performance. Some merging events may be unavoidable given the sensor's dynamics and the target trajectories, but in these cases, the merged measurement JPDF tracker will be able to hold tracks on the targets through the merging event.

Due to the high computational cost of the merged measurement JPDF, it is infeasible to compute the full covariance update step of the FVI given there are $(N^{T-1})/(N-1)$ total nodes in the search tree where N is the number of actions at each node and T is the planning horizon. Therefore, we seek an algorithm that is computationally similar to the FVI algorithm, while also equipping the algorithm with the merged measurement model information. We do so in the following way. When expanding a node, the estimated bearing of each target is calculated. The pair-wise merging probability (5) is evaluated for every feasible pair of targets. If the merging probability $P_u > \eta$ for some $\eta \in [0, 1]$, then the pair is assumed merged at that node. In simulation and experiment, we use $\eta = 0.5$ resulting in the most likely merging graph \mathcal{V} (alg. 2:8). After all pairwise evaluations of (5) have been calculated, targets are separated into groups according to \mathcal{V} and measurements are simulated for each group (alg. 2:9). With this information, the correct data association is known and the covariance updates can be performed for each target (alg. 2:10). The full FVI with merged measurement algorithm is described in Algorithm 2. As described in [18], the most computationally demanding step of the FVI algorithm is the covariance filtering update (alg. 2:10). By only considering a single merging and association event in the planner, we are still only performing a single covariance update for each target as opposed to the many required by a full merged measurement JPDF update.

We then control the vehicle using a receding horizon control approach described in Figure 4. Starting from the initial ownship state and target belief we invoke the merged measurement FVI planner to generate the next control action

Algorithm 2 Forward Value Iteration with Merged Measurement Model

Require: $\mathbf{p}_0, \hat{\mathbf{x}}_0 = [\hat{\mathbf{x}}_0^{(1)}, \dots, \hat{\mathbf{x}}_0^{(N)}], \Sigma_0 = \text{diag}\{\Sigma_0^{(1)}, \dots, \Sigma_0^{(N)}\}$

- 1: $S_0 \leftarrow \{(\mathbf{p}_0, \hat{\mathbf{x}}_0, \Sigma_0)\}, S_t \leftarrow \emptyset$ for $t = 1, \dots, T$
- 2: Predict target trajectory of length T : $\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_T$
- 3: Linearize the observation model: $H_\tau(\cdot) \leftarrow H(\cdot, \mathbf{x}_{t+\tau}), \tau = 0, \dots, T$
- 4: **for** $t = 1 : T$ **do**
- 5: **for all** $(\mathbf{p}, \hat{\mathbf{x}}, \Sigma) \in S_{t-1}$ **do**
- 6: **for all** $u \in \mathcal{U}$ **do**
- 7: $\mathbf{p}_t \leftarrow f(\mathbf{p}, u)$
- 8: Evaluate $P_u(\mathbf{x}_t^{(i)}, \mathbf{x}_t^{(j)})$ for each target pair $(i, j) \in N$ and merge targets if $P_u > \eta$. This defines a merging graph \mathcal{V}
- 9: Generate predicted measurements $\hat{z}_1, \dots, \hat{z}_M$ according to \mathcal{V}
- 10: Update $\Sigma_t(\Sigma_{t-1}, \mathbf{d}(\mathcal{V}), \hat{z}_1, \dots, \hat{z}_M)$ for each target using the correct data association from the previous line
- 11: $S_t \leftarrow S_t \cup \{(\mathbf{p}_t, \hat{\mathbf{x}}_t, \Sigma_t)\}$
- 12: **end for**
- 13: **end for**
- 14: **end for**
- 15: **return** $\min \{\log \det(\Sigma) | (\mathbf{p}, \hat{\mathbf{x}}, \Sigma) \in S_T\}$

sequence u_t . After the vehicle is moved, the sensors collect data which is then passed to the merged measurement JPDAF. The resulting state estimates are then passed back to the FVI algorithm and the process is repeated while the overall time budget T is not satisfied.

V. RESULTS

In this section we present simulation results of our tracking and planning algorithms as well as experiments with real ground vehicles equipped with cameras. In both simulation and experiment, the vehicle is controlled with a set of motion primitives $\{(v, \omega) | v = v_{\text{tran}}, \omega \in \{0, \pm v_{\text{tran}}/R_{\text{turn}}\} \text{ rad/s}\}$ where v_{tran} is the constant translation speed and ω is the turning rate given turning radius R_{turn} . Therefore, at every planning step, the vehicle has the option to go straight, turn left, or turn right. A differential drive motion model is used for (17) when predicting the trajectory during planning. This same model is used to move the robot in simulation whereas the appropriate motor commands were applied in experiment to get the desired speed and turn rate. It should be noted that any motion model can be used in (17) as long as it is known ahead of time. Simulations were performed in Python. A NVIDIA Jetson AGX Xavier development kit was utilized onboard the ground vehicles. The robotics operation system (ROS) [24] was utilized to interface the Stereo Labs ZED camera sensors, motor controller, and tracking and planning algorithms. The nodes in ROS were programmed in Python and C++.

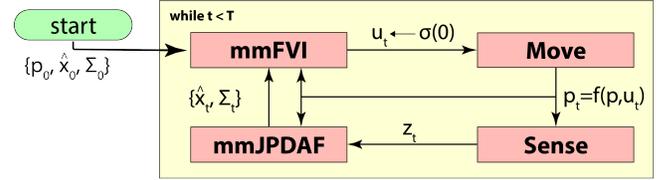


Fig. 4: High level flowchart for the overall control system. mmFVI corresponds to Algorithm 2, mmJPDAF corresponds to Algorithm 1.

A. Simulation Results

Here, we present the results of running the nonlinear EKF bearing-only merged measurement tracking algorithm coupled with our planning algorithm described in Section IV. We ran Monte-Carlo simulations of a scenario involving four targets. As a means for comparison, we compare our results to the approach found in standard FVI or its variants [18], [19], [25]. In these approaches, EKF updates are performed when predicting target beliefs in the planning process as governed by (18). Common in these approaches is the assumption that the data association is known. Therefore in order to compare to our method, a data association strategy is needed to perform the tracking. We implemented a simple nearest neighbors algorithm that selects the closest measurement to the predicted target measurement and performs the standard measurement update with the nearest measurement.

The results of 100 Monte-Carlo simulations of the scenario is shown in Figure 5. The average root mean squared error (RMSE) of the targets with our merged measurement tracker coupled with our merged measurement planner is compared with a nearest neighbors EKF tracker using FVI planning. Representative trajectories of the two planning methods are also shown for a single run in 5a. It can be seen that our algorithm drives the sensor behind the targets initially so that they are well separated in bearing. By comparison, the standard FVI algorithm drives the sensor to the left side of the map initially, which given the target trajectories, causes them to overlap in bearing and therefore creates merged measurements. The nearest neighbor algorithm cannot handle these merging events well and tracks tend to diverge or coalesce leading to the large errors shown in Figure 5b.

B. Experimental Results

In this subsection, we present experimental results showcasing our method in use on a real ground vehicle as it tracks other identical ground vehicles while optimizing its trajectory on-line. We utilized Aion Robotics M6 UGV's equipped with two cameras facing back to back. A Vicon motion capture system was utilized to measure ground truth position and heading of the ownship and target vehicles. A pre-trained YOLOv4 object detection neural network [26] was utilized to generate bearing contacts for each frame of the image stream¹. To ensure real time performance we used the tkDNN library [27] and the YOLOv4 tiny model with the default configuration. For each frame, every detected object is converted into a bearing measurement relative to the sensor by using the bounding boxes and RGB-D information provided by the cameras. Note

¹The NN model used in these experiments can be found at <https://github.com/AlexeyAB/darknet#pre-trained-models>

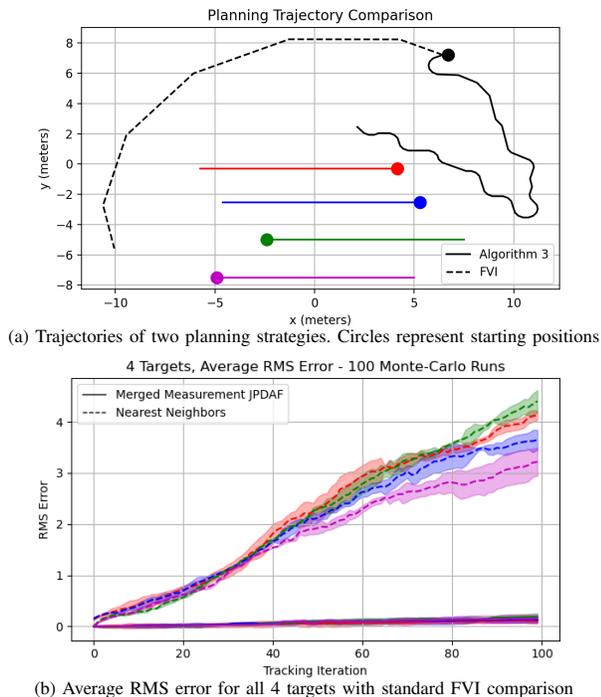


Fig. 5: Four target simulation results.

that with this approach we naturally experienced dropped contacts, noisy contacts from objects other than our targets, and merged contacts due to obstructions. Post processing was done to ensure occurrences of these behaviors were within realistic limits. Specifically, to increase the rate of merged measurements, thresholds were set on the amount of detection overlap and bearing separation of objects. A single merged contact would be returned if two objects shared more than 50 % overlap of bounding boxes or had less than ten degrees of separation. The merged contact is taken as a measurement of the group center as given by the merged measurement model (9).

We ran two scenarios, each tracking two other ground vehicles. In the first scenario one target is stationary while the other target moves straight at constant speed (see fig. 6.a). In the second scenario both targets are mobile, moving forward at constants speeds (see fig. 6.b). The second scenario is designed such that the two tracks cross each other. We also compared our method with that of a nearest neighbor EKF tracker with FVI planning. In ten trial runs of scenario 1, our method successfully held tracks on the correct vehicles nine times. By contrast, the nearest neighbor filter with FVI planning was only successful one out of ten times. In ten trial runs of scenario 2, our method successfully held tracks on the vehicles ten out of ten times (however in four scenarios the tracks switched vehicles). By contrast, the nearest neighbor method never successfully maintained tracks on both vehicles. See video attachment for demonstration of experiment scenarios. The RMSE is plotted for representative trials in Figure 6 utilizing both our method and nearest neighbors with EKF filtering for comparison. Figures (see fig. 6.a and 6.b bottom) shows two representative trajectories from scenario 1 and 2 respectively. In the both examples, the merging of the measurements causes the track for target 2 to coalesce onto the track of target 1 in

the nearest neighbor EKF trial in Figure 6.b which explains the large jump in RMSE for target 2.

A plot of the sensor performance through an example run is shown in Figure 7. From tracking iteration 30 up to around 120, the two targets are very close in relative bearing, so merging occurs. This is counted as a missed detection. However, even when the targets are well separated in bearing, there are many missed detections still present. There is also significant noise in the bearing measurement through the experiment, namely around iteration 30 where the bearing error is nearly 10 degrees.

VI. CONCLUSION

In this paper we proposed a tracking and planning algorithm for a bearing-only sensor in the presence of merged measurements. We developed an EKF-like tracking approach to an existing merged measurement tracker and then incorporated this information into our planning strategy. This allows the sensor to anticipate merging events and plan accordingly. We presented simulation results as well as real experiments with ground vehicles equipped with camera sensors.

ACKNOWLEDGMENTS

We would like to thank Kyle Crandall for help setting up the motor controller on-board the LUGVs. We would also like to acknowledge Loy McGuire and Dan Lofaro with their assistance in setting up the motion capture system.

REFERENCES

- [1] W. Jiang and Z. Yin, "Combining passive visual cameras and active imu sensors for persistent pedestrian tracking," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 419–431, 2017.
- [2] R. Henschel, T. Von Marcard, and B. Rosenhahn, "Accurate long-term multiple people tracking using video and body-worn imus," *IEEE Transactions on Image Processing*, vol. 29, pp. 8476–8489, 2020.
- [3] G. Ferri, A. Munafò, A. Tesei, P. Braca, F. Meyer, K. Pelekanakis, R. Petrocchia, J. Alves, C. Strode, and K. LePage, "Cooperative robotic networks for underwater surveillance: an overview," *IET Radar, Sonar & Navigation*, vol. 11, no. 12, pp. 1740–1761, 2017.
- [4] A. Wolek, J. McMahon, B. R. Dzikowicz, and B. H. Houston, "Tracking multiple surface vessels with an autonomous underwater vehicle: Field results," *IEEE Journal of Oceanic Engineering*, pp. 1–14, 2020.
- [5] D. Svensson, M. Ulmke, and L. Hammarstrand, "Multitarget sensor resolution model and joint probabilistic data association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3418–3434, 2012.
- [6] W. Koch and G. Van Keuk, "Multiple hypothesis track maintenance with possibly unresolved measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 883–892, 1997.
- [7] M. Beard, B.-T. Vo, and B.-N. Vo, "Bayesian multi-target tracking with merged measurements using labelled random finite sets," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1433–1447, 2015.
- [8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.
- [9] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, 1983.
- [10] G. Ferri, A. Munafò, R. Goldhahn, and K. LePage, "A non-myopic, receding horizon control strategy for an auv to track an underwater target in a bistatic sonar scenario," in *53rd IEEE Conference on Decision and Control*, pp. 5352–5358, IEEE, 2014.
- [11] S. Kemna, M. J. Hamilton, D. T. Hughes, and K. D. LePage, "Adaptive autonomous underwater vehicles for littoral surveillance," *Intelligent service robotics*, vol. 4, no. 4, p. 245, 2011.

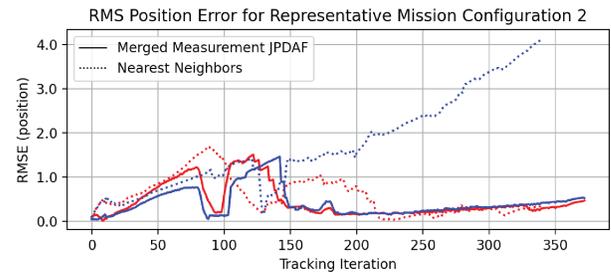
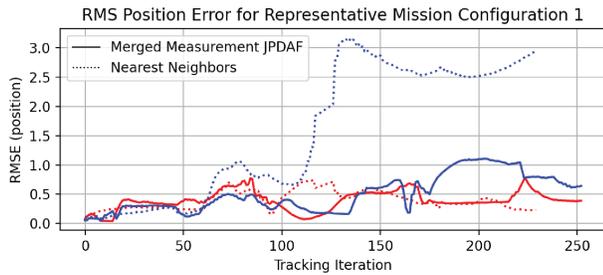
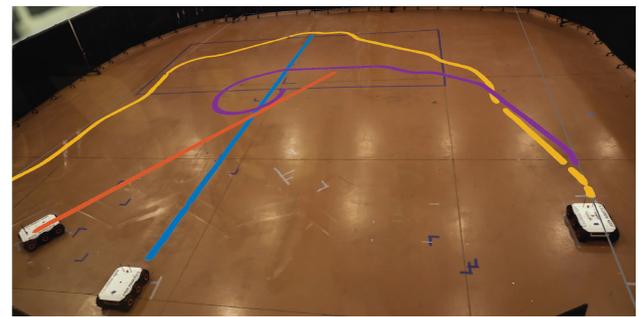
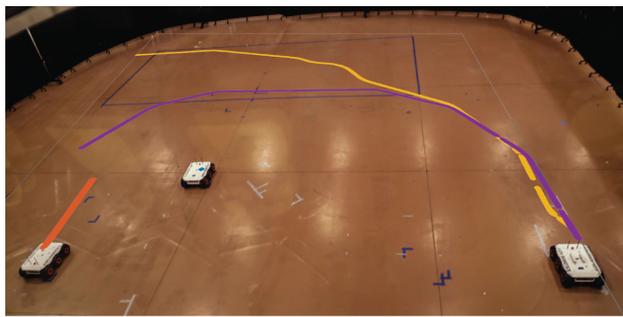


Fig. 6: Both figure a and b (top) show representative results from field experiments. Figure a shows results with one moving target while figure b shows results with two moving targets. Orange and blue trajectories are target vehicles. Purple trajectory represents Nearest Neighbor Extended Kalman Filter approach. Yellow trajectory represents the Merged Measurement JPDAF approach. Figure a and b (bottom) show the RMS Error for the two experiments respectively where the dashed lines are errors associated with tracks using the nearest neighbor approach and the solid lines are errors associated with the Merged Measurement JPDAF approach.

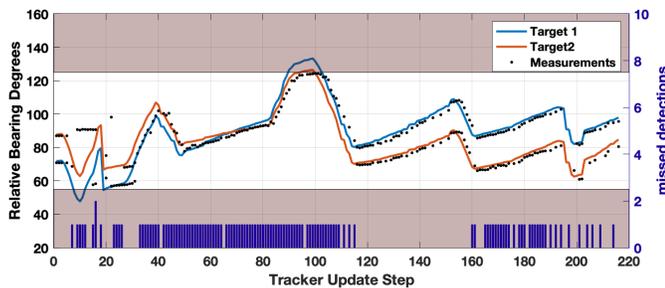


Fig. 7: Contact data associated with measurements in part one of the video attachment. Solid red and blue lines are the true bearing to target. Black dots are contacts generated from the neural network. Contacts at 0 degrees represent missed detections.)

- [12] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.
- [13] R. P. Mahler, "Multitarget bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic systems*, vol. 39, no. 4, pp. 1152–1178, 2003.
- [14] B.-n. Vo, M. Mallick, Y. Bar-shalom, S. Coraluppi, R. Osborne III, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–15, 1999.
- [15] F. E. Daum and R. J. Fitzgerald, "Importance of resolution in multiple-target tracking," in *Signal and Data Processing of Small Targets 1994*, vol. 2235, pp. 329–338, International Society for Optics and Photonics.
- [16] C. Freundlich, Y. Zhang, A. Z. Zhu, P. Mordohai, and M. M. Zavlanos, "Controlling a robotic stereo camera under image quantization noise," *The International Journal of Robotics Research*, vol. 36, no. 12, pp. 1268–1285, 2017.
- [17] L. Calkins, R. Khodayi-mehr, W. Aquino, and M. M. Zavlanos, "Sensor planning for model-based acoustic source identification," in *2020 American Control Conference (ACC)*, pp. 2679–2684, IEEE.
- [18] J. Le Ny and G. J. Pappas, "On trajectory optimization for active sensing in gaussian process models," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC)*, pp. 6286–6292, 2009.
- [19] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Information acquisition with sensing robots: Algorithms and error bounds," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6447–6454, IEEE, 2014.
- [20] S. Pathak, A. Thomas, and V. Indelman, "Nonmyopic data association aware belief space planning for robust active perception," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*.
- [21] P. Dames and V. Kumar, "Autonomous localization of an unknown number of targets without data association using teams of mobile sensors," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 3, pp. 850–864, 2015.
- [22] S. Thrun, "Probabilistic robotics," *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.
- [23] D. F. Crouse, Y. Bar-Shalom, P. Willett, and L. Svensson, "The jpdaf in practical systems: Computation and snake oil," in *Signal and Data Processing of Small Targets 2010*, vol. 7698, p. 769813, International Society for Optics and Photonics, 2010.
- [24] Stanford Artificial Intelligence Laboratory et al., "Robotic operating system."
- [25] B. Schlotfeldt, D. Thakur, N. Atanasov, V. Kumar, and G. J. Pappas, "Anytime planning for decentralized multirobot active information gathering," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1025–1032, 2018.
- [26] A. Bochkovskiy, C.-Y. Wang, and H. M. Liao, "Yolov4: Optimal speed and accuracy of object detection. arxiv," *arXiv preprint arXiv:2004.10934*, pp. 1–17, 2020.
- [27] M. Verucchi, G. Brilli, D. Sapienza, M. Verasani, M. Arena, F. Gatti, A. Capotondi, R. Cavicchioli, M. Bertogna, and M. Solieri, "A systematic assessment of embedded neural networks for object detection," in *25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1, pp. 937–944, IEEE, 2020.