

Information geometry-based action decision-making for target tracking by fixed-wing unmanned aerial vehicle: From algorithm design to theory analysis

Yunyun Zhao, Xiangke Wang , Yirui Cong  and Lincheng Shen

Abstract

In this article, we study the ground moving target tracking problem for a fixed-wing unmanned aerial vehicle equipped with a radar. This problem is formulated in a partially observable Markov process framework, which contains the following two parts: in the first part, the unmanned aerial vehicle utilizes the measurements from its radar and employs a Kalman filter to estimate the target's real-time location; in the second part, the unmanned aerial vehicle optimizes its trajectory in a real-time manner so that the radar's measurements can include more useful information. To solve the trajectory optimization problem, we proposed an information geometry-based partially observable Markov decision process method. Specifically, the cumulative amount of information in the observation is represented by Fisher information of information geometry, and acts as the criterion of the partially observable Markov decision process problem. Furthermore, to guarantee the real-time performance, an important trade-off between the optimality and computation cost is made by an approximate receding horizon approach. Finally, simulation results corroborate the accuracy and time-efficiency of our proposed method and also show our advantage in computation time compared to existing methods.

Keywords

Fixed-wing UAV, information geometry, action decision-making, partially observable Markov decision process, three-dimensional observation, target tracking

Date received: 3 January 2018; accepted: 3 June 2018

Topic: Mobile Robots and Multi-Robot Systems

Topic Editor: Nak-Young Chong

Associate Editor: Euntai Kim

Introduction

Guiding unmanned aerial vehicle (UAV) to detect and track a suspicious ground target is an important requirement in many intelligence, surveillance, target acquisition, and reconnaissance (ISTAR)¹ problems. Different from predefined trajectory tracking problems, tracking a moving target is more challenging, since it requires the UAV to respond promptly to the random movement of the target. Our article focuses on the problem of a UAV tracking a moving ground target in an uncertain environment. In many

College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, People's Republic of China

Corresponding author:

Xiangke Wang, College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, People's Republic of China.

Email: xkwang@nudt.edu.cn



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

practical scenarios, the accurate observation and tracking of a target is achieved through adequate maneuver of the UAV equipped with a sensor. Therefore, the UAV needs to make a movement strategy based on the state of the target. However, the information of the target location obtained by the UAV sensor is usually incomplete or imperfect. In this regard, the task of target tracking mainly consists of two aspects²: one is the estimation of the state of the target based on the measurements obtained by the sensor, and the other is the adjustment of the position/pose of the UAV based on the prediction of the target state to obtain better measurements.

For target state estimation, different strategies have been developed in the past decade. In order to estimate and predict the target state more accurately, the classical methods such as Kalman filter (KF),³⁻⁷ particle filter,⁸⁻¹¹ and their modifications were widely used. For instance, an unscented KF was utilized to track an underwater submarine target/moving ship.^{6,12} Tang and Ozguner proposed the particle-filter-and-hospitality-map algorithm (PF-HMap)⁸ to deal with the general target tracking maintenance problem with regional and intermittent measurements. In these abovementioned methods, the classical KF is one of the most widely used methods for estimation and tracking, due to its optimality, simplicity, and tractability.

For position/pose adjustment, one possible method is to track the estimated location of the target. A novel algorithm combining the tangent vector field guidance (TVFG) path-planning approach and the Lyapunov vector field guidance (LVFG) algorithm¹³ was developed. Given the target position and current UAV dynamic state, this method is theoretically possible to obtain the shortest path with UAV operational constraints. In addition, based on the division of two kinds of possible path patterns, that is, the z type (sinusoidal type) and the whirling type, the ground target pursuit algorithm^{14,15} generates waypoints step by step and steers the UAV to the latest waypoint. Most of these planning methods are based on precomputed vector fields or alternative paths, which do not respond well to system uncertainties.

The other method for position/pose adjustment is based on the decision-making framework, which can better handle the system uncertainties. In this framework, the transition model of the tracking system is usually assumed to be Markovian. Then, the tracking policy selection problem can be formulated as a partially observed Markov decision process (POMDP), in which the state is only partially observed and one seeks to design a control policy which maps the state probability distributions to actions and maximizes the accuracy of target location. With its decision-making ability in an uncertain environment,^{16,17} the POMDP framework has been widely used in a variety of real-world scenarios. Prentice and Roy¹⁸ addressed the problem of trajectory planning with imperfect state information, and model it as a linear-Gaussian POMDP.

Assuming that the UAV's state and kinematics are known, Ragi and Chong^{3,4} employed the POMDP, combining with the motion constraints of the UAV, to design the guidance algorithm of UAV tracking a ground target. However, solving POMDP optimally is proven to be PSPACE-hard.¹⁹ A large number of literature focuses on various heuristic or approximate solution techniques.²⁰ Some literature use the receding horizon theory to approximate a POMDP. Sunberg et al.²¹ proposed a receding horizon control approach to solve the information space dynamic programming of POMDPs. In the POMDP framework, the tracking decision problem was transformed into an optimization problem, and the choice of the optimization criterion plays a crucial role,²² which directly determines the speed and performance of the solution. Different from the approach^{3,4} that used the trace of covariance matrix, we use the cumulative quantity of information from information geometry (IG) as the reward criterion of POMDP to evaluate the performance of the tracking strategy in essence and simplify the calculation.

IG is proposed by Rao,²³ and axiomatized by Chentsov.²⁴ It constructs a distance representation between statistical distributions. It deals with families of parameterized probability densities which carry a metric structure. This structure is derived via the well-known Fisher information metric. Recently, significant attention has been drawn in the area of IG. Costa et al.²⁵ presented the Fisher information distance as a measure of difference between two probability distribution functions. The Fisher information distance is related to the information of the target estimation. Therefore, the target tracking problem is transformed into searching for the strategy that maximizes the accuracy of the target estimation. There have been many works on how IG methodology can be applied to the tracking process as it does in the signal processing.²⁶⁻²⁹ In their study, Akselrod et al.²⁷ proposed a MDP model to formulate the collaborative sensor management for multi-target tracking decision processes, where the objective function is based on the Fisher information measure. However, in that study, the Fisher information is merely approximated by the posterior covariance matrix from the KF, which is not an exact derivation. Actually, it did not analyze the source and significance of the Fisher information in essence. In target tracking, the noise covariance depends on the performance of the sensor. That means the covariance matrix of noise matches to the Riemannian manifold of positive definite matrix. Wang et al.²⁹ derived a complete form of relationship between the accuracy of the target location and the optimal sensor position, based on the maximum determinant of the Fisher information matrix (FIM) criterion, where the theory of IG to study the problem of bearing-only tracking was employed.

The above summarizes the research results about the construction of the tracking decision-making framework, the selection of the evaluation function, and the solution

of the optimal decision. However, even though IG is a power tool in analyzing the statistics of the moving target, few studies provide rigorous analysis and design of the IG-based UAV ground target tracking problem. Previous work by Zhao et al.³⁰ has shown the effectiveness of the IG method in the target tracking decision in two-dimensional (2-D) space by simulation results. In this article, we extend the problem of target tracking with an airborne radar to general three dimensions with new simulation results and more detailed analysis. The iterative form of FIM about the 3-D bearing-and-range radar, which takes into account the uncertainties of predicted target states, is derived. Then, the Fisher information distance on the Riemann manifold is regarded as the basis for POMDP, and the convergence of the POMDP approximate solution algorithm is proved theoretically. The detailed contributions of the proposed algorithm are as follows:

1. A novel IG-based POMDP frame is provided for the UAV to track a moving ground target in a 3-D uncertain environment. In this frame, the state of the target is observed by the 3-D bearing-and-range radar.
2. The optimization criterion based on the 3-D observation model in the POMDP is derived based on Fisher information in the view of the IG, which is the key to optimize the action policies in the UAV-to-target problem.
3. An approximate receding horizon control is developed to obtain an acceptable control strategy in the trade-off between the optimization and the computation cost. We also prove the convergence of approximation algorithm theoretically.

The rest of this article is organized as follows. In the ‘‘Problem formulation’’ section, the system model is given and the target tracking problem is formulated. The framework of the POMDP and its criterion of accumulative information are presented in ‘‘Target tracking decision-making based on IG’’ section, where the FIM is derived by iterative calculation via the predicted target state. In ‘‘Approximate receding horizon approach for POMDP’’ section, we introduce our proposed approximate receding horizon approach for the POMDP, and analyze the performance of the algorithm. ‘‘Simulation results’’ section presents the simulation and results, which is followed by the conclusion.

Problem formulation

In this article, the target tracking system is composed of a UAV and a target vehicle. The mission of the UAV is to observe and track the moving target on the ground. The UAV is equipped with a radar that can obtain the bearing and range measurements of the tracked object with limited precision and reliability.

Fixed-wing UAV model

The fixed-wing UAV dynamics augmented by the autopilot is a high dimensional, highly nonlinear, and extremely complex system. In our work, the UAV is supposed to fly at a constant altitude h , and the unicycle model is adopted to describe kinematics of fixed-wing UAV.

The UAV autopilot controls bank angle ϕ and forward acceleration a directly. The UAV state, which is a part of the world states, is $\mathbf{s}^u = (x^u, y^u, v^u, \theta^u)^T$, where (x^u, y^u) denotes the position of the UAV in the inertial coordinate system (CS) and (v^u, θ^u) represents the UAV’s speed and course angle. The UAV is assumed to be fully observable, and the discrete kinematic model of the UAV is as follows

$$\begin{cases} x_{k+1}^u = x_k^u + v_k^u T \cos\theta_k^u, \\ y_{k+1}^u = y_k^u + v_k^u T \sin\theta_k^u, \\ v_{k+1}^u = [v_k^u + a_k T]_{V_{\min}^u}^{V_{\max}^u}, \\ \theta_{k+1}^u = \theta_k^u + (gT \tan(\phi_k) / v_k^u) \end{cases} \quad (1)$$

where $[v]_{V_{\min}^u}^{V_{\max}^u} = \max\{V_{\min}^u, \min(V_{\max}^u, v)\}$, V_{\max}^u and V_{\min}^u are the upper and lower bounds of the ground speed of the fixed-wing UAV, g is the acceleration of gravity, and T is time step.

Target model

In this system, the target is on the ground and its exact state is not available. The state is given by $\mathbf{s}^t = (x^t, y^t, \dot{x}^t, \dot{y}^t)^T$, including the location (x^t, y^t) and velocity (\dot{x}^t, \dot{y}^t) . The mobility of the ground target is modeled as follows, which is commonly used in the literature

$$\mathbf{s}_{k+1}^t = \mathbf{A}\mathbf{s}_k^t + \mathbf{v}_k, \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \quad (2)$$

where \mathbf{A} is the state transition matrix with the following form

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

and \mathbf{Q}_k represents the process noise covariance matrix at time k .

The states of the target are measured by the active radar mounted on the UAV. The radar can measure the range and bearing information determined by relative state $\mathbf{s}^o = (x, y, \dot{x}, \dot{y})^T$, which is represented by $(x, y) = (x^t - x^u, y^t - y^u, \dot{x}^t - \dot{x}^u, \dot{y}^t - \dot{y}^u)$. Next, it is necessary to illustrate the observation model.

Observation model

In this work, the UAV only installs one radar as the measuring sensor. The airborne radar for target tracking

provides measurement of the target in the sensor CS, which is a polar CS with range r and bearing φ . The measurement model with noise at time k is

$$\mathbf{o}_k = \begin{bmatrix} r_k \\ \varphi_k \end{bmatrix} = \boldsymbol{\eta}(\mathbf{s}_k^o) + \boldsymbol{\omega}_k \quad (4)$$

where $\boldsymbol{\eta}(\cdot)$ is the observation function. Then, $\boldsymbol{\eta}(\mathbf{s}_k^o)$ indicates the true position of the target in the polar CS of the sensor and $\boldsymbol{\omega}_k$ is the measurement error. It is usually assumed that the measurement error follows a zero-mean normal distribution in the sensor CS, that is

$$\mathbf{o}_k | \mathbf{s}_k^o \sim N(\boldsymbol{\eta}(\mathbf{s}_k^o), \mathbf{C}(\mathbf{s}_k^o))$$

For the radar, the error-free target position in the sensor polar CS is represented as

$$\boldsymbol{\eta}(\mathbf{s}_k^o) = \begin{bmatrix} \eta_1(\mathbf{s}_k^o) \\ \eta_2(\mathbf{s}_k^o) \end{bmatrix} = \begin{bmatrix} \sqrt{x_k^2 + y_k^2 + h^2} \\ \arctan(y_k/x_k) \end{bmatrix} \quad (5)$$

As for an active sensor, the measurement error is dependent on the signal-to-noise ratio which is proportional to the fourth power of distance. Then the covariance matrix of the measurement is

$$\mathbf{C}(\mathbf{s}_k^o) = \begin{bmatrix} r_k^4 \sigma_r^2 & 0 \\ 0 & \sigma_\varphi^2 \end{bmatrix} \quad (6)$$

where σ_r and σ_φ are the standard deviations of the range and bearing measurements, respectively. In equation (6), r_k^4 indicates the effect of the fourth power decay of the echo signal strength on the variance of the distance measurement. Strictly speaking, the power decay should be related to the actual distance rather than the distance measurement. Since the actual distance is not available at the UAV, we use distance measurement r_k instead.

This study considers a UAV decision-making problem, in which the goal is to design an algorithm to control a UAV for target tracking. Specifically, the UAV motion model is simplified as equation (1). It is mounted with a sensor that measures the relative position of the target. The observation model (equation (4)), combined with the assumed target model (equation (2)), is used to estimate and predict the states of the target. Based on these states the UAV makes action decision. The values of actions are limited to the maximum and minimum range. The objective is to obtain better observation for a more accurate estimation of the target state.

Tracking in Cartesian coordinate

In the target tracking problem, the target's movement is described in Cartesian coordinate, while the measurement is available physically in the sensor polar CS. Thus, it is necessary to convert the measurements from the polar CS to the Cartesian coordinate. Specifically, in the Cartesian

CS (selecting the UAV as the origin), the measurement model is converted to the following form

$$\mathbf{o}'_k = \mathbf{H}\mathbf{s}_k^o + \boldsymbol{\omega}'_k \quad (7)$$

where the measurement matrix is

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (8)$$

In equation (7), \mathbf{o}'_k is the relative target position measured by the sensor, and $\boldsymbol{\omega}'_k$ stands for the resulting measurement error.

Defining $\boldsymbol{\varphi}(\cdot)$ as the transition function from polar CS to Cartesian CS, that is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{\varphi}\left(\begin{bmatrix} r \\ \varphi \end{bmatrix}\right) = \begin{bmatrix} r \cos\varphi \\ r \sin\varphi \end{bmatrix}$$

where (x, y) and (r, φ) are the coordinates in the polar CS and the Cartesian CS, respectively. Then, the relative position from the target to the UAV in Cartesian coordinate, defined as $\mathbf{H}\mathbf{s}_k^o$, has the following form

$$\mathbf{H}\mathbf{s}_k^o = \boldsymbol{\varphi}(\mathbf{o}_k - \boldsymbol{\omega}_k) \quad (9)$$

By Taylor series expansion of $\boldsymbol{\varphi}(\mathbf{o}_k - \boldsymbol{\omega}_k)$ around the noisy measurement \mathbf{o}_k , we obtain

$$\mathbf{H}\mathbf{s}_k^o = \boldsymbol{\varphi}(\mathbf{o}_k) - \mathbf{J}(\mathbf{o}_k)\boldsymbol{\omega}_k + \mathcal{O}(\boldsymbol{\omega}_k) \quad (10)$$

where $\mathcal{O}(\boldsymbol{\omega}_k)$ stands for the higher order (≥ 2) terms, and the Jacobian matrix $\mathbf{J}(\mathbf{o}_k)$ is evaluated at the noisy measurement \mathbf{o}_k

$$\mathbf{J}(\mathbf{o}_k) = \frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{o}_k} = \frac{h}{r_k} \begin{bmatrix} \cos(\varphi_k) & -r_k \sin(\varphi_k) \\ \sin(\varphi_k) & r_k \cos(\varphi_k) \end{bmatrix} \quad (11)$$

Considering $\boldsymbol{\varphi}(\mathbf{o}_k) = \mathbf{o}'_k$ and the equation (10), the exactly converted measurement model (equation (7)) can be written as

$$\mathbf{o}'_k = \mathbf{H}\mathbf{s}_k^o + \underbrace{\mathbf{J}(\mathbf{o}_k)\boldsymbol{\omega}_k - \mathcal{O}(\boldsymbol{\omega}_k)}_{\boldsymbol{\omega}'_k} \quad (12)$$

The standard approach usually ignores $\mathcal{O}(\boldsymbol{\omega}_k)$ in the exact model (equation (12)), and treats $\boldsymbol{\omega}'_k$ approximately as zero-mean with covariance

$$\mathbf{R}(\mathbf{s}_k^o) = \mathbf{J}(\mathbf{o}_k)\mathbf{C}(\mathbf{s}_k^o)\mathbf{J}(\mathbf{o}_k)^T \quad (13)$$

Target tracking decision-making based on IG

The decision-making problem of UAV under uncertain environment is modeled as a POMDP. In the POMDP framework, the decision-making program obtains measurements from the airborne radar and makes the action decisions for the UAV. For more measurement information of the target, we should predict the states of the target and make appropriate action decisions based on a certain

criterion. In this article, we represent the reward criterion by the accumulative information, which is derived from the FIM in the IG. This section introduces the POMDP-based target tracking algorithm, especially the calculation of the decision-making criterion. In the second part of this section, we describe the traditional criterion and analyze its shortcomings. Then we propose the decision-making criterion based on the IG in the last part of this section. The key point is that we use the iterative calculation form of the FIM for the range-and-bearing radar to redefine the decision-making criterion in essence.

The framework of POMDP-based decision-making

POMDP is a stochastic process controlled by a decision-maker. In the target tracking problem, since the system state transition is a random process with the Markov property, and the states of the target cannot be obtained directly, we use the POMDP to select a sequence of actions for the UAV to reduce the uncertainty of target localization. In general, an infinite horizon POMDP is defined by a tuple $M = \langle \mathcal{S}, U, T, \Omega, \mathcal{O}, R, b_0 \rangle$. The following defines the key components of the POMDP with respect to a target tracking problem.

- *States.* The state set \mathcal{S} is a discrete Borel space and the state at time k , including the state of the UAV s_k^u and the state of the target s_k^t , and the calculated relative state of the target to the UAV s_k^o .
- *Actions.* In this system, the action of the UAV is the control quantity $\mathbf{u}_k \in U$. Specifically, the action decision at time k is given by $\mathbf{u}_k = [a_k, \phi_k]^T$.
- *State transition probabilities.* T is a set of conditional transition probabilities between states. The UAV state transition function is the kinematic equation (equation (1)). The target state is unknown and approximated by a constant velocity model (equation (2)).
- *Observations and observation probabilities.* Ω is the set of possible observation, and \mathcal{O} is the set of conditional probabilities of the observation when an action is performed. In the tracking system, the UAV is assumed to be fully observable. We use a radar to measure the target state. The observation law in the sensor CS is equation (4) and the linearized conversion in the Cartesian CS is equation (7). We use the latter, when we estimate the current target state based on historical data. However, when we predict the accuracy of target location, we use the former, which is the research subject of this article.
- *Belief state.* Belief state is the distribution of the real state. Specially, the state distribution at time $k = 0$ is b_0 . Since the UAV states are fully observable, the belief state is $b_k(s^u) = \delta(s^u - s_k^u)$. At the same time, the probability distribution of target positions is assumed to be Gaussian, and can be expressed

approximately as $b_k(s^t) \sim N(\mu_k, \mathbf{P}_k)$, where μ_k and \mathbf{P}_k are derived from the standard KF. If an optimal action policy exists in POMDP, there is an optimal action sequence that only depends on the belief state feedback.³¹ Then we can make decision according to the belief state.

- *Reward function.* The real-valued reward function R defines the reward of the action. It is used to compare different alternative action policies. In the target tracking problem, the core of evaluation is the accuracy of the target location. Some literatures describe it as the trace of covariance. In our work, for the purpose of better reflecting the essential characteristics of radar measurement, we propose the IG method to represent the reward function.

Remark 1. Partially observable Markov process. At each time period, the system is in some state $s \in \mathcal{S}$. The agent takes an action $\mathbf{u} \in U$, which causes the transition of the system to state s' with probability $T(s'|s, \mathbf{u})$. At the same time, the agent receives an observation $\mathbf{o} \in \Omega$ which depends on the new state of the system with probability $\mathcal{O}(\mathbf{o}|s', \mathbf{u})$. Finally, the agent receives a reward equal to $R(s, \mathbf{u})$. Then the process repeats.

The goal of the POMDP for the agent is to choose actions at each time step that maximize its long-run average expected reward

$$J_\infty^\pi(s) = \liminf_{H \rightarrow \infty} \frac{1}{H} \mathbb{E} \left[\sum_{k=0}^{H-1} R(s_k, \mathbf{u}_k) | s_0 = s \right] \quad (14)$$

where H is the time horizon. In our target tracking system, the aim of decision-making is to find the optimal control policies for the UAV. This strategy is to make the UAV better able to observe and track the target, so as to maximize the accuracy of the target position estimation.

A policy is a sequence $\pi = \{\pi_k\}$ of probability π_k on the action space given the state (belief state) and action history satisfying the constraint. We denote by Δ the set of all policies. Once we have a POMDP model and a set Δ of admissible policies, we need to express a performance index or objective function. It is a function to assess the system's performance when a given policy $\pi \in \Delta$ is used and the initial state (or belief state) of the system is s (or b). Then we are able to select the optimal strategy. Unlike the actual state, the belief state is completely observable. It is necessary to represent the expected reward based on the belief state b

$$R'(b, \mathbf{u}) = \int R(s, \mathbf{u}) b(s) ds \quad (15)$$

This reward $R'(b, \mathbf{u})$ then represents the expected reward function of POMDP. The objective is transformed into

$$J_\infty^\pi(b) = \liminf_{H \rightarrow \infty} \frac{1}{H} \mathbb{E} \left[\sum_{k=0}^{H-1} R'(b_k, \mathbf{u}_k) | b_0 = b \right] \quad (16)$$

according to the nominal belief state optimization.⁴ The objective function is approximated as follows

$$J_{\infty}^{\pi}(b) \approx \liminf_{H \rightarrow \infty} \frac{1}{H} \left[\sum_{k=0}^{H-1} R(\hat{s}_k, \mathbf{u}_k) \middle| \hat{s}_o = \int sb(s) ds \right] \quad (17)$$

where $\hat{s}_1, \hat{s}_2, \dots$ is a nominal belief state sequence. The nominal target belief state sequence can be obtained from the target state transfer law (equation (2)) with exactly zero-noise sequence.

Covariance-based conventional criterion function

In some literatures,^{3,4} the belief of the target mentioned above can be identified with the state of the tracker (μ_k, \mathbf{P}_k) , which are the posteriori state estimation and the posteriori error covariance matrix, and are given by

$$\begin{cases} \mu_{k+1} = \mathbf{A}\mu_k \\ \mathbf{P}_{k+1|k} = \mathbf{A}\mathbf{P}_k\mathbf{A}^T + \mathbf{Q}_k \\ \mathbf{P}_{k+1} = [\mathbf{P}_{k+1|k} + \mathbf{H}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1}]^{-1} \end{cases} \quad (18)$$

The reward function is defined to represent the uncertainty of the target location, which is usually represented by the mean-squared error between the tracker and the target. Then, the objective function is as follows

$$J_{\infty}^{\pi}(b) \approx \liminf_{H \rightarrow \infty} \frac{1}{H} \left[\sum_{k=0}^{H-1} (-\text{tr} \mathbf{P}_{k+1}) \right] \quad (19)$$

The Kalman filtering equation is a linearized conversion of the actual system model. The measurement model in the sensor CS is equation (4), and the observation model in the Cartesian CS is in the form of equation (7). In contrast with the long-run reward from the trace of the Kalman filtering error covariance matrix (equation (19)), the FIM in the IG processes the original measurement data of the radar. It is derived from the measurement model in the natural sensor CS directly and has a clearly physical meaning, which reflects the volume of information from the measurement data. The greater the cumulative information is, the more accurate the measurement would be. Hence the Fisher information from the IG can better evaluate the accuracy of the predicted states. We use the Fisher information distance (which will be explained later) on the statistical manifold as the basis of the POMDP.

Cumulative information in IG

IG²³ offers a comprehensive result about statistical models by regarding them as geometrical objects. From the perspective of IG, the set of belief states forms the statistical manifold of a particular geometric structure. The information is defined on the Fisher information divergence between the current belief state and the belief state after a

measurement has been made. The cumulative information which the sensor may acquire from the target is characterized by the Fisher information distance. In a discrete measurement sampling scenario, the sum of the determinant of FIM is used to approximate the information distance.³²

In this article, the determinant of the FIM is used to characterize the volume of information obtained by measurements. The performance index is the expected long-run average reward, therefore, it is a feasible way to predict the information each time through iterations. The following introduces the iteration process of the discrete FIM of the radar measurements.

We define \mathbf{G}_k to express the FIM at time k , then the reward of the POMDP is $R(s_k, \mathbf{u}_k) = |\mathbf{G}_{k+1}|$, where $|\cdot|$ denote the determinant of the matrix. When we predict the volume of information in the coming period of time, the FIM at time k is defined^{5,23} by

$$\mathbf{G}_k = \mathbb{E} \left[(\nabla_s \ln p(\mathbf{o}_{1:k}|s_k)) (\nabla_s \ln p(\mathbf{o}_{1:k}|s_k))^T \right] \quad (20)$$

where $p(\mathbf{o}_{1:k}|s_k)$ is the batch measurement likelihood. In this tracking scenario, the batch measurement likelihood is defined as equation (21)

$$p(\mathbf{o}_{1:k}|s_k) = \prod_{m=1}^k \frac{1}{\sqrt{2\pi} |\mathbf{C}(s_k, m)|} \exp \left(-\frac{1}{2} (\mathbf{o}_m - \boldsymbol{\eta}(s_k, m))^T \mathbf{C}^{-1}(s_k, m) (\mathbf{o}_m - \boldsymbol{\eta}(s_k, m)) \right) \quad (21)$$

where $\boldsymbol{\eta}(s_k, m)$ denotes error-free target position of the sensor polar, evaluated at time m in terms of the state s_k at k . That is to say, $\boldsymbol{\eta}(s_k, m), m = 1, 2, \dots, k$ is given by

$$\boldsymbol{\eta}(s_k, m) = \begin{bmatrix} \eta_1(s_k, m) \\ \eta_2(s_k, m) \end{bmatrix} = \begin{bmatrix} \sqrt{\Delta x_{k,m}^2 + \Delta y_{k,m}^2 + h^2} \\ \arctan \left(\frac{\Delta x_{k,m}}{\Delta y_{k,m}} \right) \end{bmatrix} \quad (22)$$

Besides, $\mathbf{C}(s_k, m)$ is the measurement noise covariance evaluated at time m in terms of the state s_k at k , which is given by

$$\mathbf{C}(s_k, m) = \begin{bmatrix} r_{k,m}^4 \sigma_r^2 & 0 \\ 0 & \sigma_{\varphi}^2 \end{bmatrix} \quad (23)$$

and

$$r_{k,m}^2 = \Delta x_{k,m}^2 + \Delta y_{k,m}^2 + h^2$$

In the above equations, considering the linear motion target model described previously, we have

$$\begin{cases} \Delta x_{k,m} = x_k^t - (k-m)Tx_k^t - x_m^u \\ \Delta y_{k,m} = y_k^t - (k-m)Ty_k^t - y_m^u \end{cases} \quad (24)$$

Then the FIM at time k can be calculated in a recursive form as follows

$$\mathbf{G}_k = \mathbf{G}_{k-1} + \Delta \mathbf{G}_k \quad (25)$$

where

$$\begin{aligned} [\Delta \mathbf{G}_k]_{ij} = & \begin{bmatrix} \frac{\partial \boldsymbol{\eta}(s_k, k)}{\partial s_i} \end{bmatrix}^T \mathbf{C}^{-1}(s_k, k) \begin{bmatrix} \frac{\partial \boldsymbol{\eta}(s_k, k)}{\partial s_j} \end{bmatrix} \\ & + \frac{1}{2} \text{tr} \left(\mathbf{C}^{-1}(s_k, k) \frac{\partial \mathbf{C}(s_k, k)}{\partial s_i} \mathbf{C}^{-1}(s_k, k) \frac{\partial \mathbf{C}(s_k, k)}{\partial s_j} \right) \end{aligned} \quad (26)$$

In equation (26), $[\Delta \mathbf{G}_k]_{ij}$ represents the i th row and the j th column of elements in the matrix. The specific derivation process is given in the Online Supplementary material.

Since the sensor collects measurements only at discrete points, the accumulative information should only consider those points when the sensor takes measurements. It is assumed that the decision-making points are consistent with the measurement points. Therefore, the sum of determinant of the FIM²⁹ is used to approximate the accumulative information (s_1^u, s_2^u, \dots) , that is

$$\mathcal{D}(s_1, s_2, \dots) \approx \sum_{k=1}^{\infty} |\mathbf{G}_k| \quad (27)$$

The reward function based on the IG is

$$J_{\infty}^{\pi}(b) \approx \liminf_{H \rightarrow \infty} \frac{1}{H} \left[\sum_{k=0}^{H-1} |\mathbf{G}_{k+1}| \right] \quad (28)$$

The larger the volume of information is, the more accurate the estimation of the target state will be and the better the tracking performance the algorithm will have. Our decision-making algorithm aims to find the optimal sequence of actions which maximizes the predicted cumulative information. From equation (20), we obtain that the FIM is determined by the relative position between the UAV and the target. However, the states of the target are unobservable. Thus, when we make decision, we use the predicted belief states to present the reward of the POMDP, that is, the predicted determinant of the FIM. That is to say, we use

$$\mu_{k+1} = \mathbf{A}\mu_k$$

to predict the belief state of the target, and $\mu = \int sb(s) ds$ is used to replace the real state.

Approximate receding horizon approach for POMDP

Since it is intractable to solve the policy for maximizing the objective function (equation (17)) exactly, we use an fixed finite-horizon POMDP to create a policy to solve the

infinite-horizon POMDP approximatively in this article, which is called ‘‘approximate receding horizon approach.’’

Algorithm design

The idea of the approximate receding horizon approach is that at the current time t , we obtain the optimal policy sequence of the POMDP over a finite horizon $[t, t + H]$, and at each time t , we predict the future states over the horizon length H based on the current state of the system. The finite horizon H indicates the number of predicted steps. Once an optimal sequence is found, the first action is applied as a control command to the UAV and the whole process repeats. We can take the process at time 0 as an example for analysis. That is, the current state is assumed to be s_0 . Then the problem boils down to find the sequence of actions over a time horizon H to maximize the expected cumulative reward.

The specific algorithm is as follows.

Algorithm 1. The approximate receding horizon algorithm

```

1 Initialization ( $s_0^u, b_0(s^t)$ );
2 for  $k = 1$  to Termination do
   Input : Radar-to-target measurement sequence  $\mathbf{o}_k$ ,
          UAV state sequence  $s_k^u$ ;
3    $b_k(s^t) \leftarrow \text{KalmanFilter}(\mathbf{o}_k, s_k^u)$ ;
4    $\mu_0 \leftarrow b_k(s^t)$ ;
5    $\mathbf{G}_i \leftarrow 0$ ;
6   Solve the optimal programming:
7    $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{H-1} \leftarrow \text{Fmincon}(f \leftarrow -\sum_{i=1}^H |\mathbf{G}_i|)$ 
8   for  $i = 1$  to  $H$  do
9      $s_i^u \leftarrow \mathbf{T}(s_{i-1}^u, \mathbf{u}_{i-1})$ , ref. (1);
10     $\mu_i \leftarrow \mathbf{A}\mu_{i-1}$ ;
11     $s_i^t \leftarrow \mu_i$ ;
12     $\mathbf{G}_i \leftarrow \mathbf{G}_{i-1} + \text{FisherInfor}(\mu_i, s_i^u)$ ,
       ref. (25);
13   end
   Output: UAV action policy  $\mathbf{u}_0$ ;
14 end
```

Algorithm analysis

In this section, we analyze the performance of the approximate receding horizon approach for infinite-horizon average reward.

When we say the algorithm is stable, we mean that the difference between the value function and the optimal value function is bounded. Then the proposed algorithm is stable and the following theorem can be obtained.

Theorem 1. By defining the receding H -horizon control policy as $\pi_H \in \Delta$ with $H < \infty$, we have

$$\sup_b |V_{\infty}^{\pi_H} - V_{\infty}^*| \leq \frac{\|R'\|}{1 - \alpha} \cdot \alpha^n \quad (29)$$

Proof. Recall that the state of the UAV is observable and controllable, and the state of the target obeys normal distribution.

The expectation of the system state at time k is

$$E(\mathbf{s}_k) = (\mathbf{s}_k^u, \mu_k) \quad (30)$$

Then we have $E(\mathbf{s}_k) \in \mathcal{S}$, since the real state space \mathcal{S} is a Borel space.

Define $\hat{\mathcal{S}} = E(\mathcal{S})$, $\mathbf{K} = \{(\hat{\mathbf{s}}, \mathbf{u}(\hat{\mathbf{s}})) | \hat{\mathbf{s}} \in \mathcal{S}, \mathbf{u} \in U\}$. Since the state space is continuous and compact, the system satisfies the ergodicity conditions. In other words, there exists a state expectation $\hat{\mathbf{s}}^* \in \mathcal{S}$ such that the transition probability $p(\hat{\mathbf{s}}^* | \mathbf{k}) = 1$ for all $\mathbf{k} \in \mathbf{K}$.

The ergodicity condition can be expressed in another way.³³ That is, there exists a constant $\alpha < 1$ such that

$$\sup_{\mathbf{k}, \mathbf{k}'} \|p(\cdot | \mathbf{k}) - p(\cdot | \mathbf{k}')\| \leq 2\alpha$$

where the sup is over all \mathbf{k} and \mathbf{k}' in \mathbf{K} , and $\|\cdot\|$ denotes the variation norm for signed measures.

Recall the one-step reward $R'(b_k, \mathbf{u}_k)$ of the POMDP based on the IG. The reward is the accumulative information when the system state transforms from \mathbf{s}_k to \mathbf{s}_{k+1} .

Thus we have

$$R'(\mathbf{b}_k, \mathbf{u}_k) = |\mathbf{G}_{k+1}| \quad (31)$$

where \mathbf{G}_{k+1} is deduced from equation (25). The determinant of the FIM is:

$$|\mathbf{G}_{k+1}| = \frac{8r_{k+1}^2 \sigma_r^2 + 1}{r_{k+1}^6 \sigma_r^2 \sigma_\varphi^2} \quad (32)$$

The smallest distance between the target and the UAV is the height of the UAV which is h . From this it can be seen that the reward is bounded, that is

$$|R'| \leq \frac{8h^2 \sigma_r^2 + 1}{h^6 \sigma_r^2 \sigma_\varphi^2}$$

The above specification shows that the system satisfies the condition under which Hernández-Lerma³³ show that

$$0 \leq V_\infty^* - V_\infty^{\pi_H} \leq \frac{\|R'\|}{1 - \alpha} \cdot \alpha^n \quad (33)$$

We can see that the receding horizon approach provides a good approximation for the optimal infinite-horizon average reward and the error approaches zero geometrically with α .

The decision-making of UAV needs high real-time performance, so we tend to choose algorithm which is lower in the complexity of computation time. Next, the article analyzes the complexity of the computation time of the objective function in our algorithm (equation (28)), and compares it with that based on KF (equation (19)).

Using the reward derived from the IG to calculate the objective function, we should obtain \mathbf{G}_k from iterative computations of equations (39) to (41) in the Online

Table 1. Required operations for the IG-based objective function and the KF-based objective function.

	Addition	Multiplication	Other operations
IG	30H-2	59H	0
KF	273H-1	635H	1H sin, 1H cos

IG: information geometry; KF: Kalman filter.

Supplementary material. In the same way, using the reward derived from the KF to compute the objective function, we need to obtain \mathbf{P}_k from equation (18) and compute the measurement covariance \mathbf{R}_k in the Cartesian CS. According to the above analysis, the required operations for the IG-based objective function and the KF-based objective function during time horizon H are listed in Table 1. In these operations, we ignore the calculation of state transitions of the UAV and the target, which are the same in both algorithms.

It is clearly shown in Table 1 that the required operations, either addition and multiplication or other operations of the IG-based objective function (equation (28)), are much less than those of the KF-based objective function (equation (19)), in other words, it can greatly improve computational efficiency to use information accumulation as the objective function of the optimal decision-making.

Remark 2. Several notes on operations. (1) The operations of matrix inversion in equation (18) are counted based on the adjoint matrix inversion method. (2) In the following simulation, the MATLAB command *fmincon* is used to minimize the objective function. (3) Due to the randomness of the tracking system, the difference of initial conditions, and the limitation of iteration steps, the computation time of optimal strategy is not linear with the computational complexity of the objective function and the tracking error also has a certain randomness.

Simulation results

This section presents the simulations to verify the performance of the proposed algorithm, and compares the simulation results with other method. In the comparative algorithm, the trace of the covariance in the KF is used as the objective function.

The simulation is implemented in MATLAB R2016b (maci64), where the action decision is obtained from the MATLAB command *fmincon*. The speed of the UAV is limited between 25 m/s and 35 m/s, and the fixed flight height is set to 100 m. To corroborate the effectiveness of our algorithm for different target movements and sensor parameters, we design four groups of typical simulations first.

In these four experiments, the time horizon H is set to 6,³ which means that in every time step, the UAV plans 6 future time steps in advance, but only the first action is

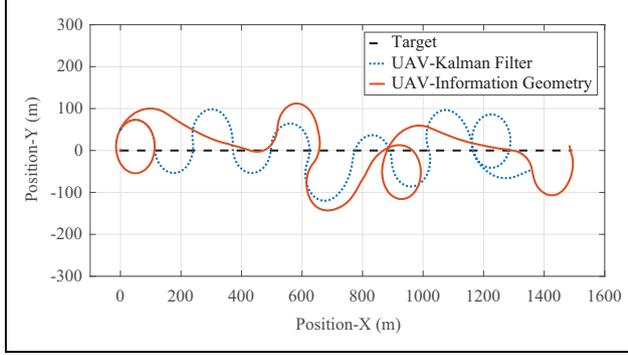


Figure 1. The trajectories of a UAV tracking a rectilinear motion target. UAV: unmanned aerial vehicle.

implemented. In each simulation group, we carry out 200 Monte Carlo simulation experiments. Each experiment performs 100 steps and takes 0.5 s per step. In the simulations, the root-mean-square error (RMSE) between the target and the tracker is employed to measure the location estimation accuracy for different algorithms (our algorithm and a KF-based algorithm; the KF-based algorithm means the performance criterion is based on the trace of the corresponding covariance matrix). Additionally, the computation time is taken into account for comparing the real-time performance of different algorithms.

In the first two groups of simulation experiments, the target is subjected to rectilinear motion, and the average linear velocity of the target is 15 m/s. The initial position of the UAV is (0, 50), and the initial position of the target is (0, 0). Figure 1 shows the trajectories of the UAV for target tracking.

In the first simulation group, the sensor measurement standard deviation parameters are $\sigma_r = 10^{-4}/\text{m}$, $\sigma_\varphi = 5 \times 10^{-3}\pi$ rad. Figure 2(a) and (b) compares the two algorithms in RMSE and computation time, respectively. From Figure 2, we can see that our algorithm has a very close performance compared to the KF-based algorithm in the sense of location RMSE. Nevertheless, our algorithm is much faster than the KF-based algorithm, and the computation time of our algorithm is usually half than that in the KF-based algorithm.

In the second simulation group, the sensor measurement standard deviation parameters are $\sigma_r = 8 \times 10^{-4}/\text{m}$, $\sigma_\varphi = 10^{-2}\pi$ rad. Figure 3 shows the performance of the tracking. From Figure 3(a), we can see that our algorithm gradually has a better RMSE performance; and after the 80th step, our algorithm outperforms the KF-based algorithm. From Figure 3(b), we can see that the computation time for our algorithm is usually half of that in the KF-based algorithm.

For the first two simulation groups, when the sensor measurement noise is larger, our proposed method performs better. Moreover, the decision-making takes only half of the computation time which is required by the

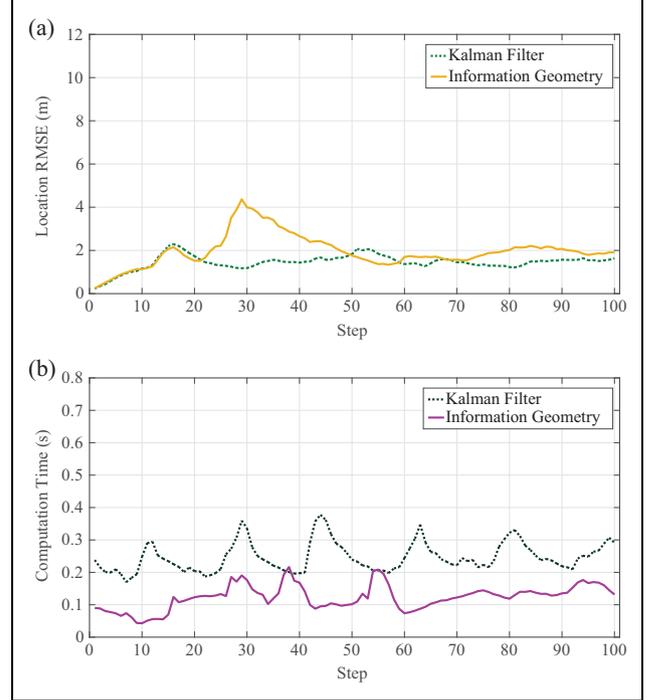


Figure 2. The performance of group I: (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error.

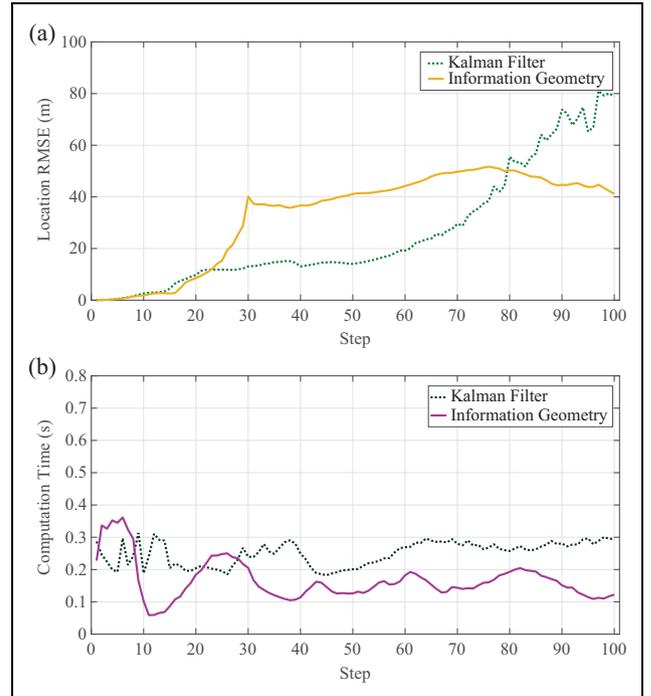


Figure 3. The performance of group II: (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error.

KF-based algorithm, which is significant in the practical application of UAVs.

In the next two simulation groups, the target moves in a circle, and the angular velocity of the target is 0.1 rad/s,

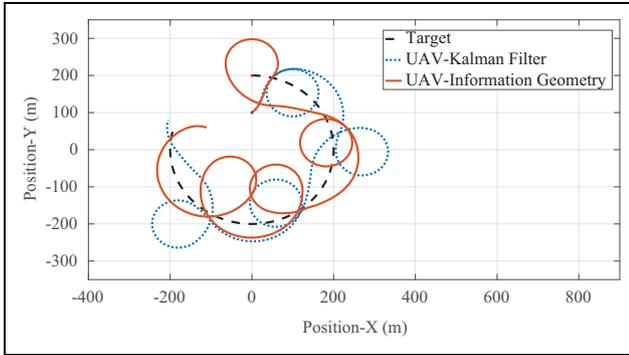


Figure 4. The trajectories of a UAV tracking a circular motion target. UAV: unmanned aerial vehicle.

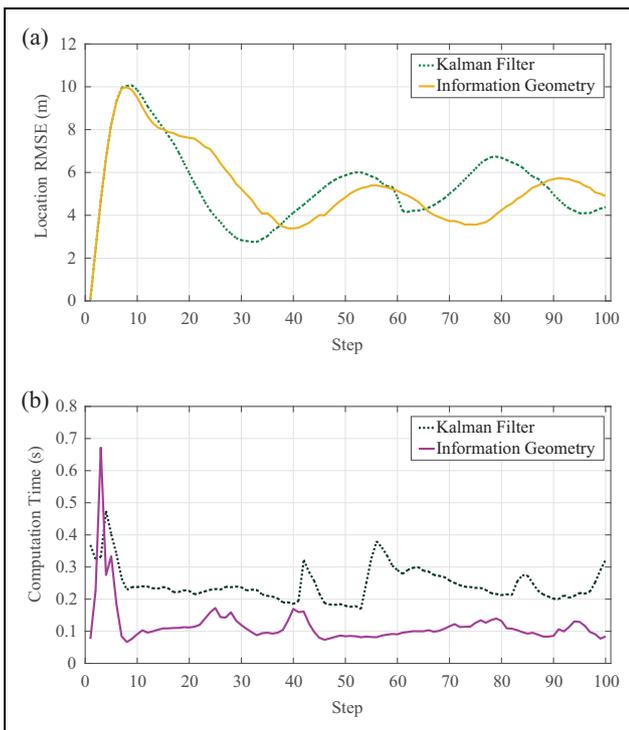


Figure 5. The performance of group III: (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error.

and the radius of motion is 200 m. The UAV's initial position is $(0, 100)$, and the initial position of the target is $(0, 200)$. Figure 4 shows the trajectories of the UAV and the target.

In the third group of simulations, the sensor measurement standard deviation parameters are set to $\sigma_r = 10^{-4}$ m, $\sigma_\varphi = 5 \times 10^{-3}\pi$ rad. The following also illustrates the simulation results of tracking a target moving circularly in Figure 5(a) and (b).

In the fourth group of simulations, the sensor measurement standard deviation parameters in group III are $\sigma_r = 8 \times 10^{-4}$ m, $\sigma_\varphi = 10^{-2}\pi$ rad. Figure 6(a) and (b) compares the two algorithms in RMSE of tracking and average computation time.

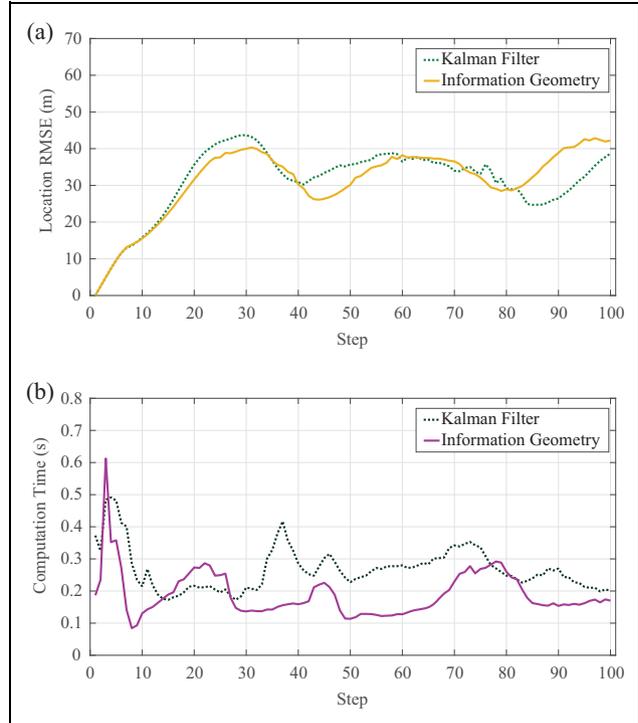


Figure 6. The performance of group IV: (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error.

As it shows in the last two scenarios, our method produces better tracking performance and is more time-efficient in solving.

The simulations indicate that the IG-based algorithm can obtain strategies faster in different sensor measurement precision and target movement forms. As shown in the previous section where the computational complexity of the objective function is analyzed, the decision-making algorithm based on the IG can greatly save the computation time. Besides, when the observation error increases, the tracking accuracy decreases.

In addition, we also present partial numerical results. For each set of experiments, we calculate the average distance error and average computation time. We summate every step distance error over the simulation runtime, and the mean of these errors (from each Monte Carlo run) is called the average tracking error. Similarly, for every step of the simulation runtime, we record the average of computation time, which is called the average computation time. The data statistics are listed in Table 2, and the better data are bolded. Obviously, our algorithm is more time-efficient in solving in each simulation. It is clear that the calculations of inverse matrix in KF cost more time than our method. It is also important to emphasize that the decision time for each step is shorter than the predefined time (0.5 s) per step. That is to say, the algorithm can meet the requirement of real-time calculation. However, in view that IG method takes shorter for decision-making, there is a greater degree of pre-improvement of the performance of

Table 2. Statistical results of average computation time and average tracking error.

Group	Computation Time		Tracking error		Group	Computation time		Tracking error	
	KF	IG	KF	IG		KF	IG	KF	IG
Group I: rectilinear motion, $\sigma_r = 10^{-4}/\text{m}$, $\sigma_\varphi = 5 \times 10^{-3}\pi \text{ rad}$.	0.2663	0.1397	0.9396	2.1303	Group III: circular motion, $\sigma_r = 10^{-4}/\text{m}$, $\sigma_\varphi = 5 \times 10^{-3}\pi \text{ rad}$.	0.2772	0.1421	5.1901	5.8166
	0.2502	0.1383	1.4695	1.5368		0.2536	0.1653	4.7843	5.4185
	0.2461	0.1149	1.5470	1.6527		0.2524	0.1295	5.2428	4.2371
	0.2377	0.1144	1.6872	1.8515		0.2386	0.1312	5.1264	4.8794
	0.2237	0.1175	1.0837	2.5165		0.2471	0.1326	4.5247	5.0728
	0.2258	0.1326	1.3376	1.7106		0.2642	0.1289	5.6331	5.4632
	0.2329	0.1194	1.2946	1.6813		0.2613	0.1171	5.0692	5.4448
	0.2264	0.1119	1.3023	2.0482		0.2497	0.1198	5.1437	4.5239
	0.2314	0.1030	0.9930	1.1756		0.2338	0.0910	5.3624	5.4249
	0.2310	0.1203	1.2539	1.2345		0.2273	0.1090	5.1987	5.2313
Group II: rectilinear motion, $\sigma_r = 8 \times 10^{-4}/\text{m}$, $\sigma_\varphi = 10^{-2}\pi \text{ rad}$.	0.2758	0.1956	12.1068	20.2923	Group IV: circular motion, $\sigma_r = 8 \times 10^{-4}/\text{m}$, $\sigma_\varphi = 10^{-2}\pi \text{ rad}$.	0.2805	0.1775	31.8603	25.7812
	0.2415	0.1720	16.1810	89.6004		0.2595	0.1934	30.2535	25.0373
	0.2483	0.1628	17.9907	24.6986		0.2580	0.1735	22.3440	29.4811
	0.2353	0.1500	18.2345	23.9607		0.2574	0.2027	34.4343	31.0254
	0.2318	0.1740	19.6474	23.1907		0.2360	0.1651	24.2302	40.2282
	0.2367	0.1693	11.5808	29.1510		0.2461	0.1739	26.1126	27.7403
	0.2255	0.1738	8.8416	25.7922		0.2497	0.1649	23.2740	22.4048
	0.2594	0.1649	12.5577	21.2700		0.2426	0.1734	28.6919	25.3280
	0.2889	0.1641	30.6833	28.5152		0.2452	0.1635	30.4491	17.5406
	0.2686	0.1654	19.0282	23.3245		0.2522	0.1677	25.4797	26.2637

KF: Kalman filter; IG: information geometry.

decision-making by the short decision period. At the same time, we notice that the tracking error of the IG-based algorithm does not always perform better than the KF-based algorithm. The main reason is that, in this discrete measurement sampling scenario, the information distance is approximated by the sum of the determinant of FIM, which sometimes reduces the estimation accuracy for some scenarios. However, as shown in Figure 5(a) and (b) and Table 2, the IG-based algorithm performs better for high-order motion system, such as the circular motion target. Nevertheless, the IG-based method is more time-efficient than the KF-based method, which is more suitable for practical applications, especially when the computational resource is limited.

In order to study the effect of different prediction horizons on computation time and tracking performance, we also carry out two groups of experiments. In these two groups, the movement of the target is different and each group including 200 times Monte Carlo simulations at different time horizons. The average computation time and the location RMSE of tracking for these 200 experiments with different time horizons are presented in Figures 7 and 8.

These figures reflect that the computation time reduces and the tracking performance gets worse with the time horizon decreasing. Therefore, when we select the time horizon, it need to make a trade-off between the computation load and the tracking performance.

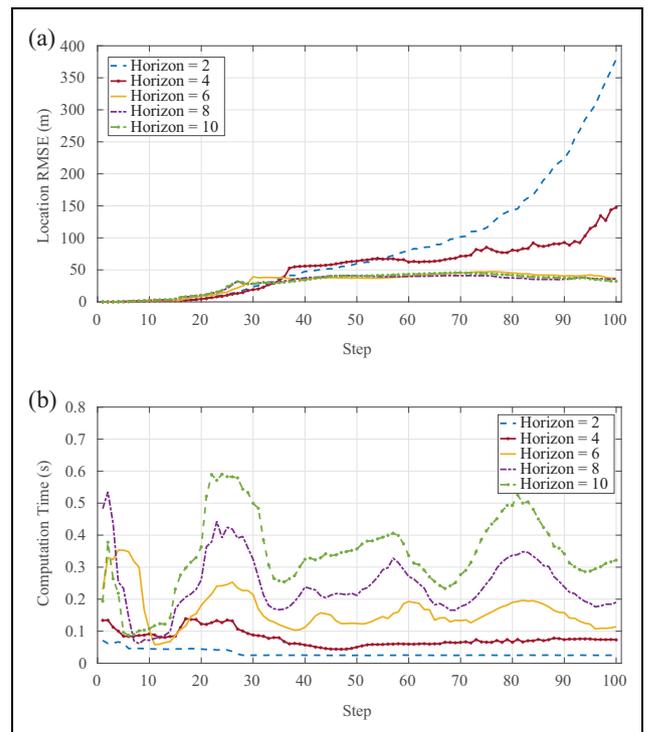


Figure 7. The UAV tracks the target of linear motion with different time horizons, (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error; UAV: unmanned aerial vehicle.

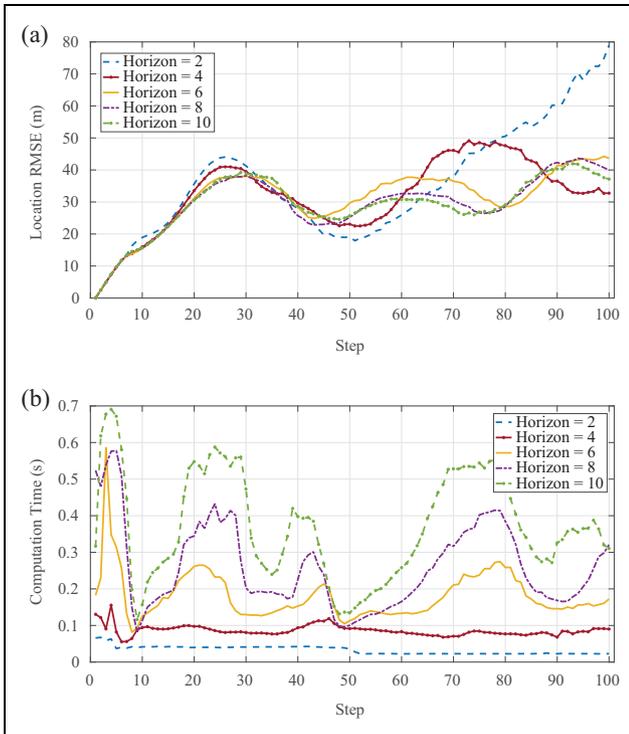


Figure 8. The UAV tracks the target of circular motion with different time horizons (a) the RMSE of tracking and (b) the computation time. RMSE: root-mean-square error; UAV: unmanned aerial vehicle.

Conclusion and future work

In this article, we have studied the moving ground target tracking problem of a fixed-wing UAV. More specifically, a POMDP-based action decision-making method has been proposed, which gives the optimal sequence to maximize the target information observed by the radar. In this method, we have introduced the FIM as the criterion of the proposed method with the aid of IG. Simulation results corroborate the effectiveness of our proposed method; and show that compared to the classical KF-based method, our method has higher time-efficiency in computations. In our future work, we would extend our method to multi-target tracking scenarios.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Xiangke Wang  <http://orcid.org/0000-0003-2666-1360>
Yirui Cong  <http://orcid.org/0000-0003-0621-6331>

Supplemental material

Supplementary material for this article is available online.

References

1. Rahman AH, Malik AS, Kumar JR, et al. Design configuration of a generation next main battle tank for future combat. *Def Sci J* 2017; 67(4): 343.
2. Capitan Fernandez J, Martinez-de Dios JR, Maza I, et al. Ten years of cooperation between mobile robots and sensor networks. *Int J Adv Robot Syst* 2015; 12(6): 70.
3. Ragi S and Chong EK. UAV path planning in a dynamic environment via partially observable Markov decision process. *IEEE T Aero Elec Sys* 2013; 49(4): 2397–2412.
4. Ragi S and Chong EK. UAV guidance algorithms via partially observable Markov decision processes. In: Ragi S and Chong EK (eds) *Handbook of unmanned aerial vehicles*. Berlin: Springer, 2015, pp. 1775–1810.
5. Wang X, Morelande M, and Moran B. Target motion analysis using single sensor bearings-only measurements. In: *2nd international congress on image and signal processing (CISP'09)* (eds P Qiu, C Yiu, H Zhang and X Wen), Tianjin, China, 17–19 October 2009, pp. 2094–2099. IEEE.
6. Allotta B, Caiti A, Costanzi R, et al. A new AUV navigation system exploiting unscented Kalman filter. *Ocean Engineering* 2016; 113: 121–132.
7. Hausman K, Müller J, Hariharan A, et al. Cooperative multi-robot control for target tracking with onboard sensing. *Int J Robot Res* 2015; 34(13): 1660–1677.
8. Tang Z and Ozguner U. PF-HMap: a target track maintenance approach for mobile sensor platforms with intermittent and regional measurements. In: *45th conference on decision and control* (ed T Parisini), San Diego, USA, 13–15 December 2016, pp. 6757–6762. IEEE.
9. Teuliere C, Eck L, and Marchand E. Chasing a moving target from a flying UAV. In: *IEEE/RSJ international conference on intelligent robots and systems (IROS)* (ed N Amato), San Francisco, USA, 25–30 September 2011, pp. 4929–4934. IEEE.
10. Lee DY, Shim SW, Hwang MC, et al. Target tracking using adaptive coarse-to-fine particle filter. In: *AIAA Guidance, navigation, and control conference*, Grapevine, Texas, 9–13 January 2017, p. 1245. American institute for aeronautics and astronautics (AIAA).
11. Wang J, Zhang Y, Lu J, et al. A framework for moving target detection, recognition and tracking in UAV videos. *ACII* 2012; 137: 69–76.
12. Li D, Ji D, Liu J, et al. A multi-model EKF integrated navigation algorithm for deep water AUV. *Int J Adv Robot Syst* 2016; 13(1): 3.
13. Chen H, Chang K, and Agate CS. UAV path planning with tangent-plus-Lyapunov vector field guidance and obstacle avoidance. *IEEE T Aero Elec Sys* 2013; 49(2): 840–856.
14. Steven A, Quintero P, Ludkovski M, et al. Stochastic optimal coordination of small UAVs for target tracking using

- regression-based dynamic programming. *J Intell Robot Syst* 2016; 82(1): 135.
15. Fu X, Feng H, and Gao X. UAV mobile ground target pursuit algorithm. *J Intell Robot Syst* 2012; 68(3–4): 359–371.
 16. Kochenderfer MJ. *Decision making under uncertainty: theory and application*. Cambridge: MIT Press, 2015.
 17. Galceran E, Cunningham AG, Eustice RM, et al. Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction: theory and experiment. *Autonomous Robots* 2017; 41(6): 1367–1382.
 18. Prentice S and Roy N. The belief roadmap: efficient planning in linear pomdps by factoring the covariance. In: Kaneko M and Nakamura Y (eds) *Robotics research*. Heidelberg: Springer, 2010, pp. 293–305.
 19. Papadimitriou CH and Tsitsiklis JN. The complexity of Markov decision processes. *Math Oper Res* 1987; 12(3): 441–450.
 20. Poupart P, Malhotra A, Pei P, et al. Approximate linear programming for constrained partially observable Markov decision processes. In: *The Twenty-Ninth AAAI conference on artificial intelligence*, Austin, Texas, USA, 25–30 January 2015, pp. 3342–3348. AAAI Press.
 21. Sunberg Z, Chakravorty S, and Erwin RS. Information space receding horizon control for multisensor tasking problems. *IEEE T Cybernetics* 2016; 46(6): 1325–1336.
 22. Wang P, Ma L, and Xue K. Multitarget tracking in sensor networks via efficient information-theoretic sensor selection. *Int J Adv Robot Sys* 2017; 14(5): 1–9.
 23. Rao CR. Information and the accuracy attainable in the estimation of statistical parameters. In: *Breakthroughs in statistics* (eds S Kotz and NL Johnson), Berlin: Springer, 1992, pp. 235–247.
 24. Chentsov NN. Statistical decision rules and optimal inferences. *Transactions of Mathematics Monograph* 1982; 53.
 25. Costa SI, Santos SA, and Strapasson JE. Fisher information distance: a geometrical reading. *Discrete Appl Math* 2015; 197: 59–69.
 26. Opitz F. Information geometry and tracking. In: *14th International radar symposium (IRS)* (ed H Rohling), IEEE, Dresden, Germany, 19–21 June 2013, pp. 313–318.
 27. Akselrod D, Goldman C, Sinha A, et al. Collaborative sensor management for multitarget tracking using decentralized markov decision processes. In: *Proceedings of SPIE* (ed OE Drummond), Orlando, US, 17–21 April 2006, Vol. 6236, 1–12. SPIE Publications.
 28. Pilté M and Barbaresco F. Tracking quality monitoring based on information geometry and geodesic shooting. In: *17th international radar symposium (IRS)* (eds A Kurowska and J Misiurewicz), Krakow, Poland, 10–12 May 2016, pp. 1–6. IEEE.
 29. Wang X, Cheng Y, Morelande M, et al. Bearings-only sensor trajectory scheduling using accumulative information. In: *12th international radar symposium (IRS)* (ed H Rohling), Leipzig, Germany, 7–9 September 2011, pp. 682–688. IEEE.
 30. Zhao Y, Wang X, Kong W, et al. Decision-making of UAV for tracking moving target via information geometry. In: *35th Chinese control conference (CCC)*, IEEE, Chengdu, China, 27–29 July 2016, pp. 5611–5617.
 31. Chong EK, Kreucher CM, and Hero AO. Partially observable Markov decision process approximations for adaptive sensing. *Discrete Event Dyn S* 2009; 19(3): 377–422.
 32. Wang X, Cheng Y, and Moran B. Bearings-only tracking analysis via information geometry. In: *13th Conference on Information fusion (FUSION)*, Edinburg, United Kingdom, 26–29 July 2010, pp. 1–6. IEEE.
 33. Hernández-Lerma O. *Adaptive Markov control processes*. Vol. 79. Berlin: Springer Science & Business Media, 2012.