Proceedings of the 2017 IEEE
International Conference on Robotics and Biomimetics
December 5-8, 2017, Macau SAR, China

A Hierarchical Decision-making Algorithm of UAV for Collision Avoidance in Target Tracking

Yunyun Zhao, Xiangke Wang, Lincheng Shen
College of Mechatronic Engineering and Automation
National University of Defense Technology (NUDT)
Hunan, China
Email: xkwang@nudt.edu.cn

Weiwei Kong
Naval Academy of Armament
Beijing, China
Email: kww_ kww@163.com

***Abstract*—Effective target tracking and obstacle avoidance strategies are essential to the success of unmanned aerial vehicle (UAV) missions. In this paper, we consider the UAV action decision-making problem for tracking a ground moving target, while avoiding dynamic ellipsoidal obstacle detected en route. The studied problem was formulated as a Partially Observable Markov Decision Process (POMDP), and then in this framework, we get the optimal or sub-optimal strategy to achieve the target tracking in the barrier environment. We first improved the conditions and methods of collision detection based on collision cone, and use common penalty function method to ensure tracking safety. Further taking into account that the UAV needs to react more quickly to obstacles than target tracking, we propose an alternative hierarchical decision-making approach. For the entire tracking process, the total information gain in the UAV-to-target observation is used as a tracking optimization criterion in case of no obstacle. Simulations are presented to demonstrate the comparison between the proposed approach with the improved penalty function approach.The results show that our proposed algorithm is superior.**

I. Introduction

Unmanned aerial vehicle (UAV) has been extensive used in military applications as well as in civilian. In target tracking, the observability of target is achieved by adequate maneuvers of the sensor mounted on the UAV. Since the sensing capabilities of UAVs are constrained by the UAV dynamics and the onboard sensor performances, one of the major research challenges is UAV action decision-making. At the same time in the complex environment, there may be sudden threats or obstacles during target tracking, then flight space appears no-fly zone. In this paper, we collectively referred to as the obstacle area. It is necessary to consider the obstacle avoidance in the process of target tracking. In the following, we introduce the research foundation of the related papers from two aspects: target tracking and obstacle avoidance.

In the recent past, partially observable Markov decision process (POMDP) has been used to design decision-making algorithms for UAVs in tracking problems, as the dynamic of the target is assumed to follow a Markov process and the UAV state and dynamics are known and can be controlled. The approach in [1, 2] uses POMDP incorporating the dynamic constraints on UAV motion for designing a UAV guidance algorithm for target tracking, but it differs from the method in this study in the cost criterion. In [4], in order to minimize the location uncertainty of a mobile target, the author proposes an offline policy iteration algorithm to find an optimal UAV path in a coarse discretized state space, followed by an online policy iteration algorithm that applies a finer grid MDP to the region of interest to find the final optimal policy for fixed-wing UAV. The composed method can improve the real-time performance.

For such tracking target missions in an obstacle environment, however, sense-and-avoid capability (i.e., the ability to detect potential obstacles and avoid them) is a critical requirement. Collision-avoidance algorithms are usually local, fast, reactive, and carried out online, ensure the safety of the vehicle from unexpected and unforeseen obstacles/collisions. Many techniques for collision avoidance have been proposed in the recent literature, such as potential field approaches [5], optimal approaches [6, 7], Lyapunov Guidance Vector Field method [8, 9], graph search algorithms [10], etc. Among the collision-avoidance algorithms, geometric approaches are effective in the midair collision situations because these algorithms can deal with an obstacle by using the geometry of relative kinematics between the vehicle and the obstacle. The collision cone [11, 12], velocity obstacles [13], and geometric optimization approaches [14] are representative of geometric approaches and solve the situations reactively by updating the current states.

In the study, the decision system includes target tracking and obstacle avoiding. Due to the different avoidance strategy, we come up with two decision-making algorithm to solve this tracking target while avoiding obstacle problem. For the target tracking part, we use a uniform method, which define the criterion function based on Fisher information to decrease the computation cost and the location uncertainty. According to the relevant knowledge of information geometry, we derive the Fisher information matrix (FIM) through the model of radar detection. When the system detects the obstacles, we propose and compare two collision avoidance algorithm. Firstly, we improved the condition and method of collision detection for traditional penalty function method. Then, on this basis, we proposed a more rapid and effective obstacle avoidance strategy. In this approach, we use the geometric policy, similar to the geometric guidance laws, to move away from the threat of obstacles as soon as possible, without considering the tracking effect. The following is the structure of this paper.

In the next section, we describe the problem of UAV decision-making for target tracking via a track-while-scan (TWS) radar system. The UAV action selection is driven by POMDP formulation and accumulative information, where we use the distribution of the predicted target state to figure up the

978-1-5386-3742-5/17/$31.00 © 2017 IEEE

accumulative information. The obstacle avoidance approach, including the improved penalty function method and the proposed approach, is presented in Section III. And Section IV discussed the simulations and results, which is followed by the conclusions in Section V.

II. The Tracking Decision-making

We use POMDP to find the optimal policy of a UAV based on the current states of the UAV and the target. It is to select a sequence of actions for the UAV that maximize the total information gain to reduce the uncertainty of target localization.

We assumed that the moving targets on the ground follows the two-dimensional motion. The motion of the UAV is a simplified kinematic model [2]. The 2-D position coordinate of the UAV is varied by applying the controls forward acceleration (which controls the speed) and bank angle (which controls the heading). The values of these control variables are restricted to lie within certain minimum and maximum limits. The UAV is mounted with an onboard TWS radar that generates the position measurements of the target. These measurements are corrupted by random errors that are spatially varying, i.e., we assume the measurement error covariance of a target is proportional to the distance from the UAV to the target, depending on the locations of UAV and the target [2].

A. The Formulation of the System based on POMDP

An infinite horizon POMDP is defined by a tuple $M = \langle \boldsymbol{S}, \boldsymbol{U}, T, \boldsymbol{\Omega}, O, R, b_0 \rangle$. The followings define the key components of POMDP with special respect to UAV tracking a target.

States. The state space $\boldsymbol{S}$ at time k includes the UAV state $\boldsymbol{s}_k^u = (x_k^u, y_k^u, v_k^u, \theta_k^u)^{\mathrm{T}}$, the target state $\boldsymbol{s}_k^t = (x_k^t, y_k^t, \dot{x}_k^t, \dot{y}_k^t)^{\mathrm{T}}$, and the relative target state to the UAV $\boldsymbol{s}_k^o = (x_k, y_k)$ can be calculated.

Actions. The action at time k is the control command of the UAV, including the forward accelerations and the bank angles, given by $\boldsymbol{u}_k = [a_k, \phi_k]^{\mathrm{T}}$.

Belief state. The state distribution at time $k = 0$ is called the belief state b_k. Since the UAV states are fully observable, the belief state is $b_k(\boldsymbol{s}^u) = \delta(\boldsymbol{s}^u - \boldsymbol{s}_k^u)$. The probability distribution of target position is assumed to be Gaussian distribution, which can be expressed approximately as $b_k(\boldsymbol{s}^t) \sim \mathcal{N}(\boldsymbol{\mu}_k, \mathbf{P}_k)$, where $\boldsymbol{\mu}_k$ and $\mathbf{P}_k$ is derived from the standard Kalman filter [15]. In our system, if an optimal choice of action decisions exists, then there exists an optimal sequence of actions that depend only on the belief-state feedback [16].

Reward function. The real-valued reward function R defines the value to the agent actions.

State-transition law. The transfer function T of the system consists of the following three parts. The UAV state transition function is the kinematic equation (18). The target state is unknown and approximated by the constant velocity model (2).

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

where $[v]_{V_{\min}}^{V_{\max}} = \max\{V_{\min}, \min(V_{\max}, v)\}$, $V_{\min}$ and $V_{\max}$ are the minimum and the maximum limits on the ground speed of the UAV. g is the acceleration due to gravity, and T is the length of the time step.

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

where $\mathbf{A}$ is the state transition matrix, the specific form of which is:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

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

Observations and observation law. $\boldsymbol{\Omega}$ is the set of possible observation, and O is the set of observation function that describes what is observed when an action is performed. The CS is polar with range r, bearing φ. The noisy measurements of the target measured at time k satisfy

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

For TWS radar, the error-free target position in the sensor polar CS can be represented as

$$\begin{aligned} \boldsymbol{\eta}(\boldsymbol{s}_k^o) &= \begin{bmatrix} \eta_1(\boldsymbol{s}_k^o) \\ \eta_2(\boldsymbol{s}_k^o) \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan\left(y_k/x_k\right) \end{bmatrix} \end{aligned} \tag{5}$$

We assume the measurement error covariance $\mathbf{C}(\boldsymbol{s}_k^o)$ of a target is proportional to the distance from the UAV to the target [2]. The specific form is

$$\mathbf{C}(\boldsymbol{s}_k^o) = \begin{bmatrix} \hat{r}_k^2\sigma_r^2 & 0 \\ 0 & \hat{r}_k^2\sigma_\varphi^2 \end{bmatrix} \tag{6}$$

where $\hat{r}_k\sigma_r$ and $\hat{r}_k\sigma_\varphi$ are the standard deviations of the range and bearing, respectively. Apparently, $\hat{r}_k$ is the range between the UAV and the target in the actual 2-dimensional space, which is equal to the measured range :

$$\hat{r}_k = r_k.$$

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

$$J_{\infty}^{\pi}(\boldsymbol{s}) := \liminf_{H\to\infty} \frac{1}{H} E\left[\sum_{k=0}^{H-1} R(\boldsymbol{s}_k, \boldsymbol{u}_k) \,|\, \boldsymbol{s}_0 = \boldsymbol{s}\right], \quad (7)$$

where H is the time horizon.

B. Tracking Criterion Function: Cumulative Information

Information geometry [17] provides comprehensive results about statistical models simply by considering them as geometrical objects and the statistical structures as geometrical structures. The amount of information which the sensor may acquired from the target at a particular time is characterized by the information distance. In a discrete measurement sampling scenario, the information distance is given by the sum of the FIM determinants. In this paper, we address the POMDP optimization problem using the criterion of Fisher information distance on a statistical manifold.

Fisher information matrix at time k is defined [17, 18] by

$$\mathbf{G}_k = E\Big[(\nabla_x \ln p(\mathbf{z}_{1:k}\,|\mathbf{x}_k)\,(\nabla_x \ln p(\mathbf{z}_{1:k}\,|\mathbf{x}_k))^T\Big] \quad (8)$$

where $p(\mathbf{z}_{1:k}\,|\mathbf{x}_k)$ is the batch measurement likelihood defined as formula (13),

where $\boldsymbol{h}(\mathbf{x}_k, i)$ denotes the measurement model (5) evaluated at time i in terms of the state $\mathbf{x}_k$ at k. We can derive its derivative (14).

Then, the Fisher information matrix can be calculated in a recursive form as (9).

$$\mathbf{G}_k = \mathbf{G}_{k-1} + \left(\frac{4}{\hat{r}_k^2} + \frac{1}{\hat{r}_k^2\sigma_r^2}\right)\Delta\mathbf{G}_{rk} + \frac{1}{\hat{r}_k^2\sigma_\varphi^2}\Delta\mathbf{G}_{\varphi k} \quad (9)$$

where

$$\Delta\mathbf{G}_{rk} = \begin{bmatrix} \frac{{x_k}^2}{\hat{r}_k^2} & \frac{x_k y_k}{\hat{r}_k^2} \\ \frac{x_k y_k}{\hat{r}_k^2} & \frac{{y_k}^2}{\hat{r}_k^2} \end{bmatrix} \quad (10)$$

$$\Delta\mathbf{G}_{\varphi k} = \begin{bmatrix} \frac{{y_k}^2}{\hat{r}_k^4} & -\frac{x_k y_k}{\hat{r}_k^4} \\ -\frac{x_k y_k}{\hat{r}_k^4} & \frac{{x_k}^2}{\hat{r}_k^4} \end{bmatrix} \quad (11)$$

The sum of determinants of FIM [19] at H sampling locations will be used to approximate the accumulative information along the UAV trajectory $(\boldsymbol{x}_1^o, \boldsymbol{x}_2^o, ..., \boldsymbol{x}_H^o)$, i.e.,

$$f_{tracking} \approx \sum_{i=1}^{H} |\mathbf{G}_i| \quad (12)$$

Our decision-making process is to find the optimal sequence of actions which maximizes the amount of cumulative information. We use the approximate rolling horizon procedure to obtain the optimal action sequence.

III. The Overall Flow of Reactive Collision Avoidance with Static Obstacles

In the common operational scene, there are obstacles or constraints, such as danger zone, no-flight zone, to restrict the UAV flight path. We assume that the UAV is capable of receiving obstacle information at any decision-making cycle by its own sensors. Without loss of generality obstacles are modeled as circles determined by their center position and radius. More specifically, during a decision-making cycle, assume an obstacle circle (no-flight zones) appear unexpectedly. These obstacle circle parameters are assumed to be available, i.e., center coordinates (x^O, y^O), radius of no-fly zone r_{no-fly}^O, and radius of danger zone r_{danger}^O. The UAV can not stay in a dangerous area for a long time and can never enter the no-fly zone. The detection range of the UAV, r^D, is greater than the threat radius of the obstacle.

In the event of obstacle avoidance, the UAV needs to determine when to take avoidance action. The collision cone approach is an effective tool for detecting a collision. The construction of the collision cone is shown in Figure. 1

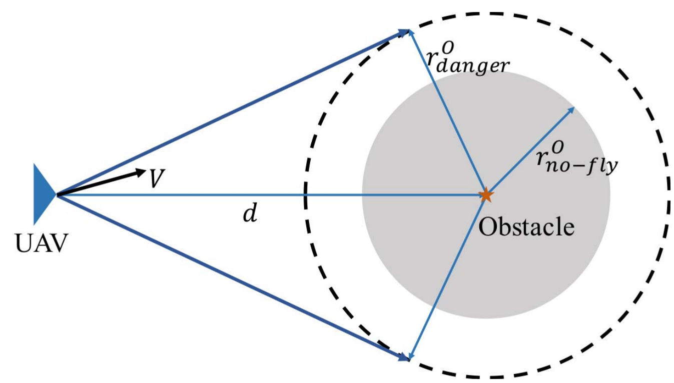

Figure 1: Construction and analysis of the collision cone

We define d as relative distance between the UAV and the obstacle, and V as the velocity of the UAV. A collision cone is constructed by dropping tangents from the UAV to the danger circle. If the velocity vector V lies within the collision cone, the UAV will violate danger zone, even no-fly zone in due course and result in collision.

A. Penalty Function Method for Collision Avoidance

This section shows how to avoid collisions between UAV and obstacle in the same context of the POMDP framework. To avoid collisions, a penalty term, that is increases as the UAV come close to obstacle, is included in the objective function. The new objective function is written as follows:

$$f_H = f_{tracking} - \alpha f_{obstacle} \quad (15)$$

where α is the weight of the penalty function, and $f_{obstacle}$ is a collision penalty function corresponding to the UAV. In our system, the obstacles are modeled as circular regions in the plane. Any continuous function can be used as the collision penalty function provided that the function increases as the UAV approaches the obstacle. The penalty function is given below:

$$f_{obstacle} = \begin{cases} C_1 - (d - r_{danger}^O)^6, & V \text{ in cone } \& \ d > r_{danger}^O \\ C_2, & d \le r_{danger}^O \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

$$p\left(\mathbf{z}_{1:k}\left|\mathbf{x}_k\right.\right)=p\left(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_k\left|\mathbf{x}_k\right.\right)=\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi|\mathbf{C}(\mathbf{x}_k,i)|}}\exp\left(-\frac{1}{2}(\mathbf{z}_i-\boldsymbol{h}(\mathbf{x}_k,i))^T\mathbf{C}^{-1}(\mathbf{x}_k,i)(\mathbf{z}_i-\boldsymbol{h}(\mathbf{x}_k,i))\right) \tag{13}$$

$$\begin{aligned}\nabla_x \ln p\left(\mathbf{z}_{1:k}\left|\mathbf{x}_k\right.\right)=\sum_{i=1}^{k}\Bigg(&-\frac{2}{\hat{r}_i}\nabla_x\hat{r}_i+\frac{(r_i-h_1(\mathbf{x}_k,i))^2}{\hat{r}_i^3\sigma_r^2}\nabla_x\hat{r}_i-\frac{(r_i-h_1(\mathbf{x}_k,i))}{\hat{r}_i^2\sigma_r^2}\nabla_x h_1(\mathbf{x}_k,i)\\&+\frac{(\varphi_i-h_2(\mathbf{x}_k,i))^2}{\hat{r}_i^3\sigma_\varphi^2}\nabla_x\hat{r}_i-\frac{(\varphi_i-h_2(\mathbf{x}_k,i))}{\hat{r}_i^2\sigma_\varphi^2}\nabla_x h_2(\mathbf{x}_k,i)\Bigg)\end{aligned} \tag{14}$$

where C_1 and C_2 are constants, and C_2 is sufficiently large. In the first part, the sixth power is used to match the objective function of the tracking.

In this method, the approximate rolling horizon method is still used to solve the optimal action when the obstacle is found. The difference is that the objective function of the optimal programming adds a penalty item.

B. Reactive Hierarchical Decision-making for Collision Avoidance

As long as the UAV detects an obstacle, it is necessary to respond quickly to the it to avoid collision or approaching. The algorithm mentioned above is a commonly used method, that is, by adding a penalty item to the objective function to change the action strategy. We propose a reactive hierarchy decision-making approach for collision avoidance. Next, the algorithm is described in detail.

Assuming that the target is detected, the state of the UAV is $\boldsymbol{s}_0^u$. The kinematic model of the UAV shows that the current action (a_0,ϕ_0) will affect the state of the next moment $\boldsymbol{s}_1^u$, thus further changing the position of the UAV (x_2^u,y_2^u). That is:

$$\left\{\begin{array}{l} x_2^u=x_0^u+v_0^uT\cos\theta_0^u+v_1^uT\cos\theta_1^u\\ y_2^u=y_0^u+v_0^uT\sin\theta_0^u+v_1^uT\sin\theta_1^u\end{array}\right. \tag{17}$$

where

$$\left\{\begin{array}{l} v_1^u=v_0^u+a_0T\\ \theta_1^u=\theta_0^u+\left(gT\tan(\phi_0)/\nu_k^u\right)\end{array}\right. \tag{18}$$

Taking into account the decision cycle is small enough, in the next movement cycle, we only consider the collision avoidance regardless of the tracking strategy. Once the UAV no longer meets the collision condition, the system recovers the target tracking. Specifically, when the collision is predicted, the objective function of the decision-making becomes:

$$f_{H=1}=dis\left((x_2^u,y_2^u)-(x^O,y^O)\right) \tag{19}$$

where dis represent the distance between two points. Obviously in order to avoid obstacles, it sacrifices a certain tracking accuracy. However, the proposed approach effectively improve the efficiency of the calculation and the effectiveness of obstacle avoidance, and the physical meaning is clear, as result, we do not have to consider the problem of parameter selection. these above are very meaningful in engineering applications.

C. The Overall Flow of the Algorithm

Before simulation verification, we describe this target tracking while obstacle avoidance system in the form of pseudo-code. In the algorithmic process, we illustrate the important algorithmic steps and highlight our works. The specific algorithm process is as follows:

Algorithm 1: Target Tracking while Avoidance Algorithm

Input : Radar-to-target measurement sequence $\boldsymbol{o}_1,\boldsymbol{o}_2,\ldots$,
UAV state sequence $\boldsymbol{s}_1^u,\boldsymbol{s}_2^u,\ldots$, obstacle state $\boldsymbol{s}^O$

Initialization ($\boldsymbol{o}_1$, $b_1(\boldsymbol{s}^t)$);
for $k=2$ **to** $Termination$ **do**
  $b_k(\boldsymbol{s}^t)$ $\leftarrow$ KalmanFilter ($\boldsymbol{o}_k,\boldsymbol{s}_k^u$);
  Penalty Function Algorithm:
  $f\leftarrow 0$;
  for $i=1$ **to** $Horizon$ **do**
    $\alpha\leftarrow 0$;
    if $\boldsymbol{s}_i^u\in$ CollisionCone ($\boldsymbol{s}_i^u,\boldsymbol{s}^O$) **then**
      $\alpha\leftarrow C_1$
    end
    $f\leftarrow f+$ FisherInfor($b_{i+1}(\boldsymbol{s}^t),\boldsymbol{s}_{i+1}^u$) +
    αPenaltyFun($\boldsymbol{s}_{i+1}^u,\boldsymbol{s}^O$);
  end
  Penalty Function Algorithm End;
  or
  Hierarchical Decision-making Algorithm:
  $f\leftarrow 0$;
  if $\boldsymbol{s}_i^u\in$ CollisionCone ($\boldsymbol{s}_i^u,\boldsymbol{s}^O$) **then**
    $f\leftarrow$Avoidance($\boldsymbol{s}_{i+2}^u,\boldsymbol{s}^O$);
    break ;
  end
  for $i=1$ **to** $Horizon$ **do**
    $f\leftarrow f+$ FisherInfor($b_{i+1}(\boldsymbol{s}^t),\boldsymbol{s}_{i+1}^u$);
  end
  Hierarchical Decision-making Algorithm End;
  ObjectiveFun $\leftarrow -f$;
  $\boldsymbol{u}_k$ $\leftarrow$ Fmincon(@ObjectiveFun);
end

Output: UAV action sequence $\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots$

IV. Simulation Results

This section shows the simulation results to demonstrate the ability of UAV avoid collision and track a target with a TWS radar using the proposed algorithm, comparing the simulation results. We compare of above two method from the aspect of tracking error, computation time cost and avoidance performance.

Table I: The results of simulation experiments

No.	Distance Error (m)		Computation Time (s)		Waypoints in Danger Zone		Waypoints in No-fly Zone	
	Penalty	Hierarchy	Penalty	Hierarchy	Penalty	Hierarchy	Penalty	Hierarchy
1	**2.4756**	2.5451	0.1421	**0.1275**	3	**1**	0	0
2	3.0079	**1.8925**	0.1842	**0.1711**	9	**2**	0	0
3	2.2992	**2.7976**	0.1505	**0.1431**	5	**1**	0	0
4	2.8602	**1.7907**	**0.1676**	0.1926	3	**1**	0	0
5	2.9607	**2.0458**	0.1750	**0.1684**	5	**1**	2	**0**
6	**1.9694**	2.9737	0.1965	**0.1328**	7	**2**	1	**0**
7	**2.1411**	2.7327	0.2034	**0.1792**	5	**2**	0	0
8	2.9469	**1.9726**	0.1701	**0.1470**	6	**3**	1	**0**
9	2.8868	**2.0443**	0.1718	**0.1552**	1	1	0	0
10	2.2992	**2.2748**	**0.1513**	0.1868	5	**1**	0	0
11	**1.6813**	4.0913	0.2170	**0.1445**	5	**0**	0	0
12	2.1869	**1.8302**	0.1687	**0.1644**	5	**1**	0	0
13	**2.3572**	2.4059	0.1599	**0.1252**	5	**2**	0	0
14	**1.9004**	2.7117	**0.1548**	0.2069	3	**0**	0	0
15	3.4115	**3.3117**	0.1981	**0.1635**	3	**2**	0	0
16	**2.3425**	3.2016	**0.1536**	0.1576	2	2	0	0
17	**1.8430**	3.3157	0.1897	**0.1843**	5	**1**	0	0
18	**1.4473**	2.8636	0.1928	**0.1586**	4	**2**	0	0
19	**2.4219**	2.8839	**0.1361**	0.1553	4	**0**	0	0
20	2.8820	**2.1467**	0.1773	**0.1464**	6	**0**	0	0

The experiments are carried out at the target moving speed of $15m/s$, and the average tracking error, the average calculation time, and the number of waypoints in the danger zone are calculated in detail. Figure 2, Figure 3 and Figure 4 are the simulation result.

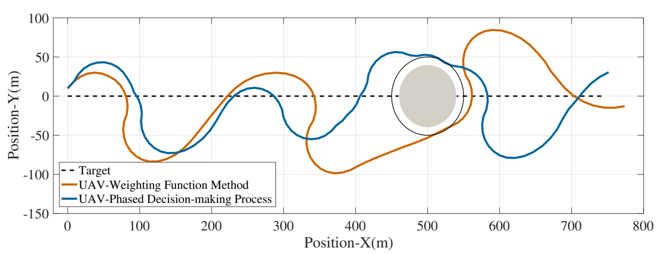

Figure 2: Trajectory of two reactive obstacle avoidance algorithm ($V_T = 15\ m/s$)

Statistical results of the 20 experiments are in the Table I. We summate these distance errors over the simulation runtime, and the mean of these errors (from each Monte Carlo run) is called the average target-location error. The average computation time is computed as follows. At every step of the simulation runtime, we record the time costs of researching the optimal strategy, and the mean of these costs (from each Monte Carlo run) is called the average computation cost. We

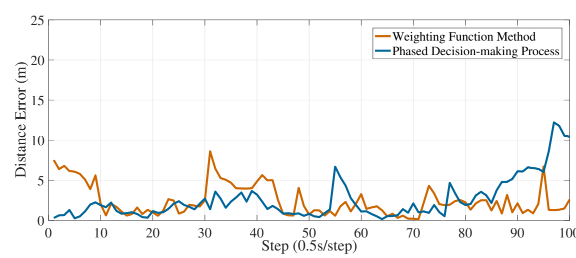

Figure 3: Distance tracking error of two avoidance algorithm ($V_T = 15\ m/s$)

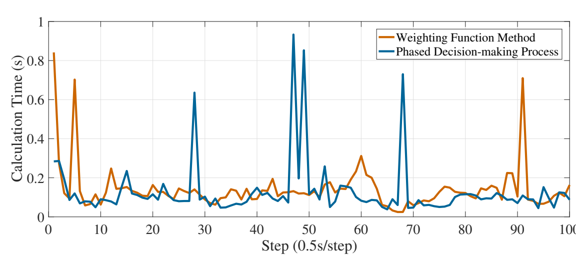

Figure 4: Computation time of two avoidance algorithm ($V_T = 15\ m/s$)

put the better indicator in bold.

From the statistical results in the Table I, it can be seen that the proposed method is generally superior to the penalty function method. Especially in the danger zone or even the no-fly zone, the waypoints is always be less than the penalty function method. The decision result of the penalty function method may even cause the following dangerous situation, shown in Figure 5. The sequence of actions obtained by the penalty function method keeps the UAV in the danger zone for a long time, even close to the no-fly zone, but our approach does not come about such a dangerous flight path. Therefore, the decision results of our proposed method are more conducive to safe flight. Besides, as long as the frequency of decision is small enough, the tracking effect is not affected by obstacle avoidance.

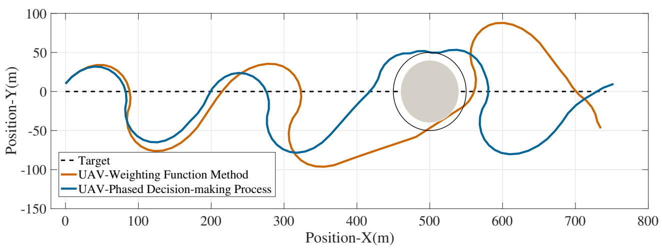

Figure 5: Trajectory of two reactive obstacle avoidance algorithm ($V_T = 15\ m/s$)

V. Conclusion

In this paper, we have shown a new UAV action decision-making method to avoid collision while tracking a ground moving target. Our method uses the theory of POMDP select an optimal sequence of UAV control actions which can maximize accumulative information of the target from the TWS radar. When the obstacle is detected, we improve the penalty function method and proposed the hierarchy decision-making approach for collision avoidance. The simulation results show that our proposed method has a better avoidance performance and less computation time cost.

Acknowledgment

This work is supported by a National Natural Science Foundation (NNSF) of China under Grant 61403406.

References

[1] S. A. Miller, Z. A. Harris, and E. K. Chong, "A pomdp framework for coordinated guidance of autonomous uavs for multitarget tracking," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–17, 2009.

[2] S. Ragi and E. K. Chong, "Uav guidance algorithms via partially observable markov decision processes," in *Handbook of Unmanned Aerial Vehicles*. Springer, 2015, pp. 1775–1810.

[3] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of markov decision processes," *Mathematics of operations research*, vol. 12, no. 3, pp. 441–450, 1987.

[4] S. S. Baek, H. Kwon, J. A. Yoder, and D. Pack, "Optimal path planning of a target-following fixed-wing uav using sequential decision processes," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 2955–2962.

[5] H. Rezaee and F. Abdollahi, "Adaptive artificial potential field approach for obstacle avoidance of unmanned aircrafts," in *Advanced Intelligent Mechatronics (AIM), 2012 IEEE/ASME International Conference on*. IEEE, 2012, pp. 1–6.

[6] C. G. Prévost, A. Desbiens, E. Gagnon, and D. Hodouin, "Uav optimal obstacle avoidance while respecting target arrival specifications," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 11 815–11 820, 2011.

[7] D. Lee, H. Lim, and H. J. Kim, "Obstacle avoidance using image-based visual servoing integrated with nonlinear model predictive control," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 5689–5694.

[8] P. Yao, H. Wang, and Z. Su, "Real-time path planning of unmanned aerial vehicle for target tracking and obstacle avoidance in complex dynamic environment," *Aerospace Science and Technology*, vol. 47, pp. 269–279, 2015.

[9] H. Chen, K. Chang, and C. S. Agate, "Uav path planning with tangent-plus-lyapunov vector field guidance and obstacle avoidance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 840–856, 2013.

[10] S. Hrabar, "Reactive obstacle avoidance for rotorcraft uavs," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 4967–4974.

[11] A. Mujumdar, R. Padhi *et al.*, "Reactive collision avoidance using nonlinear geometric and differential geometric guidance," *Journal of guidance, control, and dynamics*, vol. 34, no. 1, p. 303, 2011.

[12] H. Choi, Y. Kim, and I. Hwang, "Reactive collision avoidance of unmanned aerial vehicles using a single vision sensor," *Journal of Guidance, Control, and Dynamics*, 2013.

[13] L. R. Salazar, R. Sabatini, S. Ramasamy, and A. Gardi, "A novel system for non-cooperative uav sense-and-avoid," in *proceedings of European Navigation Conference*, 2013.

[14] J. Park and Y. Kim, "Stereo vision based collision avoidance of quadrotor uav," in *Control, Automation and Systems (ICCAS), 2012 12th International Conference on*. IEEE, 2012, pp. 173–178.

[15] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[16] E. K. Chong, C. M. Kreucher, and A. O. Hero, "Partially observable markov decision process approximations for adaptive sensing," *Discrete Event Dynamic Systems*, vol. 19, no. 3, pp. 377–422, 2009.

[17] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," in *Breakthroughs in statistics*. Springer, 1992, pp. 235–247.

[18] X. Wang, M. Morelande, and B. Moran, "Target motion analysis using single sensor bearings-only measurements," in *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on*. IEEE, 2009, pp. 1–6.

[19] X. Wang, Y. Cheng, M. Morelande, and B. Moran, "Bearings-only sensor trajectory scheduling using accumulative information," in *Radar Symposium (IRS), 2011 Proceedings International*. IEEE, 2011, pp. 682–688.