Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Optimal UAV Circumnavigation Control with Input Saturation Based on Information Geometry

Yangguang Yu * **Xiangke Wang** * **Lincheng Shen** *

* *College of Intelligence Science and Technology, National University of Defense Technology ,Changsha, 410073, China (e-mail: xkwang@nudt.edu.cn)*

Abstract: In this paper, we investigate the problem of the optimal circumnavigation around a ground moving target for a fixed-wing unmanned aerial vehicle equipped with a radar. We propose an optimal circumnavigation control law which not only achieves the circumnavigation of a UAV around a moving target, but also maximizes the utilization of the sensor information. Firstly, an optimization criterion reflecting the extent of the sensor information utilization is established based on the Fisher information. Then, based on a neural network, an optimal circumnavigation control law with input saturation is designed. The result is a nearly optimal state feedback controller that has been tuned a priori off-line. Finally, a simulation is presented to demonstrate the validity and correctness of the proposed method.

Keywords: UAV, optimal control, Fisher information, circumnavigation, input saturation

1. INTRODUCTION

With the development of the unmanned aerial vehicles (UAVs), the application of UAVs can be found in both military and civilian fields. Among a variety of applications, the surveillance and tracking of moving targets is one of the primary application domains of UAVs. To provide a better aerial monitoring for the ground target, the UAV is practically required to loiter over the target with a desired distance and circumnavigate around the target.

In the past years, a variation of circumnavigation methods have been proposed. Using the distance measurements (Shames et al. (2011)) or bearing measurements (Deghat et al. (2014)), time-varying algorithms were proposed for a robot with single-integrator dynamics. The proposed controllers consisted of an estimator to localize the target and a control law that forces the robot to circumnavigate the target. For a non-holonomic agent, Deghat et al. (2012) designed a circumnavigation algorithm assuming that the position of the target is unknown. Following their works, Cao (2015) developed a range-only control strategy for an unmanned aerial vehicle using range measurements and a sliding-mode estimator of the range rate was also designed. When the relative position of the target can be accessed, guidance laws based on Vector Fields (VF) method were also exploited (Dong et al. (2019)). The convergence and the stability of the proposed algorithms have been proved in these works. However, the optimality of the circumnavigation control algorithm was considered by none of these works.

The state of the target, which is usually obtained by the on-board sensor, is essential in the target tracking problem. As the measurement sensor data generally contains noises, a filter like Extended Kalman Filter (EKF) is usually needed in order to estimate the true state of the target. Different sensor data may contribute differently during the process of filtering. For example, for the sensors which detect the bearing information, the target state filtered with the data by circling is more accurate (see Ponda et al. (2009)). Hence, in order to quantize the contribution of the sensor data, the Fisher Information Matrix (FIM) was applied to express the amount of information, which equals to the inverse of the lower bound of the error covariance (Wang et al. (2010)) . Based on that, Chen et al. (2016) proposed a UAV guidance law by using one-step determinant of FIM. However, this work only considered one-step optimization, which may incur a local extremum.

The solution of the optimal control problem is connected with the solution of the underlying Hamilton-Jacobi-Bellman (HJB) equation. Considerable efforts have been made for developing algorithms which approximately solve this equation (e.g. Abu-Khalaf and Lewis (2005)). The neural networks were used in order to solve the value function of this HJB equation in many works, such as Li et al. (2017). Then to confront the optimal control with input saturation, Lyshevski (1998) proposed the use of nonquadratic functionals to confront constraints on inputs.

In this paper, we propose an optimal circumnavigation control law which not only achieves the circumnavigation of a UAV around a moving target, but also maximizes the utilization of the sensor information. Firstly, we formulate the optimal circumnavigation problem in which the sensor information utilization is considered based on the Fisher information. Then, a design method for the optimal circumnavigation control law with input saturation is proposed. A simulation result by comparing with the method in Dong et al. (2019) also proves the effectiveness of our method. To the best of our knowledge, it is the first time

that the UAV circumnavigation problem is optimized by infinite horizon optimal control method.

2. PROBLEM FORMULATION

In this paper, we focus on the design of a flight controller for a fixed wing UAV to circumnavigate around a moving target. Specifically, relative to the target, the UAV is expected to track a desired circle over the target with a desired radius r_d and a constant relative speed v_r (see Fig. 1). The altitude of the fixed wing UAV is assumed to be held constant and its kinematic model is described by

$$\begin{cases} \dot{x} = v\cos\theta, \\ \dot{y} = v\sin\theta, \\ \dot{\theta} = u_\theta, \\ v = u_v, \end{cases} \tag{1}$$

where (x, y) is the position of the UAV in the planar plane and θ is its heading. The UAV's linear velocity v and angular velocity ω is determined by control inputs $u = (u_v, u_\theta)$. Due to the roll angle limitation of the fixed wing UAV, the following input constraints should be satisfied:

$$|u_\theta| \le \omega_{max}. \tag{2}$$

The state of the moving target is denoted by $s_t = (x_t, y_t)^T \in \mathbb{R}^2$. The target is assumed to move with a constant linear velocity and its kinematics satisfies:

$$\begin{cases} \dot{x}_t = v_t\cos(\theta_t), \\ \dot{y}_t = v_t\sin(\theta_t), \end{cases} \tag{3}$$

where v_t and θ_t are two known constants.

Let θ_r denote the relative angle of the UAV with respect to the target, as the relative speed v_r is a constant in this paper, we have the following equations:

$$\begin{cases} v_r\cos\theta_r = v_t\cos\theta_t - v\cos\theta, \\ v_r\sin\theta_r = v_t\sin\theta_t - v\sin\theta. \end{cases} \tag{4}$$

where $v_r > v_t$.

The UAV is only equipped with one radar as the measuring sensor towards the target. The on-board radar can measure the range and bearing information determined by relative state $s_r = (x_r, y_r)^T = (x_t - x, y_t - y)^T$. The observation model of the radar sensor is illustrated as follows:

$$\zeta(t) = \kappa(s_r) + \chi(t),$$

where $\zeta(t)$ is the obtained sensor measurement at time t, $\kappa(.)$ is the observation function and $\chi(t)$ is the measurement noise. Specifically, the observation function $\kappa(.)$ is defined as in Zhao et al. (2018):

$$\kappa(s_r) = \begin{bmatrix} r \\ \varphi \end{bmatrix} = \begin{bmatrix} \sqrt{x_r^2 + y_r^2 + h^2} \\ \arctan(y_r/x_r) \end{bmatrix}, \tag{5}$$

where r and h are the distance in the three-dimensional space and the altitude of the UAV respectively. φ is the bearing between the UAV and the target.

It is usually assumed that the measurement noise follows a zero-mean Gaussian distribution (Zhao et al. (2018)), which is

$$\zeta(t)|s_r. \sim N(\kappa(s_r), C(s_r)).$$

The measurement noise depends on the signal-to-noise ratio and the covariance matrix of the measurement is

$$C(s_k^o) = \begin{bmatrix} r^4\sigma_r^2 & 0 \\ 0 & \sigma_\varphi^2 \end{bmatrix},$$

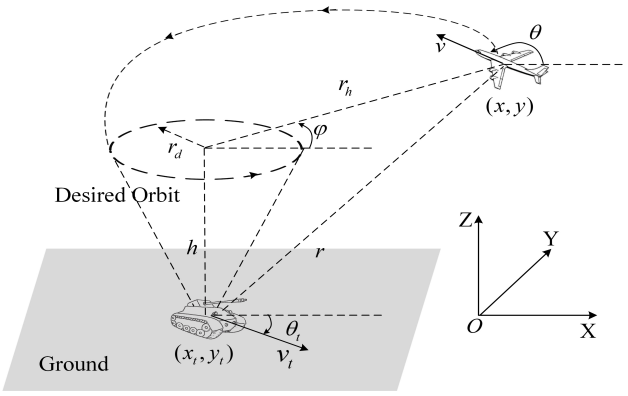

Fig. 1. The illustration of circumnavigation around a ground target

where σ_x and σ_y are the standard deviations of the measurements in the x and y direction respectively.

In order to quantify the utilization of the sensor data, the accumulated information is exploited as the optimization criterion. Given the trajectory of the target $s_t(t)$ in a given time period $[t_1, t_2]$, we expect to enforce the UAV to fly along an optimal trajectory that maximizes the accumulative information $\mathcal{D}$ with respect to the relative position $s_r(t)$, which is given by (Wang et al. (2010))

$$\begin{aligned} \mathcal{D} &\triangleq \int_{t_1}^{t_2} \sqrt{\left[\frac{ds_r(t)}{dt}\right]^T F(s_r(t)) \left[\frac{ds_r(t)}{dt}\right]} dt \\ &= \int_{t_0}^{\infty} \sqrt{[\dot{s}_t - \dot{s}]^T F(s_r(t)) [\dot{s}_t - \dot{s}]} dt, \end{aligned} \tag{6}$$

where $F(s_r(t))$ denotes the Fisher information matrix for the radar sensor whose observation model is described by (5) and its specific form is as following:

$$\begin{aligned} F_{11} &= \frac{x_r^2}{r^6\sigma_r^2} + \frac{y_r^2}{r_h^4\sigma_\varphi^2} + \frac{8x_r^2}{r^4}, \\ F_{12} = F_{21} &= \frac{x_r y_r}{r^6\sigma_r^2} - \frac{x_r y_r}{r_h^4\sigma_\varphi^2} + \frac{8x_r y_r}{r^4}, \\ F_{22} &= \frac{y_r^2}{r^6\sigma_r^2} + \frac{x_r^2}{r_h^4\sigma_\varphi^2} + \frac{8y_r^2}{r^4}, \end{aligned} \tag{7}$$

where $r_h = \sqrt{x_r^2 + y_r^2}$ denotes the loiter radius of the UAV, which is the projected relative distance between the UAV and the target onto the X-Y plane. The detail of derivation process for the formula (7) can refer to Zhao et al. (2018).

Based on the defined accumulative information, a definition of the information optimal circumnavigation control is given as follows:

Definition 1. For a system whose dynamic is given by (1), target model is given by (3) and sensor model is given by (5), if a control policy u^* satisfies the following condition:

(**a**) Driven by u^*, the loitering radius of the UAV will converge to a desired radius r_d, that is $\lim_{t\to\infty} r_h = r_d$.

(**b**) The accumulative information (6) with the control policy u^* is larger than any other control policy that satisfies condition (**a**),

then the control policy u^* is called to be the information optimal circumnavigation control.

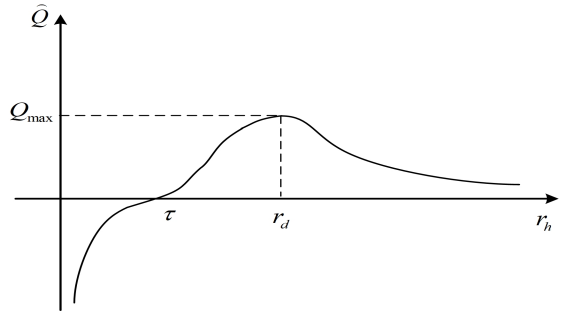

Fig. 2. The illustration of function $\hat{Q}(r_h, \eta)$ with a given η

The aim of this paper is to design an information optimal circumnavigation control law defined by Definition 1.

3. MAIN RESULTS

In this section, we firstly present the process of designing an information optimal circumnavigation control law. A neural network is then used to approximate the optimal control law and the method of tunning the weights of the network is also introduced.

3.1 Optimal Circumnavigation Control Law Design

Define the error e_r between the current loiter radius and the desired one as

$$e_r = r_h - r_d, \tag{8}$$

and a variable η as

$$\eta = \frac{\pi}{2} - (\theta_r - \varphi). \tag{9}$$

Obviously, the UAV will circumnavigate around the target with the desired loiter radius as long as $e_r \to 0$ and $\eta \to 0$. To achieve this goal, we start our analysis with the dynamics of the loiter radius error e_r.

Firstly, the dynamics of e_r is obtained as:

$$\begin{aligned}\dot{e}_r = \dot{r}_h &= \frac{x_r\dot{x}_r + y_r\dot{y}_r}{r_h} = v_r\cos\varphi cos\theta_r + v_r sin\varphi\sin\theta_r \\ &= v_r\cos(\varphi - \theta_r) = v_r\sin\eta.\end{aligned} \tag{10}$$

Then according to (4), we have

$$v^2 + v_t^2 - 2vv_t cos(\theta - \theta_t) = v_r^2. \tag{11}$$

Deriving both sides of the equations (4) and (11) by time t, it yields

$$-v_r\sin\theta_r\dot{\theta}_r + \dot{v}\cos\theta - v\sin\theta\dot{\theta} = 0, \tag{12}$$

$$v\dot{v} - \cos(\theta - \theta_t)v_t\dot{v} + \sin(\theta - \theta_t)vv_t\dot{\theta} = 0. \tag{13}$$

By eliminating $\dot{v}$, we get

$$\begin{aligned}\dot{\theta}_r &= \frac{v_t v\sin(\theta - \theta_t)\cos\theta - v\sin\theta\big(v_t\cos(\theta - \theta_t) - v\big)}{v_r\sin\theta_r\big(v_t\cos(\theta - \theta_t) - v\big)}\dot{\theta} \\ &= \frac{v}{v - v_t cos(\theta - \theta_t)}\dot{\theta}.\end{aligned}$$

Then, the dynamics of φ is obtained as

$$\begin{aligned}\dot{\varphi} &= \frac{x_r\dot{y}_r - y_r\dot{x}_r}{r_h^2} = \frac{\cos\varphi(v_r\sin\theta_r) - \sin\varphi(v_r\cos\theta_r)}{r_h} \\ &= \frac{v_r\sin(\theta_r - \varphi)}{r_h} = \frac{v_r\cos\eta}{r_h}.\end{aligned}$$

Define a state variable $\xi = (e_r, \eta)$ and an augmented state variable $\bar{\xi} = (e_r, \eta, \theta)$, we have

$$\dot{\eta} = \dot{\varphi} - \dot{\theta}_r = b_1(\xi) - b_2(\bar{\xi})u_\theta,$$

where

$$b_1(\xi) = \frac{v_r\cos\eta}{r_h}, \quad b_2(\bar{\xi}) = \frac{v}{v - v_t cos(\theta - \theta_t)}.$$

Further, the $\bar{\xi}$-dynamics is described by:

$$\begin{cases}\dot{e}_r = v_r\sin\eta, \\ \dot{\eta} = b_1(\xi) - b_2(\bar{\xi})u_\theta, \\ \dot{\theta} = u_\theta,\end{cases} \tag{14}$$

The state ξ should be stabilized by the optimal control law u_θ^*. Before we present the method to design the optimal control law u_θ^*, a definition of the admissible control policy (see Li et al. (2017)) is firstly introduced as follows:

Definition 2. (Admissible Controls). For a given system (14), a control policy $u_\theta = \mu(\bar{\xi})$ is defined to be admissible with respect to a given cost $\mathcal{J}$ on a set $\Omega \subseteq \mathbb{R}^3$, written as $\mu(\bar{\xi}) \in \mathcal{A}(\Omega)$, if u_θ is continuous, $u_\theta = 0$ when $\xi = 0$, u_θ stabilizes the state ξ and $\mathcal{J}$ is finite.

Observe that when the ξ-system is stable (i.e.,$\xi = 0$ and $\dot{\xi} = 0$), the desired control input $u_\theta = v_r/r_d \neq 0$, which implies that u_θ is not an admissible control with respect to system (14). To deal with this issue, a virtual input $\hat{u}_\theta$ is defined as

$$\hat{u}_\theta(t) = -b_2(\bar{\xi})u_\theta(t) + v_r/r_d. \tag{15}$$

Correspondingly, the dynamics for the state ξ becomes

$$\begin{cases}\dot{e}_r = v_r\sin\eta \\ \dot{\eta} = b_1 - v_r/r_d + \hat{u}_\theta.\end{cases} \tag{16}$$

Obviously, $\hat{u}_\theta$ is an admissible control with respect to the system (16). If the optimal virtual input $\hat{u}_\theta^*$ that achieves the optimal circumnavigation control defined in Definition 1 is obtained, the optimal control law u_θ^* can be obtained by

$$u_\theta^* = -(\hat{u}_\theta^* - v_r/r_d)/b_2. \tag{17}$$

Hence the question that remains is how to obtain the optimal virtual input $\hat{u}_\theta^*$, which will be illustrated in the following part.

Firstly, an optimization criterion based on the accumulative Fisher information will be constructed in the following part. Substitute (9) into (6), then the accumulative information $\mathcal{D}$ in (6) becomes:

$$\mathcal{D} = \int_0^\infty \sqrt{L(r_h, \eta)}dt, \tag{18}$$

where $L(r_h, \eta)$ is:

$$L(r_h, \eta) = \frac{v_r^2 r_h^2}{r^6\sigma_r^2}\sin^2\eta + \frac{8r_h^2 v_r^2}{r^4}\sin^2\eta + \frac{v_r^2}{r_h^2\sigma_\varphi^2}\cos^2\eta.$$

Intuitively, according to (18), more accumulative information D will be obtained in the unit time along with the decrease of the loiter radius r_h. Hence, the loiter radius of the UAV will approach to zero if the accumulative information D is taken as a to-be-maximized optimization criterion. In order to enforce the UAV to loiter around the target with a desired radius, a to-be-minimized performance index is designed based on a small variation of D as

$$\mathcal{J}(\bar{\xi}(0),\hat{u}_\theta)=\int_0^\infty \big[F_Q(r_h(s),\eta(s))+W(\hat{u}_\theta,\bar{\xi})\big]ds, \quad (19)$$
$$F_Q(r_h,\eta)=1-\frac{\hat{Q}(r_h,\eta)}{Q_{max}},$$
$$\hat{Q}(r_h,\eta)=\sqrt{L(r_h,\eta)}\tanh(r_h-\tau), \quad Q_{max}=\hat{Q}(r_d,0),$$

where $W(\hat{u}_\theta,\bar{\xi})$ is a positive definite function which to be designed, $\hat{Q}(r_h,\eta)$ is a function varied slightly from the function $\sqrt{L(r_h,\eta)}$, Q_{max} is the value of $\hat{Q}(r_h,\eta)$ when $r_h=r_d$ and $\eta=0$, τ is a bias variable which satisfies

$$\frac{d\hat{Q}(r_h,\eta)}{dr_h}(r_h=r_d,\eta=0)=0. \quad (20)$$

Further, the exact form of (20) is obtained as

$$f_m(\tau)=-\tanh(r_d-\tau)+r_d(1-\tanh^2(r_d-\tau))=0.$$

It is easy to verify that the function $f_m(\tau)$ is a monotonically decreasing function. Thus the deviation variable τ can be approximated by numerical stepwise method.

When $r_h>r_d+\epsilon$, $\hat{Q}(r_h,\eta)\approx\sqrt{L(r_h,\eta)}$ as $\rho(r_h)\approx 1$. The accumulative information performance index in the performance index (19) is nearly unchanged. The evaluated accumulative information in (19) is only changed when the loiter radius is between the interval $(0,r_d+\epsilon)$ in order to enforce the UAV to loiter with a desired radius. Fig 2 illustrates the relationship between the $\hat{Q}$ and r_h with a given η. It can be seen that the value of $\hat{Q}$ reach the maximum Q_{max} at $r_h=r_d$.

Next to confront the bounded controls, a generalized non-quadratic function is defined as

$$W(\hat{u}_\theta,\bar{\xi})=2\overline{\lambda}(\bar{\xi})R\int_0^{\hat{u}_\theta}\left(\tanh^{-1}(s/\overline{\lambda}(\bar{\xi}))\right)\mathrm{d}s, \quad (21)$$

where $\overline{\lambda}(\bar{\xi})$ is defined as

$$\overline{\lambda}(\bar{\xi})=\begin{cases}|b_2(\bar{\xi})|\omega_{max}+v_r/r_d, & \text{if } \hat{u}_\theta\geq 0,\\ |b_2(\bar{\xi})|\omega_{max}-v_r/r_d, & \text{if } \hat{u}_\theta<0.\end{cases} \quad (22)$$

To guarantee that $\overline{\lambda}(\bar{\xi})>0$, it is assumed that the maximum angular velocity ω_{max} satisfies the condition $|b_2(\bar{\xi})|\omega_{max}-v_r/r_d>0$ for $\forall\bar{\xi}\in\Omega$.

Given an admissible control policy $\hat{u}_\theta(t)$ for the system (16), assume that there exists a continuous function $V^*(\bar{\xi}(t))$ that satisfies

$$V^*(\bar{\xi}(t))=\min_{\hat{u}_\theta\in\mathcal{A}(\Omega)}\int_t^\infty\big[F_Q(r_h,\eta)+W(\hat{u}_\theta(s))\big]ds. \quad (23)$$

Based on (23), a Hamiltonian function as:

$$H(\bar{\xi},V^*,\hat{u}_\theta)=2\overline{\lambda}(\bar{\xi})R\int_0^{\hat{u}_\theta}\left(\tanh^{-1}(s/\overline{\lambda}(\bar{\xi}))\right)\mathrm{d}s + F_Q(r_h,\eta)+(V_{\bar{\xi}}^*)^T\big[F(\bar{\xi})+G(\bar{\xi})\hat{u}_\theta\big]^T, \quad (24)$$

where $V_{\bar{\xi}}^*=\partial V^*/\partial\bar{\xi}\in\mathbb{R}^3$, $F(\bar{\xi})=[v_r\sin\eta, b_1(\xi)-v_r/rd, v_r/(b_2(\bar{\xi})r_d)]^T$, $G(\bar{\xi})=[0,1,-1/b_2(\bar{\xi})]^T$.

Based on the above analysis, a main result of this paper is given as follows:

Theorem 1. Consider a UAV system with kinematics (1), sensor model (5) and target kinematics (3). If the following condition is satisfied:

$$\frac{r_h^2}{r^6\sigma_r^2}+\frac{8r_h^2}{r^4}-\frac{1}{r_h^2\sigma_\varphi^2}<0, \quad (25)$$

and the control law $u^*=(u_v^*,u_\theta^*)$ is given by

$$\begin{cases}u_v^*=v_t cos(\theta-\theta_t)+\sqrt{v_t^2cos^2(\theta-\theta_t)+v_r^2-v_t^2}\\ u_\theta^*=(\overline{\lambda}\tanh\left(1/(2\overline{\lambda}R)G(\bar{\xi})^TV_{\bar{\xi}}^*\right)+v_r/r_d)/b_2,\end{cases} \quad (26)$$

where $V_{\bar{\xi}}^*$ is the solution of (24), u^* is the optimal control law that maximizes the performance index (19) while the constraint (2) is satisfied, and the UAV will circumnavigate around the target with a desired loiter radius r_d, that is, $\lim_{t\to\infty} r_h\to r_d$.

Proof. When the condition (25) holds, it can be checked that $\hat{Q}$ reaches its maximum only when $r_h=r_d$ and $\eta=0$, where a stable circumnavigation around the target is formed. Further, observe the structure of the performance index (19) and it can be found that $\mathcal{J}(\bar{\xi}(0),\hat{u}_\theta)$ is finite only when the state ξ can be stabilized by the optimal control policy u_θ^*. When ξ is stabilized, we have $\lim_{t\to\infty} e_r\to 0$, which implies $\lim_{t\to\infty} r_h\to r_d$.

Then we will prove concisely that the optimal control law can be obtained through the method proposed in this paper. The desired velocity u_v^* of the UAV can be easily obtained according to (11). Then employing the stationary condition (see Lewis et al. (2012)) on the equation $H(\bar{\xi},V^*,\hat{u}_\theta)=0$, i.e., $\partial H/\hat{u}_\theta=0$, the optimal virtual input $\hat{u}_\theta^*$ can be obtained as

$$\hat{u}_\theta^*=-\overline{\lambda}\tanh\left(1/(2\overline{\lambda}R)G(\bar{\xi})^TV_{\bar{\xi}}^*\right). \quad (27)$$

Finally, the optimal control law u_θ^* can be acquired by (17). Obviously, $|\hat{u}_\theta^*|<\overline{\lambda}$. Combined with (17) and the definition of $\overline{\lambda}$ in (22), it can be induced that $|\hat{u}_\theta^*|<\omega_{max}$.

Due to the limitations of the length of the paper, this proof is not rigorous enough and a complete proof will be given in our another paper.

Substituting (26) into (24) yields the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\overline{\lambda}^2(\bar{\xi})R\ln\left(1-\tanh(G^TV_{\bar{\xi}}^*/(2\overline{\lambda}(\bar{\xi})R))\right) + F_Q(r_h,\eta)+(V_{\bar{\xi}}^*)^TF(\bar{\xi})=0. \quad (28)$$

Eq. (28) is a nonlinear differential equation that cannot generally be solved directly. There is currently no method for rigorously solving for the value function of this constrained optimal control problem. In the following part, a neural network is employed along with the theory of successive approximation, to solve for the value function of (28).

3.2 Neural-network-based policy iteration algorithm for solving the HJB equation

In order to solve the function $V_{\bar{\xi}}^*$ in (28), a neural-network-type structure is used to approximate the value function. To successively solve (24) and (27), V is approximated by

$$V_M(\bar{\xi})=\sum_{j=1}^M w_j\sigma_j(\bar{\xi})=(\mathbf{w}_M)^T\sigma_M(\bar{\xi}), \quad (29)$$

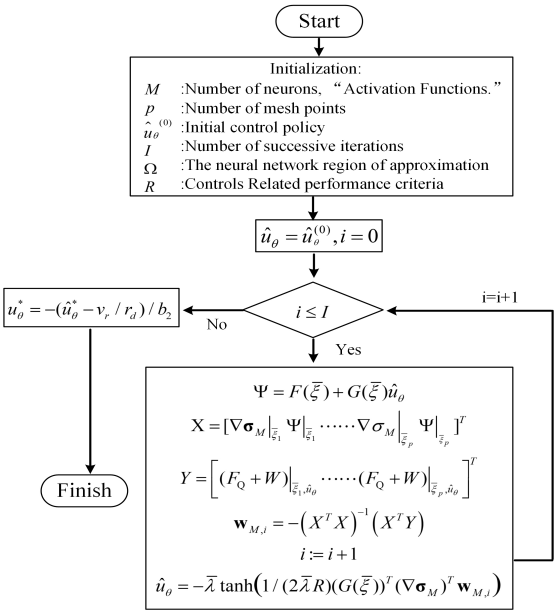

Fig. 3. The flowchart of the algorithm for nearly optimal saturated neurocontrol.

which is a neural network with the activation functions $\sigma_j(\bar{\xi}) \in C^1$ and $\sigma_M(\bar{\xi})$ is the vector of activation functions. w_j denotes the weights of the output layer and $\mathbf{w}_M$ is the weight vector. Note that when $r_h = 0$ and $\eta = 0$, there exists $V(\bar{\xi}) = 0$. To approximate $V(\bar{\xi})$ by (29), it should hold that $\sigma_j = 0$ when $r_h = 0$ and $\eta = 0$.

For $H(\bar{\xi}, V, \hat{u}_\theta) = 0$, the solution V is replaced with V_M having a residual error

$$H\left(V_M(\bar{\xi}) = \sum_{j=1}^{M} w_j \sigma_j(\bar{\xi}), \hat{u}_\theta\right) = e_M(\bar{\xi}).$$

Obviously, the residual error should be minimized. Therefore, the parameters $\mathbf{w}_M$ is tuned in order to minimize the objective

$$S = \int_\Omega |e_M(\bar{\xi})|^2 \mathrm{d}\bar{\xi},$$

The weights, $\mathbf{w}_M$, are determined by projecting the residual error onto $\mathrm{d}e_M(\bar{\xi})/\mathrm{d}\mathbf{w}_M$ and setting the result to zero for $\forall\bar{\xi} \in \Omega$ using the inner product, i.e.

$$\left\langle \frac{\mathrm{d}e_M(\bar{\xi})}{\mathrm{d}\mathbf{w}_M}, e_M(\bar{\xi}) \right\rangle = 0,$$

where $\langle f, g\rangle = \int_\Omega fg\mathrm{d}\bar{\xi}$ is a Lebesgue integral. One has

$$\langle \nabla\sigma_M\Psi, \nabla\sigma_M\Psi\rangle \mathbf{w}_M + \langle F_Q + W, \nabla\sigma_M\Psi\rangle = 0\,,$$

where $\Psi = F(\bar{\xi}) + G(\bar{\xi})\hat{u}_\theta$. Thus, $\mathbf{w}_M$ is updated as

$$\mathbf{w}_M = -\langle \nabla\boldsymbol{\sigma}_M\Psi, \nabla\boldsymbol{\sigma}_M\Psi\rangle^{-1} \langle F_Q + W,\ \nabla\boldsymbol{\sigma}_M\Psi\rangle\,.$$

Afterwards, an iterative algorithm proposed by Abu-Khalaf and Lewis (2005) is used to solve (28) and a flowchart of the computational algorithm is shown in Fig. 3. More details of the iterative algorithm can refer to Abu-Khalaf and Lewis (2005).

4. SIMULATION RESULTS

In the simulation, the UAV is controlled to circumnavigate around a moving target, whose linear velocity is set as 10 m/s, with a desired loiter radius of 50 m. The relative velocity of the UAV with respect to the target is 20 m/s and the height of the UAV is 80 m. The sensor measurement standard deviation parameters are $\sigma_r = 10^{-3}/m$ and $\sigma_\varphi = 1 \times 10^{-4}\pi$ rad. The maximum angle speed w_{max} is set as 0.8 rad/s. The following neural network is used:

$$V_M(\bar{\xi}) = \sum_{i=1}^{15}\sum_{j=1}^{5} w_{5(i-1)+j}\alpha_i\beta_j,$$

where α_i and β_j are, respectively, the i-th and j-th elements of the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which are defined as follows:

$$\begin{aligned}\boldsymbol{\alpha} =& [e_r^2, e_r\eta, \eta^2, e_r^4, e_r^3\eta, e_r^2\eta^2, e_r\eta^3, \eta^4, e_r^6,\\ & e_r^5\eta^2, e_r^4\eta^2, e_r^3\eta^3, e_r^2\eta^4, e_r^5\eta, \eta^6]^T,\\ \boldsymbol{\beta} =& [1, \theta, \theta^2, \theta^3, \theta^4].\end{aligned}$$

This is a power series neural network with 75 activation functions containing powers of the state variable of the system up to the 10th power.

The trajectories of the UAV and the target are illustrated in Fig. 4. Fig. 6 illustrates the convergence of the loiter radius r_h (red solid line). It can be observed that controlled by the designed algorithm, the loiter radius r_h converges to the desired radius 50 m. The control input $u_\theta(t)$ is limited in [-0.8,0.8] during the circumnavigation, which is illustrated in Fig 5.

To prove the validity of our algorithm, the designed method is compared with the vector field (VF) method proposed in Dong et al. (2019), which is given as:

$$u_\theta = \begin{cases} -k(\theta - \theta_d), & \text{if } |k(\theta - \theta_d)| < \omega_{max}, \\ \omega_{max}, & \text{if } |k(\theta - \theta_d)| > \omega_{max} \ \&\ \theta < \theta_d \\ -\omega_{max}, & \text{if } |k(\theta - \theta_d)| > \omega_{max} \ \&\ \theta > \theta_d \end{cases}$$

where θ_d is determined by the vector field:

$$\begin{bmatrix}\cos\theta_d \\ \sin\theta_d\end{bmatrix} = \frac{-v_r}{r_h(r_h^2 + r_d^2)} \begin{bmatrix} x_r(r_h^2 - r_d^2) + y_r(2r_d r_h) \\ y_r(r_h^2 - r_d^2) - x_r(2r_d r_h)\end{bmatrix}.$$

Fig. 6 shows the comparison of the distance variation between two methods during the process of circumnavigation. It can be observed that the loiter radius controlled by our algorithm converges faster than that of the VF method. The comparison of the accumulated Fisher information during the process of approaching the target is illustrated in Fig. 7. Obviously, the accumulated Fisher information obtained by our algorithm is higher than that of the VF method. To illustrate this better, the gained Fisher information $\hat{Q}$ in per simulation step of the two methods is illustrated in Fig 8. It can be observed that $\hat{Q}$ gained by our method is higher than the VF method when the UAV is approaching the target. When the UAV achieves the circumnavigation, $\hat{Q}$ gained by two methods are equal.

5. CONCLUSION

In this paper, we propose an optimal circumnavigation control law which not only achieves the circumnavigation of a UAV around a moving target, but also maximizes the utilization of the sensor information. By comparison with the work of Dong et al. (2019), the effectiveness of our algorithm is proved. Our algorithm can help improve the

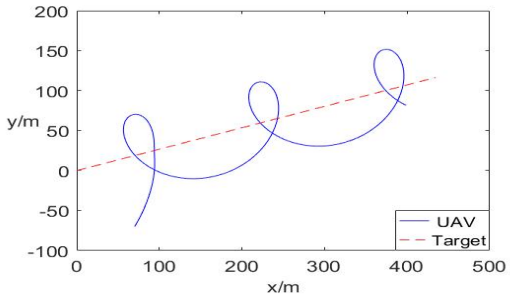

Fig. 4. The trajectories of the UAV and the target.

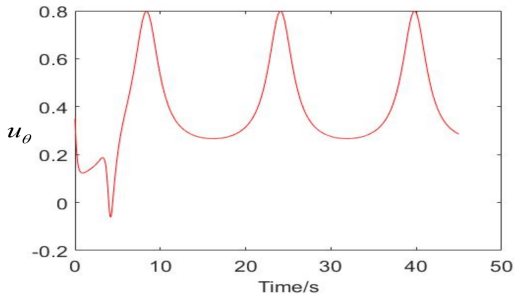

Fig. 5. The control input during the circumnavigation.

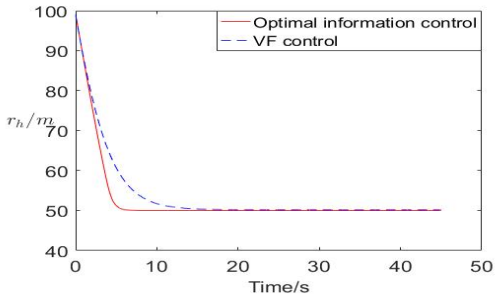

Fig. 6. The illustration of the relative distances r_h controlled by two methods.

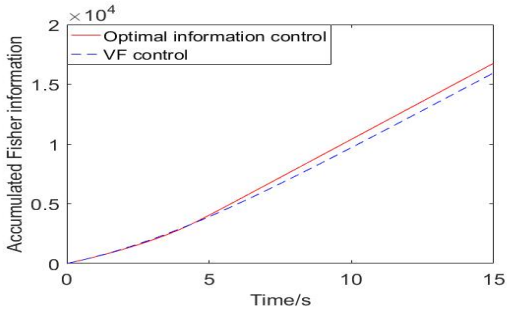

Fig. 7. The illustration of the accumulated Fisher information controlled by two methods.

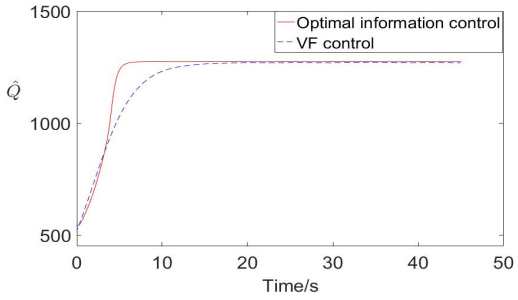

Fig. 8. The illustration of $\hat{Q}$ gained in per simulation step.

estimation accuracy of the target during the process of circumnavigation and further improve the control accuracy.

REFERENCES

Abu-Khalaf, M. and Lewis, F.L. (2005). Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network hjb approach. *Automatica*, 41(5), 779–791.

Cao, Y. (2015). UAV circumnavigating an unknown target under a gps-denied environment with range-only measurements. *Automatica*, 55, 150–158.

Chen, L., Cui, R., Gao, J., and Yan, W. (2016). Cooperative guidance of multiple UAVs for target estimation based on nonlinear model predictive control. In *2016 International Conference on Advanced Robotics and Mechatronics*, 178–183. IEEE.

Deghat, M., Davis, E., See, T., Shames, I., Anderson, B.D., and Yu, C. (2012). Target localization and circumnavigation by a non-holonomic robot. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1227–1232. IEEE.

Deghat, M., Shames, I., Anderson, B.D., and Yu, C. (2014). Localization and circumnavigation of a slowly moving target using bearing measurements. *IEEE Transactions on Automatic Control*, 59(8), 2182–2188.

Dong, F., You, K., and Zhang, J. (2019). Flight control for UAV loitering over a ground target with unknown maneuver. *arXiv preprint arXiv:1906.07000*.

Lewis, F.L., Vrabie, D., and Syrmos, V.L. (2012). *Optimal control*. John Wiley & Sons.

Li, J., Modares, H., Chai, T., Lewis, F.L., and Xie, L. (2017). Off-policy reinforcement learning for synchronization in multiagent graphical games. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2434–2445.

Lyshevski, S.E. (1998). Optimal control of nonlinear continuous-time systems: design of bounded controllers via generalized nonquadratic functionals. In *Proceedings of the 1998 American Control Conference*, volume 1, 205–209. IEEE.

Ponda, S., Kolacinski, R., and Frazzoli, E. (2009). Trajectory optimization for target localization using small unmanned aerial vehicles. In *AIAA Guidance, Navigation, and Control Conference*, 6015.

Shames, I., Dasgupta, S., Fidan, B., and Anderson, B.D. (2011). Circumnavigation using distance measurements under slow drift. *IEEE Transactions on Automatic Control*, 57(4), 889–903.

Wang, X., Cheng, Y., and Moran, B. (2010). Bearings-only tracking analysis via information geometry. In *the 13th International Conference on Information Fusion*, 1–6. IEEE.

Zhao, Y., Wang, X., Cong, Y., and Shen, L. (2018). Information geometry-based action decision-making for target tracking by fixed-wing unmanned aerial vehicle: from algorithm design to theory analysis. *International Journal of Advanced Robotic Systems*, 15(4).