

# 基于深度神经网络的无人作战飞机自主空战机动决策

张宏鹏<sup>1</sup>, 黄长强<sup>1</sup>, 轩永波<sup>2</sup>, 唐上钦<sup>1</sup>

(1. 空军工程大学 航空工程学院, 陕西 西安 710038; 2. 空军研究院 航空兵研究所, 北京 100085)

**摘要:** 机动决策是决定无人作战飞机空战成败的关键因素。为提高空战获胜概率, 提出用深度神经网络进行决策。构建了36种机动动作, 通过飞行仿真, 得到由当前态势、控制量和未来态势构成的样本; 用样本训练深度神经网络, 使其能够根据当前信息快速准确预测未来态势, 设计了决策目标函数和态势评估函数, 空战过程中, 利用训练好的网络预测所有动作对应的未来态势, 根据决策目标函数从中选出最优动作; 在不同初始条件下, 分别与采用简单机动和自主机动的敌机进行空战仿真, 并对空战态势进行评估。结果表明, 所提方法在均势时能通过较少的动作获得空战胜利, 在劣势时能通过一系列机动获得优势, 且决策用时缩短了9 ms。

**关键词:** 无人作战飞机; 机动决策; 深度神经网络; 空战态势; 飞行仿真

**中图分类号:** E844 **文献标志码:** A **文章编号:** 1000-1093(2020)08-1613-10

**DOI:** 10.3969/j.issn.1000-1093.2020.08.016

## Maneuver Decision of Autonomous Air Combat of Unmanned Combat Aerial Vehicle Based on Deep Neural Network

ZHANG Hongpeng<sup>1</sup>, HUANG Changqiang<sup>1</sup>, XUAN Yongbo<sup>2</sup>, TANG Shangqin<sup>1</sup>

(1. Aeronautics Engineering College, Air Force Engineering University, Xi'an 710038, Shaanxi, China;

2. Aviation Research Institute, Air Force Research Institute, Beijing 100085, China)

**Abstract:** Maneuver decision is a critical factor which determines the success and failure of air combat for unmanned combat aerial vehicle. In order to increase the probability of winning air combat, a deep neural network (DNN) is proposed for maneuver decision. 36 kinds of maneuvers were constructed, and the samples of current situation, control quantity and future situation were acquired through flight simulations. The DNN is trained with the samples, making it capable of predicting future situation according to current information. Decision target function and situation assessment function were designed. In the process of air combat, the trained DNN is used to predict the future situations corresponding to all maneuvers, and the best maneuver is selected from all the maneuvers according to decision target function. The enemy planes which simply and autonomously maneuver were simulated, respectively, under different initial conditions, and the air combat situations were also assessed. The results show that the proposed decision method can be used to win air combat with less actions at balanced situation, and gain an edge through a series of actions at adverse situation, and the decision-making time is reduced by 9 ms.

**Keywords:** unmanned combat aerial vehicle; maneuver decision; deep neural network; air combat

收稿日期: 2019-10-14

基金项目: 国家自然科学基金项目(51579209)

作者简介: 张宏鹏(1996—), 男, 硕士研究生。E-mail: 1152951370@qq.com

通信作者: 黄长强(1961—), 男, 教授, 博士生导师。E-mail: hcqxian@163.com

situation; flight simulation

## 0 引言

机动决策是指战斗机根据空战过程中迅速变化的态势信息做出相应的机动动作。机动决策以获得有利的攻击位置或逃离敌机攻击范围为目的,是无人作战飞机(UCAV)进行自主空战所必须具备的关键技术。

国内外学者为解决战斗机自主决策问题提出了多种自主决策方法。文献[1]提出利用博弈论的方法,用矩阵的形式评估敌我双方可能选择的机动,从而选取对我方最有利的机动动作;文献[2]提出基于多级影响图的机动决策方法;文献[3]提出了一种结合人工势场<sup>[4]</sup>、滚动时域控制<sup>[5]</sup>和粒子群算法<sup>[6]</sup>的战斗机空战机动决策方法;文献[7]构造了基于滚动时域的无人机空战决策专家系统;文献[8]提出了结合贝叶斯推理与滚动时域优化的自主决策方案;文献[9]针对空战双方毁伤概率的不确定性,将鲁棒优化<sup>[10]</sup>思想引入无人机空战博弈模型中;文献[11]提出了基于统计学原理<sup>[12]</sup>的鲁棒机动决策方法;文献[13]提出了一种基于连续动作集强化学习的无人机机动决策方法;文献[14]提出了一种将多目标优化思想<sup>[15]</sup>与强化学习<sup>[16]</sup>相结合的机动决策模型;文献[17]将博弈论和直觉模糊集结合,解决了不确定环境下无人机空战的机动决策问题。上述方法中,矩阵博弈方法不适于大动作空间的机动决策,多级影响图法不满足实时性要求,滚动时域法和基于统计学原理的方法选择机动动作时优化区间有限,强化学习法采集到有效样本的难度极大且训练时间很长。同时,现有决策方法所使用的机动动作较少,飞机机动方式不够灵活。

为解决上述问题,本文提出了一种基于深度神经网络(DNN)的机动决策模型及方法。首先构建了新的机动动作库,共 36 种机动,增强了飞行的灵活性。其次,利用深度学习技术<sup>[18]</sup>和 272 160 种飞行样本,以当前态势及飞行控制量为输入,以一段较长时域后的态势为输出,构造了具有一定态势预测能力的 DNN。空战仿真时,利用该网络根据当前信息从机动动作库中选取输出态势最优的机动动作,从而增强决策智能化程度,避免强化学习法的低效率采样、滚动时域法优化区间有限的缺点,并满足较高的实时性要求。最后通过有效性仿真、与文献[11]方法的对抗性仿真和对空战过程的态势评估,

检验了本文方法有效性及优越性。

## 1 UCAV 飞行模型及机动动作库

### 1.1 UCAV 飞行模型

在研究 UCAV 机动决策的过程中,采用以法向过载、切向过载和滚转角为飞行操控量的 UCAV 运动动力学模型。为简化问题的复杂度,不考虑飞行过程中的迎角和侧滑角,将地面坐标系视为惯性系,同时忽略地球自转及公转的影响,并且不考虑地球曲率变化。具体公式<sup>[19]</sup>如下:

$$\begin{cases} \dot{x} = v \cos \gamma \cos \psi, \\ \dot{y} = v \cos \gamma \sin \psi, \\ \dot{z} = v \sin \gamma, \\ \dot{v} = g(n_x - \sin \gamma), \\ \dot{\gamma} = \frac{g}{v}(n_z \cos \mu - \cos \gamma), \\ \dot{\psi} = \frac{g}{v \cos \gamma} n_z \sin \mu, \end{cases} \quad (1)$$

式中: $x, y, z$  为 UCAV 在惯性系中的坐标; $v, \psi$  和  $\gamma$  分别为 UCAV 速度、航迹偏角和航迹倾角; $g$  为重力加速度; $\mu$  为 UCAV 滚转角; $n_x$  和  $n_z$  分别为 UCAV 切向过载和法向过载; $\dot{x}, \dot{y}, \dot{z}, \dot{v}, \dot{\gamma}$  和  $\dot{\psi}$  分别为相应物理量关于时间的导数。

### 1.2 UCAV 机动动作库

美国国家航空航天局学者设计了 7 种基本机动动作<sup>[20]</sup>:1)保持原有状态飞行;2)最大加速直飞;3)最大过载左转;4)最大过载右转;5)最大过载爬升;6)最大过载俯冲;7)最大减速飞行。在此基础上对机动动作进行扩展,具体方式为在 $(-180^\circ, 180^\circ)$ 范围内每隔  $10^\circ$  选择一个滚转角,同时使 UCAV 纵向过载和侧向过载最大。由此构成 36 种机动动作,轨迹如图 1 所示。由于考虑了滚转角速度,给定滚转角后,飞机需要从初始滚转角开始滚转,故图中轨迹组成的图形有开口,且滚转角速度越大、开口越小。

## 2 DNN 机动决策模型

### 2.1 DNN 简介

深度学习崛起于 21 世纪初,是机器学习的一个分支领域。它是从数据中学习表示的一种数学框架,强调从连续的层中进行学习<sup>[21]</sup>。深度学习中的“深度”指的并不是利用这种方法所获取的更深层

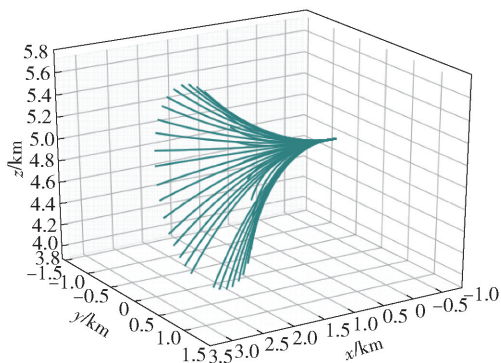


图 1 动作库

Fig. 1 Maneuver library

次的理解,而是指一系列连续的表现层<sup>[22]</sup>。在深度学习中,这些表现层是通过叫作神经网络的模型来学习得到的。DNN 的结构如图 2 所示。

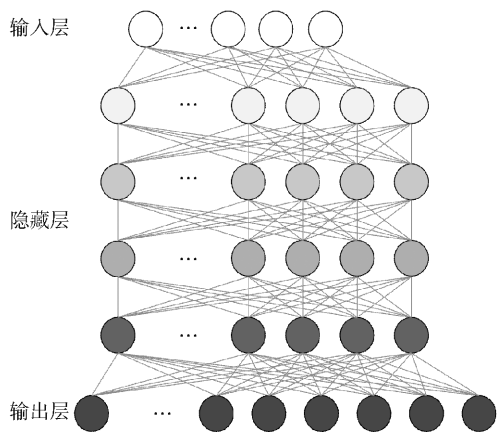


图 2 神经网络

Fig. 2 Deep neural network

## 2.2 DNN 空战机动决策模型

### 2.2.1 空战态势信息

机动决策的过程是根据态势信息选择最有利于己方机动动作的过程。我机与敌机的态势信息,直接或者间接影响双方武器作战效能的发挥<sup>[23]</sup>,从而影响空战的最终结果。态势信息主要有以下 3 种:

- 1) 距离态势:敌我距离、绝对高度和相对高度。
- 2) 速度态势:敌我绝对速度和相对接近速度。
- 3) 角度态势:我机机头指向相对于视线矢量的倾角和偏角,敌机机头指向相对于视线矢量的倾角和偏角,视线转动角速度。

### 2.2.2 DNN 决策模型

本文提出一种基于 DNN 的决策模型,包括两个部分:DNN 训练和 DNN 决策。具体如下:

DNN 训练过程如图 3 所示。以 UCAV 机头指向在水平面的投影为  $x$  轴方向, $z$  轴与水平面垂直,

$y$  轴方向按右手定则选取,此坐标系为采样坐标系。UCAV 初始时刻位于坐标原点,以初始航迹倾角  $\gamma_0$ 、初始滚转角  $\mu_0$ 、初始速度  $v_0$  和 UCAV 将要采用的动作库中某一目标滚转角  $\mu_i$  作为 DNN 输入,令 UCAV 在当前  $\gamma_0, \mu_0, v_0$  下采用  $\mu_i$  飞行一段固定时间  $T$ ,记录此时 UCAV 的航迹倾角  $\gamma_i$  和航迹偏角  $\psi_i$ ,并将其作为 DNN 输出。

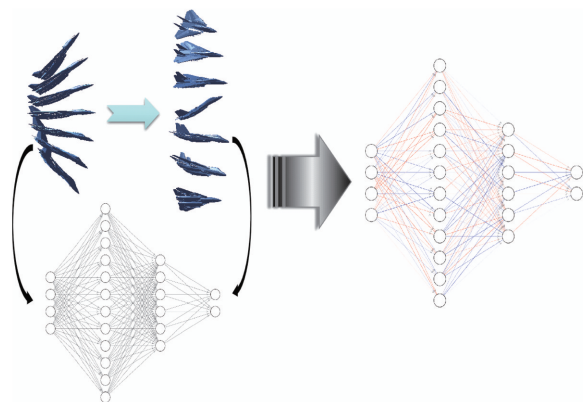


图 3 DNN 训练

Fig. 3 DNN training

通过对 4 种输入均匀采样并进行飞行仿真,得到对应的输出,由此获得大量飞行样本。再利用样本训练 DNN,使其获得根据当前状态和动作预测未来状态的能力。采样范围及间隔如表 1 所示。

表 1 采样范围及间隔

Tab. 1 Sampling range and interval

参数	采样范围	间隔
航迹倾角/(°)	(-50, 50)	5
初始滚转角/(°)	(-180, 180)	10
速度/(m·s <sup>-1</sup> )	(250, 340)	10
目标滚转角/(°)	(-180, 180)	10

DNN 决策:在每一个决策时刻  $t$ ,UCAV 根据自身状态,通过 DNN 依次计算出动作库中每个动作的输出航迹倾角  $\gamma_i^t$  和航迹偏角  $\psi_i^t$ ,其中  $i$  为动作库中每个动作的编号。并将其转换到当前的坐标系下,再根据转换后每个动作的态势,选出使决策目标函数值最大的动作。

### 2.2.3 决策目标函数

本文方法用于解决近距离空战的机动决策问题,目标是使己方快速达成导弹发射条件,不考虑使用航炮的攻击方式。因此,结合 DNN 的输出量,设计决策目标函数如下:

$$Tar = \alpha_1 |\gamma_p - \gamma_{LOS}| + \alpha_2 |\psi_p + \psi_t - \psi_{LOS}|, \quad (2)$$

式中:  $\gamma_p$  和  $\psi_p$  分别为经过 DNN 预测后得到的我机在 2.2.2 节中采样坐标系下的航迹倾角和航迹偏角;  $\alpha_1$  和  $\alpha_2$  为决策目标函数中的权重, 由于满足导弹发射条件既要考虑倾角也要考虑偏角, 因此进攻时取两个权重均为 0.5;  $\gamma_{LOS}$  和  $\psi_{LOS}$  分别为视线矢量  $R_{LOS}$  在当前坐标系下的倾角和偏角。DNN 决策的目的是从动作库中选出使  $T$  时间后的目标函数值  $Tar$  最小的动作, 由此使得 UCAV 快速达到导弹发射条件。

### 3 空战态势评估及胜负判定

#### 3.1 态势评估函数

为客观描述空战过程, 设计空战态势评估函数。现有决策方法的目的是找到使态势函数最大的动作, 而本文方法中态势函数仅用于评估我机在空战中的优劣势情况, 不参与机动决策。而敌机仍采用现有方法, 即采用使态势函数值最大的机动动作。

##### 3.1.1 角度态势

在空战中最重要的事情就是射击, 其次是战斗中的各种战术动作, 最后才是飞行能力<sup>[24]</sup>。而不论是使用航炮还是导弹进行射击, 首先要满足角度要求, 因此, 角度是空战过程中最重要的因素。角度态势函数如下:

$$Q_\gamma = \begin{cases} \frac{3}{2} \cos \gamma_r + 1 - \frac{\sqrt{3}}{2}, & |\gamma_r| < \frac{\pi}{6} \text{ rad}; \\ \frac{6}{\pi} \left( \frac{\sqrt{3}}{2} - 1 \right) \left( |\gamma_r| - \frac{\pi}{3} \right), & \frac{\pi}{6} \text{ rad} \leq |\gamma_r| \leq \frac{\pi}{2} \text{ rad}; \end{cases} \quad (3)$$

$$Q_\psi = \begin{cases} \frac{3}{2} \cos \psi_r + 1 - \frac{\sqrt{3}}{2}, & |\psi_r| < \frac{\pi}{6} \text{ rad}; \\ \frac{5}{3} - \frac{4}{\pi} |\psi_r|, & \frac{\pi}{6} \text{ rad} \leq |\psi_r| < \frac{5\pi}{12} \text{ rad}; \\ 5 - \frac{12}{\pi} |\psi_r|, & \frac{5\pi}{12} \text{ rad} \leq |\psi_r| < \frac{\pi}{2} \text{ rad}; \\ -1, & \frac{\pi}{2} \text{ rad} \leq |\psi_r| \leq \pi \text{ rad}. \end{cases} \quad (4)$$

式中:  $Q_\gamma$  为倾角态势函数;  $Q_\psi$  为偏角态势函数;  $\gamma_r$  和  $\psi_r$  分别为机头相对于视线的倾角和偏角;  $(x_a, y_a, z_a)$  和  $(x_b, y_b, z_b)$  分别为我方和敌方的三维坐标, 我方计算角度态势值时视线矢量  $R_{LOS} = (x_b - x_a, y_b - y_a, z_b - z_a)$ , 敌方计算角度态势值时视线矢量  $R_{LOS} = (x_a - x_b, y_a - y_b, z_a - z_b)$ 。

##### 3.1.2 距离态势

影响距离态势的主要因素是导弹的射程, 因此

设计距离态势函数为

$$Q_R = \begin{cases} 1, & R < R_m; \\ e^{-\frac{(R-R_m)^2}{2\sigma^2}}, & R_m \leq R. \end{cases} \quad (5)$$

式中:  $Q_R$  为距离态势函数;  $R_m$  为导弹射程;  $\sigma$  为其标准偏差;  $R$  为两机直线距离。当敌机位于导弹射程之内时, 距离态势值恒为 1; 当敌机位于射程之外时, 距离越大, 态势值越小。

空战过程中飞机的实时飞行高度也十分重要, 高度太低容易出现坠机事故, 太高不利于飞机性能发挥。因此设计高度态势函数:

$$Q_H = \begin{cases} -1, & H < H_l \text{ 或 } H > H_u; \\ 0, & H_l \leq H \leq H_u. \end{cases} \quad (6)$$

式中:  $Q_H$  为高度态势函数;  $H$  为飞行高度;  $H_l$  为飞行高度下限;  $H_u$  为飞行高度上限。

综合考虑角度、距离和高度因素, 用加权求和的方式计算空战态势函数值, 公式如下:

$$Q = \omega_1 Q_\gamma^a + \omega_2 Q_\psi^a + \omega_3 Q_\gamma^b + \omega_4 Q_\psi^b + \omega_5 Q_R^a + \omega_6 Q_H^a, \quad (7)$$

式中:  $Q$  为空战态势函数;  $Q_\gamma^a$  和  $Q_\psi^a$  分别为我机的倾角和偏角态势函数值;  $Q_\gamma^b$  和  $Q_\psi^b$  分别为敌机的倾角和偏角态势函数值;  $Q_R^a$  和  $Q_H^a$  分别为我机的距离和高度态势函数值;  $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$  为相应态势函数的权重。

#### 3.2 空战胜负判定

为验证本文方法的优越性, 敌机采用文献[11]中的自主决策算法。敌机空战机动动作库与我机相同, 如 1.2 节所示。空战胜利首先必须满足导弹发射条件, 如(8)式所示。

$$\begin{cases} R < R_f, \\ |\psi_r^a| < \frac{\pi}{6} \text{ rad}, \\ |\gamma_r^a| < \frac{\pi}{6} \text{ rad}, \\ Q_b < Q_a, \end{cases} \quad (8)$$

式中:  $R_f$  为可发射导弹的两机最大直线距离;  $\gamma_r^a$  和  $\psi_r^a$  分别为我机机头指向相对于视线矢量  $R_{LOS}$  的倾角和偏角;  $Q_a$  和  $Q_b$  分别为我机和敌机的态势函数值。首先必须满足导弹发射条件, 即(8)式中前 3 项; 其次, 为验证本文方法的有效性, 满足导弹发射条件后, 我机态势函数值  $Q_a$  必须大于敌机态势函数值  $Q_b$  才算获得胜利, 如(8)式中第 4 项所示。空战胜负判定框图如图 4 所示, 其中  $\gamma_r^b$  和  $\psi_r^b$  分别为敌机机头指向相对于视线矢量  $R_{LOS}$  的倾角和偏角。

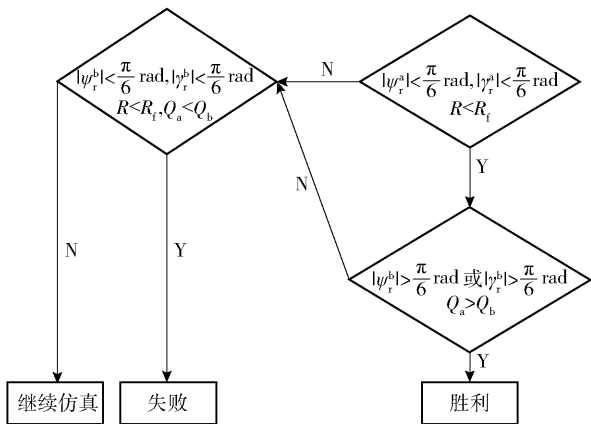


图 4 胜负判定

Fig. 4 Judgement of victory or defeat

### 4 网络训练

#### 4.1 DNN 结构及参数

采样范围及间隔如表 1 所示,使用(1)式的UCAV 运动动力学方程进行仿真实验,共得到 272 160 组样本。由于样本数量较多,因此采用 DNN 构建从当前飞行状态和控制量到未来飞行状态的映射,提高网络预测的准确性、减少网络训练时间。同时为降低过拟合风险,加入丢弃层<sup>[16]</sup>。训练算法为 RMSProp,损失函数设为样本的均方误差,学习率取 0.001。网络结构如表 2 所示。

表 2 网络结构  
Tab. 2 Structure of network

神经层序号	神经层类型	激活函数	输出大小/个
1	全连接层	双曲正切	512
2	丢弃层		512
3	全连接层	修正线性单元	512
4	丢弃层		512
5	全连接层	修正线性单元	256
6	丢弃层		256
7	全连接层	双曲正切	2

#### 4.2 DNN 训练结果

图 5 所示为 DNN 训练过程中的训练集和验证集的预测准确率变化曲线,其中训练次数为 50。从图 5 中可以看出,由于训练样本采样均匀且网络有一定深度,因此网络预测准确率较高。

### 5 仿真实验

仿真实验共分为两个部分:第 1 部分为一种初始态势,第 2 部分为 3 种初始态势。首先,为验证本

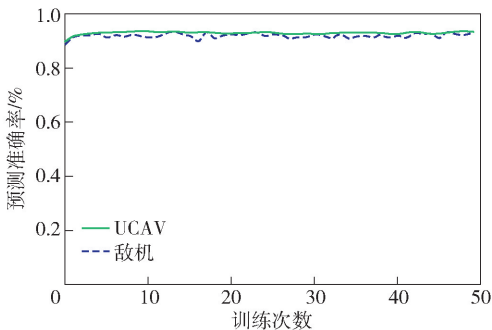


图 5 预测准确率

Fig. 5 Accuracy of prediction

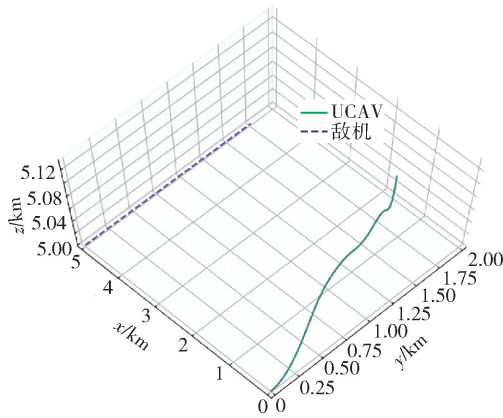
文方法的有效性,在一种初始态势下,敌机采用直线飞行策略;其次,为验证本文方法的对抗性,敌机采用 3.2 节所示的自主决策算法,在 3 种不同的初始态势下进行仿真实验。

通用参数设置如下:初始飞行速度为 250 m/s,最大飞行速度为 400 m/s,最小飞行速度为 90 m/s;  $R_r$  为 4 000 m,  $R_m$  为 5 000 m,  $\sigma$  为 100 m;  $H_l$  为 500 m,  $H_u$  为 12 000 m,初始飞行高度为 5 000 m;最大飞行时间为 100 s;决策周期为 1 s;初始滚转角始终为 0°。胜负判定方法如图 4 所示。

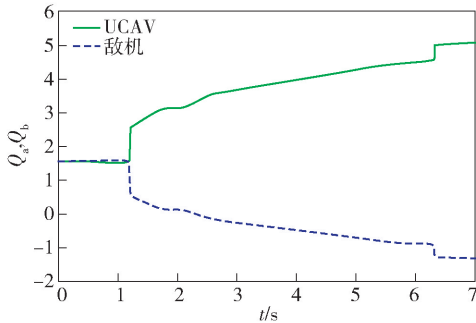
#### 5.1 敌机匀速直线运动

敌机初始位置为 (0 m, 5 000 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 0°;UCAV 初始位置为 (0 m, 0 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 0°。空战仿真轨迹、态势函数值变化情况以及滚转角变化情况如图 6 所示。其中,UCAV 轨迹及相应参数变化用实线表示,敌机轨迹及相应参数变化用虚线表示。在决策过程中,采用最大过载机动,并以最大滚转角速度改变滚转角。3 个控制量均没有突变,因此可以保证轨迹的平滑性。

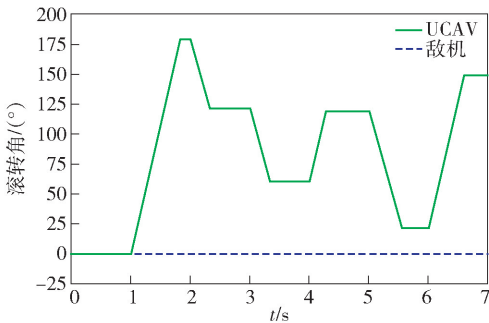
图 6(c)反映了决策过程中机动动作的变化情况,飞机在每一个整数秒时刻进行决策,之后根据决策结果以最大速度改变滚转角。若飞机在下一个决策时刻之前已经完成要做的机动,则保持此时的滚转角不变;若在下一个决策时仍未完成上一个动作,则不再继续该动作,转而做出新的机动。从图 6(c)中可以看出:敌机滚转角始终为 0°,处于匀速直线运动状态;UCAV 在第 1.0 s 时采用 180°的滚转角机动,完成动作后角度保持不变;在第 2.0 s 时减小滚转角度,采用 120°滚转机动;在第 3.0 s 时采用 60°滚转角机动;在第 4.0 s 时根据网络计算结果,又采用 120°滚转机动;在第 5.0 s 时 UCAV 减小滚转角至 20°;第 6.0 s 时又将滚转角增大至 150°;最终在



(a) 空战仿真轨迹  
(a) Trajectory of air combat simulation



(b) 态势函数数值变化情况  
(b) Change of situation function value



(c) 滚转角变化情况  
(c) Change of roll angle

图 6 敌机匀速直线运动

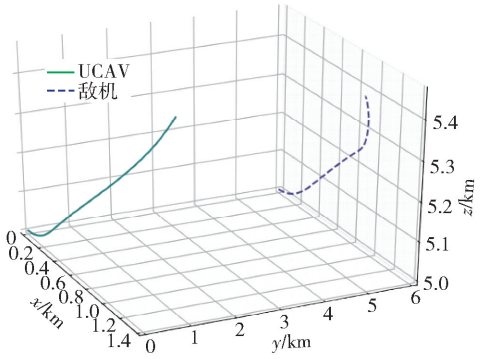
Fig. 6 Uniform straight line motion of enemy aircraft

第 7.0 s 时达到胜利条件。

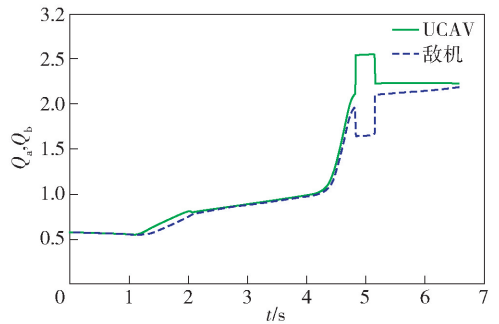
5.2 敌机自主机动

态势 1: 敌机初始位置为(0 m, 6 000 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 0°; UCAV 初始位置为(0 m, 0 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 0°。空战对抗结果如图 7 所示。

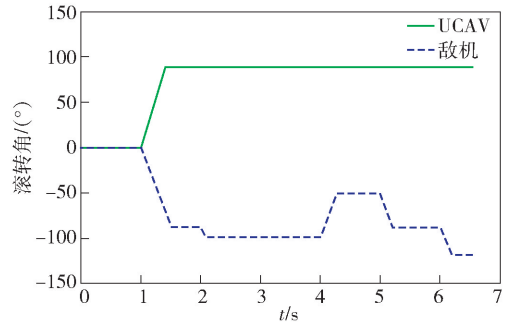
从图 7 中可以看出, 空战双方初始时刻处于均势状态, 双方态势函数值相同。由滚转角变化情况可知, 与敌机采用匀速直线运动时的决策结果不同, 态势 1 下 UCAV 在每个决策时刻计算出的最优动作均为 90° 滚转机动, 因此 UCAV 在完成该机动后控制量不再改变。而敌机除第 3.0 s 外其他每个决策



(a) 空战仿真轨迹  
(a) Trajectory of air combat simulation



(b) 态势函数数值变化情况  
(b) Change of situation function value



(c) 滚转角变化情况  
(c) Change of roll angle

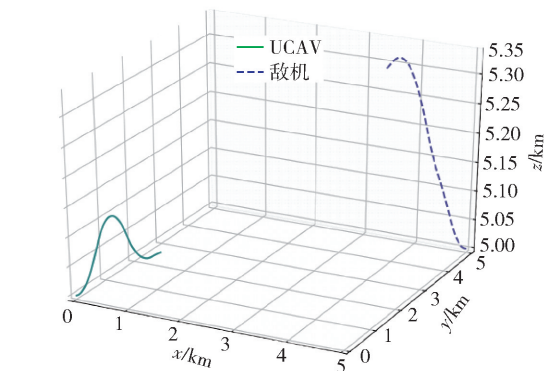
图 7 敌机自主机动态势 1

Fig. 7 Autonomous maneuver of enemy aircraft at Situation 1

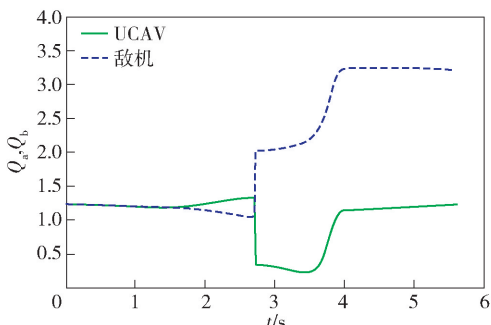
时刻都选择了新的动作。最后, 由于 UCAV 率先达成导弹发射条件且态势函数值大于敌机, 因此获得胜利。同时, UCAV 只用了一个机动动作, 而敌机多次改变动作, 表明本文方法决策效率较高。

同时, 从敌机 4 次小角度的动作调整可以看出, 所设计的动作库能满足机动决策需要, 传统的动作库中不同动作的滚转角通常相差 45° 或 90°, 使得飞行不够灵活。

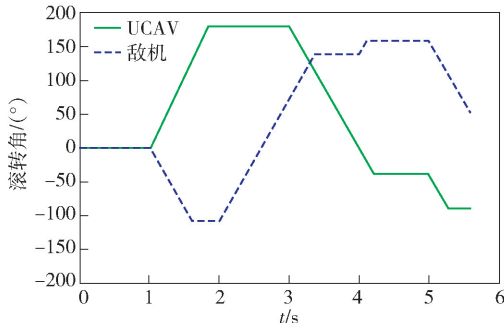
态势 2: 敌机初始位置为(5 000 m, 5 000 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 -135°; UCAV 初始位置为(0 m, 0 m, 5 000 m), 航迹倾角为 0°, 航迹偏角为 45°。空战对抗结果如图 8 所示。



(a) 空战仿真轨迹  
(a) Trajectory of air combat simulation



(b) 态势函数数值变化情况  
(b) Change of situation function value



(c) 滚转角变化情况  
(c) Change of roll angle

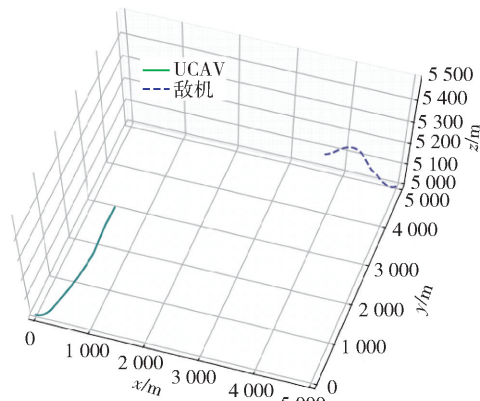
图 8 敌机自主机动态势 2

Fig. 8 Autonomous maneuver of enemy aircraft at Situation 2

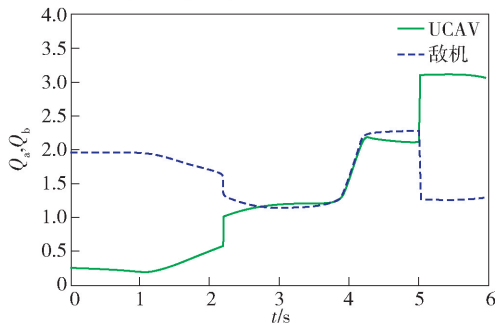
从图 8 中可以看出,初始时刻空战双方相向而行。UCAV 在第 1.0 s 时选择滚转角为  $170^\circ$  的左转机动,而敌机选择滚转角为  $110^\circ$  的左转机动。由于 UCAV 滚转角选择较大,因此其态势函数值经过短暂上升后迅速下降。由于转弯较大,因此敌机在第 2.0 s 时便选择向右滚转,而 UCAV 则在第 3.0 s 时才选择向右滚转。最终在第 5.6 s 时,双方均达成导弹发射条件,但由于此时 UCAV 的态势函数值小于敌机,被判为失败。

态势 3: 敌机初始位置为  $(5\ 000\ \text{m}, 5\ 000\ \text{m}, 5\ 000\ \text{m})$ ,航迹倾角为  $0^\circ$ ,航迹偏角为  $-135^\circ$ ;UCAV

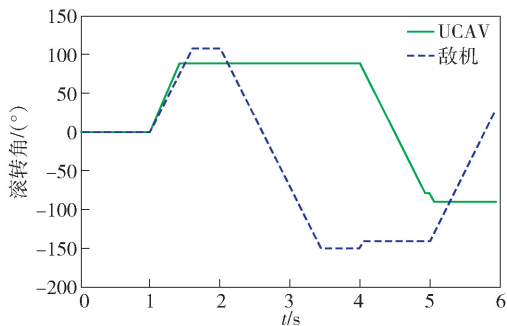
初始位置为  $(0\ \text{m}, 0\ \text{m}, 5\ 000\ \text{m})$ ,航迹倾角为  $0^\circ$ ,航迹偏角为  $0^\circ$ 。空战对抗结果如图 9 所示。



(a) 空战仿真轨迹  
(a) Trajectory of air combat simulation



(b) 态势函数数值变化情况  
(b) Change of situation function value



(c) 滚转角变化情况  
(c) Change of roll angle

图 9 敌机自主机动态势 3

Fig. 9 Autonomous maneuver of enemy aircraft at Situation 3

根据态势 3 可知,空战开始时敌机机头直接指向 UCAV,而 UCAV 机头相对于视线的偏角为  $45^\circ$ ,敌机处于有利位置。由图 9 可知,UCAV 在第 1.0 s 和 4.0 s 时经过两次机动,最终转劣势为优势,在第 6.0 s 时达成胜利条件,且态势函数值高于敌机,因此获得空战胜利。

### 5.3 逃逸机动

空战机动可分为进攻和逃逸两种,因此,可通过修改决策目标函数中的权重,选择使机头指向与视

线矢量的夹角最大机动,以达到逃逸的目的。另外,空战态势可简单分为有利和不利两种,但选择进攻还是逃逸不应根据态势。例如,态势 3 中 UCAV 初始时处于不利状态,但若选择逃逸则空战结果为敌机获胜。因此设计逃逸机动判断条件,若满足条件,则选择逃逸,否则选择进攻。逃逸条件如下:

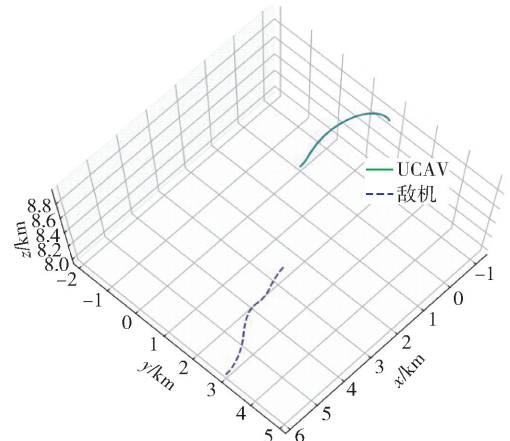
$$\begin{cases} |\psi_r^b| < \psi_e^b, \\ |\gamma_r^b| < \gamma_e^b, \\ |\psi_r^a| > \psi_e^a, \\ |\gamma_r^a| < \gamma_e^a, \end{cases} \quad (9)$$

式中: $\gamma_e^a$  和  $\psi_e^a$  分别为逃逸时 UCAV 应满足的倾角和偏角阈值, $\gamma_e^a = 60^\circ$ ,  $\psi_e^a = 90^\circ$ ;  $\psi_e^b$  和  $\gamma_e^b$  分别为逃逸时敌机应满足的倾角和偏角阈值; $\gamma_r^b = \psi_r^b = 60^\circ$ 。

设置满足上述条件的初始态势:敌机初始位置为(6 000 m, 3 000 m, 8 000 m),航迹倾角为  $0^\circ$ ,航迹偏角为  $180^\circ$ ;UCAV 初始位置为(0 m, 0 m, 8 000 m),航迹倾角为  $0^\circ$ ,航迹偏角为  $180^\circ$ 。此时 UCAV 处于被尾追状态。图 10 为 UCAV 仅采用攻击的仿真结果。

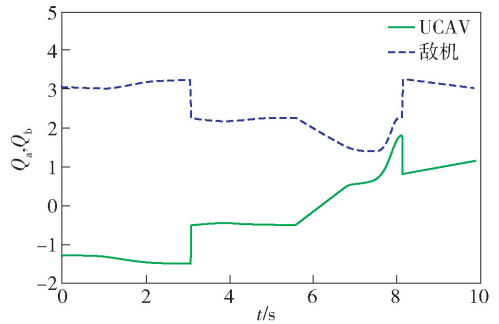
从图 10(b)可以看出,初始时刻双方态势函数值差距较大,空战过程中,UCAV 态势函数值有所上升,但始终低于敌机。图 10(c)显示了 UCAV 的机动情况。UCAV 首先进行  $-80^\circ$  的滚转机动,第 3.0 s 时又将滚转角调整为  $-90^\circ$ 。由图 10(c)可以看出由于仅采用攻击,UCAV 始终朝着减小与视线夹角的方向进行机动,但初始时刻的劣势使其无法仅通过攻击的方式获得胜利,最终在第 9.8 s 时敌机获胜。图 11 为 UCAV 采用混合方式的空战仿真结果,即双方态势满足 (9) 式时,UCAV 采用逃逸机动,不满足时则采用攻击机动。其中,逃逸机动时  $\alpha_1 = -0.6$ ,  $\alpha_2 = 0.4$ ,进攻时  $\alpha_1 = \alpha_2 = 0.5$ ,让决策目标函数值最小,可以达到使 UCAV 指向与视线矢量的夹角减小的目的,从而快速满足导弹发射条件。逃逸时  $\alpha_2 = 0.4$ ,此时让决策目标函数值最小,可以使  $\psi_{\text{LOS}}$  减小,从而为进攻做准备;若  $\alpha_2$  取负,则  $\psi_{\text{LOS}}$  增大,导致我方在水平方向上始终处于不利地位,而  $\alpha_1 = -0.6$ ,此时让决策目标函数值最小,可以使  $\gamma_{\text{LOS}}$  增大,有助于在竖直方向上逃出敌方攻击范围。

此次仿真时长为 56.8 s,说明采用逃逸机动可以提高 UCAV 生存时间。空战过程中 UCAV 首先进行逃逸,采用  $-140^\circ$  的滚转机动向斜下方飞行,因为向下方飞行可以获得重力帮助使 UCAV 在竖直方向上更快逃出对方攻击范围,同时也能保证在水



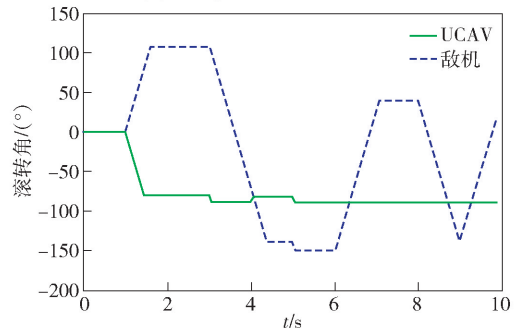
(a) 空战仿真轨迹

(a) Trajectory of air combat simulation



(b) 态势函数值变化情况

(b) Change of situation function value



(c) 滚转角变化情况

(c) Change of roll angle

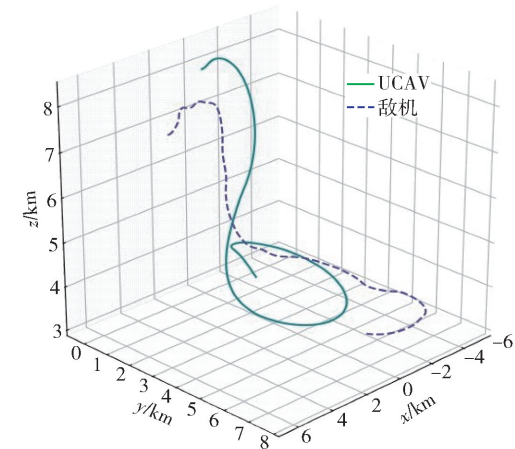
图 10 攻击机动

Fig. 10 Attacking maneuver

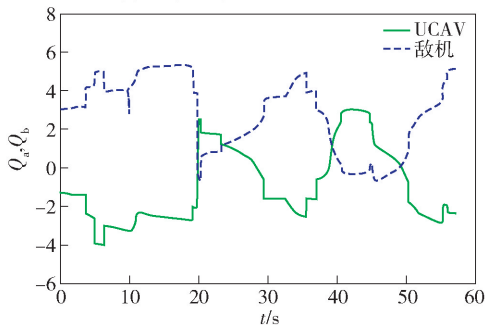
平方向逃出对方攻击范围。从图 11(a)中可看出 UCAV 有明显的高度下降,而由于敌机高度下降太晚,因此 UCAV 在第 14.0 s 时逃出对方攻击范围并转入进攻状态。第 14.0 s 以前的动作均为逃逸机动,以后的动作均为进攻机动。进攻时,UCAV 不再下降高度,而是水平盘旋以占据有利位置。最终在第 56.8 s 时双方均满足导弹发射条件,但由于 UCAV 态势函数值较低,因此失败。

用两种方法分别进行 1 000 次决策再计算其平均决策时间,结果如图 12 所示。

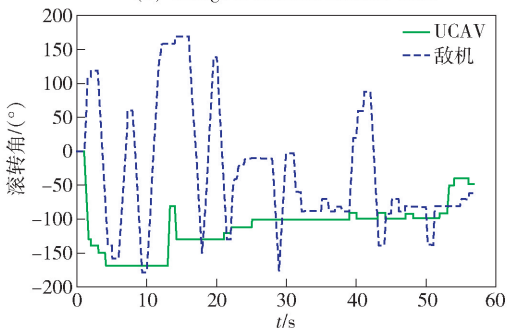
由图 12 可知,本文方法平均决策时间为 0.063 s,



(a) 空战仿真轨迹  
(a) Trajectory of air combat simulation



(b) 态势函数数值变化情况  
(b) Change of situation function value



(c) 滚转角变化情况  
(c) Change of roll angle

图 11 混合机动

Fig. 11 Hybrid maneuver

对比方法为 0.072 s。对于速度为 400 m/s 的飞机来说,0.063 s 内大约飞过了 25 m,相对于空战双方的距离来说可以忽略不计,可认为决策过程中双方态势基本不变,因此本文方法满足实时性要求。

## 6 结论

本文针对自主空战机动决策问题,提出了基于 DNN 的机动决策方法。通过分析网络训练结果、仿真实验结果和决策用时,可以得到主要结论如下:

1) 网络训练结果表明,本文采样方法合理,网络参数设置合理,网络对未来态势的预测准确性较

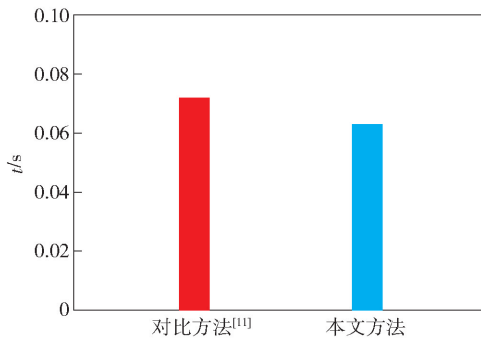


图 12 决策用时

Fig. 12 Time cost of decision-making

高,避免了强化学习的低效率采样。

2) 所设计的动作库能满足决策要求;由相应的态势函数数值变化情况和滚转角变化情况可知,在敌机采取简单直线机动时,本文方法能取得胜利,验证了该方法的有效性。

3) 在敌机采取自主机动的情况下,本文决策方法在均等态势下能达成导弹发射条件,在劣势情况下能通过多次机动转为优势态势并获得胜利,验证了该方法的优越性;同时决策用时较对比方法缩短了约 9 ms。

为使本文提出的决策方法能应对更多不利态势,仍需要进一步扩展机动动作并优化网络结构。

## 参考文献 (References)

[1] AUSTIN F, CARBONE G, FALCO M, et al. Game theory for automated maneuvering during air-to-air combat [J]. Journal of Guidance, 1990, 13(6):1143 - 1149.

[2] VIRTANEN K, KARELAHATI J, RAIIVIO T. Modeling air combat by a moving horizon influence diagram game[J]. Journal of Guidance, Control, and Dynamics, 2006, 29(5):1080 - 1091.

[3] 张涛,于雷,周中良,等. 基于混合算法的空战机动决策[J]. 系统工程与电子技术, 2013, 35(7):1445 - 1450.

ZHANG T, YU L, ZHOU Z L, et al. Decision-making for air combat maneuvering based on hybrid algorithm[J]. Systems Engineering and Electronics, 2013, 35(7):1445 - 1450. (in Chinese)

[4] DUAN H B, ZHANG Y P, LIU S Q. Multiple UAVs/UGVs heterogeneous coordinated technique based on receding horizon control and velocity vector control[J]. Science China, 2011, 54(4):869 - 876.

[5] ZHANG B, GAO X G, RU W. Route planning of UCAV based on artificial potential field method[C]//Proceedings of the International Conference on Management Science and Intelligent Control. Honolulu, HI, US: IEEE, 2011:99 - 102.

[6] LI Y, BAI B D, ZHANG Y N. Improved particle swarm optimization algorithm for fuzzy multi-class SVM[J]. Journal of Systems

- Engineering and Electronics, 2010, 21(3):509-513.
- [7] 傅莉, 谢怀福, 孟光磊, 等. 基于滚动时域的无人机电战决策专家系统[J]. 北京航空航天大学学报, 2015, 41(11):1994-1999.
- FU L, XIE H F, MENG G L, et al. An UAV air-combat decision expert system based on receding horizon control[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(11):1994-1999. (in Chinese)
- [8] HUANG C Q, DONG K S, HUANG H Q, et al. Autonomous air combat maneuver decision using Bayesian inference and moving horizon optimization[J]. Journal of Systems Engineering and Electronics, 2018, 29(1):86-97.
- [9] 王昱, 章卫国, 傅莉, 等. 基于鲁棒优化的无人机电战博弈决策[J]. 系统工程与电子技术, 2015, 37(11):2531-2535.
- WANG Y, ZHANG W G, FU L, et al. Game decision making of UAV aerial combat based on robust optimization[J]. Systems Engineering and Electronics, 2015, 37(11):2531-2535. (in Chinese)
- [10] 徐鸣, 马龙华, 顾江萍, 等. 一种改进的多目标鲁棒优化方法[J]. 控制与决策, 2013, 28(8):1178-1182.
- XU M, MA L H, GU J P, et al. An improved robust multi-objective optimization control method[J]. Control and Decision, 2013, 28(8):1178-1182. (in Chinese)
- [11] 国海峰, 侯满义, 张庆杰, 等. 基于统计学原理的无人作战飞机鲁棒机动决策[J]. 兵工学报, 2017, 38(1):160-167.
- GUO H F, HOU M Y, ZHANG Q J, et al. UCAV robust maneuver decision based on statistics principle[J]. Acta Armamentarii, 2017, 38(1):160-167. (in Chinese)
- [12] SUN T Y, TSAI S J, LEE Y N, et al. The study on intelligent advanced fighter air combat decision support system[C]//Proceedings of the International Conference on Information Reuse and Integration. Waikoloa, HI, US: IEEE, 2006:39-44.
- [13] 孙楚, 赵辉, 王渊, 等. 基于强化学习的无人机自主机动决策方法[J]. 火力与指挥控制, 2019, 44(4):142-149.
- SUN C, ZHAO H, WANG Y, et al. UCAV autonomic maneuver decision-making method based on reinforcement learning[J]. Fire Control and Command Control, 2019, 44(4):142-149. (in Chinese)
- [14] 杜海文, 崔明朗, 韩统, 等. 基于多目标优化与强化学习的空战机动决策[J]. 北京航空航天大学学报, 2018, 44(11):2247-2256.
- DU H W, CUI M L, HAN T, et al. Maneuvering decision in air combat based on multi-objective optimization and reinforcement learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(11):2247-2256. (in Chinese)
- [15] MIRJALILI S, SAREMI S, MIRJALILI S M, et al. Multi-objective grey wolf optimizer: a novel algorithm for multi-criterion optimization[J]. Expert Systems with Applications, 2016, 47:106-119.
- [16] 李磊. 基于值函数逼近与状态空间分解的增强学习方法研究[D]. 长沙:国防科学技术大学, 2011.
- ZUO L. Research on reinforcement learning based on value function approximation and state space decomposition[D]. Changsha: National University of Defense Technology, 2011. (in Chinese)
- [17] 李世豪, 丁勇, 高振龙. 基于直觉模糊博弈的无人机电战机动决策[J]. 系统工程与电子技术, 2019, 41(5):1063-1070.
- LI S H, DING Y, GAO Z L. UAV air combat maneuvering decision based on intuitionistic fuzzy game theory[J]. Systems Engineering and Electronics, 2019, 41(5):1063-1070. (in Chinese)
- [18] 伊恩·古德费洛, 约书亚·本吉奥, 亚伦·库维尔. 深度学习[M]. 赵申剑, 黎彧君, 符天凡, 等, 译. 北京:人民邮电出版社, 2017:126-188.
- GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. ZHAO S J, LI Y J, FU T F, et al, translated. Beijing: Press of People's Post and Telecommunications, 2017:126-188. (in Chinese)
- [19] WILLIAMS P. Three-dimensional aircraft terrain-following via real-time optimal control[J]. Journal of Guidance, Control, and Dynamics, 1990, 13(6):1146-1149.
- [20] FRED A, GIRO C, MICHAEL F, et al. Automated maneuvering decisions for air-to-air combat[C]//Proceedings of the Guidance, Navigation and Control Conference. Monterey, CA, US: AIAA, 1987:659-669.
- [21] 韩立群. 神经网络[M]. 北京:北京邮电大学出版社, 2006:51-66.
- HAN L Q. Artificial neural network[M]. Beijing: Beijing University of Posts and Telecommunications Press, 2006:51-66. (in Chinese)
- [22] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification[C]//Proceedings of International Conference on Computer Vision. Santiago, Chile: IEEE, 2015:1026-1034.
- [23] MCGREW J S, HOW J P, WILLIAMS B. Air-combat strategy using approximate dynamic programming[J]. Journal of Guidance, Control, and Dynamics, 2010, 33(5):1641-1654.
- [24] SHAW R L. Fighter combat tactics and maneuvering[M]. Annapolis, MD, US: Naval Institute Press, 2007:10-13.