

# 基于蒙特卡洛树搜索方法的空战机动决策

何 旭, 景小宁, 冯 超

(空军工程大学航空航天工程学院, 西安, 710038)

**摘要** 针对空战中双方的机动决策问题, 考虑角度、距离、速度变化以及高度优势函数, 以赢取最大空战优势期望为目标, 建立了空战双方机动决策模型, 并引入蒙特卡洛树搜索方法进行求解, 搜索策略采取极大极小搜索算法。针对双机对抗, 通过仿真给出空战态势及空战优势变化曲线, 结果表明: 我机在整个空战过程中绝大部分是处于优势地位的, 验证了使用蒙特卡洛树搜索方法进行空战机动决策的可行性, 同时对算法的实时性进行了仿真, 计算机动作所用时间众数位于 50~100 ms 区间内, 达到了实时性要求。

**关键词** 空战; 机动决策; 空战优势; 蒙特卡洛树搜索

**DOI** 10.3969/j.issn.1009-3516.2017.05.007

**中图分类号** V325 **文献标志码** A **文章编号** 1009-3516(2017)05-0036-06

## Air Combat Maneuver Decision Based on MCTS Method

HE Xu, JING Xiaoning, FENG Chao

(Aeronautics and Astronautics Engineering College, Air Force Engineering University,  
Xi'an 710038, China)

**Abstract:** Aimed at both sides' maneuver decision problems in air combat that changes of angle, distance, velocity and altitude dominant function are considered to make maximum air combat superiority expectation as target, the Monte Carlo tree search method is introduced. The search strategy is used by the mini-max search algorithm. Against the confrontation between the two aircraft, the simulation results of air combat situation and advantage change curves show that our fighter is in the dominant position for the most part in the whole process in air combat by using MCTS method. At the same time, the real-time performance of the algorithm is simulated. The results show that the model time of the algorithm computing maneuver action is located at 50~100 ms range to meet the needs of real-time.

**Key words:** air combat; maneuver decision; air combat superiority; MCTS

空战机动决策是指空战过程中对作战态势进行分析, 选择消灭对方保存自己的攻防策略<sup>[1]</sup>, 可以显著减轻驾驶员负担、提高空战效能<sup>[2]</sup>。

目前, 对于空战机动决策有很多研究: 张立鹏等采用了模糊推理的方法<sup>[3]</sup>, Nusyirwan I F 等使用智

能微分进行求解<sup>[4]</sup>, Kai V、万伟等使用了影响图<sup>[5-6]</sup>, 顾佼佼等采取了 Memetic 算法<sup>[7]</sup>, 但模糊推理对控制量操作不灵活, 微分对策求解困难, 且计算时间长不满足实时性要求。本文采用蒙特卡洛树搜索(Monte Carlo tree search, MCTS<sup>[8-9]</sup>)方法建立

**收稿日期:** 2017-03-09

**基金项目:** 国家自然科学基金(71501184); 航空科学基金(20155196022)

**作者简介:** 何 旭(1993—), 男, 江苏徐州人, 硕士生, 主要从事航空兵任务规划与作战效能评估研究。E-mail: hexupaper@163.com

**引用格式:** 何旭, 景小宁, 冯超. 基于蒙特卡洛树搜索方法的空战机动决策 [J]. 空军工程大学学报(自然科学版), 2017, 18(5):36-41. HE Xu, JING Xiaoning, FENG Chao. Air Combat Maneuver Decision Based on MCTS Method [J]. Journal of Air Force Engineering University (Natural Science Edition), 2017, 18(5):36-41.

作战飞机运动学和动力学方程,构建合适的空战优势函数,选取空战优势最大的机动动作进行机动,并通过仿真验证了决策的实时性、有效性和可行性。

## 1 空战机动决策建模

空战机动动作决策的实质就是在当前战场环境下,选择合适的基本机动动作,使得空战优势函数达到最大。

### 1.1 基本机动动作库

空战机动决策集合主要分为“典型战术动作决策集模型(典型战术动作库)”与“基本机动动作决策集模型(基本机动动作库<sup>[10]</sup>)”。本文主要使用文献[10]中的7种基本机动动作作为空战决策的机动动作集合。

设空战过程中机动动作状态空间为  $E = \{E_1, E_2, \dots, E_7\}$ , 分别对应7种飞行状态。由于整个空战过程就是敌我合理选择基本机动动作,所以整个空战过程可以描述为一个多叉树上从根节点到叶节点的一条路径。见图1,空战就是敌我双方依次在空战多叉树上进行机动动作选择。

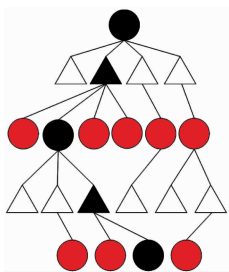


图1 空战机动动作选择过程

Fig. 1 Air combat maneuver selection

### 1.2 作战飞机运动学和动力学方程

选择相应的机动动作之后,需要根据机动动作对应的控制量  $\mathbf{u} = \{\phi, u, n\}$  求解出作战飞机的状态量  $\mathbf{x} = \{x, y, z, v, \gamma, \psi\}$ , 其中  $\phi$  表示滚转角,  $u$  表示油门位置,  $n$  表示过载,  $x, y, z$  表示飞机的空间坐标,  $v$  表示速度,  $\gamma, \psi$  表示航迹倾角和航迹偏角。

假设作战飞机在运动过程中:①地球曲率影响忽略不计;②重力加速度  $g$  忽略作战飞机高度变化的影响;③战机质量不变。基于以上假设,作战飞机在航迹坐标系上的质点运动学方程为<sup>[11]</sup>:

$$\begin{aligned} \dot{x} &= v \cos \gamma \cos \psi \\ \dot{y} &= v \cos \gamma \sin \psi \\ \dot{z} &= v \sin \gamma \end{aligned} \quad (1)$$

式中:  $\dot{x}, \dot{y}, \dot{z}$  表示速度在3个坐标轴上的分量。则联立作战飞机运动学方程:

$$\begin{aligned} \dot{v} &= g(u - \sin \gamma) \\ \dot{\gamma} &= \frac{g}{v} (n \cos \phi - \cos \gamma) \\ \dot{\psi} &= \frac{gn \sin \phi}{v \cos \gamma} \end{aligned} \quad (2)$$

可根据作战飞机控制量  $\mathbf{u}$  得到其状态量  $\mathbf{x}$ 。

### 1.3 空战优势函数

对作战飞机机动动作的选择通过依据当前空战态势构建对应的空战优势函数来完成,使我机对敌机的空战优势达到最大。本文采取文献[12~14]空战优势函数的解决方法,对角度优势函数、距离优势函数、速度变化优势函数、高度优势函数建模,将这4种函数值加权求和得到空战优势函数。下面给出具体求解公式。

#### 1.3.1 角度优势函数

使用 Simpson 公式定量描述角度优势函数:

$$S_a = \frac{180^\circ - (A_{TA} + A_{OT})}{180^\circ} \quad (3)$$

式中:  $A_{TA}$  表示目标角;  $A_{OT}$  表示进入角。

#### 1.3.2 距离优势函数

相对距离在空战过程中是一个十分重要的参数。我机距离优势函数的计算,需要建立包括最大攻击距离  $r_{\max}$ 、最佳攻击距离(不可逃逸区)  $r_{\text{opt}}$  和最小攻击距离  $r_{\min}$  在内的机载武器攻击包线的索引表。计算公式为:

$$S_r = \frac{\cos(\theta) + 1}{2}$$

$$\theta = \begin{cases} \pi \left( \frac{r_{\text{opt}} - r}{r_{\text{opt}} - r_{\min}} \right), & r_{\min} \leq r \leq r_{\text{opt}} \\ \pi \left( \frac{r - r_{\text{opt}}}{r_{\max} - r_{\text{opt}}} \right), & r_{\text{opt}} < r \leq r_{\max} \end{cases} \quad (4)$$

我机距离优势处于范围  $[-1, 1]$  之间,当本机被敌机攻击时,  $S_r = -1$ ; 当敌机处于本机最佳攻击距离内,  $S_r = 1$ 。

#### 1.3.3 速度变化优势函数

文献[12~13]采取速度优势函数,但对于近距格斗而言,使用速度变化进行衡量更有意义,速度变化越激烈的战机越不容易被敌方瞄准打击,所以使用过载  $n$  作为速度变化优势。

$$S_v = 0.5 + (\eta_{\text{range}} - 0.5) \max \left\{ 0, 1 - \frac{|n|}{e_g} \right\} \quad (5)$$

式中:  $e_g$  为预设常数;  $\eta_{\text{range}}$  为最大范围(令其为1)。

#### 1.3.4 高度优势函数

导弹射程随两机之间的高低位置而受影响:两机携带有同型导弹,位于高位的飞机,其导弹的最大发射距离要大于位于低位的飞机。并且,位于高位的飞机可以利用高度优势进行俯冲,快速进入攻击区并退出战斗。文献[1]等默认我机对敌机的高差

越大越有优势,但若高差太大,导弹的命中概率势必会下降,因而调整高度优势函数解决此问题,高度优势函数为:

$$S_{\Delta h} = \begin{cases} 0.5 - \frac{1}{2} \left( \frac{\Delta h}{500} \right)^2, & \Delta h < 0 \\ 1 - \frac{1}{2} \sqrt{\left( \frac{500 - (\Delta h)}{500} \right)}, & \Delta h \geq 0 \end{cases} \quad (6)$$

式中: $\Delta h$ 表示我机相对于敌机的高差。

综上,得到空战决策过程的优势函数为:

$$S = W_a S_a + W_r S_r + W_v S_v + W_{\Delta h} S_{\Delta h} \quad (7)$$

式中: $W_a, W_r, W_v, W_{\Delta h}$ 分别为4种优势函数在空战优势中的权系数,要求 $\sum W_i = 1, (i = a, r, v, \Delta h)$ 。

在不同作战态势下,空战优势函数各参数的重要程度不同,对此,通过调整权系数来体现,第*i*空战优势函数参数的权重计算公式<sup>[15]</sup>如下:

$$W_i = W_{i0} + k_i e^{-\left(\frac{r}{1.5R}\right)^2} \quad (8)$$

式中: $W_{i0}$ 为各参数的初始权重; $k_i$ 为其调节系数, $\sum k_i = 0$ ;  $r$ 为敌我距离; $R$ 为空空导弹射程。

#### 1.4 空战态势

根据空战优势函数,空战态势与进入角( $A_{OT}$ )、视线角( $A_{TA}$ )、相对距离( $r$ )、过载( $n$ )、相对高度( $\Delta h$ )有关。若当前时刻为*t*决策时刻,则*t*决策时刻对应的空战态势为 $\mathbf{x}(t) = \{A_{OT}(t), A_{TA}(t), r(t), n(t), \Delta h(t)\}$ ;相应的,根据空战态势 $\mathbf{x}(t)$ ,可以得到此空战态势下的空战优势 $S(t)$ ;机动动作的状态空间为 $E = \{E_1, E_2, \dots, E_7\}$ ,使用 $a(t)$ 表示在第*t*决策时刻选择的机动动作。

假设已知*t*-1决策时刻的空战态势 $\mathbf{x}(t-1)$ 以及机动动作 $a(t-1)$ ,现在需要对*t*决策时刻的机动动作进行决策。传统方法主要以试探法为主,在*t*决策时刻分别选取不同的机动动作,估计相应的空

战态势,得到空战优势 $S(t)$ ,选择使得空战优势 $S(t)$ 最大的机动动作 $c_i$ 作为该时刻的机动动作。本文设计一种新的机动动作决策方法,假如作战飞机此时选择机动动作 $c_i$ ,则作战飞机在*t*时刻的空战优势为 $\{S(t) | a(t) = c_i\}$ 。设定正整数*k*,在*t*时刻选择机动动作 $c_i$ 的条件下,作战飞机在*t+k*时刻选择机动动作 $c_j$ 的空战优势为 $\{S(t+k) | a(t) = c_i, a(t+k) = c_j\}$ ,对于任意动作 $c_j$ ,在*t*时刻选择机动动作 $c_i$ 的前提下,未来*k*步所有决策过程的空战优势期望为

$$E[S(t) | a(t) = c_i] =$$

$$\sum_{j=1}^7 \sum_{m=1}^k [S(t+m) | a(t) = c_i, a(t+m) = c_j] \quad (9)$$

得到*t*时刻最大空战优势期望对应的机动动作:

$$c_i = \arg \max \{E[S(t) | a(t) = c_i]\} \quad (10)$$

这就是本文选择机动动作的根本原则。由于求解未来*k*步所有决策过程的空战优势期望计算量较大,使用穷举方法不现实,故本文使用蒙特卡洛树搜索的方式来进行求解。

## 2 基于蒙特卡洛树搜索方法的空战机动动作选择

空战机动动作选择是通过试探未来的机动动作来确定下一步决策的空战优势概率。蒙特卡洛树搜索(Monte Carlo tree search, MCTS)是一种平衡历史收益和未来开发的树搜索策略。其基本原理是:首先随机选择机动动作策略,然后通过预期收益来更新原来选择策略的价值,设定随机策略选择的概率,与先前计算出的策略价值成正比,通过大量反复随机模拟,让最好的策略出现。简单的说,MCTS分为4个部分<sup>[16]</sup>,其原理见图2。

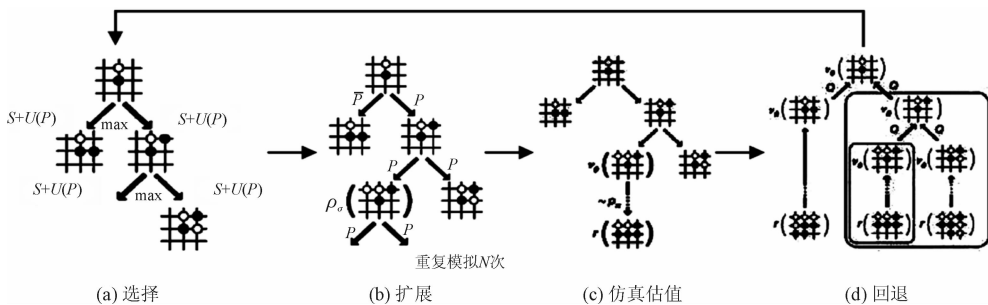


图2 蒙特卡洛树搜索算法

Fig. 2 The Monte Carlo tree search algorithm

本文基于蒙特卡洛树搜索来进行空战机动动作决策。具体决策流程如下:

**步骤1** 选择部分。假设当前为第*t*-1决策时刻,其空战态势为 $\mathbf{x}(t-1)$ ,机动动作为 $a(t-1)$ ,需

要对第  $t$  决策时刻的机动动作进行决策。首先依次选择其中基本动作,分别计算选择不同基本机动动作之后得到的空战优势  $\{S(t) | a(t) = c_i\}, i \in 1, 2, \dots, 7$ 。将选取不同机动动作得到的空战优势作为每个机动动作节点的值;若蒙特卡洛算法没有达到终止条件且有未探索的节点时,该节点就是可以扩展的;否则终止算法,转步骤 4;

**步骤 2 扩展部分。**若蒙特卡洛搜索算法没有达到终止条件,则可以继续向下选择机动动作,将新的机动动作添入蒙特卡洛树中;

**步骤 3 仿真部分。**在新得到的机动动作节点上,计算当前的空战态势;

**步骤 4 反馈更新。**将更新得到的节点的值反馈回根节点,沿途更新节点的统计评估值。

蒙特卡洛搜索框架使用 UCT 搜索策略对空战机动动作节点进行搜索。从根节点出发,记前非叶子节点为  $n$ ,节点  $n$  的空战优势值为  $S(n)$ 。对于每一个节点  $n$  的孩子  $n_i \in ch(n)$ ,采用选择策略得到节点的一个评估值  $r_i$ ,根据评估值对节点的下一步进行选择,这个过程不断重复,直到满足终止条件(当向下搜索超过 5 步时,或者整体没有未探索的节点时)。如果当前节点为评估值最大节点(由于采用极大极小搜索,所以选择的点只可能是最大节点或最小节点),则评估值计算公式使用正号,并进行下一次选择;如果当前节点为评估值最小节点时,则评估值计算公式用负号,在整个机动决策树中使用下式选择评估值  $r_i$  最小的值进行搜索。评估值  $r_i$  计算:

$$r_i = v_i \pm \omega c \sqrt{\frac{2 \ln \sum_i T_i}{T_i}} + \sigma \quad (11)$$

式中: $v_i$  是以节点  $n_i$  为根节点时所有已知结果的节点态势优势值  $S(n_i)$  的平均数,反映了观测节点  $n_i$  能够提供的回报值的期望; $T_i$  是节点  $n_i$  的访问次数; $\sum_i T_i$  是节点  $n$  的访问次数; $\omega c$  是一个人工给定的常数,用于调节回报值和未探索节点的平衡性; $\sigma$  为一较大正数,用来保证评估值非负。而最后节点  $n$  的每一个节点的孩子  $n_i \in ch(n)$  的评估值  $r_i$  就是下一步节点的平均优势期望。这里需要对获胜概率归一化。则第  $t$  决策时刻选择不同机动动作的未来空战优势的平均期望为:

$$E[S(t) | a(t) = c_i] = \frac{r}{|\sum_i r_i|} \quad (12)$$

因为机动动作决策树仅仅选择了一步,即我方的机动动作,这时的空战态势表示的是我方下一步决策时刻和敌方当前决策时刻的空战态势,当再向

下搜索一次时,才能更新敌方的空战态势。由于 MCTS 中采用的具体搜索策略是极大极小搜索算法,其本质在于尽可能提高我方空战优势而减少敌方的空战优势,故这是一种可行的决策方法。

### 3 实验仿真

本文通过仿真来验证方法的有效性与实时性。实验硬件配置为 Intel(R) Xeon(R) CPU E5-2643,主频 3.50 GHz,内存 4 GB。软件环境为 Microsoft Windows 7 32 位操作系统,在 Matlab2014 上运行。

#### 3.1 有效性验证

作战想定如下:以敌我双方一对一空战对抗为背景,目标机依照设定的路径进入空战,而我方作战飞机依据本文确定的机动动作决策方法进行机动。空战过程设定的决策步长为 0.2 s。我机初始位置为  $[10, 10, 5]$  km,初始速度为 180 m/s,航向角为  $300^\circ$ ,俯仰角为  $-30^\circ$ ,滚转角为  $-50^\circ$ ;敌机初始位置为  $[8, 7, 5]$  km,初始速度为 170 m/s,航向角为  $110^\circ$ ,俯仰角为  $0^\circ$ ,滚转角为  $-90^\circ$ 。双方的空战态势见图 3,其中红方表示我机,蓝方表示敌机。

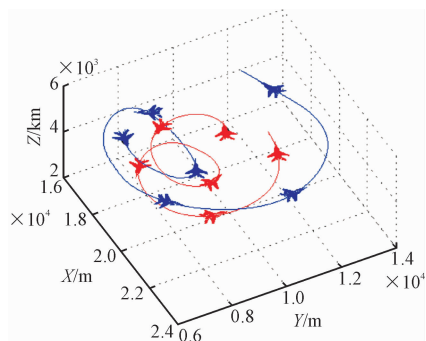


图 3 敌我双方空战过程三维航迹图

Fig. 3 A three-dimensional chart of air combat between the enemy and mine

图 4 给出整个空战过程中敌我双方的空战态势变化曲线。根据本文给出的空战优势函数,可以给出红蓝双方在整个空战过程空战优势变化曲线见图 5(红方用虚线表示,蓝方用实线表示)。从空战优势变化曲线可以看出,我机在整个空战过程中绝大部分是处于优势地位的。该决策方法可行。

#### 3.2 实时性验证

虽然实验仿真 1 验证了蒙特卡洛树搜索方法的可行性,但若运行计算时间过长,算法就失去了实战价值。所以进行实验仿真 2,讨论算法是否满足空战实时性。实验仿真 1 中,红方一共决策 330 次,花费时间 66 s,本实验仿真对这 330 次决策过程中每次决策所用的时间进行统计,做出其统计频数图,见图 6。

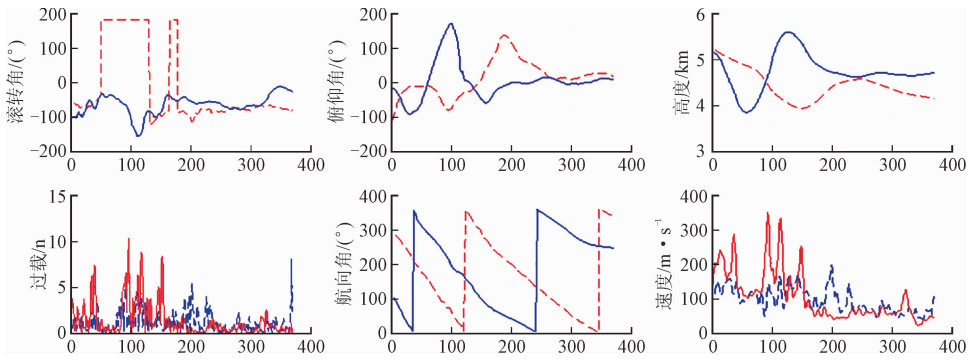


图4 空战态势变化曲线

Fig. 4 Air combat posture change curve

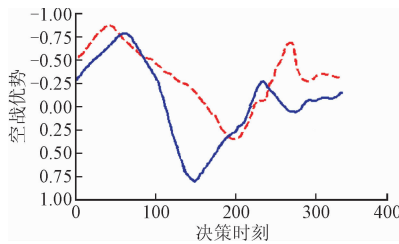


图5 空战优势变化曲线

Fig. 5 The change of air combat superiority curve

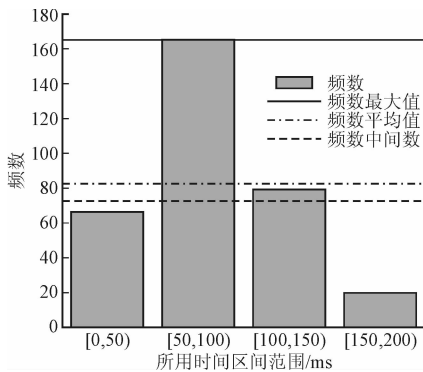


图6 算法所用时间频数统计图

Fig. 6 The figure of using time frequency statistics for the algorithm

从图6可以看出,算法计算机动动作所用时间的众数位于50~100 ms区间内,可以认作算法是满足实时性要求的。当实验环境改为机载平台时,理论上而言时间不会增加,满足要求。

## 4 结语

将作战飞机的机动过程划分为若干基本机动动作,建立飞机运动学和动力学方程,根据控制量求解作战飞机在空战过程中的状态量。构建了合适的空战优势函数,将对方法的可行性与实时性进行了验证。仿真实验表明,该方法是有本文采用蒙特卡洛搜索方法,不仅可以估计未来一步的机动动作,而且可以估计未来若干步的机动动作,从而选择使得

空战优势期望最大的机动动作。

## 参考文献(References):

- [1] 阮毓巍,寇英信,徐安,等. 基于二元模糊比较法的单步空战机动决策研究[J]. 指挥控制与仿真, 2012, 34(5): 10-13.  
RUAN C W, KOU Y X, XU An, et al. Research on Single Step Manuevering Decision in Air Combat Based on Binary Fuzzy Comparison Method [J]. Command Control & Simulation, 2012, 34(5): 10-13. (in Chinese)
- [2] JAMES S, JONATHAN P, BRIAN W, et al. Air combat strategy using approximate dynamic programming [J]. Journal of Guidance, Control, and Dynamics, 2010, 33(5): 1641-1654.
- [3] 张立鹏,魏瑞轩,李霞. 无人作战飞机空战自主战术决策方法研究[J]. 电光与控制, 2012, 19(2): 92-96.  
ZHANG L P, WEI R X, LI X. Autonomous tactical decision-making of UCAVs in air combat [J]. Electronics Optics & Control, 2012, 19(2): 92-96. (in Chinese)
- [4] NUSYIRWAN I F, BILL C. Factorial analysis of a real time optimization for pursuit-evasion problem [C]// Proc. of the 46th AIAA Aerospace Sciences Meeting and Exhibit, 2008: 195-198.
- [5] KAI V, JANNE K, TUOMAS R. Modeling air combat by a moving horizon influence diagram game [J]. Journal of Guidance, Control, and Dynamics, 2006, 29(5): 1080-1091.
- [6] 万伟,姜长生,吴庆宪. 单步预测影响图法在空战机动决策中的应用[J]. 电光与控制, 2009, 16(7): 13-17.  
WAN W, JIANG C S, WU Q X. Application of One-Step Prediction Influence Diagram in Air Combat Manuevering Decision [J]. Electronics Optics & Control, 2009, 16(7): 13-17. (in Chinese)
- [7] 顾佼佼,赵建军,刘卫华. 基于博弈论及 Memetic 算

- 法求解的空战机动决策框架 [J]. 电光与控制, 2015, 22(1): 20-23.
- GU J J, ZHAO J J, LIU W H. Air Combat Maneuvering Decision Framework Based on Game Theory and Memetic Algorithm [J]. Electronics Optics & Control, 2015, 22(1): 20-23. (in Chinese)
- [8] COULOM R. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search [C]//In 5th International Conference on Computers and Games. Turin, Italy: Springer Berlin Heidelberg, 2006: 72-83.
- [9] TESAURO G, GALPERING G. On-line Policy Improvement Using Monte-Carlo Search [C]// Denver, CO, USA; in Advances in Neural Information Processing(NIPS), 1996: 1068-1074.
- [10] AUSTIN F. Automated Aneuvering Decisions for Air-to-air Combat [R]. AIAA 87-2393.
- [11] VIRTANENK, EHTAMOH, RAIVIOT, et al. Visual Interavtive Aircraft Trajectory Optimization [J]. IEEE Transactions on Systems, Man, and Cybernetics-part C: Applications and Reviews, 1999, 29(3): 2280-2285.
- [12] 顾佼佼, 刘卫华, 姜文志. 基于攻击区和杀伤概率的视距内空战态势评估 [J]. 系统工程与电子技术, 2015, 37(6): 1306-1312.
- GU J J, LIU W H, JIANG W Z. WVR Air Combat Situation Assessment Model Based on Weapon Engagement Zone and Kill Probability [J]. Systems Engineering and Electronics, 2015, 37(6): 1306-1312. (in Chinese)
- [13] JOSEPH A K, BLACKSBURG, VIRGINIA. The Analysis of a Generic Air-to-Air Missile Simulation Model [R]. ADA, 2003.
- [14] 李望西, 黄长强, 王勇, 等. 三维空间空战态势评估角度优势建模与仿真 [J]. 电光与控制, 2012, 19(2): 21-25.
- LI W X, HUANG C Q, WANG Y, et al. Modeling and Simulation of Air Combat Situation Assessment's Angel Superiority in Three Dimensional Space [J]. Electronics Optics & Control, 2012, 19(2): 21-25. (in Chinese)
- [15] 张涛, 于雷, 周中良, 等. 基于变权重伪并行遗传算法的空战机动决策 [J]. 飞行力学, 2012, 30(5): 470-474.
- ZHANG T, YU L, ZHOU Z L, et al. Decision-making for Air Combat Maneuvering Based on Variable Weight Pseudo-parallel Genetic Algorithm [J]. Flight Dynamics, 2012, 30(5): 470-474. (in Chinese)
- [16] 陶九阳, 吴琳, 胡晓峰. AlphaGo 技术原理分析及人工智能军事应用展望 [J]. 指挥与控制学报, 2016, 2(2): 114-120.
- TAO J Y, WU L, HU X F. Principle Analysis on ALphaGo and Perspective in Millitary Application of Artificial Intelligence [J]. Journal of Command and Control, 2016, 2(2): 114-120. (in Chinese)

(编辑: 徐敏)