

文章编号: 2097-1974(2025)02-0001-10

DOI: 10.7654/j.issn.2097-1974.20250201

机动突防/精确打击在线制导方法

陈万春, 王楠翔, 龚晓鹏
(北京航空航天大学宇航学院, 北京, 102206)

摘要: 为了应对空空导弹机动突防精确打击在线制导问题, 推导了非零阶制导系统脱靶量的解析解, 将所得结果推广至比例导引之外的先进制导律, 并以此为基础分析了不同机动形式下的突防逃逸效果, 从理论上解释了阶跃机动和蛇形机动脱靶量的典型特征, 对最优机动策略的影响因素展开研究, 通过合理调整该策略的机动时刻和方向, 提升了规避拦截的效果。同时考虑突防打击一体化智能制导律, 提出了一种涵盖自适应步长疏化采样、融合模型的先验知识和失效场景中补差训练的ASTRAL深度强化学习算法。仿真结果表明, 通过ASTRAL算法训练的进攻弹智能体能够有效规避防御弹并在突防过程中保持对目标的攻击能力, 最终生成的制导律拥有良好的工程可行性。

关键词: 突防; 脱靶量; 解析解; 最优机动; 深度强化学习

中图分类号: TJ765; V448 **文献标识码:** A

Online Guidance Law of Maneuvering Penetration and Precision Strike

CHEN Wanchun, WANG Nanxiang, GONG Xiaopeng
(School of Astronautics, Beihang University, Beijing, 102206)

Abstract: To address the issues of penetration and precision attacking for air-to-air missiles, an analytical solution for the miss distance of a non-zero-order guidance system is derived. The solution is then extended to advanced guidance laws beyond proportional navigation guidance. Based on these results, the effectiveness of evasion under different maneuvering strategies is analyzed which theoretically explaining the typical characteristics of miss distances in step and serpentine maneuvers and factors influencing the optimal maneuvering strategy are studied. Through reasonable adjustment of maneuver timing and direction, the effectiveness of evasion is improved. Furthermore, considering the integration of penetration and attacking through intelligent guidance laws, an adaptive-stepsize sparsified sampling prior model theory infusion remedial training agent learning (ASTRAL) is introduced. Simulation results show that the missile agent trained with the ASTRAL can effectively evade defender while maintaining its attacking capability during penetration. In addition, the generated guidance law demonstrates good engineering feasibility.

Keywords: penetration; miss distance; analytical solution; optimal maneuver; deep reinforcement learning

0 引言

现代空战具有高度复杂、动态且充满不确定性的特点, 攻防对抗异常激烈, 战场态势瞬息万变。在作战过程中, 战机通常通过周期性机动、置尾逃逸、发射防御弹等方式来摆脱导弹的追击。这些方式可以破坏导弹的锁定, 增加导弹脱靶的可能性, 而由于对进攻导弹的特性深入分析不足, 在机动突防策略选择上存在一定的盲目性。因此, 研究导弹的高阶制导特性, 从解析解角度探讨脱靶量的影响因素, 揭示不同机动突防策略的有效性, 对于未来空战突防打击一体化的理论支撑具有重要意义。与此同时, 战机发射防

御弹拦截来袭导弹成为空战中的一种新型作战方式, 形成了载机、导弹与防御弹三方参与博弈的复杂局面。因此, 提出适合进攻导弹突破防御弹拦截并精确打击载机的攻击策略, 提升对这一复杂战场场景的应对能力, 显得尤为必要。

脱靶量是评估导弹制导律与进攻导弹机动突防/精确打击策略有效性的关键指标。解析解(或闭式解)由于其物理意义清晰、表达简洁高效, 具有明显的优势。然而, 大部分现有研究通过仿真来计算脱靶量, 对解析方法的探讨较为有限^[1]。同时, 当前脱靶量解析研究中的导弹模型多采用简化的理想化动力学

或一阶延迟系统,这往往导致脱靶量被低估^[2]。脱靶量解析解的研究起源于比例导引制导律(Proportional Navigation Guidance, PNG),对于零延迟导弹追踪非机动目标的情形,Guelman^[3-4]对纯比例导引(Pure Proportional Navigation, PPN)进行了定性研究,并推导出真比例导引(True Proportional Navigation, TPN)的解析解。此外,假设使用PPN的零延迟导弹模型,推导出了针对机动目标的侧向加速度的解析解^[5]和攻击非机动目标的导弹与目标相对距离的解^[6]。对于使用TPN攻击机动目标的导弹,其解析解以剩余距离为函数形式给出^[7]。此外,对于机动和非机动目标,广义比例导引(Generalized Proportional Navigation, GPN)的解析解也有所讨论^[8-9]。Moon等^[10]提出了一种基于解析解的终端加速度稳定制导律,用于拦截非机动目标并约束攻击角。然而,大多数现有研究未充分考虑导弹动力学的复杂性,过度依赖简化的一阶系统模型,从而可能显著低估脱靶量。同时,过去的研究主要集中在加速度或导弹与目标相对距离的解析解推导上,鲜少涉及脱靶量的解析解,因此未能全面评估目标机动性和制导系统阶数对脱靶量的影响。

除了对脱靶量解析解的研究不足外,基于这些解的机动策略评估同样存在欠缺。Zarchan^[11]探讨了目标机动频率、幅度、拦截弹时间常数及有效导引比对线性制导系统脱靶量的影响,分析了通过滤波估计机动频率对制导律的影响^[12]。为进一步提高蛇形机动的效果,Yanushevsky^[13-14]采用频域分析研究了最优机动频率。Rusnak等^[15]基于传统蛇形机动的改进,提出了螺旋变频机动,并提出了一种新的制导律作为对策。此外,桶滚机动作为三维空间中最具代表性的周期性机动,通常由飞机执行以规避来袭导弹。近似解析解的研究讨论了桶滚机动的特点,并通过仿真验证了大过载桶滚机动可与最优机动产生类似效果^[16]。基于虚拟滑动目标或二阶锥规划的螺旋机动策略随后也被提出,能够在攻击静止目标时提高抗拦截能力^[17-18]。然而,程序化机动策略未考虑拦截弹的运动信息,导致目标无法有效适应高度动态的战斗环境,进而可能导致规避失败。从而研究者们提出了基于微分对策^[19-20]、最优化^[21-22]及人工智能^[23-24]的机动策略。尽管这些先进机动策略能有效应对复杂战场环境,但它们通常需要大量的测量数据和计算资源。因此基于解析解研究机动效果的不足,限制了对弹目交战过程的深入理解,进一步凸显了提出脱靶量解析解

对机动突防/精确打击策略的重要理论意义。

三体博弈问题的研究可以追溯到20世纪70年代Boyell^[25-26]对于舰船防御鱼雷或导弹的研究。Boyell在简化条件下,假设所有参与者沿平行航线移动且不作机动,分析了防御弹的最小发射距离和最佳发射时机,这种假设在实际应用中并不现实,且仅从几何角度进行了静态分析。进入21世纪后,随着技术进步与需求的增长,三体博弈问题的研究再次受到关注。Rusnak^[27-28]将这一问题类比为“二队三人”的博弈,采用多目标优化和微分对策理论进行建模与求解,并通过阶跃函数模拟防御弹拦截成功后的消失情况。Ratnoo等^[29]从视线(Line-of-Sight, LOS)制导的角度提出了一种策略,使防御弹始终位于目标和进攻弹之间,即便在防御弹速度与机动性较差的情况下,也能有效拦截,还对比了目标不作机动时,防御弹与进攻弹在不同经典制导律下的交战结果^[30]。Yamasaki等^[31-33]基于LOS制导提出了三角拦截制导律和机载视线指令制导律等。

在研究抽象化的一般性场景时,通常将其称为目标-攻击者-防御者(Target-Attacker-Defender, TAD)博弈问题。Garcia等^[34]分析了这一博弈的动态特性,重点探讨了攻击者占优区域的划分以及防御者的最优拦截策略。在后续研究中,他们进一步完善了攻击者的最优策略,确保这些策略与Hamilton-Jacobi-Isaacs方程的一致性^[35]。Liang等^[36]引入了界栅分析方法,用于划分攻击者与目标-防御者之间的获胜区域,并通过显式策略和几何分析制定相应的战略决策。随后的研究探讨了针对快速、非自杀式攻击者的阶段性策略,并设计了基于捕获半径的战术^[37]。然而,现有的空中追逃博弈建模方法忽略了如转弯半径和系统对命令的动态响应(如振荡或延迟)等关键因素。这种过于简化的建模方法未能充分利用动力学的普遍原理,且忽视了飞行器的特性,导致所得到的策略通常缺乏现实可行性。为了更贴近实际作战环境,近年来的研究还考虑了能力边界的强约束,使用卡尔曼滤波在不完美信息条件下进行状态估计^[38-39],同时提出了多飞行器协同攻击主动防御目标的策略^[40]。此外,深度强化学习在追逃博弈中的应用也取得了显著进展。Wang等^[41]利用强化学习方法研究了多个追击者合作捕捉高级躲避者的问题,成功克服了微分对策或最优控制在复杂不确定环境中的合作追击难题。English等^[42]则将强化学习算法应用于TAD粒子追逃场景,攻击者能够感知防御者的存在并进行规避,最终

成功捕获目标，表现出比传统导引算法更优的性能，并分析了逃逸边界，该模型较为简单，未涉及动力学特征。因此，建立具备动力学特征的智能突防打击一体制导律具有重要意义。

1 问题描述

基于现代空战机动突防/精确打击任务的典型战斗场景，图1将交战场景提炼为一个惯性坐标系 x_iOy_i 中的几何表示。该模型包含了进攻弹 M 、目标 T 、防御弹 D 3个参与者。每个飞行器的飞行状态由速度倾角 γ 、飞行速度 v 和实际加速度 a 描述，分别用下标 M 、 T 和 D 表示是进攻弹、目标、防御弹的相关量。

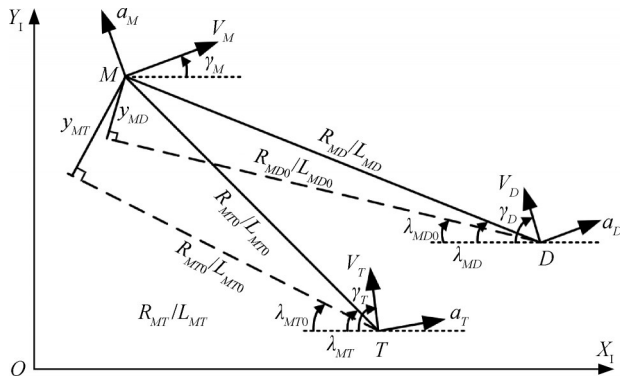


图1 三体博弈交战几何
Fig.1 T、M、D engagement geometry

交战过程分为两个相互关联但又独立的部分：进攻弹对目标的精确打击和进攻弹对防御弹的机动突防。在这类交战中， R 为各个飞行器之间的距离， L 为视线， λ 为视线角。交战开始时的初始条件用下标0表示。如图1所示，进攻弹-目标和进攻弹-防御弹之间交战的非线性运动学方程可写为

$$\dot{R}_{Mi} = -V_i \cos(\gamma_i - \lambda_{Mi}) - V_M \cos(\gamma_M + \lambda_{Mi}) \quad (1)$$

$$\dot{\lambda}_{Mi} = -\frac{1}{R_{Mi}} (V_i \cos(\gamma_i - \lambda_{Mi}) + V_M \sin(\gamma_M + \lambda_{Mi})) \quad (2)$$

式中 M 为敌方进攻弹； i 为我方目标或者防御弹，即 $i \in \{T, D\}$ 。速度倾角的变化率可以表示为

$$\dot{\gamma}_j = \frac{a_j}{V_j}, j \in \{M, T, D\} \quad (3)$$

此外，假设每个飞行器的动力学模型可以表述成任意阶的线性系统：

$$\dot{x}_j = A_j x_j + B_j u_j, j \in \{M, T, D\} \quad (4)$$

$$a_j = C_j x_j + d_j u_j, j \in \{M, T, D\} \quad (5)$$

式中 x_j 为飞行器的内部状态量； a_j 作为系统的输出表示飞行器垂直速度方向的实际加速度； u_j 作为系统

的输入表示不同建模方式下相应的控制指令。

下面将从脱靶量解析解和智能突防打击一体制导策略两方面研究导弹机动突防/精确打击在线制导方法。

2 基于脱靶量解析解的机动突防策略分析

2.1 制导系统建模

脱靶量是评估导弹制导律和目标规避机动有效性的基本指标。然而，先前的大多数研究依赖于仿真来获得脱靶量，对解析方法的关注有限。尽管随着计算能力的提高，暴力仿真变得更加可行，但解析解仍然有着独特的优势：a) 解析解可以帮助研究人员更深入理解潜在机制；b) 解析解允许解析推导，可以为所研究的问题提供合适的数学表示。这些解析表达可以揭示变量之间的关系，发现隐藏的模式，更为直观地解释结果；c) 解析解通常能更高效、更简洁地表示问题，从而形成更简单的模型和计算过程。这在计算资源有限且需要实时决策的航空航天场景中尤为重要。

图2为高阶制导系统的寻的回路框图，在线化交战几何的基础上涵盖了考虑寄生回路的导引头、噪声滤波器、制导律、飞控系统（即三回路驾驶员）以及弹体动力学（考虑舵机动态响应）等多个环节。制导律采用比例导引（Proportional Navigation, PN），造成制导系统脱靶量的主要因素除了拦截弹自身的响应延迟和过载饱和之外，还包括来自目标的突防机动和对目标的初始瞄准误差（即航向误差）。

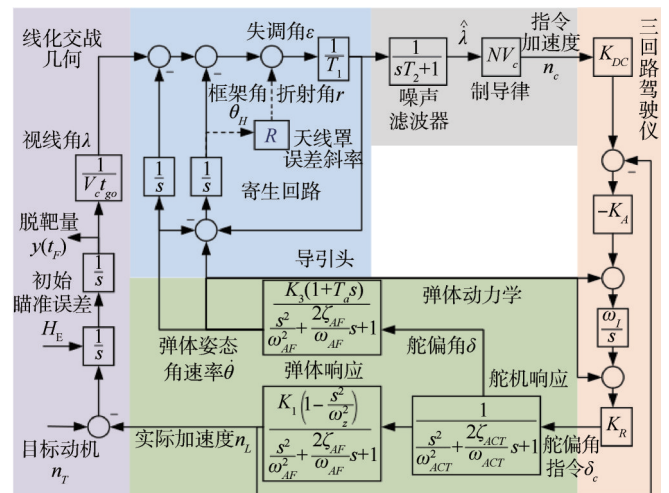


图2 交战寻的回路框图

Fig.2 Diagram of the engagement homing loop

对于这样一个包括不同时间尺度的复杂动态系统,直接对其进行分析是非常困难的,因此通过适当的建模实现简洁性和代表性的平衡是十分重要的。简化为理想动力学模型,即不考虑响应延迟,是当前制导律设计常采用的方式,但不能客观评价制导系统的性能。二项式形式的传递函数是一种能够较好地反映高阶制导系统特征的建模方式。图3给出了不同阶次的二项式系统与图2中的高保真模型在目标5g机动下的脱靶量曲线,由图3可以看出,五阶系统较好地逼近了实际的脱靶量结果,这说明基于二项式形式制导系统推导的脱靶量解析解是可靠的。同时可以看出,虽然低阶系统在脱靶量的量级上误差较大,但整体变化趋势基本与高阶系统相同,因此低阶系统的结果也是有参考价值的。

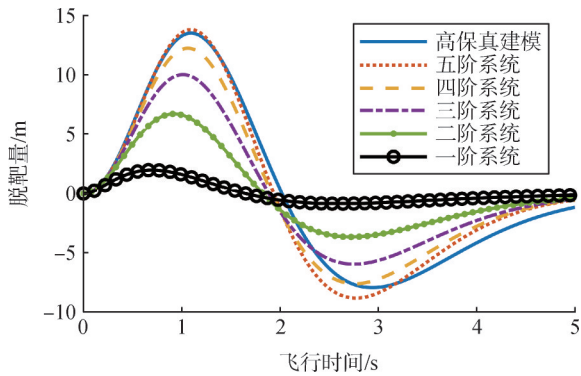


图3 不同阶次建模下的脱靶量曲线

Fig.3 Miss distance curves under different order modeling

2.2 阶跃机动

阶跃机动是一种广泛应用的机动策略,常用于评估导弹制导系统的性能。下面给出基于阶跃机动推导的脱靶量解析解并分析阶跃机动的特性。

在一阶系统建模的基础上,可以通过对奇异微分方程的求解推导出不要求导引比为整数情况下的脱靶量解析解。由航向误差和目标机动导致的脱靶量的解析解可分别表示为

$$\begin{aligned}\hat{M}_{H_e} &= \hat{t}_f e^{-i} F(2 - N, 2, \hat{t}_f) \\ \hat{M}_{n_r} &= \frac{1}{2} \hat{t}_f^2 e^{-i} F(3 - N, 3, \hat{t}_f)\end{aligned}\quad (6)$$

式中 $F(a, \gamma, x)$ 为第一类合流超几何函数。通过式(6)可直接解算出在指定导引比 $N (N > 2)$ 和归一化飞行时间 \hat{t}_f 时归一化的脱靶量。实际上通过分析式(6)及伴随系统,可以发现将 \hat{M}_{n_r} 求导就可以得到 \hat{M}_{H_e} 。

当导引比 N 为正数时,第一类合流超几何函数退

化为Laguerre多项式的形式,此时脱靶量表达式可写为

$$\begin{aligned}\hat{M}_{H_e} &= \hat{t}_f e^{-i} \sum_{k=0}^{N-2} (-1)^k \frac{C_{N-2}^k}{(k+1)!} \hat{t}_f^k \\ \hat{M}_{n_r} &= \hat{t}_f^2 e^{-i} \sum_{k=0}^{N-3} (-1)^k \frac{C_{N-3}^k}{(k+2)!} \hat{t}_f^k\end{aligned}\quad (7)$$

综合考虑目标机动和航向误差的影响可以得到不同初始条件下的脱靶量等高线,如图4所示。通过对式(6)或式(7)中两个脱靶量来源进行组合,可以直接得到图5中线化模型下的结果。通过对比可以发现,在大航向误差的影响下脱靶量结果会出现一定偏差,但整体而言仍然保留了核心特征,结果比较相近。

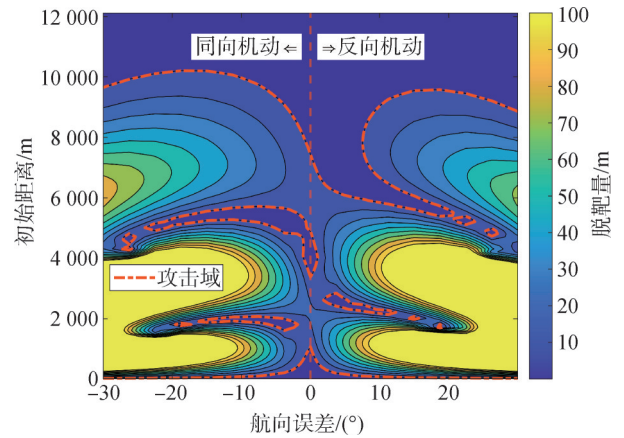


图4 非线性模型下阶跃机动导致的脱靶量

Fig.4 Miss distance caused by step maneuver under nonlinear model

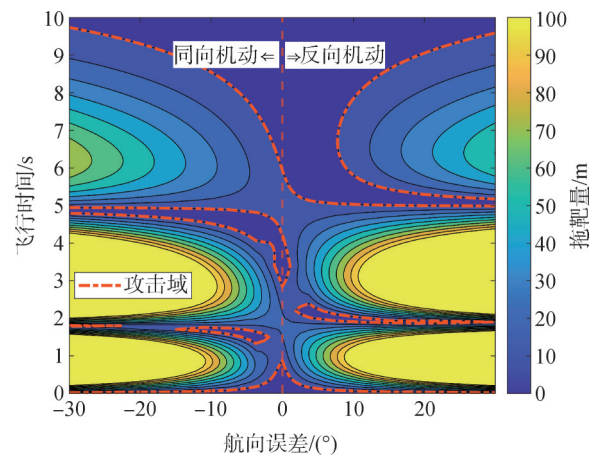


图5 线化模型下阶跃机动导致的脱靶量

Fig.5 Miss distance caused by step maneuver under linearized model

对于二项式形式的高阶系统,通过频域分析的方法可以求得脱靶量表达式为

$$\hat{M}_{n_r}(\hat{t}) = \hat{t}^2 e^{-n\hat{t}} \sum_{i=0}^{N-3} C(n, N, \hat{t}) \quad (8)$$

不同阶制导系统结果的差别集中在累加项 $\sum_{i=0}^{N-3} C(n, N, \hat{t})$ 上。如果将目标阶跃机动下的零控脱靶量表示为 $\frac{1}{2} n_r t^2$ ，那么还可以从式 (8) 中发现，脱靶量的表达式实际上包含着零控脱靶量和一个指数衰减项 $e^{-n\hat{t}}$ 的乘积。从物理意义上讲，这个指数衰减项可被视为制导律和导弹动态响应综合作用的体现。

2.3 蛇形机动

蛇形机动是另一种简单且常用的突防机动形式，在不同初始距离和航向误差下的脱靶量结果如图6所示。随着初始距离增大，飞行时间延长，脱靶量结果逐渐进入稳态，攻击域规律性地反复出现，这说明虽然蛇形机动不会像阶跃机动那样在较长的飞行时间内脱靶量收敛到零，但仍然有撞上拦截弹的可能。

通过频域分析可知，蛇形机动的脱靶量表达式由瞬态分量和稳态分量组成，其中瞬态分量以指数衰减。稳态正弦输出的幅值和相位可以通过下式获得：

$$G_w(j\omega_r) = \frac{(j\omega_r)^{N-2}}{(j\omega_r + n/T)^N} \cdot \text{Exp} \left[\sum_{k=2}^n \frac{n^N N (n - j\omega_r T)^{k-1}}{n^{n-k+1} (n^2 + \omega_r^2 T^2)^{k-1} (k-1)} \right] \quad (9)$$

由此可得到脱靶量的稳态结果，进一步结合航向误差的脱靶量就可以得到线化模型下的结果。

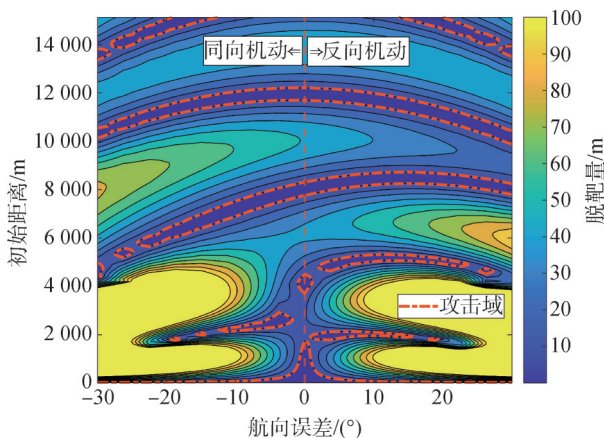


图6 非线性模型下蛇形机动导致的脱靶量

Fig.6 Miss distance caused by weave maneuver under nonlinear model

2.4 最优机动

尽管阶跃机动和蛇形机动执行起来相对简单，但由于脱靶量随交战飞行时间波动，可能存在直接撞上

导弹的潜在风险。因此，有必要推导一种最大化脱靶量的最优机动策略，以确保成功规避风险。基于线化交战模型，最优机动的推导可转化为终端时间固定和控制量存在不等式约束下的最优控制问题求解。

最优机动表现为 bang-bang 控制的形式，即在机动能力边界间切换加速度指令^[43]。加速度指令的方向取决于协态变量的符号，基于协态变量与伴随分析中脱靶量灵敏度函数的等价性，阶跃机动导致的脱靶量导数可以用以替代协态变量来确定切换时刻。因此，最优机动可表示为

$$u^*(\hat{t}) = -\text{sign}(\dot{M}_{n_r}(\hat{t}_r - \hat{t})) \quad (10)$$

通过式 (6) 或式 (7) 中的脱靶量解析解可以实现机动方向切换的在线计算，从而使得最优机动突防具有在线使用的潜力，而不必依赖于求解两点边值问题或离线数值优化。

综合考虑突防机动和航向误差的影响，可得到图7中的结果，可以看出，基于线化模型推导的最优机动突防策略在非线场景下仍然能获得较好的突防结果。而且在不存在航向误差的情况下，最优机动导致的脱靶量随着初始距离和飞行时间的增加而变大，突防效果显著优于阶跃机动和蛇形机动。

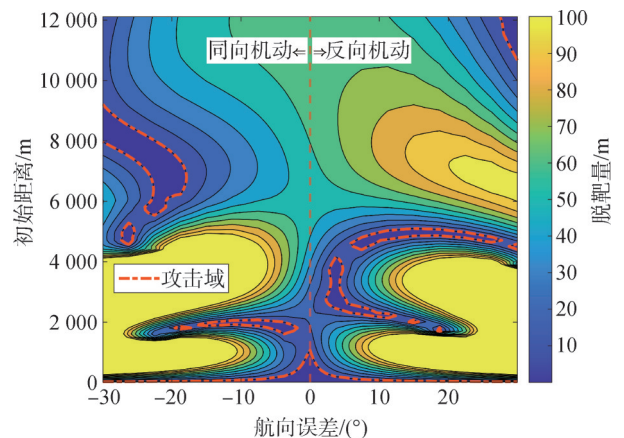


图7 非线性模型下最优机动导致的脱靶量

Fig.7 Miss distance caused by optimal maneuver under nonlinear model

但是在航向误差的影响下，盲目决策机动方向可能会影响最优机动效果。图8给出了一个具体算例，可以看到由于拦截弹航向误差的影响，不同的初始机动方向将得到完全不同的结果。因此采用如图9所示的根据战场信息灵活决策机动方向的策略，能如图10所示显著压缩拦截弹的攻击域，保证突防效果。

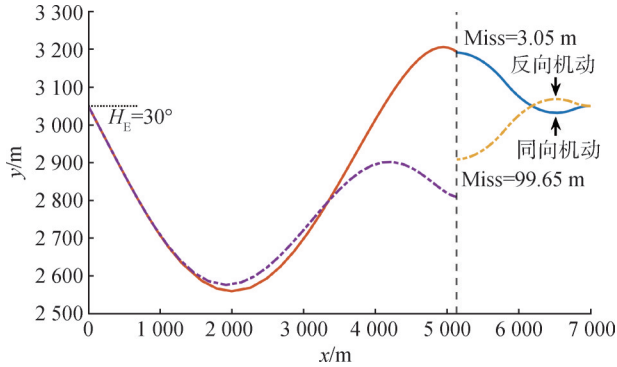


图 8 选择不同机动方向对最优机动效果的影响

Fig.8 Influence of different maneuvering directions on the optimal maneuver effect

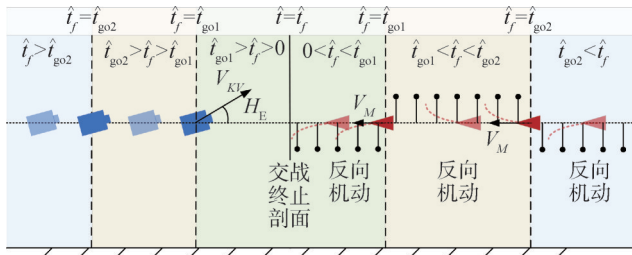


图 9 最优机动方向示意

Fig.9 Diagram of the optimal maneuvering direction

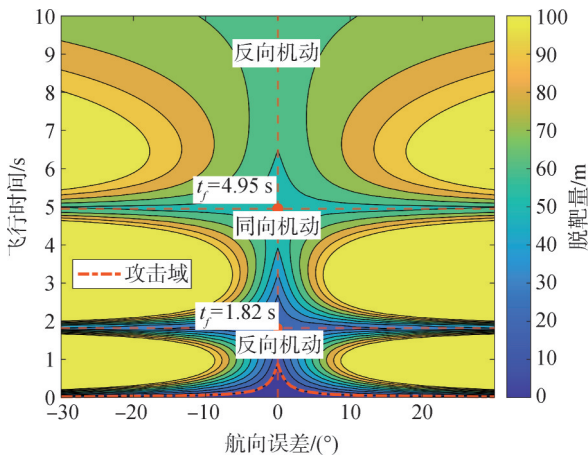


图 10 调整初始机动方向后的最优机动导致的脱靶量

Fig.10 Miss distance caused by optimal maneuver after adjusting the initial maneuver direction

2.5 面向先进制导律的脱靶量分析

很多先进制导律都是在比例导引的基础上增加偏置项得到的，比如增强比例导引 (Augmented Proportional Navigation, APN) 其实相当于是在比例导引中增加了一个考虑目标机动的偏置项 $Nn_r/2$ 。经过一系列框图变换，增强比例导引的伴随系统框图可表示为图 11，图中 $A_{PN}=1$ 。如果设图 11 中 $A_{PN}=0$ ，则和比例导引的制导框图等价。

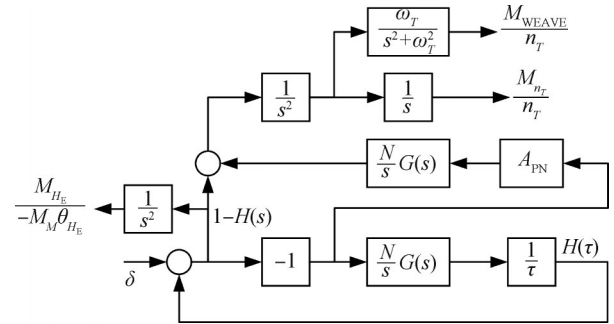


图 11 增强比例导引的伴随框图

Fig.11 Adjoint block diagram of APN

可推导得到采用增强比例导引的一阶制导系统因目标阶跃机动引起的脱靶量解析解如下：

$$\begin{aligned} \hat{M}_{n_r}^{A_{PN}}(t) &= f_1(\hat{t}) + f_2(\hat{t}) \\ f_1(\hat{t}) &= e^{-\hat{t}^2} \sum_{i=0}^{N-3} C_{N-3}^i (-\hat{t})^{N-3-i} \frac{1}{(N-1-i)!} \\ f_2(\hat{t}) &= \frac{N}{2} e^{-\hat{t}^2} \sum_{i=0}^{N-3} C_{N-3}^i (-\hat{t})^{N-3-i} \frac{1}{(N-i)!} \end{aligned} \quad (11)$$

通过式 (11) 可以解算出针对增强比例导引的最优机动的机动方向切换时刻，所得到的仿真结果如图 12 所示。从仿真结果可以看出，和比例导引情况下的结果类似，随着飞行时间增加，由阶跃机动引起的脱靶量趋于零，而由蛇形机动引起的脱靶量达到正弦形式的稳态，而且阶跃机动和蛇形机动在某些飞行时间点也都存在撞击导弹的风险。相比之下，由最优机动引起的脱靶量随交战时间持续增加，且显著高于其他机动的脱靶量。这说明在面向增强比例导引这样的先进制导律时，基于脱靶量解析解推导的最优机动策略依然是有效率的。

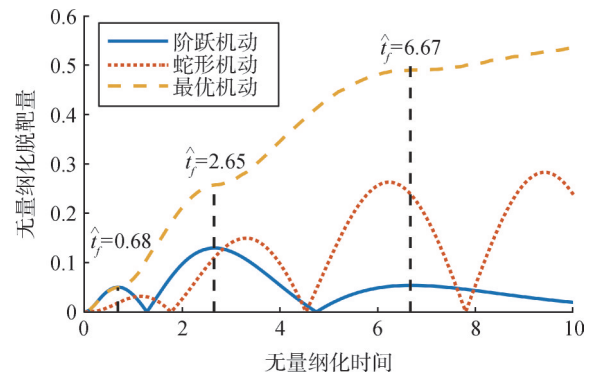


图 12 增强比例导引下不同机动形式导致的脱靶量对比

Fig.12 Comparison of miss distances caused by different maneuvers under APN

但另一方面也发现，导弹在采用两种制导律时阶跃机动脱靶量曲线的极值点并不相同，甚至从图 13 可以看出，两条曲线的极值点和零点近似相反，这也

就意味着两种情况下最优机动切换方向的时刻存在差异。类似现象在图14蛇形机动中也存在，也就意味着相同的机动过程在一种情况下脱靶量能达到最大，但在另一种情况下却恰好是脱靶量最小。

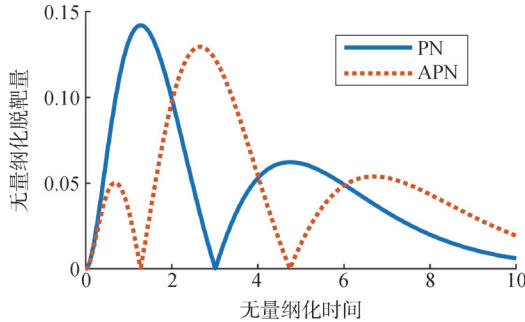


图13 阶跃机动在PN和APN下脱靶量对比

Fig.13 Comparison of miss distance of step maneuver under PN and APN

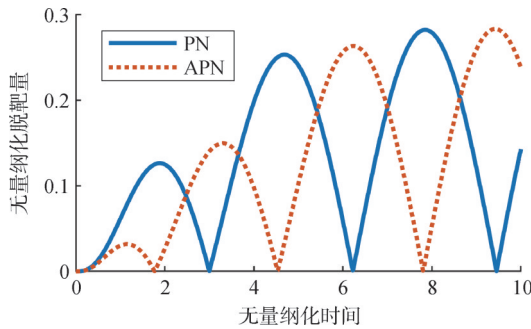


图14 蛇形机动在PN和APN下脱靶量对比

Fig.14 Comparison of miss distance of weave maneuver under PN and APN

3 基于深度强化学习的突防打击一体化制导策略

3.1 智能体训练

基于脱靶量解析解的最优机动策略只考虑了机动突防，并没有考虑突防之后对目标的攻击，因此有可能在大机动突防之后丧失攻击目标的条件。同时，解析法往往考虑理想动力学模型，对于延迟响应等因素考虑不足，而数值法会带来较大的计算负担，对拓展到多体博弈、三维交战等场景潜力不足。因此下面给出一种基于深度强化学习算法的突防打击一体化制导策略。

选用PPO作为智能体训练的算法，智能体的动作输出为突防弹的指令加速度。观测量选取为

$$o_t = \{R_k, \dot{R}_k, \dot{\lambda}_k\}, k \in \{MT, MD\} \quad (12)$$

即观测量中同时考虑了两组交战过程 (MT为突防弹打击目标，MD为突防弹规避拦截弹)，且均将距离、接近速度和视线角速率视为策略输入。奖励函数

的形式取为

$$\begin{aligned} f_1(x) &= e^{-\beta_1|x|} + e^{-\beta_2|x|} \\ f_2(y) &= \alpha [(y \leq R_1) + (y \leq R_2) + (y \leq R_3)] \quad (13) \\ R_t &= f_1(R_{MT}) + f_2(R_{MT}) - f_1(R_{MD}) - f_2(R_{MD}) \end{aligned}$$

训练环境设置见表1。智能体训练的学习曲线见图15，智能体在约 5×10^5 步左右时的累积回报便达到了200左右，此后稳定上升，直至 2×10^6 步后逐步上升到约500左右，直至最终得到一稳定收敛的策略。

表1 训练环境设置

Tab.1 Training environment settings

设置项目	拦截弹	突防弹
横向位置x/km	[20,30]	0
纵向位置y/km	0	[18,22]
飞行速度/(m·s ⁻¹)	1000	1800
最大过载/g	30	20
响应时间常数/s	0.3	0.3

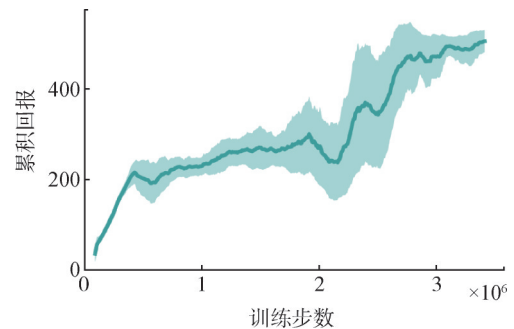


图15 智能体训练学习曲线

Fig.15 Learning curve of agent training

采用自适应步长疏化采样技术时，拦截弹拦截突防弹的脱靶量 R_{MD}^* 和突防弹打击地面目标的脱靶量 R_{MT}^* 在整个训练过程中的变化如图16所示。为了更明显地展示训练效果，这里将纵坐标轴设置为对数格式。在训练刚开始的阶段，智能体还不具备突防能力，因此 R_{MD}^* 在刚开始的值很小。随着智能体策略开始更新，拦截弹的脱靶量迅速增大，但突防弹的脱靶量变化并不显著，这意味着突防弹仅仅是通过机动避免被拦截，并不能兼顾对地面目标的打击。此后随着训练的开展，拦截弹的脱靶量稳定在 10^2 量级，也就是说突防弹能够评估拦截弹的威胁程度，以相对固定的绕飞距离避免自己被拦截。这避免了由于过度机动而给打击地面目标的交战过程带来额外的困难。与此同时可以观察到，突防弹打击地面目标的脱靶量 R_{MT}^* 随着训练的进行不断减小，直至精确到 10^0 即米级以下。这意味着智能体成功获得一个稳定突防并精确打击目标的策略。

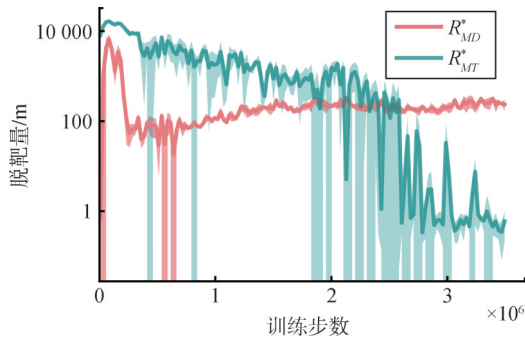


图 16 脱靶量随训练变化曲线

Fig.16 Miss distance changing with training

3.2 策略效果验证

本节对强化学习训练得到的策略进行仿真验证。单次仿真的交战轨迹如图 17 所示，仿真结果表明，拦截弹在 11.79 s 以 237.13 m 的脱靶量脱靶，之后，突防弹则在 18.00 s 成功以 0.10 m 的脱靶量命中目标。

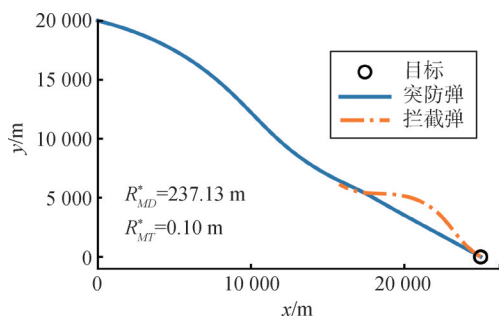


图 17 交战轨迹

Fig.17 Engagement trajectory

该场景下两枚导弹的加速度如图 18 所示，突防弹在接近拦截时刻加大了机动幅度，以最大机动能力规避了拦截弹，而拦截弹则在突防机动的干扰下出现了加速度饱和，最终导致了脱靶。可以看出，突防机动类似 bang-bang 形式，这也增加了拦截弹的加速度变化范围，给拦截制造了困难。而突防弹在突防成功之后迅速改变加速度的方向，调整至瞄准地面目标的飞行方向。

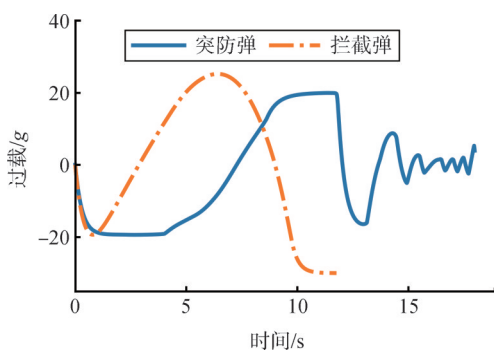


图 18 突防弹与拦截弹加速度

Fig.18 Acceleration of penetration missile and interceptor

该过程通过图 19 中展示的零控脱靶量的结果也能得到印证。拦截弹的零控脱靶量 z_{MD} 直至脱靶时仍有约 200 m 左右，而突防弹的零控脱靶量 z_{MT} 在突防后迅速稳定在零附近，即突防弹已进入打击地面目标的拦截三角形。

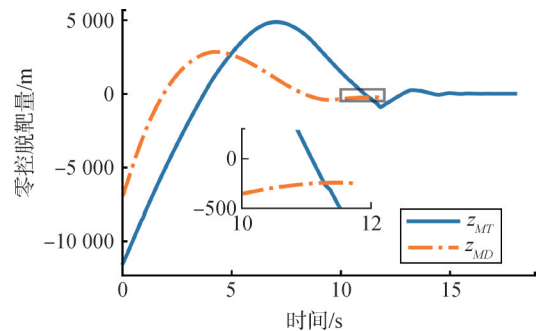


图 19 两组交战过程的零控脱靶量

Fig.19 Zero-effort-miss during the two engagement process

在 [15, 25] km 范围内不断改变突防弹的初始位置，进行蒙特卡罗仿真，所得突防弹弹道结果如图 20 所示。为了能够更清楚地展示突防弹道的特点，这里省去了拦截弹的弹道。可以看出突防弹道先是用大过载纠正攻击目标的初始航向误差，这在图 18 中表现为前 5 s 内 $-20g$ 的机动过载。但突防弹会主动将航向误差过度修正，从而为接下来拦截弹接近时切换机动方向制造脱靶预留调整空间。在突防成功后弹道方向重新调整，直至以接近直线的弹道轨迹命中目标。这一过程同时兼顾了机动突防和精确打击，因此该场景和单纯的最优机动并不完全相同，机动策略也有相应调整。

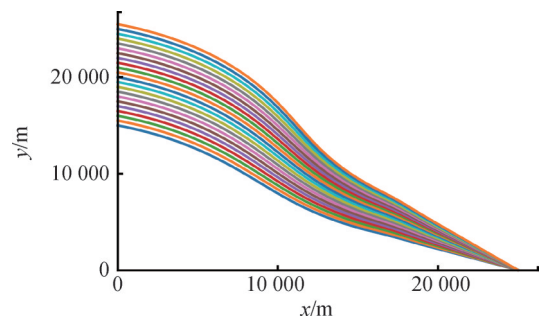


图 20 蒙特卡罗仿真突防弹道轨迹

Fig.20 Monte Carlo simulation of penetration trajectory

在图 21 中给出了蒙特卡罗仿真的零控脱靶量曲线，其结果和图 19 类似。突防弹并不要求针对地面目标的零控脱靶量迅速归零，也不单纯拉大拦截弹的零控脱靶量，而是兼顾突防和打击的需求，在突防之后可以用较小的过载修正对地面目标的偏差。

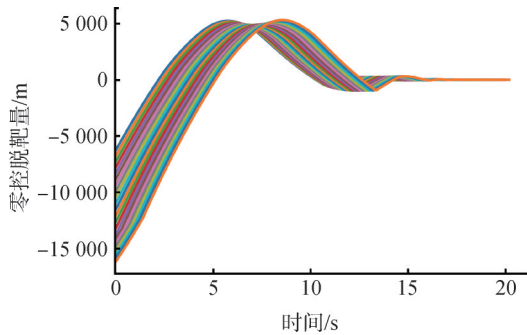


图 21 蒙特卡罗仿真零控脱靶量

Fig.21 Monte Carlo simulation of zero-effort-miss

另外需要指出的是,智能体在训练过程中设置的高度范围为 $[18, 22]$ km,所以蒙特卡罗仿真中有 60% 的场景是训练时没有遇到的,这也同时印证了该强化学习策略的泛化能力。最终用于在线部署的策略神经网络本质上是一系列矩阵运算和激活函数运算,没有迭代循环及收敛性问题,具有在线应用的潜力。

4 结束语

本文针对现有高阶拦截弹制导系统在脱靶量解析解上存在形式复杂且可靠性不足的问题,提出了一种非零阶制导系统的无量纲线性模型,并通过伴随系统进行了深入分析。首次推导了带有一阶延迟和二项式高阶特性的系统的脱靶量解析解。与传统模型相比,本研究在分析中考虑了实际响应延迟,从而增强了脱靶量表达式在工程应用中的实用性。在比例导引制导律的基础上,本文进一步扩展了该解析解的应用,探讨了增强型比例导引的脱靶量解析解。仿真结果表明,在特定条件下,阶跃机动与蛇形机动的逃逸效果呈现出对立的趋势,解析解为这一现象提供了理论依据。此外,本文还研究了最优机动策略的影响因素,通过根据脱靶量解析解调整机动方向的方式,压缩了拦截弹攻击域,显著提升了突防打击效果。

同时,针对导弹-目标-防御弹的以突防打击一体化为任务目标的三体博弈场景,将 ASTRAL 算法应用于三体博弈中的进攻弹制导问题,成功设计出一种智能博弈制导律,能在面对目标与防御弹的复杂策略组合时,为进攻弹提供平滑且适应性强的制导指令。

参 考 文 献

[1] LEE Y I, KIM S H, TAHK M J. Analytic solutions of optimal angularly constrained guidance for first-order lag system[J]. Journal of Aerospace Engineering, 2013, 227(5): 827-837.
 [2] ZARCHAN P. Tactical and strategic missile guidance[M]. Reston:

American Institute of Aeronautics and Astronautics, 2012.
 [3] GUELMAN M. A qualitative study of proportional navigation[J]. IEEE Transactions on Aerospace and Electronic Systems, 1971, AES-7(4): 637-643.
 [4] GUELMAN M. The closed-form solution of true proportional navigation[J]. IEEE Transactions on Aerospace and Electronic Systems, 1976, AES-12(4): 472-482.
 [5] MAHAPATRA P R, SHUKLA U S. Accurate solution of proportional navigation for maneuvering targets[J]. IEEE Transactions on Aerospace and Electronic Systems, 1989, 25(1): 81-89.
 [6] BECKER K. Closed-form solution of pure proportional navigation [J]. IEEE Transactions on Aerospace and Electronic Systems, 1990, 26(3): 526-533.
 [7] YUAN P J, CHERN J S. Solutions of true proportional navigation for maneuvering and nonmaneuvering targets[J]. Journal of Guidance, Control, and Dynamics, 1992, 15(1): 268-271.
 [8] YANG C D, YEH F B, CHEN J H. The closed-form solution of generalized proportional navigation[J]. Journal of Guidance, Control, and Dynamics, 1987, 10(2): 216-218.
 [9] YUAN P J, HSU S C. Solutions of generalized proportional navigation with maneuvering and nonmaneuvering targets[J]. IEEE Transactions on Aerospace and Electronic Systems, 1995, 31(1): 469-474.
 [10] MOON H B, JUNG Y K, RA W S. Terminal acceleration stabilizing guidance law for impact angle constrained interception of a non-maneuvering target[J]. International Journal of Control, Automation and Systems, 2015, 13(6): 1410-1422.
 [11] ZARCHAN P. Proportional navigation and weaving targets[J]. Journal of Guidance, Control, and Dynamics, 1995, 18(5): 969-974.
 [12] ZARCHAN P. Tracking and intercepting spiraling ballistic missiles [C]. San Diego: Position Location and Navigation Symposium, 2000.
 [13] YANUSHEVSKY R. Analysis of optimal weaving frequency of maneuvering targets[J]. Journal of Spacecraft and Rockets, 2004, 41(3): 477-479.
 [14] YANUSHEVSKY R. Frequency domain approach to guidance system design[J]. IEEE Transactions on Aerospace and Electronic Systems, 2007, 43(4): 1544-1552.
 [15] RUSNAK I, PELED-EITAN L. Guidance law against spiraling target [J]. Journal of Guidance, Control, and Dynamics, 2016, 39(7): 1694-1696.
 [16] IMADO F, UEHARA S. High-g barrel roll maneuvers against proportional navigation from optimal control viewpoint[J]. Journal of Guidance, Control, and Dynamics, 1998, 21(6): 876-881.
 [17] HE L, YAN X. Adaptive terminal guidance law for spiral-diving maneuver based on virtual sliding targets[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(7): 1591-1601.
 [18] HE L, YAN X, TANG S. Spiral-diving trajectory optimization for hypersonic vehicles by second-order cone programming[J]. Aerospace Science and Technology, 2019(95): 105427.
 [19] IMADO F, KURODA T. Family of local solutions in a missile-aircraft differential game[J]. Journal of Guidance, Control, and

- Dynamics, 2011, 34(2): 583-591.
- [20] TURETSKY V, SHIMA T. Pursuit-evasion guidance in a switched system[J]. SIAM Journal on Control and Optimization, 2018, 56(4): 2613-2633.
- [21] YANG Z, ZHOU D, PIAO H, et al. Evasive maneuver strategy for UCAV in beyond-visualrange air combat based on hierarchical multi-objective evolutionary algorithm[J]. IEEE Access, 2020(8): 46605-46623.
- [22] CARR R W, COBB R G, PACHTER M, et al. Solution of a pursuit-evasion game using a nearoptimal strategy[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(4): 841-850.
- [23] QIU X, GAO C, JING W. Maneuvering penetration strategies of ballistic missiles based on deep reinforcement learning[J]. Journal of Aerospace Engineering, 2022, 236(16): 3494-3504.
- [24] JIANG L, NAN Y, LI Z H. Realizing midcourse penetration with deep reinforcement learning[J]. IEEE Access, 2021(9): 89812-89822.
- [25] BOYELL R. Defending a moving target against missile or torpedo attack[J]. IEEE Transactions on Aerospace and Electronic Systems, 1976, AES-12(4): 522-526.
- [26] BOYELL R. Counterweapon aiming for defense of a moving target [J]. IEEE Transactions on Aerospace and Electronic Systems, 1980, AES-16(3): 402-408.
- [27] RUSNAK I. The lady, the bandits and the body guards: a two team dynamic game[J]. IFAC Proceedings Volumes, 2005, 38(1): 441-446.
- [28] RUSNAK I. Guidance laws in defense against missile attack[C]. Eilat: 2008 IEEE 25th Convention of Electrical and Electronics Engineers, 2008.
- [29] RATNOO A, SHIMA T. Line-of-sight interceptor guidance for defending an aircraft[J]. Journal of Guidance, Control, and Dynamics, 2011, 34(2): 522-532.
- [30] RATNOO A, SHIMA T. Guidance strategies against defended aerial targets[J]. Journal of Guidance, Control, and Dynamics, 2012, 35(4): 1059-1068.
- [31] YAMASAKI T, BALAKRISHNAN S. Triangle intercept guidance for aerial defense[C]. Reston: AIAA Guidance, Navigation, and Control Conference, 2010.
- [32] YAMASAKI T, BALAKRISHNAN S N. Intercept guidance for cooperative aircraft defense against a guided missile[J]. IFAC Proceedings Volumes, 2010, 43(15): 118-123.
- [33] YAMASAKI T, BALAKRISHNAN S N, TAKANO H. Modified command to line-of-sight intercept guidance for aircraft defense[J]. Journal of Guidance, Control, and Dynamics, 2013, 36(3): 898-902.
- [34] GARCIA E, CASBEER D W, PACHTER M. Pursuit in the presence of a defender[J]. Dynamic Games and Applications, 2019, 9(3): 652-670.
- [35] GARCIA E, CASBEER D W, PACHTER M. The complete differential game of active target defense[J]. Journal of Optimization Theory and Applications, 2021, 191(2-3): 675-699.
- [36] LIANG L, DENG F, PENG Z, et al. A differential game for cooperative target defense[J]. Automatica, 2019(102): 58-71.
- [37] LIANG L, DENG F, LU M, et al. Analysis of role switch for cooperative target defense differential game[J]. IEEE Transactions on Automatic Control, 2021, 66(2): 902-909.
- [38] SHAFERMAN V, SHIMA T. Cooperative multiple-model adaptive guidance for an aircraft defending missile[J]. Journal of Guidance, Control, and Dynamics, 2010, 33(6): 1801-1813.
- [39] 程涛, 周浩, 董晓飞, 等. 多飞行器突防打击一体化微分对策制导律设计[J]. 北京航空航天大学学报, 2022, 48(5): 898-909.
- CHENG Tao, ZHOU Hao, DONG Xiaofei, et al. Differential game guidance law design for integration of penetration and strike of multiple flight vehicles[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(5): 898-909.
- [40] LIU F, DONG X, LI Q, et al. Cooperative differential games guidance laws for multiple attackers against an active defense target [J]. Chinese Journal of Aeronautics, 2022, 35(5): 374-389.
- [41] WANG Y, DONG L, SUN C. Cooperative control for multi-player pursuit-evasion games with reinforcement learning[J]. Neurocomputing, 2020(412): 101-114.
- [42] ENGLISH J T, WILHELM J P. Defender-aware attacking guidance policy for the target-attacker-defender differential game[J]. Journal of Aerospace Information Systems, 2021, 18(6): 366-376.
- [43] SHINAR J, STEINBERG D. Analysis of optimal evasive maneuvers based on a linearized two-dimensional kinematic model[J]. Journal of Aircraft, 1977, 14(8): 795-802.

作者简介

陈万春 (1964—), 男, 博士, 教授, 主要研究方向为飞行力学、导弹总体设计、导弹制导与控制。

王楠翔 (1999—), 男, 博士研究生, 主要研究方向为末段突防打击一体化。

龚晓鹏 (1995—), 男, 博士研究生, 主要研究方向为智能制导与控制、攻防博弈对抗。