

文章编号:1671-6833(2019)04-0020-05

基于改进鸽群优化的机器人自抗扰控制方法

海星朔,徐炳辉,任 羿,崔京京

(北京航空航天大学 可靠性与系统工程学院,北京 100191)

摘 要:提出一种基于演化博弈理论的鸽群优化算法(EGTPIO),用来自动调节自抗扰控制器(ADRC)的参数.这种新的组合控制方法可以作为执行机构用来提高可变形地面移动机器人(DGMR)的性能. EGTPIO 不但保持了鸽群优化(PIO)的优点,而且通过演化博弈过程确定地图和指南针算子与地标算子的执行权重,有效降低局部最优概率,提升 ADRC 控制器的快速性和稳定性.仿真结果表明:EGTPIO 在最优解数量、质量和收敛速度方面的优越性使 ADRC 快速达到稳态,提高 DGMR 的机动性能.

关键词:可变形地面移动机器人;自抗扰控制;局部最优;演化博弈;鸽群优化

中图分类号:TP242. 6 文献标志码:A doi:10. 13705/j.issn.1671-6833. 2019.04. 012

0 引言

仿生智能计算作为一种研究复杂系统的新方法,被广泛用于优化计算、机器学习、自动控制等领域<sup>[1-3]</sup>.受自然界中鸽子归巢行为启发,Duan 等<sup>[4]</sup>首次提出鸽群优化算法(pigeon-inspired optimization, PIO).PIO 针对鸽子寻找目标的两个阶段,提出地图和指南针算子(map and compass operator)和地标算子(landmark operator)两种计算模型.PIO 因其在收敛速度和搜索效率方面具有优势,已得到广泛应用<sup>[5-8]</sup>.然而,算子的独立计算限制了参数之间的相互协调与最优分配,导致算法易陷入局部最优.

在演化博弈理论(evolutionary game theory, EGT)<sup>[9]</sup>中,演化稳定策略(evolutionary stable strategy, ESS)和复制动态(replicator dynamics, RD)分别表征演化博弈的稳定状态和向这种稳定状态的动态收敛过程,将 EGT 与 PIO 相结合形成新的算法 EGTPIO,将提升原始 PIO 算法的搜索效率,解决算法易陷入局部最优的问题.

作为一种新型非线性反馈控制方法,ADRC 不依赖被控对象模型<sup>[10]</sup>,具有超调小、精度高以及抗干扰能力,被广泛用于控制工程领域<sup>[11-13]</sup>.但是繁琐的参数整定工作不仅耗时,影响控制器的控制效果,还会限制被控对象的性能.

笔者将 EGTPIO 算法用于 ADRC 控制器参数的动态调节,将使控制器参数整定工作更加准确、高效.更进一步,将这种全新的控制器用于自主设计开发的可变形地面移动机器人(deformable ground mobile robot, DGMR)的姿态控制,通过对系统阶跃响应的超调量、上升时间、稳态误差进行分析,建立系统性能评价指标,对所提出的方法进行验证.仿真结果显示:笔者提出的方法能够提升 ADRC 控制器的性能,使 DGMR 的姿态控制更加高效.

1 演化博弈鸽群优化

1.1 鸽群优化

PIO 中,设  $D$  维搜索空间的  $N$  只鸽子的位置和速度分别为  $\mathbf{X}_i = [x_{i1}, x_{i2}, \cdots, x_{iD}]$  和  $\mathbf{V}_i = [v_{i1}, v_{i2}, \cdots, v_{iD}]$ ,其中  $i = 1, 2, \cdots, N$ ,地图和指南针算子如式(1)所示:

$$\begin{cases} \mathbf{V}_i^{N_c} = \mathbf{V}_i^{N_c-1} \cdot e^{-R \times N_c} + rand \cdot (\mathbf{X}_{gbest} - \mathbf{X}_i)^{N_c-1}; \\ \mathbf{X}_i^{N_c} = \mathbf{X}_i^{N_c-1} + \mathbf{V}_i^{N_c}, \end{cases} \quad (1)$$

式中: $R$  是地图和指南针因数; $N_c$  是目前迭代次数; $rand$  是随机数; $\mathbf{X}_{gbest}$  是  $N_c - 1$  次迭代后比较所有鸽子得到的全局最优位置.当循环次数达到要求后,鸽群的位置和速度由地标算子计算.地标算子中,每次迭代后  $N$  会减半以舍弃适应度函数较差的鸽子地标算子,如式(2)所示.

收稿日期:2019-01-06;修订日期:2019-03-16  
基金项目:装备预研领域基金资助项目(61400020101)  
通信作者:任羿(1973—),男,辽宁台安人,北京航空航天大学教授,主要从事可靠性系统工程研究,E-mail:renyi@buaa.edu.cn.

$$\begin{cases} X_{\text{center}}^{N_{\text{C}}-1} = \frac{\sum_{i=1}^{N_{\text{C}}-1} X_i^{N_{\text{C}}-1} F(X_i^{N_{\text{C}}-1})}{N_{\text{C}}-1 \sum_{i=1}^{N_{\text{C}}-1} F(X_i^{N_{\text{C}}-1})}; \\ N_{\text{C}} = \frac{N_{\text{C}}-1}{2}; \\ X_i = X_i^{N_{\text{C}}-1} + \text{rand} \cdot (X_{\text{center}}^{N_{\text{C}}-1} - X_i^{N_{\text{C}}-1}), \end{cases} \quad (2)$$

式中:  $X_{\text{center}}$  为鸽群的参考位置;  $\text{rand}$  是随机数。

## 1.2 演化博弈理论

演化博弈理论 (evolutionary game theory, EGT) 将博弈理论分析和生物进化过程分析相结合<sup>[14]</sup>。演化稳定策略 (evolutionary stable strategy, ESS)<sup>[15]</sup> 基于策略在迭代过程中的适应性, 用以表现策略的动态调整过程。假设多参与者博弈中存在  $n$  种纯策略, 令  $p_i$  为使用策略  $i$  的参与者比例, 状态向量  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  表示参与者的概率集合。当状态  $\mathbf{q} \neq \mathbf{p}$  令  $(1 - \varepsilon)\mathbf{p} + \varepsilon\mathbf{q}$  满足  $F(\mathbf{q} | \mathbf{p}) < F(\mathbf{p} | \mathbf{p})$  ( $\varepsilon$  为足够小的正数) 时, 状态  $\mathbf{p}$  是一个 ESS。  $F(i | \mathbf{p})$  表示每个参与者单位时间内使用策略  $i$  的收益是  $\mathbf{p}$  的函数。

模仿者动态 (replicator dynamics, RD) 描述了  $\mathbf{p}$  的变化, 设线性博弈过程收益矩阵为  $\mathbf{A}$ , 则  $F(\mathbf{q} | \mathbf{p}) = \mathbf{q}\mathbf{A}\mathbf{p}$ 。动态方程用以表达某一策略的适应度与平均适应度的差异, 可由式(3)表示。

$$\dot{\mathbf{p}}_i = \mathbf{p}_i [F(i | \mathbf{p}) - F(\mathbf{p} | \mathbf{p})], \quad (3)$$

式中:  $\mathbf{p}_i = \frac{n_i}{N}$  为策略  $i$  的比例,  $N = \sum n_i$  为所有参与者数量,  $n_i = r_i n_i$ ,  $r_i$  为策略变化率。

## 1.3 演化博弈鸽群优化

PIO 由两个独立迭代循环部分组成, 段海滨等<sup>[7]</sup>引入导航工具过渡因子将两个算子进行整合。然而, 由于两种算子的配置权重问题导致算法中参数无法相互协调。鉴于 ESS 和 RD 的优势, 可将 PIO 中的鸽子抽象为博弈参与者, 地图和指南针算子与地标算子为参与者使用的两种纯策略, 通过 ESS 和 RD 计算算子的最优配置, 记为  $y_1$  和  $y_2$ 。根据式(3), 可建立鸽群的 RD 方程:

$$\dot{\mathbf{y}}_i = \mathbf{y}_i (\mathbf{a}_i \mathbf{y} - \mathbf{y}^T \mathbf{A} \mathbf{y}), \quad (4)$$

其中, 向量  $\mathbf{a}_i$  表示  $\mathbf{A}$  的第  $i$  列,  $\mathbf{A}$  包含鸽群的全部适应度信息, 如式(5)所示。

$$\mathbf{A} = \begin{pmatrix} a(s_1) & \frac{a(s_2) + a(s_1)}{2} \\ \frac{a(s_1) + a(s_2)}{2} & a(s_2) \end{pmatrix}, \quad (5)$$

其中,  $a(s_i)$  中  $i = 1, 2$ , 表示一只鸽子使用策略  $i$

的收益, 其计算式:

$$a(s_i) = \frac{1}{t} \sum_{j=1}^t Y_i^j \cdot f(X^j), \quad (6)$$

式中:  $t$  为迭代次数;  $Y_i^j$  为一组 ESS 值。此时, 根据笔者的设计, EGPIO 的迭代方程:

$$\begin{aligned} \mathbf{V}_i^{N_{\text{C}}} &= \mathbf{V}_i^{N_{\text{C}}-1} \cdot e^{-R \times N_{\text{C}}} + \text{rand} \cdot \text{tr} \cdot y_1 \cdot \\ &(\mathbf{X}_{\text{gbest}} - \mathbf{X}_i)^{N_{\text{C}}-1} + \\ &\text{rand} \cdot \text{tr} \cdot y_2 \cdot (\mathbf{X}_{\text{center}}^{N_{\text{C}}-1} - \mathbf{X}_i^{N_{\text{C}}-1}), \end{aligned} \quad (7)$$

式中:  $\text{tr}$  表示导航工具过渡因子, 通常取常数 2;  $y_1$  和  $y_2$  是式(5)的一组解, 且规定  $y_1 + y_2 \approx 1$ 。

结构上, 演化博弈鸽群优化将鸽群优化中两个独立的循环迭代过程进行了整合, 这使得算法无需通过人为设定迭代次数进行计算, 大大提高了搜索效率; 同时, 演化博弈过程为两个算子的作用分配了权值, 使得算法具备一定的适应性, 降低算法陷入局部最优的概率, 提升了搜索的准确度。但是, 算法的局限性表现在它的搜索结果完全依赖于演化博弈中的模仿者动态方程的平衡解, 如果方程无法输出平衡解, 算法无效。

## 2 仿真实验与结果分析

### 2.1 DGMR 的 ADRC 模型

ADRC 由跟踪微分器 (tracking differentiator, TD)、扩张状态观测器 (extended state observer, ESO) 和非线性状态误差反馈控制律 (nonlinear state error feedback law, NSEFL) 组成<sup>[16]</sup>。其基本结构如图 1 所示。

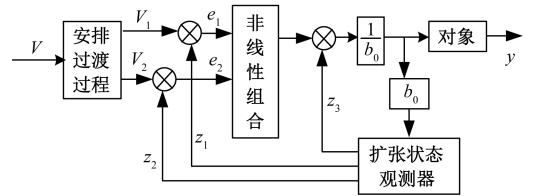


图1 自抗扰控制器结构

Fig.1 Structure of ADRC controller

笔者以自主设计开发的一款可变形地面移动机器人 (DGMR) 模型为研究对象, 通过 ADRC 对其运动系统核心部分——推杆变形机构加以控制, 其控制过程如图 2 所示。

由图 2 可知, 传感器对 DGMR 姿态角信号进行监测并将其传递给控制器, 控制器经过分析和计算给予驱动器信号以驱动电机, 连接可动轮与固定点之间的连杆因此绕固定点转动, 使可动轮的位置发生改变, 进而改变 DGMR 姿态。其中, DGMR 推杆变形机构模型参数如表 1 所示。

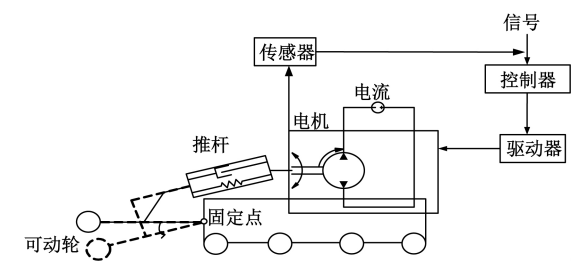


图 2 DGMR 运动系统推杆变形机构控制过程图  
Fig.2 Control chart of pusher deformation mechanism  
in motion system of DGMR

表 1 可变形地面移动机器人推杆变形机构模型参数取值

Tab.1 Parameter values of pusher deformation mechanism for DGMR

描述		取值	描述		取值
电机	输入电压 $V_a/V$	24	质量 $M/kg$		2
	齿轮半径 $r1/m$	0.025	连杆 旋转半径 $r2/m$		0.100
	感抗 $R/\Omega$	1	弹性系数 $k2/(N \cdot m^{-1})$		8
推杆	阻尼 $b/(N \cdot s \cdot m^{-1})$	4	驱动系数 $K$		1
	弹性系数 $k1/(N \cdot m^{-1})$	8	系统传感器系数 $a$		1

式中:  $v_0$  为系统输入(角度);  $v_1$  为  $v_0$  的跟踪信号;  $v_2$  为  $v_1$  的微分值;  $r_0$  和  $h$  为控制器参数.非线性函数为:

$$\begin{cases} d = rh^2; \\ a_0 = hx_2; \\ y = x_1 + a_0; \\ a_1 = \sqrt{d(d + 8r|y|)}; \\ a_2 = a_0 + \text{sign}(y)(a_1 - d)/2; \\ a = (a_0 + y)fsg(y, d) + a_2(1 - fsg(y, d)) \\ fhan(x_1, x_2, r, h) = -r\left(\frac{a}{d}\right)fsg(a, d) - \\ r\text{sign}(a)(1 - fsg(a, d)), \\ fsg(x, d) = \frac{\text{sign}(x + d) - \text{sign}(x - d)}{2}. \end{cases} \quad (9)$$

ESO 如式(10)~(12)所示.

$$\begin{cases} e_1 = z_1 - y; \\ z_1 = z_2 + h(z_2 - \beta_{01}e_1); \\ z_2 = z_2 + h(-\beta_{02}fe_1), \end{cases} \quad (10)$$

$$\begin{cases} e_3 = z_3 - z_2; \\ z_3 = z_3 + h(z_2 - \beta_{01}e_3); \\ z_4 = z_4 + h(-\beta_{03}fe_3 + bu), \end{cases} \quad (11)$$

$$\begin{cases} e_5 = z_5 - z_4; \\ z_5 = z_5 + h(z_2 - \beta_{01}e_5 + bu); \\ z_6 = z_6 + h(-\beta_{02}fe_5), \end{cases} \quad (12)$$

其中,

通过对系统物理模型分析,可使用键合图<sup>[17]</sup>对系统结构进行描述,随后根据键合图规则<sup>[18]</sup>将模型转换成 Simulink 方块图,最后在 Simulink 中完成模型的搭建.笔者拟使用 3 个二阶 ESO 串联形成 1 个三阶 ADRC 控制器对 DGMR 进行控制,由此可得,安排过渡过程如式(8)所示.

$$\begin{cases} e = v_1 - v_0; \\ fh = fhan(e_1, v_2, r_0, h); \\ v_1 = v_1 + hv_2; \\ v_2 = v_2 + hf, \end{cases} \quad (8)$$

$fe_i = fal(e_i, \alpha, \delta) = \begin{cases} \frac{e_i}{\delta^{\alpha-1}}, & |e_i| \leq \delta; \\ |e_i|^{\alpha} \text{sign}(e_i), & |e_i| \geq \delta, \end{cases}$  为非线性函数;  $y$  为系统的输出反馈信号;  $z_1, z_3, z_5$  为状态  $e$  的估计值;  $z_2, z_4, z_6$  为系统总扰动的估计值;  $\beta_{01}$  和  $\beta_{02}$  分别影响状态变量  $e$  和系统总扰动的估计值;  $b$  为补偿因子.

NSEFL 如式(13)所示:

$$\begin{cases} e_1 = v_1 - z_1; \\ e_2 = v_2 - z_2; \\ u_0 = k(e_1, e_1, p), \end{cases} \quad (13)$$

式中:  $e_1$  和  $e_2$  为状态误差;  $u_0 = -fhan(e_1, ce_2, r, h_1)$  为误差反馈控制量;  $p$  为一组参数.

2.2 适应度函数评价

将 DGMR 模型中的两个 ADRC 控制器的控制参数取相同值,因此共有来自安排过渡过程、扩张状态观测器和误差状态反馈律的 7 个参数  $C(r_0, r, h, \alpha, \beta_{01}, \beta_{02}, c)$  需进行调节. 本工作利用系统阶跃响应的超调量  $f_o$ 、上升时间  $f_{tr}$  和稳态误差  $f_e$  作为阶跃响应的评价指标,设计适应度函数及相应评价指标计算方法为:

$$\begin{cases} fitness = w_1f_o + w_2f_{tr} + w_3f_e; \\ f_o = \frac{\min_{t>0} \theta(t) - \theta_c}{\theta(0) - \theta_c}; \\ f_{tr} = t_2 \mid \theta(t_2) = 0.9 \times (\theta(0) - \theta_c) - t_1 \mid \theta(t_1) = 0.1 \times (\theta(0) - \theta_c); \\ f_e = \int_{t>t_{set}} |\theta(t) - \theta_c|, \end{cases} \quad (14)$$

式中： $w_1$ 、 $w_2$ 、 $w_3$  分别为评价指标对应的权值， $w_1 = 200$ ， $w_2 = 100$ ， $w_3 = 30$ ，笔者取较大值以保证响应的稳定性。以上 3 种性能指标越小，性能越好，因此以最小化适应度函数为优化目标。

2.3 基于 EGTPIO 的控制器参数优化

基于 EGTPIO 优化算法的参数优化流程如下。

步骤 1 DGMR 模型参数初始化,确定模型目标姿态角度  $\theta=0$  rad,仿真步长为 0.01 s,仿真采样时间为 30 s。

步骤 2 初始化参数.包括 EGTPIO 算法参数、种群以及适应度函数的权值。

步骤 3 计算每只鸽子适应度函数.利用鸽子的位置计算适应度函数值,更新鸽群中最优个体的位置。

步骤 4 分别单独执行地图和指南针算子与地标算子,并利用式(4)~(6)计算收益,得到一组 ESS。

步骤 5 根据式(7),利用上一步得到的 ESS 值执行 EGTPIO。

步骤 6 如果循环次数小于最大迭代次数,返回 Step 3,否则输出优化结果。

2.4 仿真结果与分析

为验证 EGTPIO 算法的有效性,在相同初始条件下,分别利用粒子群 (particle swarm optimization, PSO)、PIO、含有导航工具过渡因子的鸽群优化<sup>[7]</sup> (CPIO)和 EGTPIO 算法对 DGMR 系统 ADRC 控制器参数进行优化,不同算法的参数取值如表 2 所示,算法参数优化的进化曲线如图 3 所示。

表 2 算法参数取值

Tab.2 Parameters values of different algorithms			
算法名称	参数	描述	取值
PSO	$Num$	粒子数目	20
	$\omega$	惯性权重	$[0.1,1]$
	$c1$	认知学习因子	2
	$c2$	社会学习因子	2
PIO	$N_{C1max}$	地图和指南针迭代次数	20
	$N$	鸽子数目	20
	$R$	地图和指南针因数	0.3
CPIO	$tr$	导航工具过渡因子	2
	$N_{dec}$	每轮递减的鸽子数目	1
EGTPIO	$Y$	演化稳定策略初始值	(0.5,0.5)

图 3(a)为系统在阶跃信号为 0.4 rad 下的不同算法适应度曲线变化 20 次的平均值;图 3(b)为系统在阶跃信号为 0.6 rad 下的不同算法适应度曲线变化 20 次平均值。从进化曲线可以看到,

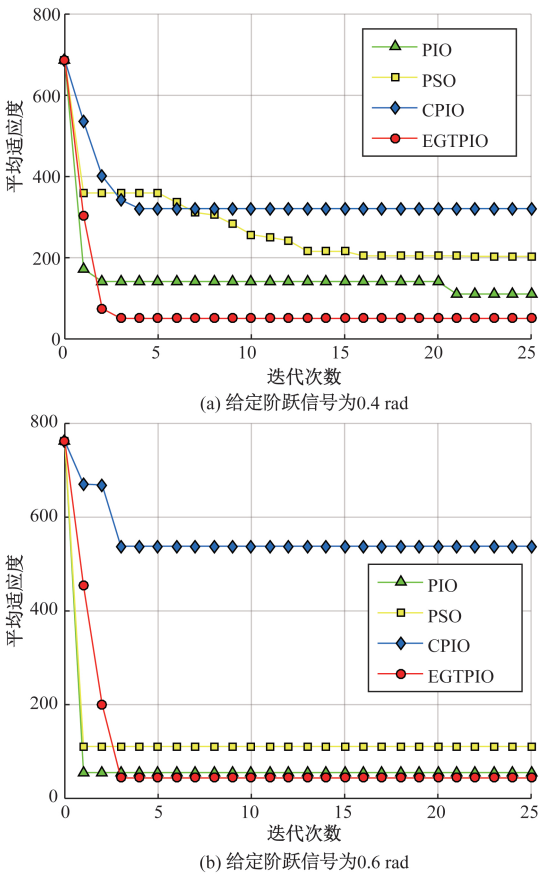


图 3 自抗扰控制器参数优化收敛曲线

Fig.3 Convergence map of parameter optimization in ADRC

EGTPIO、PIO 和 PSO 在参数优化中表现较好,都找到了较优的解,而 CPIO 则陷入局部最优,最终得到的解与其他 3 种算法相比,质量差距较明显。相比于 PSO、PIO,EGTPIO 在阶跃信号为 0.4 rad 时能够在较短时间内搜索到最优解,有着较为理想的收敛速度和搜索效率;在 0.6 rad 时虽然收敛较缓慢,但是能够找到最优解。

通过图 4 可直观看到,4 种算法优化下的系统两种阶跃信号的响应曲线 EGTPIO 在超调量、上升时间方面较其他 3 种方法都有着较明显的优势;除此之外,EGTPIO 作用下的稳态误差较其他集中算法也要小。

3 结论

(1)针对 PIO 中两个算子独立计算的问题,提出一种基于演化博弈理论的鸽群优化算法.将传统 PIO 中两个独立循环进行合并,并利用鸽群演化博弈过程对参数相互协调与最优配置进行整体改进。

(2)将演化博弈鸽群优化应用于自主设计开发的可变形地面移动机器人自抗扰控制模型,针

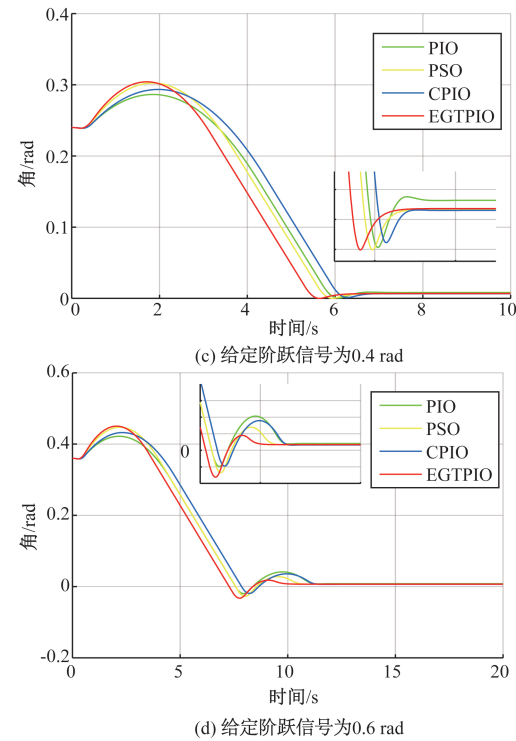


图 4 算法优化结果的系统阶跃响应对比  
Fig.4 Comparative result of step response of system for different algorithms

对 ADRC 参数过多且参数物理意义不明确、难以整定的缺陷,使用对控制器参数进行优化. 相同初始条件下对比 4 种算法的效果,仿真结果表明:相比于 PSO、PIO、CPIO、EGTPIO 的搜索质量、效率均明显.

(3)系统时域响应进一步验证了 EGTPIO 的优良性能,证明其可通过优化控制器参数提升可变形移动机器人的整体性能.

参考文献:

[1] 王守娜,刘弘,高开周. 一种应用于函数优化问题的多种群人工蜂群算法[J]. 郑州大学学报(工学版), 2018, 39(6):30-35.

[2] 孟祥萍,王圣镔,王欣欣. 基于蚁群算法和轮盘算法的多 Agent Q 学习[J]. 计算机工程与应用, 2009, 45(16):60-62.

[3] 梁静,宋慧,瞿博阳,等. 基于改进粒子群算法的路径优化问题研究[J]. 郑州大学学报(工学版), 2014, 35(1):34-38.

[4] DUAN H B, QIAO P X. Pigeon-inspired optimization; a new swarm intelligence optimizer for air robot path

planning[J]. International journal of intelligent computing and cybernetics, 2014, 7(1):24-37.

[5] DENG Y M, DUAN H B. Control parameter design for automatic carrier landing system via pigeon-inspired optimization[J]. Nonlinear dynamics, 2016, 85(1):97-106.

[6] ZHANG S J, DUAN H B. Gaussian pigeon-inspired optimization approach to orbital spacecraft formation reconfiguration[J]. Chinese journal of aeronautics, 2015, 28(1):200-205.

[7] 段海滨,邱华鑫,范彦铭. 基于捕食逃逸鸽群优化的无人机紧密编队协同控制[J]. 中国科学(技术科学), 2015, 45(6):559-572.

[8] 段海滨,杨之元. 基于柯西变异鸽群优化的大型民用飞机滚动时域控制[J]. 中国科学(技术科学), 2018, 48(3):277-288.

[9] SMITH J M. Evolution and the theory of games[M]. Cambridge:Cambridge University Press, 1982.

[10] HAN J Q. From PID to active disturbance rejection control [J]. IEEE transactions on industrial electronics, 2009, 56(3):900-906.

[11] ZHANG H, ZHAO S, GAO Z Q. An active disturbance rejection control solution for the two-mass-spring benchmark problem[C]//2016 American Control Conference. Boston.MA: IEEE, 2016:1566-1571.

[12] 马光远,沈为群,刘旺开,等. 自抗扰控制在高低温环境模拟系统中的应用[J]. 计算机测量与控制, 2015,23(11):3644-3647.

[13] 吴云洁,李琛,马征. 基于自抗扰滑模的三轴挠性卫星姿态控制[J]. 系统仿真学报, 2015, 27(8):1831-1837.

[14] SMITH J M. The theory of games and the evolution of animal conflicts[J]. Journal of theoretical biology, 1974, 47(1):209-221.

[15] TAYLOR P D, JONKER L B. Evolutionary stable strategies and game dynamics[J]. Mathematical biosciences, 1978,40(1/2):145-156.

[16] 韩京清. 自抗扰控制技术[M]. 北京:国防工业出版社, 2008.

[17] KARNOPP D C, MARGOLIS D L, ROSENBERG R C. 系统动力学: 机电系统的建模与仿真[M]. 刘玉庆,译.北京:国防工业出版社, 2012.

[18] 黄洪钟,祖旭,张旭. 基于 Matlab/Simulink 的键合图仿真[J]. 大连理工大学学报, 2003, 43(5):632-635.

## Robot Active Disturbance Rejection Control Based on an Enhanced Pigeon-inspired Optimization

HAI Xingshuo, XU Binghui, REN Yi, CUI Jingjing

(School of Reliability System Engineering, Beihang University, Beijing 100191, China)

**Abstract:** In this paper, an evolutionary game theory based pigeon-inspired optimization (EGTPIO) algorithm was proposed to automatically tune the parameters of active disturbance rejection control (ADRC) controllers. The novel approach combining EGTPIO and ADRC was subsequently used to improve the performance of the deformable ground mobile robot (DGMR). Based on our design, EGTPIO not only could keep the advantage of fast convergence of PIO, but also determine the proportion of the two operators through the process of group evolution. Thus, the probability of results falling into local optimum was reduced and the speed and stability of the controller was improved. Simulation results showed that the superiority of EGTPIO in the quantity, quality and convergence speed of optimal solutions made ADRC reach stable state quickly and improved the maneuverability of DGMR.

**Key words:** deformable ground mobile robot; active disturbance rejection control; local optima; evolutionary game theory; pigeon-inspired optimization