



# 开放环境中基于自适应分布控制的群体机器人多目标搜索方法

谭莹<sup>1,2,3\*</sup>, 陈人龙<sup>1,2</sup>, 翟天一<sup>1</sup>

1. 北京大学智能学院, 北京 100871;

2. 机器感知与智能教育部重点实验室, 北京 100871;

3. 北京大学人工智能研究院, 北京 100871

\* E-mail: [ytan@pku.edu.cn](mailto:ytan@pku.edu.cn)

收稿日期: 2021-08-14; 接受日期: 2021-11-26; 网络版发表日期: 2023-03-06

国家重点研发计划(编号: 2018AAA0102301)和国家自然科学基金(批准号: 62076010)资助项目

**摘要** 多目标搜索问题是群体机器人一个重要的研究方向. 现有工作多集中在带边界空间内的多目标搜索问题, 而在开放环境中, 探索机制会导致群体分散性过强而减弱探索能力. 本文通过引入自适应扩散回归策略, 在带有假目标的开放环境中, 提出了具有高鲁棒性和适应性的群体机器人多目标搜索算法. 文中首先从初始状态和处理假目标两方面对现有的主流群体机器人多目标搜索算法进行优化; 基于机器人分布控制, 本文对自适应群体机器人粒子群优化算法进行优化, 提出基于自适应分布控制的群体机器人粒子群优化算法; 其次, 基于概率有限状态机搜索算法(PFSMS)对开放环境中的多目标搜索算法进行进一步的探索, 本文以搜索时间为切入点, 在PFSMS原有三种状态的基础上, 添加回归状态作为附加状态, 提出了基于自适应分布控制的概率有限状态机搜索算法(DPFSMS). 当智能体的探索时间超过阈值时, 智能体的速度由回归分量和扩散/搜索分量构成. DPFSMS算法给出了在无边界开放环境中的搜索策略, 通过限制群体的扩散速度来自适应地调整智能体在无适应度值区域的运动随机性. 最后, 本文将DPFSMS算法与现有方法进行了对比, 在对比实验中DPFSMS算法取得了目前最好的效果.

**关键词** 群体机器人, 多目标搜索, 开放环境, 假目标, 群体智能

## 1 引言

群体智能的灵感来源于自然界中动物的群体协作<sup>[1,2]</sup>. 相较于传统多智能体系统<sup>[3]</sup>, 基于群体智能的群体机器人要求不存在主服务器或单个个体控制其他个体的行为, 而是利用大量简单机器人的自适应协同控

制来实现目标任务<sup>[4,5]</sup>, 灵活性、可扩展性较强<sup>[6]</sup>. 群体机器人<sup>[7]</sup>由于其个体消耗的代价较低, 非常适用于有着大范围、长时间、高危险性的任务<sup>[8]</sup>, 如定位<sup>[9]</sup>、搜索<sup>[10]</sup>、救援、监视<sup>[11]</sup>等<sup>[12]</sup>. 从这些应用中, 我们抽象出了具有一般性的多目标搜索问题, 这一抽象问题的搜索策略也可以应用于其他众多领域, 是群

**引用格式:** 谭莹, 陈人龙, 翟天一. 开放环境中基于自适应分布控制的群体机器人多目标搜索方法. 中国科学: 技术科学, 2023, 53: 395–404  
Tan Y, Chen R L, Zhai T Y. Adaptive intra-group cooperative control based multiple-target search method for swarm robotics in boundless environment (in Chinese). Sci Sin Tech, 2023, 53: 395–404, doi: [10.1360/SST-2021-0376](https://doi.org/10.1360/SST-2021-0376)

体机器人研究的重要课题<sup>[13]</sup>。

与一般的觅食任务类似, 多目标搜索问题的搜索空间非常大。在大空间中, 良好的探索能力和合理的分布范围显得极为重要。此外, 目标的分布范围广, 而机器人只能感知目标产生的局部适应度信息, 意味着机器人必须具有良好的探索开发能力<sup>[14]</sup>。多目标搜索任务对算法的适应性和灵活性要求较高, 能体现群体智能算法的优势, 是群体智能算法应用研究的主要方向。

近年来, 多目标搜索问题受到了广泛关注。Shi等人<sup>[15]</sup>受到行人行为的启发, 提出了基于行人行为的群体机器人搜索算法, 考虑了多目标搜索中的诸多现实约束, 包括通信范围、无中心协调等限制。Yang等人<sup>[17]</sup>在机器人粒子群优化算法(RPSO)<sup>[16]</sup>的基础上引入了增强探索的操作, 避免了群体陷入局部极值, 动态调整多样性。Luo等人<sup>[18]</sup>提出了一种基于机器人链模型和消除机制的目标搜索方法, 在保证机器人连通性的前提下, 有效可靠地完成多目标搜索任务。

群体机器人协同机制的设计方法有基于学习的设计和基于行为的设计两种。基于学习的多目标搜索算法多是利用强化学习算法解决多智能体的多目标搜索问题<sup>[19,20]</sup>。本文主要关注基于行为的多目标搜索算法。

基于行为的多目标搜索算法大致可以分成两类。一类关注于目标密集的搜索环境, 从现有的群优化算法(人工萤火虫群优化算法(GSO)<sup>[21]</sup>、粒子群优化算法(PSO)<sup>[22]</sup>等)中得到启发。例如, 从PSO算法发展而来的自适应群体机器人粒子群优化算法<sup>[23]</sup>(Adaptive Robotic Particle Swarm Optimization, A-RPSO)以及从烟花算法<sup>[24]</sup>(Firework Algorithm, FWA)发展而来的分组爆炸策略<sup>[25]</sup>(Group Explosion Strategy, GES), 另一类则是关注于目标比较松散、空白区域较多的搜索环境, 借鉴random walk策略<sup>[26]</sup>, 如Intermittent Search<sup>[27]</sup>等。

A-RPSO算法在机器人粒子群优化算法(RPSO)<sup>[16]</sup>的基础上进行了改进。其中, 历史分量的权值 $w$ 不再为一个常量, 而是由相邻两次迭代中最佳适应度值的差、当前迭代中最佳适应度与平均适应度值的差两项所共同决定<sup>[23]</sup>。

GES算法<sup>[25]</sup>将群体分成有数量上限的组, 以组中适应度最好的个体作为烟花, 其余个体作为火花, 以此决定组中个体的速度行为。在一次迭代后, 再次在

组中寻找适应度值最高的个体, 将其作为新的爆炸中心。

PFSMS算法是在三种基于random walk的独立搜索策略的基础上发展而来的。三种独立搜索策略分别是惯性机制、广域搜索、精炼搜索<sup>[28]</sup>。PFSMS算法将个体行为组织为一个有三种状态的概率有限状态机, 其状态分别是扩散、搜索、处理目标, 状态之间依概率相互转换。

图1中黑色实线代表概率决策, 红色虚线代表确定性决策, 中间的“开始扩散或搜索”为虚拟状态, 用来更好地解释内部过程<sup>[28]</sup>。PFSMS算法简化了概率转移: 由于处理目标状态的状态转移只与是否发现目标有关, 因此有关该状态的状态转移应为确定值。

从以往的研究来看, 多目标搜索任务可以描述为封闭环境中有多目标, 多个智能机器人通过合作的方式寻找并处理目标, 如PFSMS算法的实验环境<sup>[28]</sup>(图2)。

而在大部分搜索问题中, 划定精确范围是一件困难的事情。与封闭环境相比, 开放环境更能模拟真实的搜索环境; 而划定大致范围进行搜索任务, 对算法有着更高的要求。在多目标搜索任务中, 由于机器人只能与附近的机器人进行通信与交互, 这就要求群体的分布不能过于分散; 同时, 群体过小的分布范围会减弱群体搜索的效率, 无法发挥群体智能的优势。因此, 机器人群体的分布是群体智能算法效果的主要影响因素之一。同时, 在现实环境中存在着许多干扰项, 会对个体的通信和搜索产生影响。本文选择在环境中添加假目标来对个体的搜索造成阻碍, 增大群体的搜索难度。

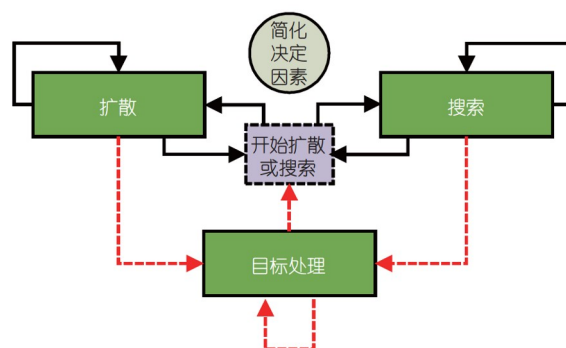


图1 概率有限状态机

Figure 1 Probabilistic finite state machine.

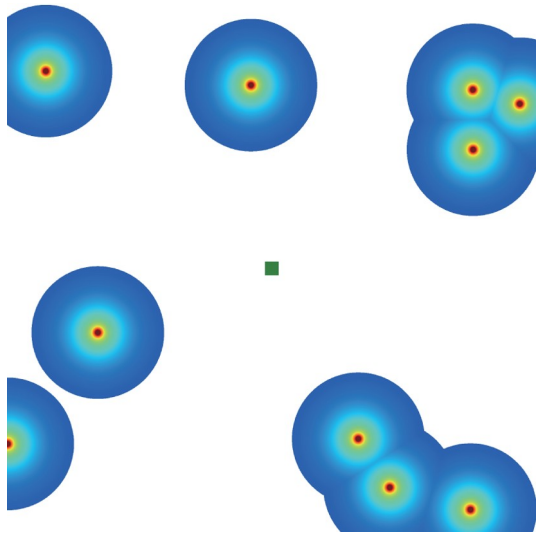


图 2  $1000 \times 1000$  的正方形仿真模拟环境(目标均匀分布在环境中, 红点代表目标, 周围的颜色代表适应度分布, 绿点代表机器人)

Figure 2  $1000 \times 1000$  square simulation environment (targets are evenly distributed in the environment, red dots represent targets, surrounding colors represent fitness distribution, and green dots denote robots).

现有工作多集中在带边界限制的空间中的多目标搜索问题<sup>[17,18,23,25,28]</sup>, 尽管这些环境的限制相对机器人的观测范围已相对较大(一般为观测范围的100倍以上), 而在无边界的开放环境中, 现有工作中的探索机制由于缺乏群体分布约束, 部分个体受探索机制引导偏离目标搜索区域, 从而导致群体分散性过强而减弱探索能力. 更进一步地, 在搜索范围内添加假目标时, 群体的搜索效率进一步下降.

本文从机器人自适应分布控制的角度出发, 改善算法的搜索性能, 并进一步应用于开放环境中. 本文通过修改A-RPSO算法的搜索逻辑并添加防聚集机制, 提出了适用于多目标搜索的DA-RPSO算法; 通过引入自适应扩散回归策略, 在带有假目标的开放环境中, 提出了具有高鲁棒性和适应性的DPFSMS算法.

DPFSMS算法给出了在开放环境中较大空白区域内的搜索策略, 通过限制群体的扩散速度来自适应地调整智能体在空白区域的运动随机性. 最后, 我们将DPFSMS算法与现有方法进行了对比, 在对比实验中DPFSMS算法取得了目前最好的效果, 验证了组内自适应分布控制机制在开放环境探索任务中的有效性.

## 2 带有假目标的开放环境设定

### 2.1 实验环境

本实验环境没有边界, 仅设置目标与假目标存在的位置. 每次迭代中, 每个个体从环境和周围获取信息, 计算速度并在环境中运动. 每次迭代的时间跨度很小, 使得个体对适应度的变化更加敏感.

### 2.2 实验假设

**规模** 环境中设置有10个目标, 6个假目标, 群体规模为50. 初始状态下所有个体分布在中心(0, 0)附近, 所使用的实验环境<sup>[29]</sup>为每个机器人提供一个局部视窗, 但不提供整个环境的视窗.

**适应度值** 本实验中适应度值用距离值计算. 由于考虑环境噪声、传感器精度等造成的误差, 本文将距离值离散化, 适应度值 =  $(\text{int})(\text{距离} \times 10)$ .

**目标** 静态半径0.05, 可观测到适应度值的范围为1, 距离值为1~10, 离目标越近, 适应度值越小. 每个目标处理都需要100次迭代, 多个个体机器人可以叠加处理目标; 目标在处理过程中保持不变, 处理完后从地图上消失. 当多个目标的可观测范围重叠时, 取最小的距离值表示该位置的适应度.

**假目标** 静态半径0.05, 可观测到适应度值的范围为1, 距离值为1~10, 离目标越近适应度值越低; 只有当被处理时才得知该目标为假目标, 其他时候与目标没有区别; 假目标在被发现后不会消失, 仍会停留在原位置.

**个体机器人** 动态半径0.03, 通信范围为1, 最高速度值为0.1, 可处理目标的范围为0.3. 个体没有先验的环境信息, 拥有有限的存储空间, 可从环境和邻近个体获取信息.

**目标分布空间** 实验环境没有边界, 目标与假目标分布在 $[-n, n]$ 的范围内(以初始位置为原点,  $x, y$ 轴取值均为 $[-n, n]$ 的矩形范围). 每次实验, 使用随机生成器在设定范围内生成地图, 10次实验求平均值, 作为算法的最终效果.

**状态速度与回归速度** 状态速度是依据当前状态(扩散状态、搜索状态)算得的速度; 回归速度是由扩散回归策略得到的速度, 当机器人处于扩散回归状态时, 该速度与状态速度共同构成机器人的最终速度.

### 3 自适应分布控制

A-RPSO算法作为解决单搜索问题的算法,其机器人群体和机器人群组都有聚集的趋势,在一段时间的搜索后会聚集在一起搜索目标;在开放环境中,由于没有边界,PFSMS算法的随机漫步策略很容易跑出目标分布范围,导致随着时间的增长,有效搜索的机器人变得越来越少.两种算法对应于分布范围的两种变化趋势,适合用于探索机器人群体的自适应分布控制.

#### 3.1 初始状态优化

初始状态如图3所示,机器人在初始位置聚集.根据之前的研究<sup>[26]</sup>可知,群体在聚集时的搜索效率明显较差,因此快速将群体分散是提高搜索效率的有效途径.

在初始状态下,为所有个体机器人设置初始速度.这样,在最初的迭代中所有个体机器人会以一定速度向外扩散.为了使个体机器人分布较为均匀,加快群体的扩散,初始状态将所有机器人按群体分为四个象限.每个个体机器人检查邻近的个体数量,选出邻近个体最少的象限,在该象限随机方向以最高速度运动.

#### 3.2 扩散回归策略

在开放环境中,当目标较为分散时,有个体机器人

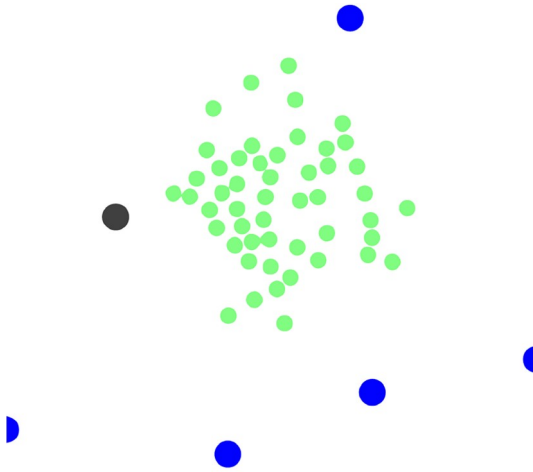


图3 随机地图初始状态的局部视窗(蓝色代表目标,黑色代表假目标,绿色代表个体机器人)

Figure 3 The local view of the initial state of the random map (blue represents the target, black represents the false target, and green represents the individual robot).

会出现跑出目标分布范围的情况,如果不加处理,这些个体基本上不可能返回到目标范围中.这就要求我们在搜索过程中对机器人的分布范围做出限制.因此,在空白区域的random walk策略不适于开放环境下机器人的搜索.

虽然不能直接限制单个机器人的搜索范围,但可以从这一角度出发,通过限制单个机器人在空白区域的扩散行为,减缓或避免群体的扩散,从而避免许多个体直接脱离搜索范围的情况.在机器人探索与扩散过程中,状态的保持与转换的概率都与迭代次数有关.本文以时间为限制量的回归算法来解决开放环境的walk策略.

回归算法以时间为限制量,当机器人处于搜索/扩散状态的时间超过一定阈值时,进入回归状态.仅当个体机器人离开空白区域或者回归时间归0时,个体机器人解除回归状态.在回归过程中,个体机器人的速度更新方程为

$$v = w \times v_{\text{back}} + (1 - w) \times v_p, \quad (1)$$

其中 $w$ 为回归项的权重,是在 $[0.6, 0.9]$ 范围内的随机数; $v_p$ 是由扩散或搜索状态得到的速度. $v_{\text{back}}$ 是回归速度,即个体机器人朝向回归中心运动的速度分量.回归方程中的回归系数 $w$ 决定了范围的扩散速度,回归系数越大,个体机器人的扩散就越慢.初始状态下,个体将起始位置作为回归中心,当某个机器人遇到目标时,将目标位置更新为回归中心.

初始时刻,探索时间从0开始计时.每次迭代,如果机器人没有进入或即将进入目标处理状态,则探索时间+1;当探索时间到达设定的回归阈值时,机器人进入回归状态;回归状态下,个体的速度由(1)式计算得到;当个体保持回归状态的时间到达阈值或个体离开空白区域,个体退出回归状态,探索时间置0,回归时间置0.

#### 3.3 DPFSMS模型

图4描述了每次迭代中个体机器人的速度计算过程.

(1) 首先检查个体机器人是否发现目标,如果发现目标则进入处理目标状态,并重置搜索时间.

(2) 检查邻近个体是否发现目标,如果发现则直接向目标运动.这一步使得整个群体在搜索过程中避免



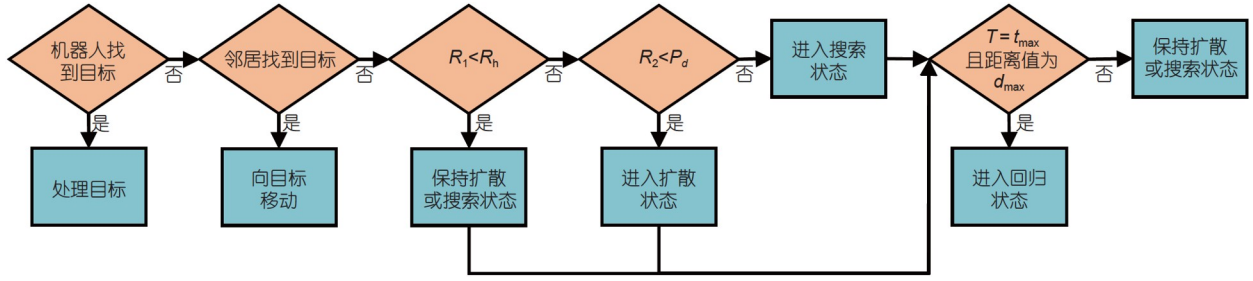


图 4 基于自适应分布控制的概率有限状态机搜索算法流程图

Figure 4 Flow chart of probabilistic finite state machine search algorithm based on adaptive distributed control.

重复搜索的行为。

(3) 检查是否继续保持当前状态.  $R_1$  是  $[0, 1]$  内的随机数,  $R_h$  初始值为  $P_i = 0.9997$ , 设  $N$  为保持当前状态的时间, 则  $P_h = P_i^N$ . 当满足  $R_1 < P_h$  时, 个体继续保持上次迭代的状态. 若保持扩散状态, 则速度等于上次迭代的速度; 若保持搜索状态, 则利用三角形梯度估计<sup>[16,28]</sup>获取速度。

(4) 如果不继续保持上次迭代的状态, 则令  $R_h = P_i$ , 利用其选择此次迭代的状态.  $R_2$  是  $[0, 1]$  内的随机数,

$$P_d = \begin{cases} 0, & N_d \leq T_d, \\ 1 - \frac{T_d}{N}, & N_d > T_d, \end{cases} \quad (2)$$

其中  $T_d = 2.3$ ,  $N_d$  为单个个体的邻居数量. 若选择扩散状态, 速度方向为邻近个体最少的方向; 若选择搜索状态, 则利用三角形梯度估计<sup>[28,30]</sup>获取速度。

(5) 确定当前次迭代的状态后, 进入回归条件的判断. 当搜索的时间到达阈值并且适应度为 0 时, 进入回归状态. 由 (3) 和 (4) 步获得的速度代入算法 1 中获得最终速度。

在本实验环境中, 假设给出的范围为  $[-n, n]$ , 则个体的搜索时间的阈值为  $n \times 30$ . 回归状态作为一种附加状态存在, 个体最终表现出的算法行为是回归+搜索状态或者回归+扩散状态, 个体的概率有限状态机仍满足图 1。

### 3.4 DA-RPSO 算法

本节将讨论组内自适应分布在 A-RPSO 算法中的应用. DA-RPSO 算法在环境中采用了上述初始状

态的优化算法, 同时本文修改 A-RPSO 算法以使得 A-RPSO 算法适用于多目标搜索环境。

#### 3.4.1 群体分组

为了将群体分组, 本文采取的方法是在初始状态下按照初始速度的象限将全部个体分为四组, 这种做法相当于将搜索空间分为四个象限, 个体分为四组分别进行搜索. 由于环境限制了个体的通信范围为 1, 因

#### 算法 1 扩散回归策略

Require:  $T_r$ : 搜索时间

$T_b$ : 回归时间

$t_{\max}$ : 时间阈值

$v_p$ : 搜索/扩散速度

$v_{\max}$ : 最大速度

$R$ : 个体适应度值

$(x_0, y_0)$ : 回归中心

$(x_a, y_a)$ : 个体位置

Ensure:  $v$ : 个体最终速度。

```

1: if  $T_r = t_{\max}$  then
2: if  $R \neq 0$  or  $T_b = t_{\max}$  then
3:  $T_r \leftarrow 0, T_b \leftarrow 0$ 
4:  $T_b \leftarrow T_b + 1$ 
5:  $v_{\text{back}} \leftarrow (x_a, y_a) - (x_0, y_0)$ 
6: if  $|v_{\text{back}}| > v_{\text{back}}$  then
7:  $v_{\text{back}} \leftarrow \frac{v_{\text{back}}}{|v_{\text{back}}|} \times v_{\max}$ 
8:  $w = \text{random}[0.6, 0.9]$ 
9:  $v \leftarrow w \times v_{\text{back}} + (1-w) \times v_p$ 
10: end if

```

此在初始状态的四组下, 个体在搜索过程中仍会采用动态分组的方式进行搜索。

### 3.4.2 个体极值与群体极值

原A-RPSO算法选用的个体极值是历史状态中的最好位置, 这在PFSMS算法中也有类似的使用, 适用于当前环境下的多目标搜索, 因此本文在计算个体极值时仍采用这种方法。

根据PSO对群体极值的定义, 群体极值应为全局状态下所有个体极值中的最优极值。在多目标搜索问题中, 如果仍采用该定义, 则随着时间的迭代, 机器人的整体分布趋于集中。因此, 本文中全局极值改为群体极值, 群体指个体机器人及其通信范围内的所有机器人构成的群组。同时, 群体极值不再是所有机器人个体极值的最优极值, 而是个体所处组内当前适应度值的极值。这样, 群体极值不再具有记忆性, 而以组为单位的群体极值也减少了个体聚集行为, 提高了多目标搜索效率。

### 3.4.3 组内自适应分布控制——防聚集机制

由于群体极值的使用, 如果不加限制, 机器人最终会聚集在一起进行搜索, 这减小了机器人群体的搜索半径, 降低搜索效率。A-RPSO算法添加了避障机制<sup>[24,31]</sup>, 但由于本文不考虑障碍物, 避障机制可用于防止个体聚集。

防聚集机制选择当前邻近个体最少的方向, 以随机速度作为防聚集分量。这使得在搜索过程中机器人群组保持相对稳定的搜索半径, 防止了组内个体聚集的情况出现。

### 3.4.4 速度更新方程

DA-RPSO的算法更新方程为

$$v_i^{t+1} = w_i^t v_i^t + c_1 r_1 (pBest_i^t - x_i^t) + c_2 r_2 (gBest_i^t - x_i^t) + c_3 r_3 v_m + c_4 m_i^t, \quad (3)$$

其中  $pBest$  是个体极值,  $gBest$  是群体极值,  $v_m$  是防聚集的速度分量,  $m_i^t$  是随机方向的单位向量,  $c_1, c_2, c_3, c_4$  是对应项的权重, 本实验中前三者都取2,  $c_4$  取1,  $r_1, r_2, r_3$  是  $[0, 1]$  之间的随机数。

DA-RPSO算法将作为后文采用的比较算法之一。

## 4 实验与结果分析

### 4.1 假目标问题

假目标和开放环境相同, 都是对现实环境进行模拟的必要因素。针对假目标问题, 本文利用机器人间的通信提高处理效率, 体现了合理的分布范围在处理任务中的价值。

根据时间和空间局部性原理, 个体机器人在最近一段时间内遇到的假目标, 在未来一段时间内更可能会遇到。个体机器人利用假目标表保存最近遇到的5个假目标的位置; 当假目标表已满时, 遇到假目标会优先覆盖最早遇到的假目标。

个体之间允许对假目标位置进行通信。由于单个机器人周围的个体所探索的范围有较大概率到达, 因此单个机器人会根据周围个体传来的假目标位置更新自己的假目标表。如果每个个体都向周围个体传自己的假目标表, 则通信成本较高, 而且一些较早记录的假目标位置对邻近个体的价值较低, 甚至可能会覆盖其原本存储的较高价值的假目标位置。因此, 每个个体只会向邻近个体传递最近遇到的假目标位置。

### 4.2 参数分析

为了测试算法在环境中的性能, 本实验首先采用随机生成的、较为均匀的地图作为实验地图, 对比算法在同一地图中的效果, 并探索Alogrithm1中时间阈值  $set\_time$  和权重  $w$  对算法效果的影响。在确定两者对实验的影响后, 采用朴素贝叶斯优化、随机搜索等方式获取参数。之后选择每个指定范围随机40张地图, 对比所有算法的效果。

GES算法在之前研究中的表现是三种算法中最差的, 在本文中对GES算法仅添加了假目标优化、初始状态优化两项基本优化, 在新环境中将作为对比算法使用。

由于机器人搜索的半径为1, 因此在目标分布范围为  $[-1, 1]$  的环境中, 机器人只处于目标搜索状态, 在该环境中的效果展示算法搜索的收敛速度; 在  $[-1, 1]$  的基础上, 逐步扩大目标的分布范围, 增大空白区域, 提高多目标的搜索难度, 测试算法的性能。

为了观察三种算法以及改进算法在环境中的表现, 首先选取  $[-1.5, 1.5]$  范围的目标分布较为均匀的固定地图作为测试环境, 该环境可以体现出较合理的空

白区域, 同时迭代次数也相对较少. 在该环境中, 本文选择设定1000次迭代为上限, 当迭代到达上限时被认为搜索失败. 图5中横坐标表示测试次数, 纵坐标代表迭代次数, 不同颜色的曲线代表不同的算法, 其中 $w = 0.9$ 和 $w = 0.6$ 分别是DPFSMS算法中在固定 $\text{set\_time} = 35$ 的情况下对应的权重 $w$ .

从图像中可以看出, PFSMS算法和DPFSMS算法基本上不会出现搜索失败的情况, A-RPSO算法偶尔会, GES算法则会多次出现. 同时, 在权重 $w$ 为0.9和0.6的情况下, DPFSMS算法的效果变化不大. 这说明在 $\text{set\_time}$ 固定的情况下,  $w$ 对算法的影响很小.

为了探索 $\text{set\_time}$ 对算法效果的影响, 实验选择 $[-3, 3]$ 范围的固定地图作为测试环境, 对比PFSMS算法以及DPFSMS算法的效果. 在该环境中本文选择的迭代次数上限为2000, 当迭代时间到达上限时视为搜索失败. 图6中横坐标表示测试次数, 纵坐标代表迭代次数, 不同颜色的曲线代表不同的算法, 其中 $w = 0$ 表示回归项系数一直为0, 代表PFSMS算法;  $\text{set\_time} = 60$ 和 $\text{set\_time} = 100$ 分别是DPFSMS算法中在固定 $w = 0.9$ 的情况下对应的时间阈值 $\text{set\_time}$ .

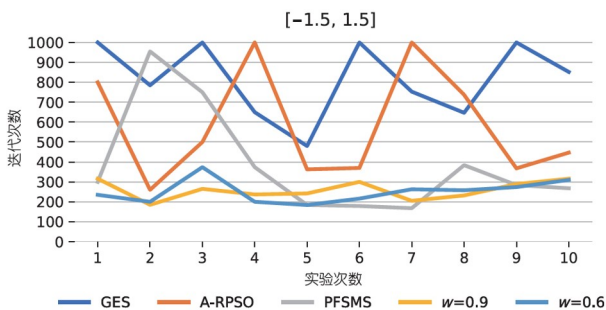


图5  $[-1.5, 1.5]$ 范围分布的效果图  
Figure 5 Effect diagram of distribution range of  $[-1.5, 1.5]$ .

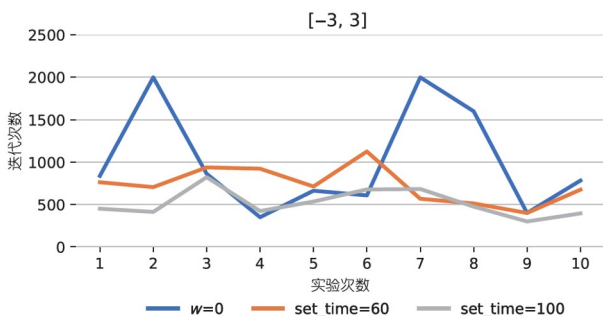


图6  $[-3, 3]$ 范围分布的效果图  
Figure 6 Effect diagram of distribution range of  $[-3, 3]$ .

从图中来看, DPFSMS算法搜索失败的概率要小于PFSMS算法. 虽然在多次随机实验中DPFSMS算法也会出现搜索失败的情况, 但是在总体下DPFSMS算法表现良好. 并且, 从图中可以看出, DPFSMS算法在 $\text{set\_time} = 100$ 时的效果明显优于 $\text{set\_time} = 60$ .

### 4.3 参数优化

针对参数优化, 本实验以 $(\text{set\_time}, w)$ 为元组测试其对算法效果的影响. 实验选择 $[-1.5, 1.5]$ 和 $[-3, 3]$ 范围的实验环境, 使用随机搜索优化策略对二元组 $(\text{set\_time}, w)$ 进行参数优化. 使用随机搜索迭代100次得到二元组, 在随机10张地图中测试, 每次测试10次取平均值.

由实验结果表1可知, 权重 $w$ 的影响很小, 而在 $[-1.5, 1.5]$ 和 $[-3, 3]$ 中,  $w$ 在 $[0.6, 0.9]$ 间波动, 在 $[0.6, 0.8]$ 间的概率较大, 与环境没有明显的关联性. 在算法实现中,  $w$ 被设置 $[0.6, 0.8]$ 间的随机数.

时间阈值 $\text{set\_time}$ 在一个相对稳定的区间内浮动, 且该值与范围所关联性较强, 对算法效果的影响较大, 因而该变量是影响算法效果的关键因素. 经过在 $[-n, n]$  ( $n = 1, 1.5, 2, 2.5, 3$ )环境中测试, 本文最终将 $\text{set\_time}$ 拟合为 $x \times n$ 的一维函数. 利用贝叶斯参数优化, 最终得到 $\text{set\_time} = 30 \times n$ , 其中 $n$ 为范围 $[-n, n]$ 中的边界值.

### 4.4 结果分析

结果分析选择在每个范围中随机生成40张地图, 每个地图中测试10次取平均值, 最终将40张地图所获数据取平均值. 横坐标代表地图范围为 $[-n, n]$ 中的 $n$ 值, 纵坐标代表平均迭代次数.

从图7中可以看出, 随着范围的扩大, DPFSMS算

表1 算法随机搜索参数优化

Table 1 Algorithm random search parameter optimization

| 参数          |                         | 结果    |       |       |
|-------------|-------------------------|-------|-------|-------|
| [-1.5, 1.5] | 时间阈值 $\text{set\_time}$ | 83    | 91    | 88    |
|             | 权重 $w$                  | 0.614 | 0.631 | 0.76  |
|             | 迭代次数                    | 619.4 | 616   | 621.3 |
| [-3, 3]     | 时间阈值 $\text{set\_time}$ | 49    | 54    | 52    |
|             | 权重 $w$                  | 0.822 | 0.672 | 0.639 |
|             | 迭代次数                    | 243.4 | 250.7 | 248.3 |

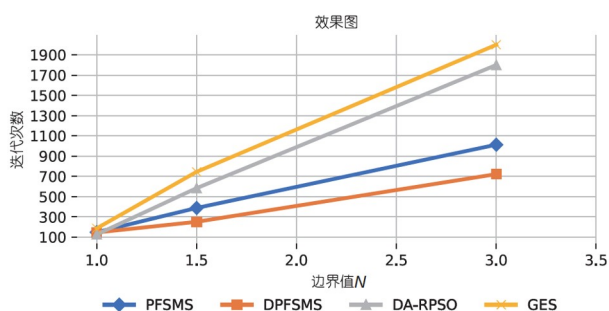


图7 算法效果对比图

Figure 7 Comparison on different algorithms.

法的效果逐渐优于PFSMS算法。之前的工作结果<sup>[28]</sup>显示PFSMS算法在所有情况下都显著优于A-RPSO算法,而图中显示在 $[-1, 1]$ 范围内效果最好的算法是改进后的A-RPSO算法。因此针对A-RPSO算法在开放环境中的优化有一定的效果,在目标分布范围较小时效果明显。

图8展示了算法在目标分布较为均匀的地图中的效果,需要单独考虑比较极端的情况。经过多次实验,DPFSMS算法在目标集中分布在范围边缘的时候效果较差。这是由于相对于设定环境范围,DPFSMS算法限制了agent的扩散速度,导致其在该环境中与PFSMS算法处于两个极端: PFSMS算法扩散过快,而DPFSMS算法扩散过慢。在该环境中本文选择的迭代次数上限为2000,当迭代时间到达上限时视为搜索失败。在目标分布在范围边缘的极端情况下,适当扩张搜索范围会提高DPFSMS算法的效果,因此在该算法中, `set_time`取值为 $30 \times 3.2$ 。从图中可以看出,在极端地图下两种算法都出现了搜索失败的情况。在10次测试中, DPFSMS算法的搜索失败现象出现2次,而PFSMS算法出现5次, DPFSMS算法的表现仍优于PFSMS算法。

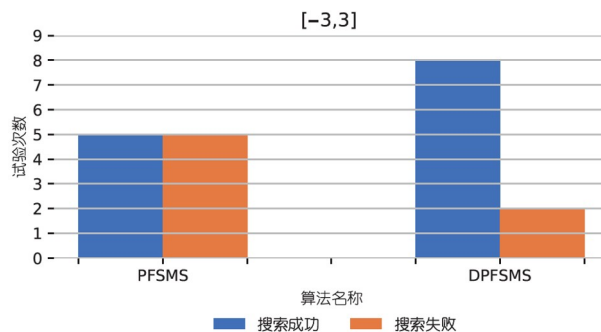


图8 极端地图下DPFSMS及PFSMS算法效果对比图

Figure 8 Comparison of DPFSMS and PFSMS in extreme maps.

## 5 总结

群体机器人多目标搜索算法落地的过程中需要考虑许多问题,如假目标、屏障、干扰源等。本文中所探究的开放环境,可以归纳为目标范围不可知的搜索问题,而添加的假目标可以看作干扰源。在该环境下的算法探索,扩充了基于行为的算法设计中要考虑的行为范畴,为多目标算法在具体环境中的应用提出了尝试和思考。

本文基于组内自适应分布控制机制,对PFSMS算法和A-RPSO算法进行了优化。在PFSMS原有三种状态的基础上,添加回归状态作为附加状态,提出了DPFSMS算法。对A-RPSO算法的机器人分布和分组进行优化,提出DA-RPSO算法。两种优化算法均相对原算法在实验中表现提升,验证了组内自适应分布控制机制的有效性。

对于开放环境中的算法探索,本文提供了一个研究方向,即选择以时间为限制的回归策略。由实验结果可知, DPFSMS算法有效弥补了PFSMS算法在开放环境中的缺陷,自适应地控制了个体机器人在空白区域的扩散速度,减少群体机器人搜索的损失。

## 参考文献

- 1 Ravary F, Lecoutey E, Kaminski G, et al. Individual experience alone can generate lasting division of labor in ants. *Curr Biol*, 2007, 17: 1308–1312
- 2 Rubenstein M, Cornejo A, Nagpal R. Programmable self-assembly in a thousand-robot swarm. *Science*, 2014, 345: 795–799
- 3 Arai T, Pagello E, Parker L E. Guest editorial advances in multirobot systems. *IEEE Trans Robot Automat*, 2002, 18: 655–661
- 4 Dudek G, Jenkin M, Milios E, et al. A taxonomy for swarm robots. In: *Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'93)*. Yokohama, 1993. 441–447
- 5 Luckcuck M, Farrell M, Dennis L A, et al. Formal specification and verification of autonomous robotic systems. *ACM Comput Surv*, 2019, 52:



1–41

- 6 Tan Y, Zheng Z. Research advance in swarm robotics. *Defence Tech*, 2013, 9: 18–39
- 7 Dorigo M, Birattari M, Brambilla M. Swarm robotics. *Scholarpedia*, 2014, 9: 1463
- 8 Şahin E. Swarm robotics: From sources of inspiration to domains of application. In: *Proceedings of the International Workshop on Swarm robotics*. Berlin, 2004. 10–20
- 9 Sinha A, Kaur R, Kumar R, et al. Cooperative control of multi-agent systems to locate source of an odor. *arXiv*: 1711.03819
- 10 Kantor G, Singh S, Peterson R, et al. Distributed search and rescue with robot and sensor teams. In: *Proceedings of the Field and Service Robotics*. Berlin, 2003. 529–538
- 11 Saska M, Chudoba J, Přeučil L, et al. Autonomous deployment of swarms of micro-aerial vehicles in cooperative surveillance. In: *Proceedings of the 2014 International Conference on Unmanned Aircraft Systems (ICUAS)*. Orlando, 2014. 584–595
- 12 Bayındır L. A review of swarm robotics tasks. *Neurocomputing*, 2016, 172: 292–321
- 13 Zheng Z, Li J, Li J, et al. Improved group explosion strategy for searching multiple targets using swarm robotics. In: *Proceedings of the 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. San Diego, 2014. 246–251
- 14 Li J, Tan Y. The multi-target search problem with environmental restrictions in swarm robotics. In: *Proceedings of the 2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*. Bali, 2014. 2685–2690
- 15 Shi H, Li J, Li Z. A distributed strategy for cooperative autonomous robots using pedestrian behavior for multi-target search in the unknown environment. *Sensors*, 2020, 20: 1606
- 16 Couceiro M S, Rocha R P, Ferreira N M F. A novel multi-robot exploration approach based on particle swarm optimization algorithms. In: *Proceedings of the 2011 IEEE International Symposium on Safety, Security, and Rescue Robotics*. Kyoto, 2011. 327–332
- 17 Yang J, Xiong R, Xiang X, et al. Exploration enhanced RPSO for collaborative multitarget searching of robotic swarms. *Complexity*, 2020, 2020: 8863526
- 18 Luo Y, Guo J, Ye G, et al. Toward target search approach of swarm robotics in limited communication environment based on robot chains with elimination mechanism. *Int J Adv Robotic Syst*, 2020, 17: 172988142091995
- 19 Busoniu L, Babuska R, De Schutter B. A comprehensive survey of multiagent reinforcement learning. *IEEE Trans Syst Man Cybern C*, 2008, 38: 156–172
- 20 Sutton R S, Barto A G. Reinforcement learning: An introduction. *IEEE Trans Neural Netw*, 1998, 9: 1054
- 21 Krishnanand K N, Ghose D. A glowworm swarm optimization based multi-robot system for signal source localization. *Design and Control of Intelligent Robotic Systems*. Berlin, Heidelberg: Springer, 2009. 49–68
- 22 Kennedy J, Eberhart R. Particle swarm optimization. In: *Proceedings of the ICNN'95-international conference on neural networks*. Perth, 1995. 1942–1948
- 23 Dadgar M, Jafari S, Hamzeh A. A PSO-based multi-robot cooperation method for target searching in unknown environments. *Neurocomputing*, 2016, 177: 62–74
- 24 Tan Y, Zhu Y. Fireworks algorithm for optimization. In: *Proceedings of the International Conference in Swarm Intelligence*. Berlin, 2010. 355–364
- 25 Zheng Z, Tan Y. Group explosion strategy for searching multiple targets using swarm robotic. In: *Proceedings of the 2013 IEEE Congress on Evolutionary Computation*. Berlin, 2013. 821–828
- 26 Bartumeus F, da Luz M G E, Viswanathan G M, et al. Animal search strategies: A quantitative random-walk analysis. *Ecology*, 2005, 86: 3078–3087
- 27 Bénichou O, Loverdo C, Moreau M, et al. Two-dimensional intermittent search processes: An alternative to Lévy flight strategies. *Phys Rev E*, 2006, 74: 020102
- 28 Li J, Tan Y. A probabilistic finite state machine based strategy for multi-target search using swarm robotics. *Appl Soft Computing*, 2019, 77: 467–483
- 29 Lowe R, Wu Y, Tamar A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments. *arXiv*: 1706.02275
- 30 Li J, Tan Y. Triangle formation based multiple targets search using a swarm of robots. In: *Proceedings of the International Conference on Swarm Intelligence*. Cham: Springer, 2016. 544–552
- 31 Wang M, Zhou S W, Zhang H Q, et al. Multi-target search of swarm robots cooperative control in an unknown environment (in Chinese). *Control Theory Appl*, 2022, 4: 750–760 [王茂, 周少武, 张红强, 等. 未知环境下群机器人多目标搜索协同控制. *控制理论与应用*, 2022, 4: 750–760]

# Adaptive intra-group cooperative control based multiple-target search method for swarm robotics in boundless environment

TAN Ying<sup>1,2,3</sup>, CHEN RenLong<sup>1,2</sup> & ZHAI TianYi<sup>1</sup>

<sup>1</sup> School of Artificial Intelligence, Peking University, Beijing 100871, China;

<sup>2</sup> Key Laboratory of Machine Perception and Intelligence (MOE), Beijing 100871, China;

<sup>3</sup> Institute for Artificial Intelligence, Peking University, Beijing 100871, China

The multi-objective search problem is one of the most important research aspects of swarm robots. The majority of existing research focuses on the multi-objective search problem in bounded space, whereas in the boundless environment, the existing work's exploration mechanism leads to excessive group dispersion and weakens the exploration ability. A multi-objective search algorithm for swarm robots with high robustness and adaptability in a boundless environment with false targets is proposed by introducing an adaptive diffusion regression strategy. First, we optimize the existing multi-objective search algorithm of mainstream swarm robots from the aspects of the initial state and false target processing. Based on the distributed control of the robot, this study optimizes the adaptive swarm robot particle swarm optimization algorithm and proposes the distributed adaptive robot particle swarm optimization algorithm. Then, based on the Probabilistic Finite State Machine based Strategy (PFSMS) algorithm, this study investigates the multi-objective search algorithm in a boundless environment, adding the return state as an additional state to the original three states of PFSMS, using exploration time as the starting point to investigate effective search algorithms in a boundless environment, and proposes An Improved Simplified Three-State PFSM (DPFSMS). When the agent's exploration time exceeds the threshold, the velocity of the agent is made up of regression and diffusion (search) components. The algorithm essentially provides a walking strategy in large blank areas of the boundless environment and reduces the randomness of the speed in this area by adaptively adjusting the agent group's diffusion speed. Finally, DPFSMS is compared against existing Multiple-Target Search Algorithms; in our experiments, the DPFSMS algorithm has achieved the best effect in the comparative experiment.

**swarm robotics, search multiple targets, boundless environment, false target, swarm intelligence**

doi: [10.1360/SST-2021-0376](https://doi.org/10.1360/SST-2021-0376)