



(12)发明专利申请

(10)申请公布号 CN 108594858 A
(43)申请公布日 2018.09.28

(21)申请号 201810779927.8

(22)申请日 2018.07.16

(71)申请人 河南大学

地址 475001 河南省开封市金明大道

申请人 河南宙合网络科技有限公司

(72)发明人 陈立家 王赞 汪晓群 薛政钢

管禹 赵瑞杰 冯帅栋 冯子凯

王敬飞 赵成伟 袁蒙恩

(74)专利代理机构 苏州知途知识产权代理事务
所(普通合伙) 32299

代理人 张锦波 陈瑞泷

(51)Int. Cl.

G05D 1/10(2006.01)

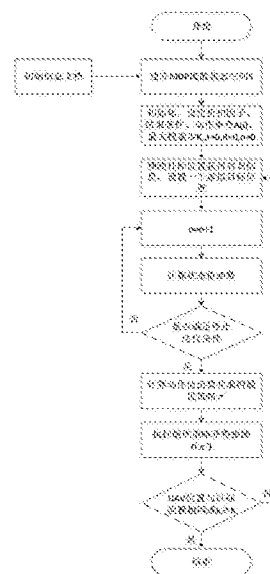
权利要求书4页 说明书11页 附图6页

(54)发明名称

马尔科夫运动目标的无人机搜索方法及装置

(57)摘要

本发明公开了马尔科夫运动目标的无人机搜索方法及装置,在接收到搜索任务后,得到马尔科夫运动目标运动过程中和无人机搜索过程中所有可能出现的状态及其概率分布;构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科夫决策的多阶段启发式策略迭代算法;获取收益最大的搜索行为策略,规划出无人机最优的搜索航迹。本发明克服了传统搜索算法对搜索目标运动规律缺乏严格数学定义而导致搜索代价高昂的缺点,根据无人机当前飞行状态和此刻马尔科夫运动目标在搜索区域的存在概率分布,确定无人机下一时刻即将搜索的虚拟目标位置,并得到收益最大的搜索行为策略,能够应用于搜索运动目标,且能够以较低的搜索成本成功搜索到目标。



1. 一种马尔科夫运动目标的无人机搜索方法,其特征在于,包括:

目标步骤,接收到搜索任务后,构建马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

无人机步骤,获取无人机搜索过程中所有可能出现的状态及其概率分布;

构建步骤,根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布,构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科夫决策的多阶段启发式策略迭代算法;

规划步骤,利用基于马尔科夫决策的多阶段启发式策略迭代算法,获取收益最大的搜索行为策略,从而规划出无人机最优的搜索航迹。

2. 根据权利要求1所述的马尔科夫运动目标的无人机搜索方法,其特征在于,所述目标步骤,具体为:

接收到搜索任务后,对目标的运动区域进行栅格化划分,利用概率论对马尔科夫运动目标建模,获取马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;所述目标的运动区域等同于无人机的搜索区域。

3. 根据权利要求1或2所述的马尔科夫运动目标的无人机搜索方法,其特征在于,所述无人机步骤,具体为:

对无人机的飞行行为进行编码并描述,对无人机在执行搜索任务过程中存在的状态进行描述,从而获取无人机搜索过程中所有可能出现的状态及其概率分布。

4. 根据权利要求1-3任一项所述的马尔科夫运动目标的无人机搜索方法,其特征在于,所述构建搜索任务下无人机行为预测的马尔科夫模型,具体为:

设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$, 该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$, 表示无人机所有可能的改变状态的动作,其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数;

设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$, 表示无人机位于某个状态下可以采取的所有动作集合;

设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合,其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下,执行可用动作 $a_x(s_n)$ 之后,状态变化到 s_m 的概率,

$$\sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1;$$

设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为:

$$MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n);$$

其中, π 为策略,表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

5. 根据权利要求4所述的马尔科夫运动目标的无人机搜索方法,其特征在于,所述规划步骤,包括:

计算步骤:用MDP折扣模型计算报酬效用函数,其中折扣因子 γ 满足: $0 < \gamma < 1$;折扣模型

的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬;

根据MDP折扣模型的最优方程,建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$, 以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$, 并根据两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

给定步骤:对协同搜索区域进行栅格化划分,确定MDP模型的离散状态空间 S , 给定参与搜索任务的无人机数目 g , g, i 为正整数, $S_{UAV(i)}$ 为第 i 架无人机当前状态, $S_{UAV(i)} \in S$, $A(S_{UAV(i)})$ 为第 i 架无人机在状态 $S_{UAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长; 给定折扣因子 γ 和策略迭代的结束条件 ε , 令迭代次数 $b=0$;

初始步骤:确定目标运动的初始位置以及每个无人机开始搜索的位置; 每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布, 从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

迭代步骤:每个无人机根据自己的虚拟目标位置, 并根据自己当前的状态 $S_{UAV(i)}$, 迭代计算各自的状态值函数 $V^{b+1}(S_{UAV(i)})$, 令迭代次数 $b=b+1$;

判断步骤:如果 $||V^{b+1}(S_{UAV(i)}) - V^b(S_{UAV(i)})|| < \varepsilon$, 则结束迭代, 进入遍历步骤; 否则, 转到迭代步骤;

遍历步骤:每个无人机根据最终得到的状态值函数 $V^{b+1}(S_{UAV(i)})$ 遍历 $A(S_{UAV(i)})$ 获得 $Q(S_{UAV(i)}, a_i)$, 最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

转移步骤:按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i , 状态由 $S_{UAV(i)}^n$ 转移到 $S_{UAV(i)}^m$; 同时, 无人机获得立即报酬 $r_i(S_{UAV(i)}, a_i)$, 此时令 $t=t+1$, 令第 i 架无人机搜索步长 $k_i = k_i + 1$;

结束步骤:若在某一个时刻 t , 第 i 架无人机位置 $S_{UAV(i)}$ 与目标当前模拟位置 S_{target} 相同, 则第 i 架无人机成功搜索到目标, 搜索任务完成, 算法结束; 若搜索步长 $\forall k_i \geq K_i, i=1, 2, \dots, n$, 则搜索任务失败, 算法结束; 所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现的位置。

6. 一种马尔科夫运动目标的无人机搜索装置, 其特征在于, 包括:

目标模块, 用于接收到搜索任务后, 构建马尔科夫运动目标的概率模型, 从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

无人机模块, 用于获取无人机搜索过程中所有可能出现的状态及其概率分布;

构建模块, 用于根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布, 构建搜索任务下无人机行为预测的马尔科夫模型, 建立基于马尔科夫决策的多阶段启发式策略迭代算法;

规划模块, 用于利用基于马尔科夫决策的多阶段启发式策略迭代算法, 获取收益最大的搜索行为策略, 从而规划出无人机最优的搜索航迹。

7. 根据权利要求6所述的马尔科夫运动目标的无人机搜索装置, 其特征在于, 所述目标模块用于:

接收到搜索任务后,对目标的运动区域进行栅格化划分,利用概率论对马尔科夫运动目标建模,获取马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;所述目标的运动区域等同于无人机的搜索区域。

8. 根据权利要求6或7所述的马尔科夫运动目标的无人机搜索装置,其特征在于,所述无人机模块用于:

对无人机的飞行行为进行编码并描述,对无人机在执行搜索任务过程中存在的状态进行描述,从而获取无人机搜索过程中所有可能出现的状态及其概率分布。

9. 根据权利要求6-8任一项所述的马尔科夫运动目标的无人机搜索装置,其特征在于,所述构建搜索任务下无人机行为预测的马尔科夫模型,具体为:

设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$, 该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$, 表示无人机所有的改变状态的动作, 其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数;

设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$, 表示无人机位于某个状态下可以采取的所有动作集合;

设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合, 其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下, 执行可用动作 $a_x(s_n)$ 之后, 状态变化到 s_m 的概率, 其中 $i, j = 1, 2, 3, \dots, \sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1$;

设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为:

$MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$;

其中, π 为策略, 表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

10. 根据权利要求9所述的马尔科夫运动目标的无人机搜索装置, 其特征在于, 所述规划模块包括:

计算单元, 用于: 用MDP折扣模型计算报酬效用函数, 其中折扣因子 γ 满足: $0 < \gamma < 1$; 折扣模型的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬;

根据MDP折扣模型的最优方程, 建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$, 以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$, 并根据两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

给定单元, 用于: 对协同搜索区域进行栅格化划分, 确定MDP模型的离散状态空间 S , 给定参与搜索任务的无人机数目 g , $S_{UAV(i)}$ 为第 i 架无人机当前状态, $S_{UAV(i)} \in S$, $A(S_{UAV(i)})$ 为第 i 架无人机在状态 $S_{UAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长; 给定折扣因子 γ

和策略迭代的结束条件 ε ,令迭代次数 $b=0$;

初始单元,用于:确定目标运动的初始位置以及每个无人机开始搜索的位置;每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布,从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

迭代单元,用于:每个无人机根据自己的虚拟目标位置,并根据自己当前的状态 $S_{UAV(i)}$,迭代计算各自的状态值函数 $V^{b+1}(S_{UAV(i)})$,令迭代次数 $b=b+1$;

判断单元,用于:如果 $\|V^{b+1}(S_{UAV(i)})-V^b(S_{UAV(i)})\|<\varepsilon$,则结束迭代,调用遍历单元,否则,调用迭代单元;

遍历单元,用于:每个无人机根据最终得到的状态值函数 $V^{b+1}(S_{UAV(i)})$ 遍历 $A(S_{UAV(i)})$ 获得 $Q(S_{UAV(i)},a_i)$,最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

转移单元,用于:按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i ,状态由 $S_{UAV(i)}^n$ 转移到 $S_{UAV(i)}^m$;同时,无人机获得立即报酬 $r_i(S_{UAV(i)},a_i)$,此时令 $t=t+1$,令第 i 架无人机搜索步长 $k_i=k_i+1$;

结束单元,用于:若在某一个时刻 t ,第 i 架无人机位置 $S_{UAV(i)}$ 与目标当前模拟位置 S_{target} 相同,则第 i 架无人机成功搜索到目标,搜索任务完成,算法结束;若搜索步长 $\forall k_i \geq K_i, i=1,2,\dots,n$,则搜索任务失败,算法结束;所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现的位置。

马尔科夫运动目标的无人机搜索方法及装置

技术领域

[0001] 本发明涉及无人机的控制领域,尤其涉及马尔科夫运动目标的无人机搜索方法及装置。

背景技术

[0002] 如今,科技的力量渗透在各行各业。作为一股新的科技力量,无人机在各行各业的应用前景备受关注。其中无人机在军事目标搜索系统中搜索与营救应用已经得到世界各国的重视。航迹规划是无人机应用于目标搜索系统中必须考虑的重要问题。

[0003] 无人机目标搜索系统的航迹规划包括两大类,即静止目标搜索系统航迹规划以及运动目标搜索系统航迹规划。当前,对于无人机静止目标搜索系统航迹规划的研究成果比较成熟,已经广泛应用于实际问题。但对于运动目标搜索系统航迹规划的研究工作相对较少。目前,对于运动目标搜索系统航迹规划问题的研究方向主要集中在对搜索算法本身的设计与应用上,而对搜索过程中目标的运动规律问题缺少严格的数学定义。另一方面,无人机独立执行任务的能力较弱,其智能化程度在很多情况下不足以满足人们的期望。在任务执行的过程中,无人机一般都是按照事先载入的程序进行,应变能力差。因此,多无人机协同参与任务是一种比较可靠的选择。这也使得无人机在搜索方面的应用与实际的应用契合的越来越紧密。

[0004] 马尔科夫决策过程(Markov Decision Processes,MDP)是一种具有决策能力的马尔科夫奖赏过程。在功能应用方面,它是用来解决在不确定性环境下序贯决策问题的理论知识。其中的不确定环境下序贯决策问题是指在一系列连续的或者离散的时刻(称为决策时刻)做出决策。人们在做决策的时候不仅要考虑决策当前的效果,也要照顾到所做的决策对长远利益的影响。MDP是一种应用非常广泛的决策过程。

[0005] 传统的无人机搜索算法,如扫描线搜索算法,只注重算法本身的设计且对搜索目标运动规律缺乏严格数学定义,存在搜索代价高昂的较大缺点。

发明内容

[0006] 为了克服现有技术的不足,本发明的目的在于提供马尔科夫运动目标的无人机搜索方法及装置,旨在解决现有的传统算法应用于无人机搜索任务时不能满足运动目标搜索航迹规划需求或者搜索代价高昂的问题。

[0007] 本发明的目的采用以下技术方案实现:

[0008] 一种马尔科夫运动目标的无人机搜索方法,包括:

[0009] 目标步骤,接收到搜索任务后,构建马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

[0010] 无人机步骤,获取无人机搜索过程中所有可能出现的状态及其概率分布;

[0011] 构建步骤,根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布,构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科

夫决策的多阶段启发式策略迭代算法;

[0012] 规划步骤,利用基于马尔科夫决策的多阶段启发式策略迭代算法,获取收益最大的搜索行为策略,从而规划出无人机最优的搜索航迹。

[0013] 在上述实施例的基础上,优选的,所述目标步骤,具体为:

[0014] 接收到搜索任务后,对目标的运动区域进行栅格化划分,利用概率论对马尔科夫运动目标建模,获取马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;所述目标的运动区域等同于无人机的搜索区域。

[0015] 在上述任意实施例的基础上,优选的,所述无人机步骤,具体为:

[0016] 对无人机的飞行行为进行编码并描述,对无人机在执行搜索任务过程中存在的状态进行描述,从而获取无人机搜索过程中所有可能出现的状态及其概率分布。

[0017] 在上述任意实施例的基础上,优选的,所述构建搜索任务下无人机行为预测的马尔科夫模型,具体为:

[0018] 设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

[0019] 设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$,该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

[0020] 设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$,表示无人机所有可能的改变状态的动作,其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数;

[0021] 设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$,表示无人机位于某个状态下可以采取的所有动作集合;

[0022] 设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合,其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下,执行可用动作 $a_x(s_n)$ 之后,状态变化到 s_m 的概率,

$$\sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1;$$

[0023] 设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

[0024] 则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为:

[0025] $MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$;

[0026] 其中, π 为策略,表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

[0027] 在上述实施例的基础上,优选的,所述规划步骤,包括:

[0028] 计算步骤:用MDP折扣模型计算报酬效用函数,其中折扣因子 γ 满足: $0 < \gamma < 1$;折扣模型的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬;

[0029] 根据MDP折扣模型的最优方程,建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$,以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$,并根据两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

[0030] 给定步骤:对协同搜索区域进行栅格化划分,确定MDP模型的离散状态空间 S ,给定

参与搜索任务的无人机数目 g, g, i 为正整数, $S_{SUAV(i)}$ 为第 i 架无人机当前状态, $S_{SUAV(i)} \in S, A(S_{SUAV(i)})$ 为第 i 架无人机在状态 $S_{SUAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长; 给定折扣因子 γ 和策略迭代的结束条件 ε , 令迭代次数 $b=0$;

[0031] 初始步骤: 确定目标运动的初始位置以及每个无人机开始搜索的位置; 每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布, 从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

[0032] 迭代步骤: 每个无人机根据自己的虚拟目标位置, 并根据自己当前的状态 $S_{SUAV(i)}$, 迭代计算各自的状态值函数 $V^{b+1}(S_{SUAV(i)})$, 令迭代次数 $b=b+1$;

[0033] 判断步骤: 如果 $||V^{b+1}(S_{SUAV(i)}) - V^b(S_{SUAV(i)})|| < \varepsilon$, 则结束迭代, 进入遍历步骤; 否则, 转到迭代步骤;

[0034] 遍历步骤: 每个无人机根据最终得到的状态值函数 $V^{b+1}(S_{SUAV(i)})$ 遍历 $A(S_{SUAV(i)})$ 获得 $Q(S_{SUAV(i)}, a_i)$, 最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

[0035] 转移步骤: 按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i , 状态由 $S_{SUAV(i)}^n$ 转移到 $S_{SUAV(i)}^m$; 同时, 无人机获得立即报酬 $r_i(S_{SUAV(i)}, a_i)$, 此时令 $t=t+1$, 令第 i 架无人机搜索步长 $k_i = k_i + 1$;

[0036] 结束步骤: 若在某一个时刻 t , 第 i 架无人机位置 $S_{SUAV(i)}$ 与目标当前模拟位置 S_{target} 相同, 则第 i 架无人机成功搜索到目标, 搜索任务完成, 算法结束; 若搜索步长 $\forall k_i \geq K_i, i = 1, 2, \dots, n$, 则搜索任务失败, 算法结束; 所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现的位置。

[0037] 一种马尔科夫运动目标的无人机搜索装置, 包括:

[0038] 目标模块, 用于接收到搜索任务后, 构建马尔科夫运动目标的概率模型, 从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

[0039] 无人机模块, 用于获取无人机搜索过程中所有可能出现的状态及其概率分布;

[0040] 构建模块, 用于根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布, 构建搜索任务下无人机行为预测的马尔科夫模型, 建立基于马尔科夫决策的多阶段启发式策略迭代算法;

[0041] 规划模块, 用于利用基于马尔科夫决策的多阶段启发式策略迭代算法, 获取收益最大的搜索行为策略, 从而规划出无人机最优的搜索航迹。

[0042] 在上述实施例的基础上, 优选的, 所述目标模块用于:

[0043] 接收到搜索任务后, 对目标的运动区域进行栅格化划分, 利用概率论对马尔科夫运动目标建模, 获取马尔科夫运动目标的概率模型, 从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布; 所述目标的运动区域等同于无人机的搜索区域。

[0044] 在上述任意实施例的基础上, 优选的, 所述无人机模块用于:

[0045] 对无人机的飞行行为进行编码并描述, 对无人机在执行搜索任务过程中存在的状态进行描述, 从而获取无人机搜索过程中所有可能出现的状态及其概率分布。

[0046] 在上述任意实施例的基础上, 优选的, 所述构建搜索任务下无人机行为预测的马尔科夫模型, 具体为:

[0047] 设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

[0048] 设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$, 该状态空间包含了无人机搜索过程

中和马尔科夫运动目标运动过程中所有可能出现的状态；

[0049] 设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$, 表示无人机所有可能的改变状态的动作, 其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数；

[0050] 设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$, 表示无人机位于某个状态下可以采取的所有动作集合；

[0051] 设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合, 其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下, 执行可用动作 $a_x(s_n)$ 之后, 状态变化到 s_m 的概率, 其中 $i, j = 1, 2, 3, \dots, \sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1$;

[0052] 设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬；

[0053] 则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为：

[0054] $MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$ ；

[0055] 其中, π 为策略, 表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

[0056] 在上述实施例的基础上, 优选的, 所述规划模块包括：

[0057] 计算单元, 用于: 用 MDP 折扣模型计算报酬效用函数, 其中折扣因子 γ 满足: $0 < \gamma < 1$; 折扣模型的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬；

[0058] 根据 MDP 折扣模型的最优方程, 建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$, 以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$, 并根据两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

[0059] 给定单元, 用于: 对协同搜索区域进行栅格化划分, 确定 MDP 模型的离散状态空间 S , 给定参与搜索任务的无人机数目 g , $s_{UAV(i)}$ 为第 i 架无人机当前状态, $s_{UAV(i)} \in S$, $A(s_{UAV(i)})$ 为第 i 架无人机在状态 $s_{UAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长; 给定折扣因子 γ 和策略迭代的结束条件 ϵ , 令迭代次数 $b=0$;

[0060] 初始单元, 用于: 确定目标运动的初始位置以及每个无人机开始搜索的位置; 每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布, 从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

[0061] 迭代单元, 用于: 每个无人机根据自己的虚拟目标位置, 并根据自己当前的状态 $s_{UAV(i)}$, 迭代计算各自的状态值函数 $V^{b+1}(s_{UAV(i)})$, 令迭代次数 $b=b+1$;

[0062] 判断单元, 用于: 如果 $||V^{b+1}(s_{UAV(i)}) - V^b(s_{UAV(i)})|| < \epsilon$, 则结束迭代, 调用遍历单元, 否则, 调用迭代单元;

[0063] 遍历单元, 用于: 每个无人机根据最终得到的状态值函数 $V^{b+1}(s_{UAV(i)})$ 遍历 $A(s_{UAV(i)})$ 获得 $Q(s_{UAV(i)}, a_i)$, 最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

[0064] 转移单元, 用于: 按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i , 状态由 $s_{UAV(i)}^n$ 转移到 $s_{UAV(i)}^m$; 同时, 无人机获得立即报酬 $r_i(s_{UAV(i)}, a_i)$, 此时令 $t=t+1$, 令第 i 架无人机搜索步长

$k_i = k_i + 1$;

[0065] 结束单元,用于:若在某一个时刻 t ,第 i 架无人机位置 $S_{UAV(i)}$ 与目标当前模拟位置 S_{target} 相同,则第 i 架无人机成功搜索到目标,搜索任务完成,算法结束;若搜索步长 $\forall k_i \geq K_i, i=1,2,\dots,n$,则搜索任务失败,算法结束;所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现的位置。

[0066] 相比现有技术,本发明的有益效果在于:

[0067] 本发明公开了马尔科夫运动目标的无人机搜索方法及装置,在接收到搜索任务后,构建马尔科夫运动目标的概率模型,得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;获取无人机搜索过程中所有可能出现的状态及其概率分布;构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科夫决策的多阶段启发式策略迭代算法;利用基于马尔科夫决策的多阶段启发式策略迭代算法,获取收益最大的搜索行为策略,从而规划出无人机最优的搜索航迹。

[0068] 本发明能够采用MDP折扣模型确定搜索任务下无人机存在状态转移概率集合和转移过程中飞行行为操作的报酬集合,计算报酬效用函数,建立搜索任务下无人机行为收益的最优方程,进行迭代计算和判断,得到收益最大的搜索行为策略,规划出无人机最优的搜索航迹。

[0069] 本发明克服了传统搜索算法,如扫描线搜索算法,只注重算法本身的设计且对搜索目标运动规律缺乏严格数学定义而导致搜索代价高昂的较大缺点,根据无人机当前飞行状态和此刻马尔科夫运动目标在搜索区域的存在概率分布,确定无人机下一时刻即将搜索的虚拟目标的位置,并得到一组无人机的搜索行为操作序列,此序列为行为收益最大的搜索行为策略,能够应用于搜索运动目标,且能够以较低的搜索成本成功搜索到目标。

附图说明

[0070] 下面结合附图和实施例对本发明进一步说明。

[0071] 图1示出了本发明实施例提供的一种搜索区域栅格划分示意图;

[0072] 图2a~2d分别示出了本发明实施例提供的马尔科夫运动目标在不同时刻的存在概率分布图;

[0073] 图3示出了本发明实施例提供的一种MDP-MSHPI算法理论框架结构示意图;

[0074] 图4示出了本发明实施例提供的一种单个无人机运用MSHPI算法搜索Markov目标的流程图;

[0075] 图5示出了本发明实施例提供的一种多个无人机运用MSHPI算法搜索Markov目标的流程图;

[0076] 图6示出了本发明实施例提供的一种马尔科夫运动目标的无人机搜索装置的结构示意图。

具体实施方式

[0077] 下面,结合附图以及具体实施方式,对本发明做进一步描述,需要说明的是,在不相冲突的前提下,以下描述的各实施例之间或各技术特征之间可以任意组合形成新的实施

例。

[0078] 具体实施例一

[0079] 本发明实施例提供了一种马尔科夫运动目标的无人机搜索方法,包括:

[0080] 目标步骤,接收到搜索任务后,构建马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

[0081] 无人机步骤,获取无人机搜索过程中所有可能出现的状态及其概率分布;

[0082] 构建步骤,根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布,构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科夫决策的多阶段启发式策略迭代算法(Multi-Stage Heuristic Strategy Iteration Algorithm,MDP-MSHPI);

[0083] 规划步骤,利用基于马尔科夫决策的多阶段启发式策略迭代算法,获取收益最大的搜索行为策略,从而规划出无人机最优的搜索航迹。本发明实施例中,无人机的数量可以是一个、两个或多个。

[0084] 本发明克服了传统搜索算法,如扫描线搜索算法,只注重算法本身的设计且对搜索目标运动规律缺乏严格数学定义而导致搜索代价高昂的较大缺点,根据无人机当前飞行状态和此刻马尔科夫运动目标在搜索区域的存在概率分布,确定无人机下一时刻即将搜索的虚拟目标的位置,并得到一组无人机的搜索行为操作序列,此序列为行为收益最大的搜索行为策略,能够应用于搜索运动目标,且能够以较低的搜索成本成功搜索到目标。

[0085] 如图1、图2所示,优选的,所述目标步骤,可以具体为:接收到搜索任务后,对目标的运动区域进行栅格化划分,利用概率论对马尔科夫运动目标建模,获取马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;所述目标的运动区域等同于无人机的搜索区域。这样做的好处是,能够对运动目标的运动规律进行模拟,获取马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布,例如可以得到某一时刻下运动目标最可能所处的位置。

[0086] 优选的,所述无人机步骤,可以具体为:对无人机的飞行行为进行编码并描述,对无人机在执行搜索任务过程中存在的状态进行描述,从而获取无人机搜索过程中所有可能出现的状态及其概率分布。这样做的好处是,能够获取无人机在搜索过程中可能出现的状态及其概率分布。

[0087] 优选的,所述构建搜索任务下无人机行为预测的马尔科夫模型的步骤,可以具体为:

[0088] 设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

[0089] 设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$,该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

[0090] 设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$,表示无人机所有可能的改变状态的动作,其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数;

[0091] 设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$,表示无人机位于某个状态下可以采取的所有动作集合;

[0092] 设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合,其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下,执行可用动作 $a_x(s_n)$ 之后,状态变化到 s_m 的概率,

$$\sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1;$$

[0093] 设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

[0094] 则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为:

[0095] $MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$;

[0096] 其中, π 为策略,表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

[0097] 优选的,所述规划步骤,可以包括:

[0098] 计算步骤:用MDP折扣模型计算报酬效用函数,其中折扣因子 γ 满足: $0 < \gamma < 1$;折扣模型的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬;

[0099] 根据MDP折扣模型的最优方程,建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$,以及

最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$,并根据

两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

[0100] 给定步骤:对协同搜索区域进行栅格化划分,确定MDP模型的离散状态空间 S ,给定参与搜索任务的无人机数目 g , g, i 为正整数, $s_{UAV(i)}$ 为第 i 架无人机当前状态, $s_{UAV(i)} \in S, A(s_{UAV(i)})$ 为第 i 架无人机在状态 $s_{UAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长;给定折扣因子 γ 和策略迭代的结束条件 ε ,令迭代次数 $b=0$;

[0101] 初始步骤:确定目标运动的初始位置以及每个无人机开始搜索的位置;每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布,从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

[0102] 迭代步骤:每个无人机根据自己的虚拟目标位置,并根据自己当前的状态 $s_{UAV(i)}$,迭代计算各自的状态值函数 $V^{b+1}(s_{UAV(i)})$,令迭代次数 $b=b+1$;

[0103] 判断步骤:如果 $||V^{b+1}(s_{UAV(i)}) - V^b(s_{UAV(i)})|| < \varepsilon$,则结束迭代,进入遍历步骤;否则,转到迭代步骤;

[0104] 遍历步骤:每个无人机根据最终得到的状态值函数 $V^{b+1}(s_{UAV(i)})$ 遍历 $A(s_{UAV(i)})$ 获得 $Q(s_{UAV(i)}, a_i)$,最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

[0105] 转移步骤:按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i ,状态由 $s_{UAV(i)}^n$ 转移到 $s_{UAV(i)}^m$;同时,无人机获得立即报酬 $r_i(s_{UAV(i)}, a_i)$,此时令 $t=t+1$,令第 i 架无人机搜索步长 $k_i = k_i + 1$;

[0106] 结束步骤:若在某一个时刻 t ,第 i 架无人机位置 $s_{UAV(i)}$ 与目标当前模拟位置 s_{target} 相同,则第 i 架无人机成功搜索到目标,搜索任务完成,算法结束;若搜索步长 $\forall k_i \geq K_i, i = 1, 2, \dots, n$,则搜索任务失败,算法结束;所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现(所处)的位置。

[0107] 这样做的好处是,能够采用MDP折扣模型确定搜索任务下无人机存在状态转移概率集合和转移过程中飞行行为操作的报酬集合,计算报酬效用函数,建立搜索任务下无人

机行为收益的最优方程,进行迭代计算和判断,得到收益最大的搜索行为策略,规划出无人机最优的搜索航迹。

[0108] 本发明的一种应用场景可以是:

[0109] 1.对目标的运动区域(同样是无人机的搜索区域)进行栅格化划分,划分后的区域模型如图1所示;利用概率论的知识对马尔科夫运动目标建模,获得马尔科夫运动目标的概率模型,马尔科夫运动目标在运动一定步数后在运动区域的存在概率如图2a~2d所示,其中图2a为初始时刻,图2b为运动一步后,图2c为运动5步后,图2d为运动10步后;

[0110] MDP-MSHPI算法理论框架结构如图3所示;

[0111] 2.对无人机的飞行行为进行编码并描述,并对无人机在执行搜索任务过程中存在的状态进行描述;

[0112] 3.构建搜索任务下无人机行为预测的马尔科夫模型,在这一步骤中需要提前定义以下内容:

[0113] 定义1:设 $T = \{1, 2, 3, \dots\}$ 表示时刻集合;

[0114] 定义2:设 $S = (s_1, s_2, s_3, \dots)$,该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

[0115] 定义3:设 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$,表示无人机处于状态 s_n 下的可以采取的所有动作集合;

[0116] 定义4:设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合,其元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下,执行可用动作 $a_x(s_n)$ 之后,无人机状态变化到 s_m 的概率,并假设 $\sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1$;

[0117] 定义5:设 $R(s_n)$ 表示报酬集合,其中任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

[0118] 定义6:设 $A = \{a_1, a_2, \dots, a_q\}$,A代表无人机的动作空间,包含了无人机所有可能的改变无人机状态的动作,其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数 $x = 1, 2, 3, \dots, q$;

[0119] 给出无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型如下:

[0120] $MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$;

[0121] 其中, π 为策略,表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射;

[0122] 4.确定无人机在该搜索任务下状态转移概率集合 $T(s_n, a_x(s_n), s_m)$,并根据搜索任务要求确定状态转移过程中的报酬集合 $R(s_n)$;

[0123] 根据无人机当前所处的状态 s_n ,建立无人机搜索任务下搜索操作收益的最优状态值函数方程;

[0124] 本发明在计算报酬效用函数时使用MDP折扣模型,即在确定一个策略并执行后,决策者会在时刻 T 依一定的概率获取一定的报酬,报酬累加起来后就是该系统的效用函数,其中的折扣模型中的折扣因子用 γ 表示,且有 $0 < \gamma < 1$,MDP折扣模型的效用函数表示为 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_x)]$,该函数表示系统在 $t=0$ 时刻从 s_n 状态下,使用策略 π 后无人机的折扣期望总报酬,其中效用函数为有界函数;

[0125] 根据MDP折扣模型的最优方程,建立无人机在状态 s_n 下执行该搜索任务过程中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$

以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$,并

根据两个最优函数模型建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{\operatorname{argmax}} Q^{\pi(t)}(s_n, a_x)$;

[0126] 5. 初始化参数;给定参与搜索任务的无人机数目 g , $S_{\text{UAV}(i)}$ 为第 i 架无人机当前状态, $S_{\text{UAV}(i)} \in S$, $A(S_{\text{UAV}(i)})$ 为第 i 架无人机在状态 $S_{\text{UAV}(i)}$ 下的动作集合; K_i 为第 i 架无人机的最大搜索步长;给定折扣因子 γ 和策略迭代的结束条件 ε ,令迭代次数 $b=0$;其中 i 为正整数;

[0127] 6. 确定目标运动的初始位置以及每个无人机开始搜索的位置;每架无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布;从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

[0128] 7. 每个无人机根据自己的虚拟目标位置,并根据自己当前的状态 $S_{\text{UAV}(i)}$,通过计算搜索操作收益的最优状态值函数方程得到 $V^{b+1}(S_{\text{UAV}(i)})$ 迭代次数 $b=b+1$;

[0129] 8. 如果 $||V^{b+1}(S_{\text{UAV}(i)}) - V^b(S_{\text{UAV}(i)})|| < \varepsilon$,则结束迭代,进行下一步骤,否则,转到步骤7;

[0130] 9. 每架无人机根据最终得到的状态值函数 $V^{b+1}(S_{\text{UAV}(i)})$ 遍历 $A(S_{\text{UAV}(i)})$ 获得 $Q(S_{\text{UAV}(i)}, a_i)$,最终求得收益最大的行为策略 $\pi_i(t+1)^*$;每架无人机按照以上步骤都能找到下一时刻的最优搜索行为策略,如图4所示;

[0131] 10. 按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i ,状态由 $S_{\text{UAV}(i)}^n$ 转移到 $S_{\text{UAV}(i)}^m$;于此同时,无人机获得立即报酬 $r_i(S_{\text{UAV}(i)}, a_i)$, $t=t+1$,第 i 架无人机搜索步长 $k_i=k_i+1$;

[0132] 11. 若在某一个时刻 t ,第 i 架无人机位置 $S_{\text{UAV}(i)}$ 与目标真实位置 S_{target} 相同,则第 i 架无人机成功搜索到目标,搜索任务完成,算法结束;若搜索步长 $\forall k_i \geq K$, $i=1, 2, \dots, n$,则搜索失败,算法结束;多架无人机协同搜索Markov运动目标的程序框图如图5所示。

[0133] 在上述的具体实施例一中,提供了马尔科夫运动目标的无人机搜索方法,与之相对应的,本申请还提供马尔科夫运动目标的无人机搜索装置。由于装置实施例基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。下述描述的装置实施例仅仅是示意性的。

[0134] 具体实施例二

[0135] 如图6所示,本发明实施例提供了一种马尔科夫运动目标的无人机搜索装置,包括:

[0136] 目标模块201,用于接收到搜索任务后,构建马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;

[0137] 无人机模块202,用于获取无人机搜索过程中所有可能出现的状态及其概率分布;

[0138] 构建模块203,用于根据无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布,构建搜索任务下无人机行为预测的马尔科夫模型,建立基于马尔科夫决策的多阶段启发式策略迭代算法;

[0139] 规划模块204,用于利用基于马尔科夫决策的多阶段启发式策略迭代算法,获取收益最大的搜索行为策略,从而规划出无人机最优的搜索航迹。

[0140] 本发明实施例克服了传统搜索算法,如扫描线搜索算法,只注重算法本身的设计且对搜索目标运动规律缺乏严格数学定义而导致搜索代价高昂的较大缺点,根据无人机当前飞行状态和此刻马尔科夫运动目标在搜索区域的存在概率分布,确定无人机下一时刻即将搜索的虚拟目标的位置,并得到一组无人机的搜索行为操作序列,此序列为行为收益最大的搜索行为策略,能够应用于搜索运动目标,且能够以较低的搜索成本成功搜索到目标。

[0141] 优选的,所述目标模块201可以用于:接收到搜索任务后,对目标的运动区域进行栅格化划分,利用概率论对马尔科夫运动目标建模,获取马尔科夫运动目标的概率模型,从而得到马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布;所述目标的运动区域等同于无人机的搜索区域。

[0142] 优选的,所述无人机模块202可以用于:对无人机的飞行行为进行编码并描述,对无人机在执行搜索任务过程中存在的状态进行描述,从而获取无人机搜索过程中所有可能出现的状态及其概率分布。

[0143] 优选的,所述构建搜索任务下无人机行为预测的马尔科夫模型,可以具体为:

[0144] 设无人机搜索任务进行中的时刻集合 $T = \{1, 2, 3, \dots\}$;

[0145] 设无人机的离散状态空间 $S = (s_1, s_2, s_3, \dots)$,该状态空间包含了无人机搜索过程中和马尔科夫运动目标运动过程中所有可能出现的状态;

[0146] 设无人机的动作空间 $A = \{a_1, a_2, \dots, a_x, \dots, a_q\}$,表示无人机所有可能的改变状态的动作,其中元素 a_x 表示第 x 个动作, q 为动作空间中的元素个数;

[0147] 设无人机处于状态 s_n 下的可行动作集合 $A(s_n) = \{a_1(s_n), a_2(s_n), a_3(s_n), \dots\}$,表示无人机位于某个状态下可以采取的所有动作集合;

[0148] 设 $T(s_n, a_x(s_n), s_m)$ 表示无人机所有状态转移概率集合,其中的任意元素 $p(s_m | s_n, a_x(s_n))$ 表示在状态 s_n 下,执行可用动作 $a_x(s_n)$ 之后,状态变化到 s_m 的概率,其中 $i, j = 1, 2, 3, \dots, \sum_{s_m \in S} p(s_m | s_n, a_x(s_n)) = 1$;

[0149] 设报酬集合 $R(s_n)$ 的任意元素 $r(s_n, a_x(s_n))$ 表示在状态 s_n 执行动作 $a_x(s_n)$ 的报酬;

[0150] 则无人机执行任意一个搜索任务下的搜索行为预测的马尔科夫模型为:

[0151] $MDP = \{S, A, T(s_n, a_x(s_n), s_m), R(s_n)\} \rightarrow \pi(s_n)$;

[0152] 其中, π 为策略,表示从状态集合到动作集合的映射, $\pi(s_n)$ 代表无人机从状态 s_n 到动作集合的映射, \rightarrow 表示输出最优策略。

[0153] 优选的,所述规划模块204可以包括:

[0154] 计算单元,用于:用MDP折扣模型计算报酬效用函数,其中折扣因子 γ 满足: $0 < \gamma < 1$;折扣模型的报酬函数 $V^\pi(s_n) = E[\sum_{t=0}^{\infty} \gamma^t r(s_n, a_x)]$ 表示在从时刻 $t=0$ 开始无人机从状态 s_n 使用策略 π 后的折扣期望总报酬;

[0155] 根据MDP折扣模型的最优方程,建立在状态 s_n 下无人机在搜索任务中搜索操作收益的最优状态值函数方程 $V(s_n) = \max_{a_x \in A(s_n)} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) V(s_m))$,以及最优动作值函数方程 $Q(s_n, a_x) = \max_{a_x} (r(s_n, a_x) + \gamma \sum_{s_m \in S} p(s_m | s_n, a_x) Q(s_m, a_x))$,并根据两个最优函数方程建立最优搜索策略函数方程 $\pi(t+1) = \underset{\pi}{argmax} Q^{\pi(t)}(s_n, a_x)$;

[0156] 给定单元,用于:对协同搜索区域进行栅格化划分,确定MDP模型的离散状态空间 S ,给定参与搜索任务的无人机数目 g , $s_{UAV(i)}$ 为第 i 架无人机当前状态, $s_{UAV(i)} \in S$, $A(s_{UAV(i)})$ 为第 i 架无人机在状态 $s_{UAV(i)}$ 下的动作集合, K_i 为第 i 架无人机的最大搜索步长;给定折扣因子 γ 和策略迭代的结束条件 ϵ ,令迭代次数 $b=0$;

[0157] 初始单元,用于:确定目标运动的初始位置以及每个无人机开始搜索的位置;每个无人机根据目标开始运动的初始位置以及目标运动启发式信息获得目标在整个区域的存在概率分布,从而确定每个无人机下一时刻即将搜索的虚拟目标位置;

[0158] 迭代单元,用于:每个无人机根据自己的虚拟目标位置,并根据自己当前的状态 $s_{UAV(i)}$,迭代计算各自的状态值函数 $V^{b+1}(s_{UAV(i)})$,令迭代次数 $b=b+1$;

[0159] 判断单元,用于:如果 $||V^{b+1}(s_{UAV(i)}) - V^b(s_{UAV(i)})|| < \epsilon$,则结束迭代,调用遍历单元;否则,调用迭代单元;

[0160] 遍历单元,用于:每个无人机根据最终得到的状态值函数 $V^{b+1}(s_{UAV(i)})$ 遍历 $A(s_{UAV(i)})$ 获得 $Q(s_{UAV(i)}, a_i)$,最终求得收益最大的搜索行为策略 $\pi_i(t+1)^*$;

[0161] 转移单元,用于:按照所求的最优策略 $\pi_i(t+1)^*$ 执行动作 a_i ,状态由 $s_{UAV(i)}^n$ 转移到 $s_{UAV(i)}^m$;同时,无人机获得立即报酬 $r_i(s_{UAV(i)}, a_i)$,此时令 $t=t+1$,令第 i 架无人机搜索步长 $k_i=k_i+1$;

[0162] 结束单元,用于:若在某一个时刻 t ,第 i 架无人机位置 $s_{UAV(i)}$ 与目标当前模拟位置 s_{target} 相同,则第 i 架无人机成功搜索到目标,搜索任务完成,算法结束;若搜索步长 $\forall k_i \geq K_i, i=1, 2, \dots, n$,则搜索任务失败,算法结束;所述目标当前模拟位置为根据马尔科夫运动目标运动过程中所有可能出现的状态及其概率分布获取的目标当前最可能出现的位置。

[0163] 本发明从使用目的上,效能上,进步及新颖性等观点进行阐述,其具有的实用进步性,已符合专利法所强调的功能增进及使用要件,本发明以上的说明及附图,仅为本发明的较佳实施例而已,并非以此局限本发明,因此,凡一切与本发明构造,装置,待征等近似、雷同的,即凡依本发明专利申请范围所作的等同替换或修饰等,皆应属本发明的专利申请保护的范围之内。

[0164] 需要说明的是,在不冲突的情况下,本发明中的实施例及实施例中的特征可以相互组合。尽管本发明已进行了一定程度的描述,明显地,在不脱离本发明的精神和范围的条件下,可进行各个条件的适当变化。可以理解,本发明不限于所述实施方案,而归于权利要求的范围,其包括所述每个因素的等同替换。对本领域的技术人员来说,可根据以上描述的技术方案以及构思,做出其它各种相应的改变以及形变,而所有的这些改变以及形变都应该属于本发明权利要求的保护范围之内。

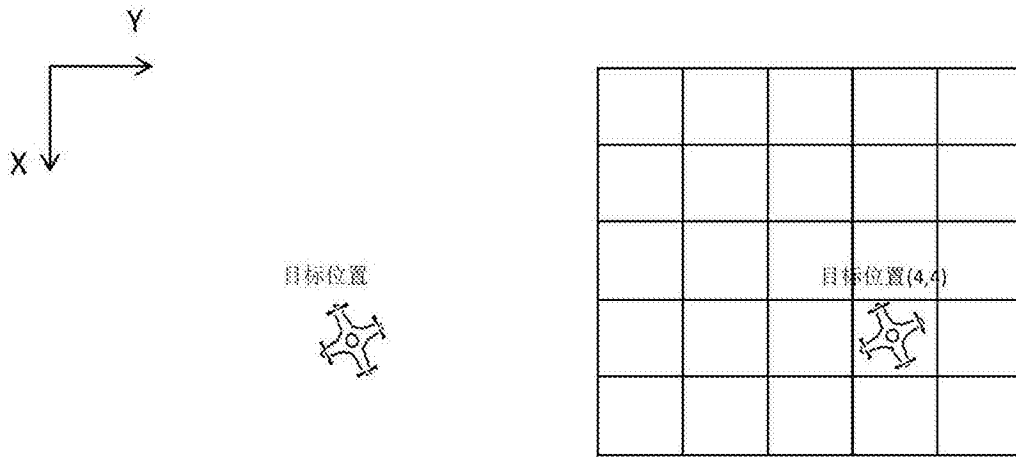


图1

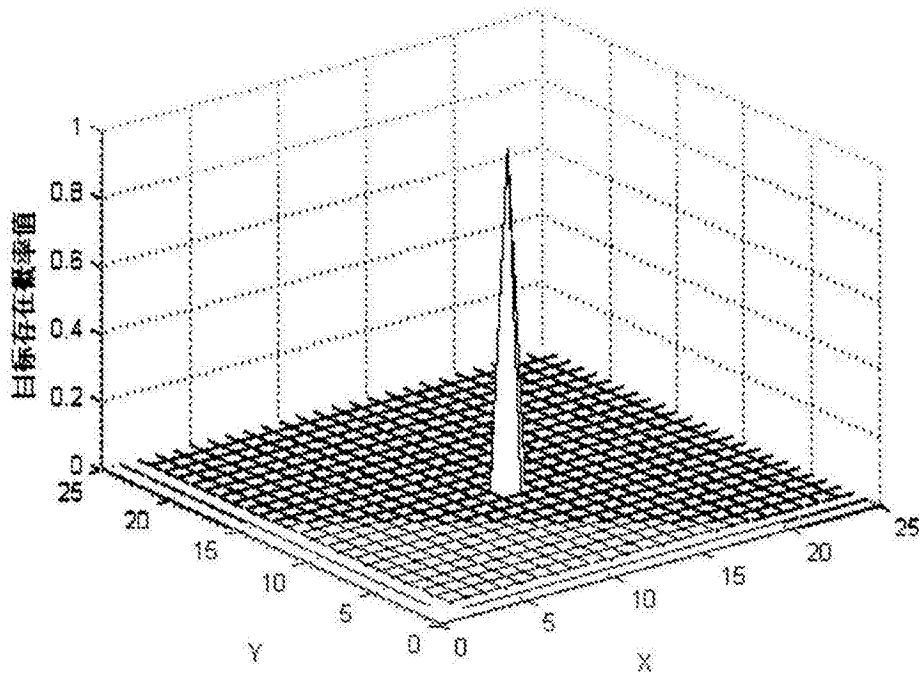


图2a

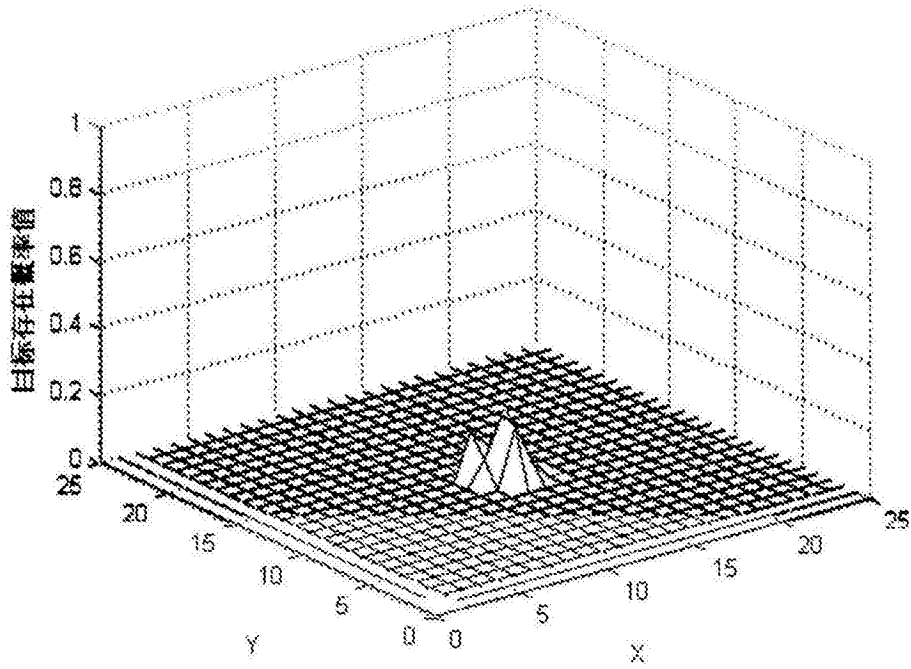


图2b

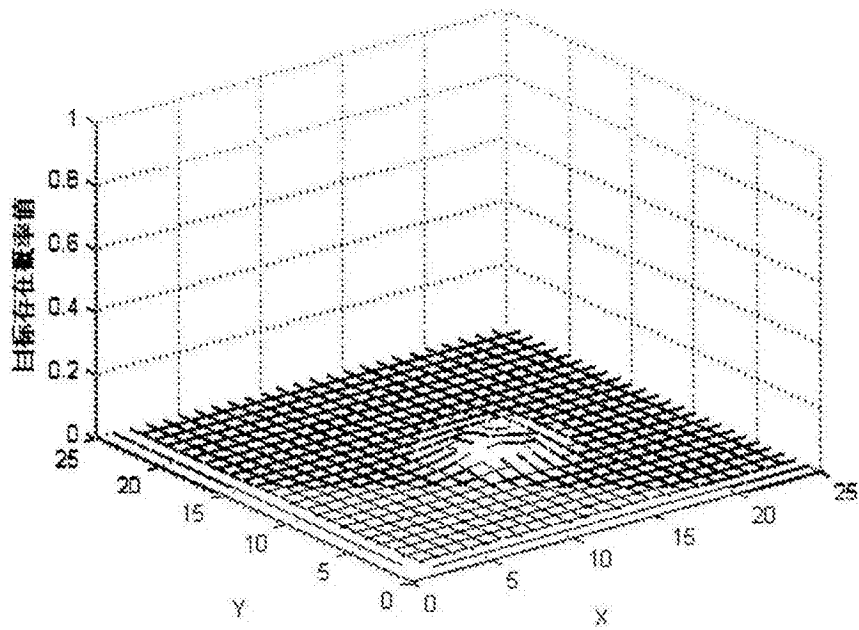


图2c

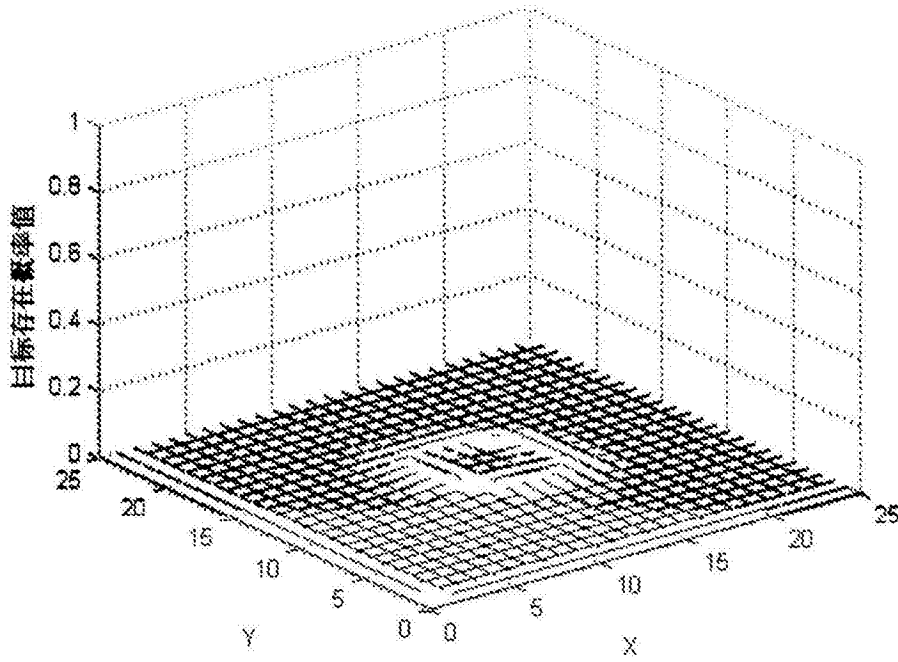


图2d

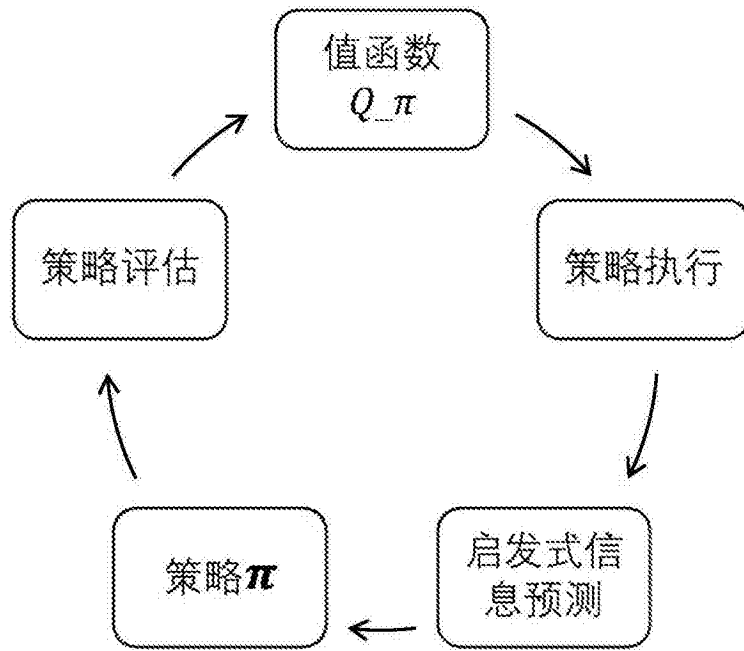


图3

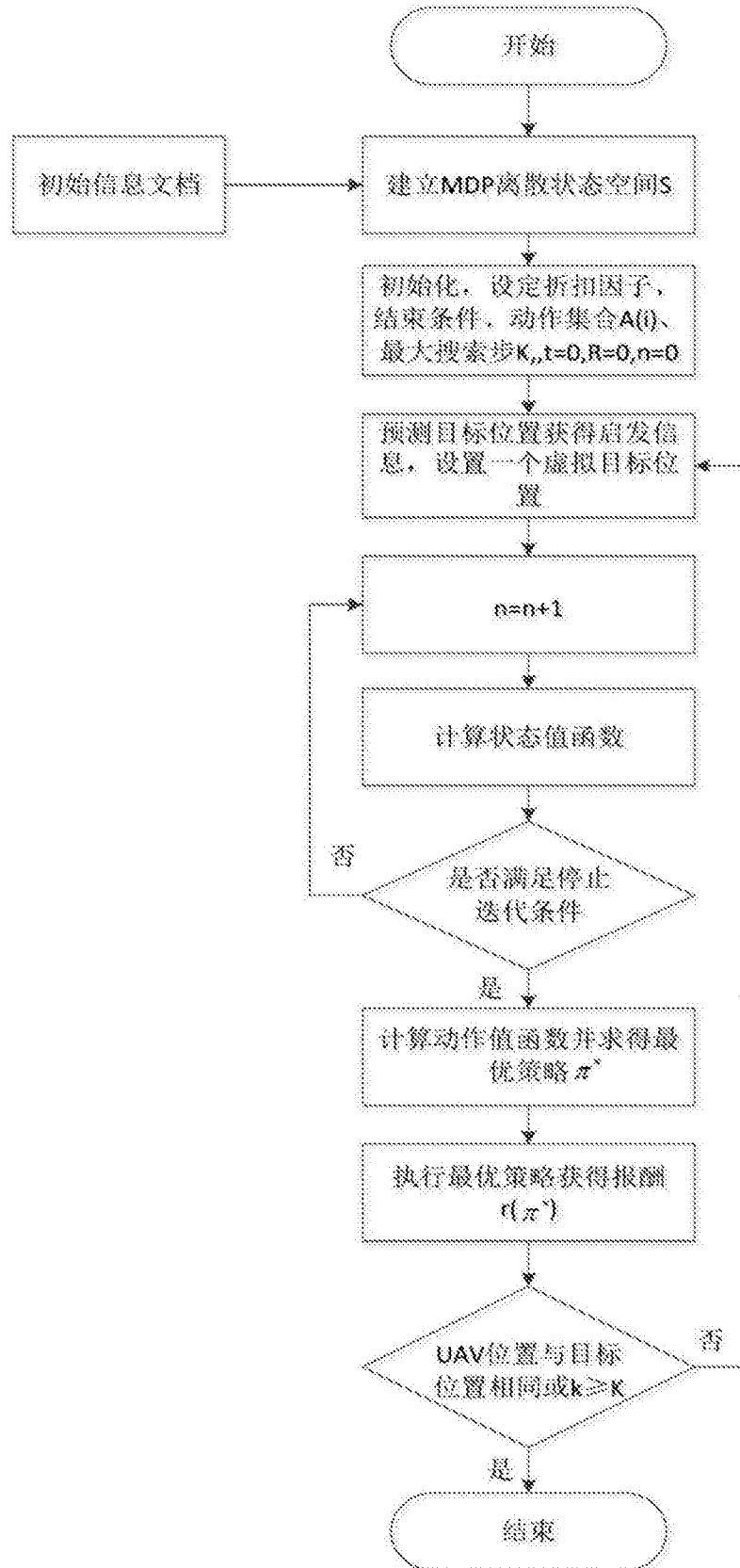


图4

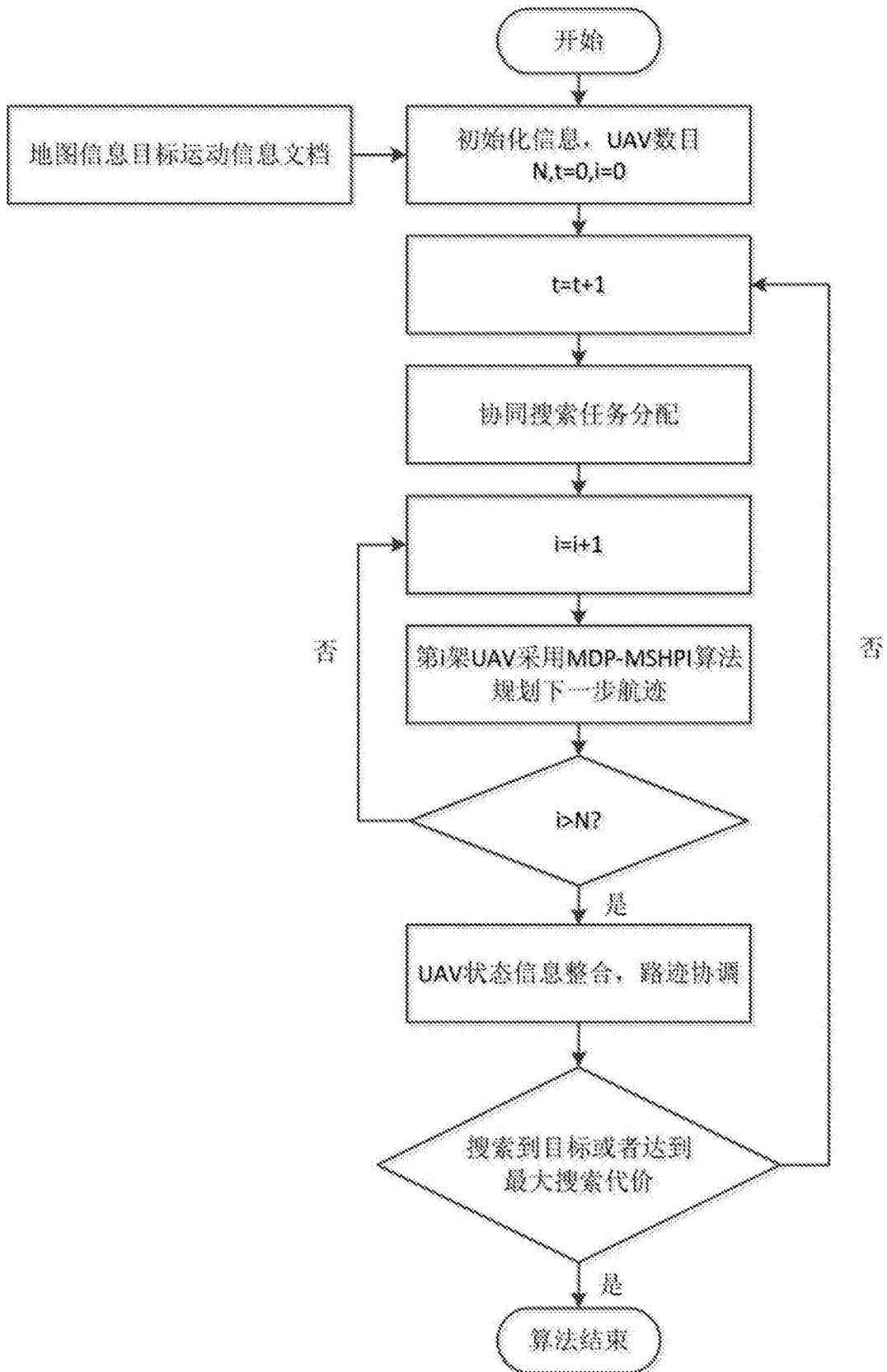


图5



图6