基于 HMDP 的无人机三维路径规划

洪 晔 房建成

(北京航空航天大学 仪器科学与光电工程学院, 北京 100191)

摘 要: 路径规划是 UAV(Unmanned Aerial Vehicle)自主飞行的重要保障. 初步建立了基于 MDP(Markov Decision Processes)的全局路径规划模型,把 UAV 的路径规划看作是给定环境模型和奖惩原则的情况下,寻求最优策略的问题;为解决算法时空开销大、UAV 航向改变频繁的缺点,提出一种基于状态聚类方法的 HMDP(Hierarchical Markov Decision Processes)模型,并将其拓展到三维规划中. 仿真实验证明:这种简单的规划模型可以有效解决 UAV 的三维全局路径规划问题,为其在实际飞行中的局部规划奠定了基础.

关 键 词: 无人机(UAV); 路径规划; 马尔可夫决策过程(MDP); 分层马尔可夫决策过程(HMDP); 仿真

中图分类号: TP 24

文献标识码: A **文 章 编 号**: 1001-5965(2009)01-0100-04

Hierarchical Markov decision processes based path planning for UAV in three-dimensional environment

Hong Ye Fang Jiancheng

(School of Instrument Science and Opto-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: The ability of path planning is an important ensure for unmanned aerial vehicle (UAV) in autonomous flight. A path planning model was based on Markov decision processes (MDP), in which the problem of path planning was regarded as looking for the best tactic through the model of environment and the principle of rewards and punishment. To solve the problem such as huge space-time spending and changing course at high frequency, the hierarchical Markov decision processes (HMDP) were introduced based on the method of clustering states. The arithmetic was also used for path planning in three-dimensional environment. The results of simulation show the HMDP model can be used to path planning for UAV in three-dimensional environment. It lays the foundation for local path planning in real flight.

Key words: unmanned aerial vehicle(UAV); path planning; Markov decision processes(MDP); hierarchical Markov decision processes(HMDP); simulation

路径规划作为无人机(UAV, Unmanned Aerial Vehicle)自主飞行的重要保障,是指依靠已知的地形(包括障碍)信息和威胁信息,在某些约束条件下,寻找到从起点到目标点的可行飞行路线.

目前,路径规划的研究方法很多[1-2],基于决策论的路径规划是近年来才出现在人工智能研究领域的[3],作为一种处理顺序决策问题的规划方法,特别适合于处理不确定情况下的规划问题. MDP(Markov Decision Process)则是其中应用最为基础的模型,在这个理论框架下,可以把 UAV 的路径规划问题看作是给定环境模型和奖惩原则的情况下,寻求最优策略的问题. 针对某型 UAV 进行低空超视距飞行的环境和飞机运动特性,本文初步建立了基于 MDP 的全局路径规划模型,由于出现时空开销大、航向改变频繁的缺点,提出

收稿日期: 2008-02-28

基金项目: 国家自然科学基金重点资助项目(60736025)

作者简介: 洪 晔(1979—),女,黑龙江哈尔滨人,讲师,hongye@buaa.edu.cn.

一种基于状态聚类方法的 HMDP(Hierarchical Markov Decision Process)模型,并将其拓展到三维全局规划中.

1　平面 MDP 模型描述

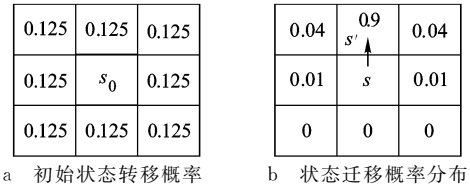
1.1　MDP 模型定义

一个 MDP 可以用一个四元组 $M=\langle S,A,T,R\rangle$ 描述[4]:

S:指包括所有环境状态的有限集合.定义 40 km×40 km 范围的环境作为 UAV 路径规划的环境状态.规划时,基于栅格法以 100 m(由实际无人机运动约束决定)间隔进行二维离散化空间建模,得到 1.6×10^5 个空间状态.

A:指包括所有动作的有限集合.定义 UAV 有 8 个可行的动作,分别为:北、东北、东、东南、南、西南、西和西北.

$T:S\times A\to\Pi(T)$ 是状态转移函数,在给定目前状态和动作的情况下,下一状态的概率分布将直接决定下一动作的输出.这里认为在没有目标时的初始状态转移概率是平均分配的,如图 1a.以目标点在正北方为例:若给出动作北,状态 s 周围栅格的状态迁移概率分布如图 1b 所示,表示给定目前状态 s 和动作北的情况下,下一状态 s' 的概率分布,已知状态分布后,再根据相应的报酬,就可以得到最优策略.需要注意的是,这只是概率分布的一种特殊情况,分布值会随着目标点、障碍情况发生变化.



a　初始状态转移概率　　b　状态迁移概率分布

图 1　状态迁移概率分配图

$R:S\times A\to\Pi(R)$ 是报酬函数,表示在给定目前状态和动作的情况下所期望的立即报酬(一般用 $R(s,a)$ 来表示在状态 s 下执行动作 a 所能得到的立即报酬).这里构造无模型的均匀表示的报酬函数模型 R_m 和 R_a:二者分别为正常飞行和遇到障碍时的报酬函数值.

$$R_m=\begin{cases}5 & \text{趋向目标}\\ -1 & \text{远离目标}\\ -10 & \text{发生碰撞}\end{cases}\tag{1}$$

$$R_a=\begin{cases}5 & \text{远离障碍}\\ -5 & \text{没有远离障碍}\\ 0 & \text{没有障碍}\end{cases}\tag{2}$$

在这个模型中,下一个状态和期望获得的立即报酬只和当前状态、所执行的动作有关,而与历史无关,这就是所谓的马尔可夫属性[5]:$t+1$ 时刻的状态和报酬只依赖于 t 时刻的状态和在 t 时刻执行的动作.

1.2　搜索策略

UAV 对动作策略进行搜索需要考虑因素:①必须对环境状态空间作充分的探索,从而能够找到最优的或者次优的策略,即探索问题;②要利用通过概率学习获得的经验进行动作选择,即利用问题.两者相互矛盾,如何合理地平衡两者进行有效的动作选择,即搜索策略问题.本文采用动态规划,使要搜索的那些评价函数最优,即无限折扣报酬期望和最大的决策序列,最优评价函数为[5]

$$V^*(s)=\max_{\pi}E\left(\sum_{t=0}^{\infty}\gamma R_t\right)\tag{3}$$

其中,γ 为折扣因子,本文取 0.9;$E(\cdot)$ 为期望值;R_t 为 t 时刻的立即报酬.对于任意状态,使评价函数最优的充要条件为[5]

$$V^*(s)=\max_a\Big(R(s,a)+\gamma\sum_{s'\in S}P(s'\mid s,a)V^*(s')\Big)\tag{4}$$

式中,P 为转移概率.式(4)为 Bellman 方程.相应地,最优策略 π^* 为

$$\pi^*(s)=\arg\max_a\Big(R(s,a)+\gamma\sum_{s'\in S}P(s'\mid s,a)V^*(s')\Big)\tag{5}$$

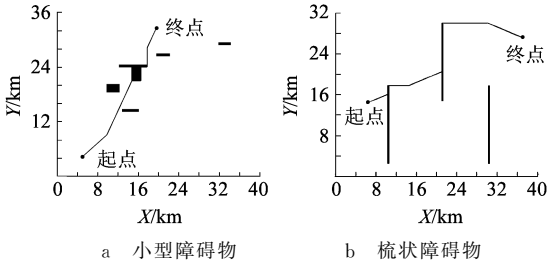
对于 Bellman 方程的解,采用函数迭代法,即直接对最优评价函数 V^* 进行搜索.设在时间步 t,系统的状态为 s,$V^*(s)$ 则按下式进行迭代:

$$V_{t+1}(s)=\max_a\Big(R(s,a)+\gamma\sum_{s'\in S}P(s'\mid s,a)V_t(s')\Big)\tag{6}$$

对两步连续迭代的评价函数的最大值进行比较,如果其差值小于指定的精度 ε,则结束迭代过程.

1.3　平面 MDP 模型下的二维路径规划

为了测试基于 MDP 模型的规划算法的性能,在各种复杂环境中,利用VC++6.0编程,在



a　小型障碍物　　b　梳状障碍物

图 2　附加不同障碍的 MDP 规划实验图

2 平面 HMDP 模型描述

2.1 状态聚类

通过上文构建状态空间的方法，可以看到，栅格的大小影响着算法的时空复杂度. 栅格的规格越小，划分环境后得到的小区域越多，则数据所占的内存空间越多，搜索时速度也就越慢. 但是栅格太大又将影响路径的精确程度. 这里构造的平面状态为 1.6×10^5 个，如果拓展到三维空间，时空开销是很大的，同时在规划中(图 2)可以看到 UAV 出现了航向频繁变化的问题，规划的转向角度也无法在实际飞行中实现. 因此引入状态空间聚类的思想.

状态聚类的思想是通过把原始状态归并为较小的集结状态集合，从而在更小维数的空间上进行路径规划，明确地说，把环境状态空间 S 划分成 m 个子集：S_1，S_2，…，S_m，$S=S_1\cup S_2\cup\cdots\cup S_m$，把每个子集作为一个集结状态，进行规划，这样迭代的次数会明显减少，有效克服维数灾和节点过多的问题.

依据标准的 MDP 模型进行状态类聚，加入分层结构. 定义由 MDP 组成完整的分层系统，它们可以分别转化成标准的 MDP[5]. 重新定义四元组 $M^n=\langle S^n,A^n,T^n,R^n\rangle$，其中 n 代表层数，根据状态设定. 定义 M^0 为初始的平面 MDP，当 $n\geqslant1$ 时，M^n 由 M^{n-1} 通过聚类状态 S^{n-1} 得到，每类状态聚类后变为一个状态，在分层的过程中，无形之间减少了空间状态数量，大大加快了搜索速度.

对于四元组的其他变量作如下定义：

$T^n=P(s_m^n\mid s_k^n,a)$，为在第 n 层上由状态 s_k^n 转移到状态 s_m^n 的概率；

$R^n=R(s_m^n,a^n,s_k^n)$，为在第 n 层上由状态 s_k^n 转移到状态 s_m^n 所得到的报酬；

每一层上的状态转移概率和报酬函数与平面 MDP 模型中计算方法相同，只是范围已经被局限在所在层的状态之间进行.

2.2 基于 HMDP 模型的二维路径规划

根据环境中的障碍信息，进行状态聚类，参照文献[6－8]中的八叉树方法，采用纵向划分状态层次的方法，如图 3 所示. 假设环境中有复杂形状的障碍物，按照纵向划分标准把初始环境状态划分为 10 个子状态，采用迭代策略，计算每一层次状态之间的转移概率，在第 1 次的搜索中规划的可能路径为 1→2→4→6→8→9→10，当底层完成后再返回到上层，再次搜索时已经摒弃了一些无关的状态，只在底层选定的状态中继续寻优. 在图 3 的例子中，在底层的路径规划实际上是路径的粗略选择，在顶层的规划其实是路径的细化和执行过程.

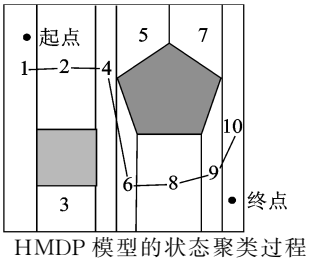
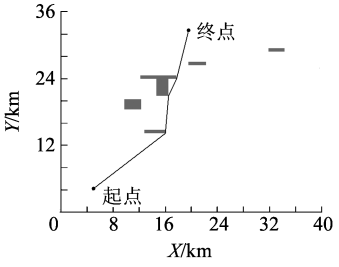
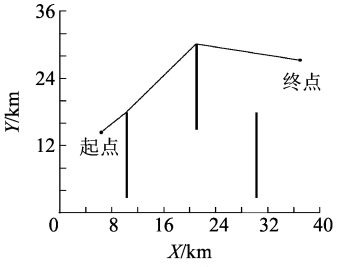


图 3 HMDP 模型的状态聚类过程

对比平面 MDP 规划附加多个障碍物的例子，进行相同条件下的仿真实验，结果如图 4 所示. 比较图 2 和图 4，将两种算法的性能比较由表 1 列出. 可以看出分层后的规划路径具有节点少、路径代价小、搜索时间快等优点. 同时减少了 UAV 的航向变化，因此规划更具有合理性.



a 小型障碍物



b 梳状障碍物

图 4 附加不同障碍的 HMDP 规划实验图

表 1 不同算法性能分析

实验方法		规划时间/ms	描述路径结点数	路径代价
附加小型障碍	MDP	90	8	854
	HMDP	60	5	706
附加梳状障碍	MDP	165	8	1 432
	HMDP	101	4	1 020

3 基于 HMDP 模型三维路径规划

通过前面的基于 MDP 规划算法的分析与研究，已经能够很好地求解 UAV 的二维全局路径规划问题. 但是，在 UAV 的实际飞行中，必须有高度方向的运动，且 UAV 的使命要素中可能包

含不同高度的规划信息,即要求 UAV 具有三维规划的能力.将 HMDP 模型拓展到三维环境中(针对某型 UAV 进行低空超视距飞行的环境 40 km×40 km×1 km),三维的空间状态并不是立体的栅格形式,考虑 UAV 的最大爬升角,定义高度方向的栅格尺寸为 50 m,三维 HMDP 模型中的转移概率和报酬函数与平面中的相同.

如图 5 所示,在三维环境坐标系(X-Y-Z)下根据障碍的高度(h_1～h_2之间),按照 2.2 节中的方法纵向聚类垂直面内的状态,在这里对障碍物进行规则化描述,均表示为长方体形式.

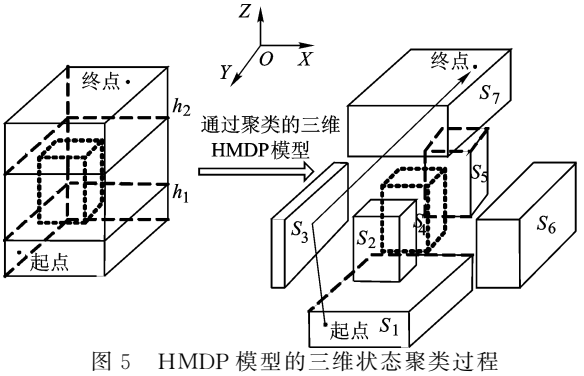


图 5　HMDP 模型的三维状态聚类过程

分层后得到的顶层规划状态为 $S_1, S_2, \cdots, S_7$,计算这些状态的转移概率,并按照 R 进行 $V^*(s)$的计算,搜索最大报酬动作,初步得到可行状态为 $S_1 \rightarrow S_3 \rightarrow S_7$,再在这 3 个状态上最后确定最优的细化路径.分层后将问题分解,避免产生由于数据量太大造成的维数灾.为了真实地描述实际环境中规划问题,利用 Matlab 生成了带有基准面的网络图,并附加了两个障碍,图 6 为三维规划图,可以看到,这种简单的 HMDP 模型是合理、有效的.

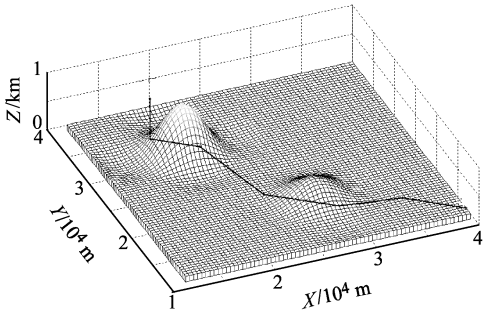


图 6　基于 HMDP 的三维路径规划图

4　结　论

带有层次的 HMDP 模型的总体思想其实仍旧是一种规划策略,一种从分层状态的概率分布到 UAV 采取最佳动作的优化匹配.划分层次后,顶层的 MDP 直观地给出了到达目标点所应走的总体路线,并且包含了 UAV 的当前状态,通过这种方法摒弃了与规划无关的状态.顶层规划完成后,再在每一层状态中搜索最优路径.同时,结合高度分层,实现了三维路径规划,为 UAV 在实际飞行中的局部规划奠定了基础.

参考文献(References)

[1] 张建英,刘暾. 基于人工势场法的移动机器人最优路径规划[J].航空学报,2007, 28(S1): 183－188
Zhang Jianying, Liu Tun. Optimized path planning of mobile robot based on artificial potential field [J]. Acta Aeronautica et Astronautica Sinia, 2002,28(S1): 183－188(in Chinese)

[2] 孙汉昌,朱华勇. 基于概率地图方法的无人机路径规划研究[J].系统仿真学报, 2006(11): 3050－3054
Sun Hanchang, Zhu Huayong. Study on path planning for UAV based on probabilistic roadmap method [J]. Journal of System Simulation, 2006(11): 3050－3054(in Chinese)

[3] Foka A F, Trahanias P E. Predictive autonomous robot navigation [C]// Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ:IEEE, 2002:490－495

[4] Roy N, Burgard W, Fox D, et al. Coastal navigation mobile robot navigation with uncertainty in dynamic environments [C]//IEEE International Conference on Robotics and Automation. Pittsburgh: [s. n.], 1999, 5(1): 35－40

[5] Baker B, Zivkovic Z, Krose B, et al. Hierarchical dynamic programming for robot path planning[C]// Proceedings of the 2001 IEEE International Conference on Robotics & Automation. Orleans:IEEE, 2002, 3(2): 46－50

[6] 史红兵,张毅彬,童若锋,等. 虚拟场景自动漫游的路径规划算法[J]. 计算机辅助设计与图形学学报,2006(11):592－598
Shi Hongbing, Zhang Yibin, Tong Ruofeng, et al. Path planning for automated navigation in virtual environment [J]. Journal of Computer-aided Design & Computer Graphics, 2006(11):592－598(in Chinese)

[7] Salomon B, Garber M. Interactive navigation in complex environments using path planning[C] // Proceedings of Symposium on Interactive 3D Graphics. California, USA: [s. n.], 2003: 41－50

[8] Peters B, Dziugys A. Numerical simulation of the motion of granular material using object-oriented techniques[J]. Computer Methods in Applied Mechanics and Engineering, 2002 (3): 193－200

(责任编辑:刘登敏)