

# 一种动态不确定环境中机器人路径规划方法<sup>①</sup>

肖国宝, 严宣辉

(福建师范大学 数学与计算机科学学院, 福州 350007)

**摘要:** 针对动态不确定环境下的机器人路径规划问题, 将部分可观察马尔可夫决策过程 (POMDP) 与人工势场法 (APF) 的优点相结合, 提出一种新的机器人路径规划方法。该方法充分考虑了实际环境中信息的部分可观测性, 并且利用 APF 无需大量计算的优点指导 POMDP 算法的奖赏值设定, 以提高 POMDP 算法的决策效率。仿真实验表明, 所提出的算法拥有较高的搜索效率, 能够快速到达目标点。

**关键词:** POMDP; APF; 路径规划; 动态不确定环境; 移动机器人

## Path Planning of Mobile Robot in Dynamic Nondeterministic Environments

XIAO Guo-Bao, YAN Xuan-Hui

(School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

**Abstract:** This paper introduces a new path planning in dynamic nondeterministic environments. We combine POMDP and APF into the new path planning which takes full account of the uncertainty of the information in real world. Based on the APF's advantage of avoiding the expensive computation, it guides the setting of POMDP's rewards value to improve the efficiency of decision making. The result of the simulation shows that the proposed algorithm has higher search efficiency and can make the robot reach the target faster.

**Key words:** POMDP; APF; path planning; dynamic nondeterministic environments; mobile robot

### 1 引言

所谓移动机器人路径规划技术<sup>[1]</sup>, 就是机器人根据自身传感器对环境的感知, 在线规划出一条安全的、可靠的运行路径, 同时高效完成作业任务。部分研究学者从环境感知的角度, 将移动机器人路径规划分为3种类型<sup>[2]</sup>: 基于环境模型的规划方法、基于事例学习的规划方法和基于行为的路径规划方法; 从机器人路径规划的目标范围看, 又可分为全局路径规划和局部路径规划; 从规划环境是否随时间变化方面看, 还可分为静态路径规划和动态路径规划。

未知环境下的机器人路径规划是智能体实现其在线实时规划的前提, 也是移动机器人导航中一个重要的问题。目前研究中, 确定性环境的导航控制方法已取得了大量的研究和应用成果, 虽然对未知环境中的导航控制已开展了一些研究, 并提出了若

干方法<sup>[3-5]</sup>, 但尚未有完善的体系结构和方法, 部分可观察马尔可夫决策过程 (partially observable Markov decision processes, POMDPs) 为解决带有这种不确定性的规划问题提供了丰富的数学框架<sup>[6,7]</sup>。它将环境的变迁看作状态空间上的 Markov 链, 用状态空间上的信念分布表示主题对当前状态的估计, 并根据主体的行动和获取的观察加以更新。在此基础上, POMDP 依据效用最大化原则, 根据回报函数和状态转换函数计算最优策略<sup>[8]</sup>。

目前国内外研究机器人路径规划的方法中, 人工势场法 (Artificial potential field, APF) 已经较为成熟<sup>[9-11]</sup>, 该方法优点在于无需大量的计算, 能够自动的产生光滑的路径, 提供一种确定性的规划信息, 并能够实现实时在线规划, 但容易陷入局部最小而导致目标不可达等问题。

① 收稿时间:2011-07-30;收到修改稿时间:2011-09-12

本文融合了 POMDP 算法与 APF 模型的优势,给出一种新的机器人路径规划方法。主要思想是,利用 POMDP 的信念分布描述主体对当前环境的认知,以适应环境的不确定性;利用 APF 计算量小和实时性强的优点,通过影响 POMDP 的回报函数快速地指导智能体的动作选择,有效降低了 POMDP 设定奖赏值有效降低了 POMDP 设定奖赏值的复杂度,使得智能体快速到达目标点,提高决策效率和对动态不确定环境的适应能力。

本文第 2 节描述了传统 POMDP 模型,包括 POMDP 模型的定义和 POMDP 算法分析;第 3 节描述了 POMDP 算法如何映射到机器人路径规划问题,并分析了震荡现象和局部陷阱问题;第 4 节描述了本文算法的具体步骤;最后,第 5 节给出了本文算法的仿真实验,并进行算法比较分析。

## 2 POMDP模型

在 POMDP 模型中,智能体必须利用随机环境中部分观察到的信息进行决策,在每个时间点上,智能体都可能是众多可能状态中的某一状态,但是由于观察到的信息不完整或者是不可能直接知道自己的当前状态,它必须利用现有的部分信息、历史动作序列和立即报酬值来采用一种策略进行决策<sup>[12,13]</sup>。

### 2.1 POMDP 模型的定义

通常,一个 POMDP 模型可以描述为一个六元组: $M = \langle S, A, T, R, \Omega, O \rangle$ <sup>[14]</sup>,其中:

$S$ , 表示智能体的状态集,是一个表示环境状态的有限集,且其集合元素是部分可观测的;

$A$ , 表示动作的有限集合;

$\Omega$ , 表示观测值的有限集合;

$T(s, a, s')$ :  $S \times A \rightarrow \Pi(S)$ , 智能体的状态转移函数,表示智能体在状态  $s$  下采用动作  $a$  时可能转移到状态  $s'$  的概率,用  $p(s' | s, a)$  表示;

$O(s', a, o)$ :  $S \times A \rightarrow \Pi(O)$ , 智能体的观察函数,表示智能体采用动作  $a$  后,处在状态时得到状态  $s'$  值  $o$  的概率,用  $p(o | s', a)$  表示;

$R(s, a)$ :  $S \times A \rightarrow R$  是智能体的报酬函数,表示在状态  $s$  下采用动作  $a$  时返回系统的报酬值;

在实际应用中, POMDP 模型用来模拟智能体和环境之间的交互。如图 1, 智能体首先通过传感器观测实时环境,取得状态信息和信度信息,并根据解决

策略决定采取一个动作;动作过后,智能体可获取一个即时奖赏值;而环境在动作的影响下,转到了一个新的状态;智能体也转到了一个新的状态,故能够获得一个即时观测概率值。如此通过智能体和环境的交互过程,以取得求解问题的整个策略。

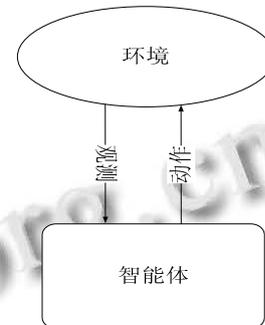


图 1 POMDP 模型中智能体与环境的交互

### 2.2 POMDP 算法分析

本文分析基于部分点的值迭代改进 POMDP 算法,此算法用基于部分点的值迭代代替原有的对全部状态的值迭代计算,减少更新的次数来加速值迭代的过程,以达到尽快收敛的目的。

ITERATOR 迭代算法:

Step1: 初始化值函数  $V = V_0, t = 0$ , 设定阈值  $\varepsilon$ ;

Step2: 更新值函数

$V_{t+1} = TV_t; t = t + 1; k = \max_{b \in S} |V_t(b) - V_{t-1}(b)|$ , 其中  $S$  为智能体的状态集

Step3: 判断  $k$  是否大于阈值  $\varepsilon$ , 若大于则终止循环跳到下一步, 若小于则跳到 Step2;

Step4: 输出  $V_{t-1}$

在机器人路径规划算法中大部分算法对不确定性的适应能力是较差的, 本文通过分析动作不确定性和环境不确定性两方面分析 POMDP 的适应能力, 并阐述了奖赏值的概念。

(1) 动作不确定性分析

POMDP 使用一个状态转换分布  $S \times A \rightarrow \Pi(S)$ , 而不是单一的结果来描述动作的效果。这在一定程度上能够体现智能体在主观上认识到了行动不确定性, 并且能够将这种认识应用于对环境状态的估计对计划选择产生影响, 从而保证对动作不确定性有一定程度的适应能力。

(2) 环境不确定性分析

在 POMDP 模型中  $M = \langle S, A, T, R, \Omega, O \rangle$ , 状态集、

观测集、状态转移函数和观测函数都是智能体对环境认识,对于不确定环境,智能体通过信念分布来估计所观测到的状态,并用观测函数来表示,从而对环境不确定性有一定程度的适应能力。

### (3) 奖赏值定义

智能体的奖赏值,通常用具体实数表示,值的大小是评价动作的好坏,能够用于改善规划性能。

## 3 问题描述

在设计 POMDP 算法主要考虑三个问题:状态空间和动作空间的表示;奖赏值的设置;根据奖赏值合适的动作选择。

### 3.1 状态空间的表示

在本文的仿真实验中,用二维坐标表示环境信息,用一个整数坐标对应状态集中的一个状态,并由一组概率值来表示栅格存在障碍物的可能性。此处进行状态空间的压缩,只考虑智能体四周的环境信息,故状态空间只填充有限的状态。

状态空间的每个元素分为障碍点和非障碍点,并赋予一组概率值 $(x, y)$ ,其中表示元素为障碍点的概率, $y$ 表示非障碍点的概率。本文的算法是在动态不确定环境中进行,故没有先验知识,所有元素的初始概率为 $(0.5, 0.5)$ ,通过智能体传感器的探测对元素进行更新。

### 3.2 动作空间的表示

本文采用理想状态下智能体可以到达四周的任意点,故动作集采用整数倍角度表示,即此处的动作集有360个元素。智能体可以在下个时刻到达四周的任意点,动作的选择依赖于智能体所获得的奖赏值决定。转换概率由仿真程序设定,此处采用的转换概率为0.9,表示智能体成功执行动作的概率为0.9。

### 3.3 奖赏值的设定

奖赏值的设置对智能体的规划能力很重要,是评价动作的好坏,能够用于改善规划性能。本文采用 APF 方法(人工势场法),根据势能越小,奖赏值越大的原则,分别给予智能体在当前状态下选取每个动作时设定一个即时奖赏值。智能体的值函数为

$$V_t^*(s) = \max_{a \in A} [R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') V_{t-1}^*(s')] \quad (1)$$

其中, $\gamma$ 是折扣因子,决定着智能体所获得奖赏值的重要性。 $V_t^*(s)$ 表示 $t$ 时刻的值函数。

### 3.4 震荡现象的判断

假设上一步智能体规划位置为 $X_i$ ,当前规划位置为 $X_j$ ,下一步规划位置为 $X_k$ ,如果三个向量满足公式(2),则认为路径发生震荡现象,反之认为是正常现象。

$$\arccos[c / (|a| \cdot |b|)] > \frac{\pi}{4} \quad (2)$$

其中, $a = X_j - X_i, b = X_k - X_j, c = a \cdot b$ 。

### 3.5 局部陷阱的判断

由于是在动态不确定环境下进行路径规划,在进行局部陷阱判断时,只考虑目标距离势能。令 $E_{\min} = \min\{E_i, i = 0, \dots, n-1\}$ ,随着智能体向目标点靠近, $E_n$ 将逐渐减小。令

$$\Delta \hat{E} = E_n - E_{\min} \quad (3)$$

显然,在无障碍物环境或障碍引起的势场改变度不大的环境下,目标距离势能必然呈现单调下降的趋势,则 $\Delta \hat{E} < 0$ , $E_{\min}$ 应随着传感器数据的变化不断更新到最小值。但在复杂的环境下,可能出现局部势能陷阱,出现 $\Delta \hat{E} > 0$ 的情况。在此根据 $\Delta \hat{E}$ 的变化来衡量智能体是否陷入局部势能陷阱,并触发智能体施加随机噪声。

## 4 基于POMDP算法与APF导向的规划算法

POMDP 导航属于反应式规划方法,此类型规划方法的优点在于能够对环境的即时变化作出响应,计算代价小,能够实时运行,可以应用于未知以及动态环境<sup>[15,16]</sup>。基于 POMDP 算法与 APF 导向的规划算法实现步骤如下:

Step1:确定滚动窗口,获取局部环境信息

此算法用于实时规划,必须实时更新环境信息,而智能体的决策只与靠近传感器的环境信息有关,此处采用滚动窗口的形式来描述局部信息,滚动窗口以智能体的参考坐标为中心,滚动窗口的大小及形状与智能体自身的外形、传感器的感知区域有关,常设计为矩形、圆形或半圆形区域<sup>[17]</sup>,智能体在每一步的行动计划中,在滚动窗口内搜索满足条件的局部优化解,即搜索最佳动作决策。

Step2:更新状态集的元素信息

将传感器的读数转化为概率(以概率表示的可信度与离散点到传感器的欧式距离成反比),并用概率论

中的贝叶斯规则组合概率, 将传感器模型产生的条件概率通过贝叶斯规则进行转化, 同一时刻不同传感器或同一传感器不同时刻的概率都可以用该方法加与融合。每个离散点的状态为障碍物或非障碍物, 则为障碍物的概率如公式 (4)。

$$p(\text{障碍物}) = \frac{\left(\frac{R-r}{R}\right) + \left(\frac{\beta-\alpha}{\beta}\right)}{2} \times \text{Max}_{\text{occupied}} \quad (4)$$

式中,  $r$  和  $\alpha$  分别表示传感器到离散点的距离和角度,  $R$  表示滚动窗口的半径,  $\beta$  表示每个传感器视野的扇形宽度半角。  $\frac{\beta-\alpha}{\beta}$  项和  $\frac{R-r}{R}$  项表示离散点离传感器越近, 可信度也越高。则  $\text{Max}_{\text{occupied}}$  表示离散点为障碍物的最大概率。

#### Step3: 计算滚动窗口边缘的势能函数

本文使用改进的人工势场函数。相较于传统的人工势场法, 改进的人工势场函数可以避免智能体在临近目标点时的抖动现象, 保证目标点为全局环境下势能最小点, 解决了目标不可达的问题。

式 (5) 和式 (6) 分别表示引力势场函数和斥力势场函数<sup>[18,19]</sup>:

$$U_G = \frac{1}{2} k_p |X - X_G|^2 \quad (5)$$

$$U_{RO} = \begin{cases} \frac{1}{2} \eta \left| \frac{1}{\rho} - \frac{1}{\rho_0} \right|^2 |X - X_G|^2, & \rho \leq \rho_0 \\ 0, & \rho > \rho_0 \end{cases} \quad (6)$$

其中,  $k_p, \eta$  是正增益系数;  $X = (x_p, y_p)$  和  $X_G = (x_G, y_G)$  分别表示智能体位置向量和目标点位置向量;  $\rho$  表示智能体到障碍物的欧式距离;  $\rho_0$  表示单个障碍物影响的最大距离范围。

#### Step4: 计算奖赏值, 确定临时目标点

在状态集中选择障碍物概率小于一定阈值的元素组成候选元素集, 根据势能越小, 奖赏值越大的原则在智能体当前状态下采用某个动作到达每个候选集状态设定一个即时奖赏值, 由公式 (1) 更新值函数。将智能体采用值函数最大的动作到达的状态确定为下一个临时目标点。临时目标点的确定决定了智能体规划路径的走向, 本文通过人工势

场法来指导智能体临时目标点的确定, 能够提高决策效率。

#### Step5: 验证临时目标点的有效性, 合理地选择动作

根据公式 (2) 来判断规划的路径轨迹是否过分震荡, 以保证智能体规划的路径具有一定的平滑度。若路径无效, 将临时目标点从候选元素集中去除, 并在剩下的候选元素集中重新选择临时目标点; 根据公式 (3) 判断智能体是否陷入势能陷阱, 若  $\Delta \hat{E} > 0$ , 则添加一个随机噪声, 让智能体摆脱陷阱。

#### Step6: 判断是否到达目标点, 否则转向 Step1。

## 5 仿真实验结果及分析

本文使用 Visual C++ 上进行仿真实验, 采用模拟声纳传感器测试, 故滚动窗口设计成圆形区域, 如图 2, 半径由声纳传感器的准确度决定, 半径的选择与环境的障碍物位置的精确度相关联, 合理的选择半径可以提高智能体避障的效率。

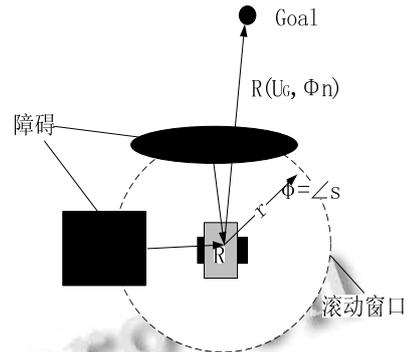


图 2 以智能体为中心的滚动窗口

仿真过程中, 参数设置如下: 折扣因子  $\gamma = 0.95$ , 障碍物影响距离  $\rho_0 = 2\text{cm}$ , 引力系数  $k_p = 1$ , 斥力系数  $\eta = 100$ , 滑动窗口半径  $r = 2\text{cm}$ ,  $\text{Max}_{\text{occupied}} = 98\%$ 。

#### (1) 简单静态环境下路径规划仿真

图 3 是智能体的运动轨迹, 可以看出智能体能够在与障碍物保持一定安全距离的条件下避开障碍物, 并快速地到达目标点。图 4 反应了运动过程中距离势能的变化趋势。距离势能的下降幅度说明了智能体能够快速到达目标点。

#### (2) U 形障碍物环境下路径规划仿真

人工势场法容易进入陷阱, 使得智能体在遇到 U 形障碍物时陷入局部最优, 无法到达目标点。此处智能体认为陷入势能陷阱时, 通过添加随机噪声, 使智能体摆脱陷阱。

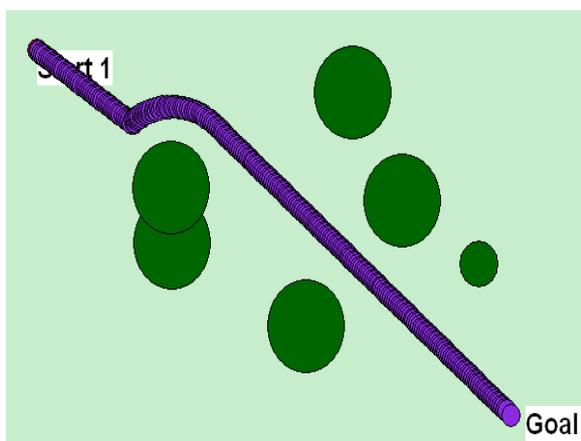


图 3 静态环境下的路径规划

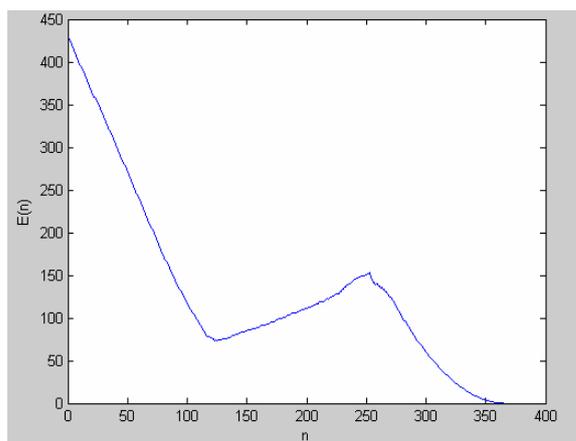


图 5 距离势能的变化曲线

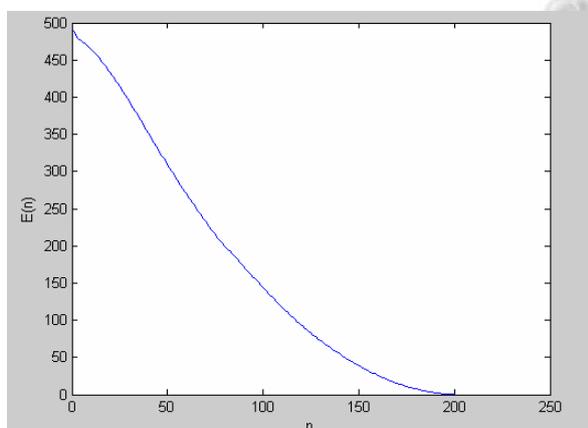


图 4 距离势能的变化曲线

图 5 是智能体在 U 形障碍物的运动轨迹，可以看出智能体能够快速地摆脱陷阱，到达目标点。图 6 是智能体在 U 形障碍物运动的距离势能，智能体通过增加距离势能摆脱陷阱，在摆脱陷阱后，智能体快速地向目标点移动。

(3) 动态不确定环境下路径规划仿真

由图 7 至图 10 可知，通过动态地改变障碍物来模拟动态不确定环境，仿真实验结果表明，智能体能够适应动态不确定环境，在保证避开障碍物的前提下获得较优的规划路径，并快速地到达目标点。

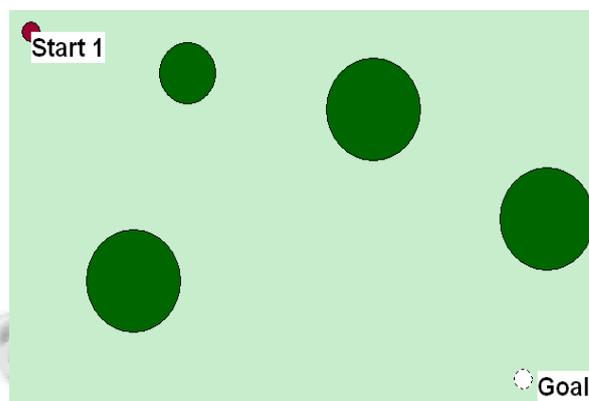


图 7 初始化环境

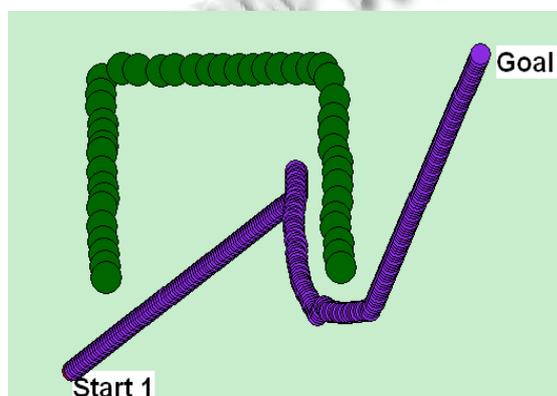


图 5 摆脱陷阱规划效果

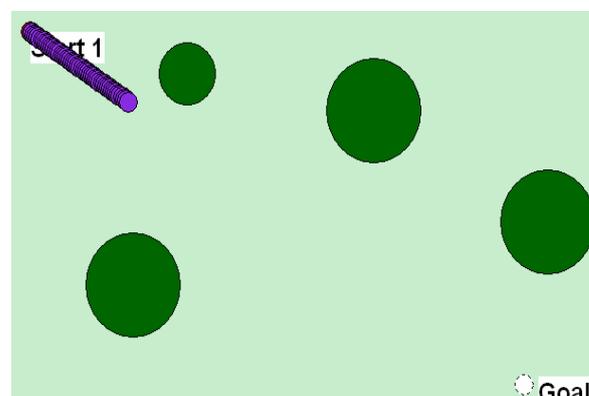


图 8 智能体启动路径规划

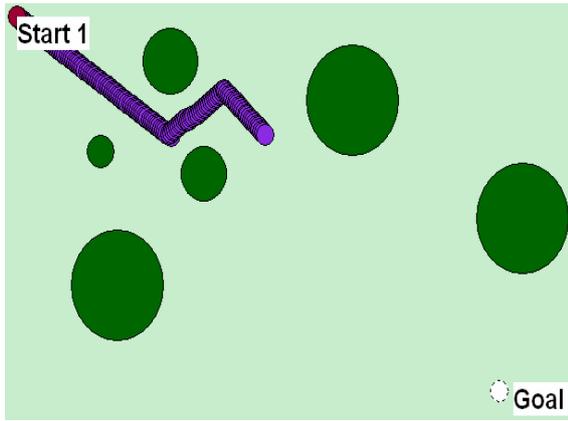


图 9 动态改变环境

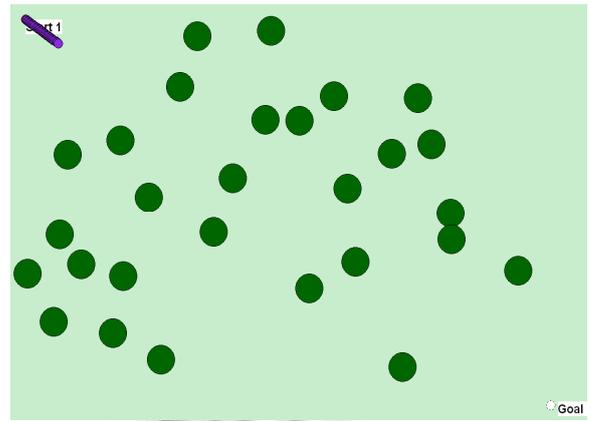


图 12 智能体开始路径规划

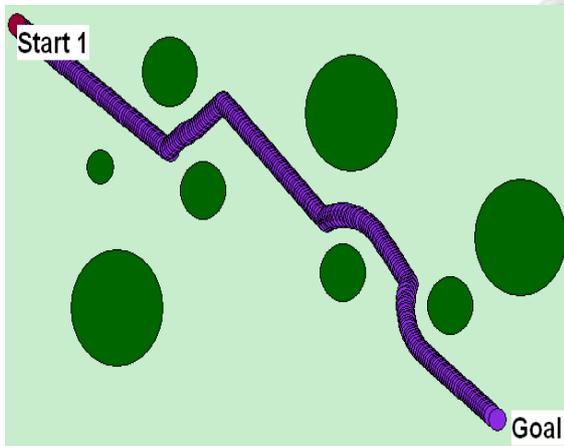


图 10 智能体到达目标点

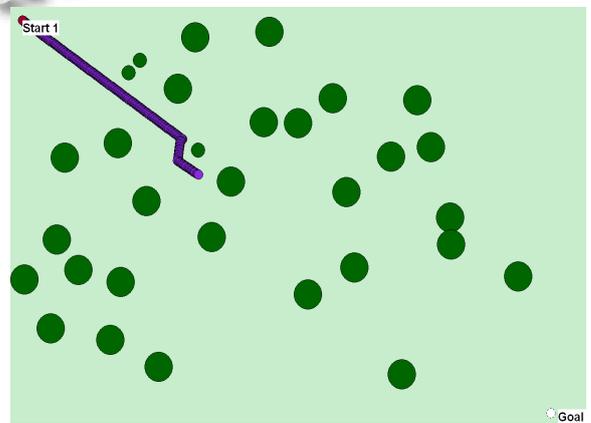


图 13 动态添加障碍物

(4) 复杂动态不确定环境下路径规划仿真

在复杂的动态不确定环境下，如图 11 至图 14，智能体采用改进的 POMDP 算法进行路径规划能够适应动态不确定环境，避开障碍物，并到达目标点。

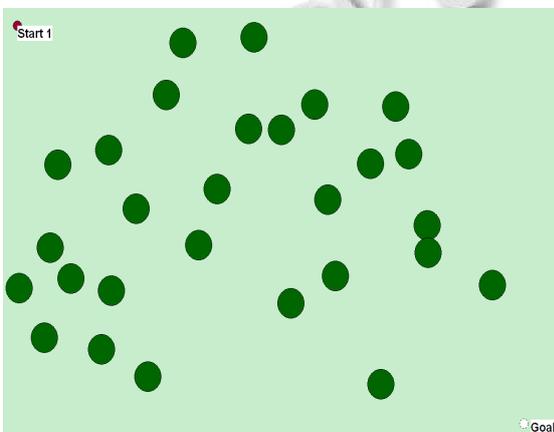


图 11 初始化复杂环境

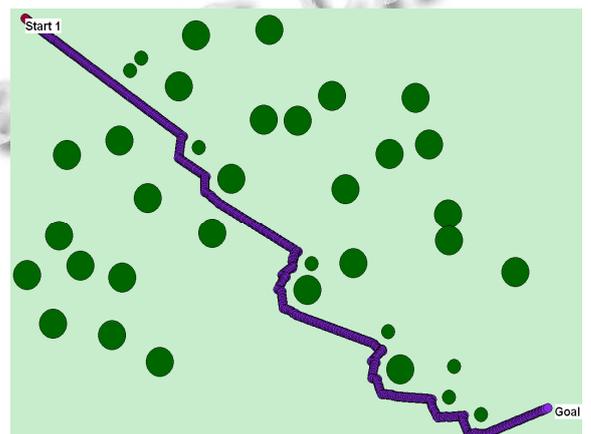


图 14 到达目标点

(5) 对比研究

从理论上可以得知，求解一个简单的有限范围的 POMDP 问题会是一个 PSPACE 完全问题<sup>[20]</sup>，它是一个比 NP 问题更复杂的问题。POMDP 算法在设定奖赏值时

经过了复杂的计算,本文通过人工势场法来降低设定奖赏值的复杂度,并通过500次重复实验来验证改进算法的有效性。传统的POMDP算法和人工势场法均能够对环境的即时变化作出响应,传统的POMDP算法的路径规划与本文算法的规划方法主要区别在于对奖赏值的设定,传统的奖赏值设定是智能体执行了整个动作序列后,得到的累积奖赏值,而不是单一动作的奖赏。

如表1,在500次重复实验中,采用POMDP算法对动态不确定环境具有更好的适应能力,实验表明PODMP算法在避障准确率上相对于人工势场法占有优势,本文算法在与传统POMDP算法障碍物相当的环境下能够有效地缩短智能体的运动轨迹长度,更快地到达目标点,对动态不确定环境有更强的适应能力,从而验证了本文算法的有效性。

表1 对比实验

算法	仿真数据			
	平均障碍物数量	平均路径长度(cm)	平均到达目标点时间(s)	避障准确率
传统 POMDP 算法	75	68.6824	34	99%
APF 算法	75	70.4562	38	98%
本文算法	77	64.0458	32	99%

## 6 结束语

本文提出了一种基于POMDP算法与APF导向的机器人路径规划方法,能够适应环境的动态不确定性,快速地在起点和终点之间找到一条较优的安全路径。POMDP和APF都能够实现在线规划,而POMDP算法过于复杂,APF对动态不确定环境适应能力不强,本文一方面利用POMDP的信念分布描述主体对当前环境的认知,以适应环境的不确定性;另一方面利用APF计算量小和实时性强的优点,快速地通过影响POMDP的回报函数为智能体提供相应的导向,有效降低了POMDP设定奖赏值的复杂度,一定条件下提高了POMDP决策的效率,结合了两种算法的优点,从而智能体在动态不确定环境下获得更有效的适应能力。

## 参考文献

- 1 朱大奇,颜明重. 移动机器人路径规划技术综述. 控制与决策, 2010,25(7):961-1967.
- 2 戴博,肖晓明,蔡自兴. 移动机器人路径规划技术的研究现状与展望. 控制工程, 2005,12(3):198-1202.

- 3 Bennet DJ, McInnes CR. Distributed control of multi-robot systems using bifurcating potential fields. *Robotics and Autonomous Systems*, 2010,58(3):256-264.
- 4 邓万,郑庆,陈琳,等. 神经网络极速学习方法研究. 计算机学报, 2010,33(2):279-287.
- 5 邢焕来,潘炜,邹喜华. 一种解决组合优化问题的改进型量子遗传算法. 电子学报, 2007,35(10):1999-2007.
- 6 Madani O, Hanks S, Condon A. On the undecidability planning and related stochastic optimization problems. *Artificial Intelligence*, 2003,147(1-2):5-34.
- 7 Kaelbling LP, Littman ML, Cassandra AR. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 1998,101:99-134.
- 8 李响,陈小平. 一种动态不确定性环境中的持续规划系统. 计算机学报, 2005,28(7):1163-1170.
- 9 刘满禄,张华,胡天链. 改进的人工势场法用于移动机器人导航. 华中科技大学学报, 2008,36(S1):177-180.
- 10 Sheng JW, Gao QH, Wei BG, Li JH. An Improved Artificial Potential Field Algorithm for Virtual Human Path Planning. *Edutainment*, 2010:592-601.
- 11 Sabattini L, Secchi C, Fantuzzi C. Arbitrarily shaped formations of mobile robots: artificial potential fields and coordinate transformation. *Auton. Robots*, 2011, 30(4): 385-397.
- 12 Majercik SM, Littman ML. Contingent planning under uncertainty via stochastic satisfiability. *Artificial Intelligence*, 2003,147(1-2):119-162.
- 13 Strehl AL, Littman ML. An empirical evaluation of interval estimation for Markov decision processes. 16th IEEE International on Tools with Artificial Intelligence Conference. Seattle: The MIT Press, 2004.531-539.
- 14 Kaelbling LP, Littman ML, Cassandra AR. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 1998,101(1):99-134.
- 15 李磊. 移动机器人系统设计与视觉导航控制研究. 北京:中国科学院研究生院, 2003.
- 16 Na YK, Se YO. Hybrid control for autonomous mobile robot navigation using neural network based behavior modules and environment classification. *Autonomous Robots*, 2003,15(2): 193-206.

(下转第45页)

一个零件,不同供应商可能给出不同的名称;同时,同一个名称可能对应不同的零部件。采用上述“搜索—排序”方式来管理本体库,将使用次数多的作为该零件的本体,将其余的名称或零件作为该本体的具体实例,继承该本体的所有属性。

#### 4 实例

本文在制造云的基础上开发了名为 CMoPL (Cloud Manufacturing oriented Parts Library, 面向云制造的零件库)的零件库系统。系统运行界面如图 7 所示,左侧为零件分类结构树,右上部分为标签列表,右下方为零件列表。用户通过浏览器即可在线查看零件三维模型。

#### 5 结论

鉴于中小企业难以承担建立零件库的费用,本文提出采用云制造模式建立零件库系统。本文对基于云制造模式的零件库运营模式进行了阐述;针对多用户的零件库系统,提出多层架构、Force 数据模型、数据隔离等方式建立系统框架;采用 X3D 进行零部件在线显示及动态交互,通过分类、标签、本体来进行零部

件的组织与管理。

中小企业通过租用来使用本系统,该零件库可以有效帮助中小企业宣传产品,对于提高企业信息化水平有很大帮助。

#### 参考文献

- 徐向荣.基于 Web 服务的分布式零件库系统开发与应用研究.杭州:浙江大学,2006.
- 苏宝华,祁国宁,顾新建,等.零件库的基本原理及其在产品中的应用.工程设计,1998,(4).
- 张霖,罗永亮,陶菲,等.制造云构建关键技术研究.计算机集成制造系统,2010,16(11):2511-2520.
- 李伯虎,张霖,王时龙,等.云制造——面向服务的网络化制造新模式.计算机集成制造系统,2010,16(1):1-7,16.
- 韩燕波,王桂玲,刘晨,等.互联网计算的原理与实践.北京:科学出版社,2010.
- web3d 官方网站.http://www.web3d.org.
- 朱立达,梁伟立,董圣广,等.基于 Web 的数控机床动态仿真系统的研究.计算机集成制造系统,2009,15(5):954-958.
- 黄建年,侯汉青.Tag 分类基本问题探究.情报理论与实践,2008,31 (3):461-465.
- Zhang CG, Xi YG. A real time path planning method for mobile robot avoiding oscillation and dead circulation. Acta Automatica Sinica, 2003,29(2):197-205.
- 王萌,王晓荣,李春贵,等.改进人工势场法的移动机器人路径规划研究.计算机工程与设计,2008,29(6):1504-1506.
- Geand SS. New potential functions for mobile robot path planning. IEEE Trans. on Robotics and Automation, 2000, 16(4):615-620.
- Burago D, de Rougemont M, Slissenko A. On the complexity of partially observed Markov decision processes. Theoretical Computer Science, 1996, 157(2): 161-183.
- 蔡自兴,贺汉根,陈虹.未知环境中移动机器人导航控制理论与方法.北京:科学出版社,2009.
- 仵博,吴敏.一种基于信念状态压缩的实时 POMDP 算法.控制与决策,2007,22(12)1417-1420.
- 王键,张汝波.基于 POMDP 模型的机器人导航控制方法.华中科技大学学报,2008,36(1):12-15.
- 刘海涛,洪炳熔,朴松昊,王雪梅.不确定性环境下基于进化算法的强化学习.电子学报,2006,34(7)1356-1360.
- 刘克.实用马尔可夫决策过程.北京:清华大学出版社,2004.20-52.

(上接第 98 页)