

# Research on Obstacle Avoidance Planning for ALV Based on XCS in Narrow Environments

Jie Shao<sup>1,2</sup>, Qingzhen Wang<sup>1</sup>

<sup>1</sup>Department of Computer Science, Zhengzhou College of Science & Technology, Zhengzhou Henan

<sup>2</sup>School of Computer Science, Nanjing University of Science and Technology, Nanjing Jiangsu

Email: sjsyxx@163.com

Received: Jan. 27<sup>th</sup>, 2017; accepted: Feb. 11<sup>th</sup>, 2017; published: Feb. 16<sup>th</sup>, 2017

---

## Abstract

Local obstacle avoidance in dynamic narrow environments, as a principal ability for ALV, plays an important role in autonomous navigation. Due to premature convergence, local optimal solution, accounting for a larger storage space and other shortcomings of genetic algorithms still exist. In order to improve the ability of obstacle avoidance for ALV, this paper presents a path obstacle avoidance planning method based on LCS for ALV in narrow environments, designs and improves special Genetic Operators. Different environments of simulation results showed that the combination of LS-SVM and learning classifier for ALV path obstacle avoidance planning was convergent, increasing ALV's ability to quickly find safe paths in narrow environments.

## Keywords

Obstacle Avoidance Planning, LS-SVM, Autonomous Land Vehicle (ALV), Accuracy-Based Learning Classifier System (XCS), Genetic Algorithm

---

# 基于XCS和LS-SVM的ALV在狭隘环境中的避碰规划

邵杰<sup>1,2</sup>, 王清珍<sup>1</sup>

<sup>1</sup>郑州科技学院信息工程学院, 河南 郑州

<sup>2</sup>南京理工大学计算机学院, 江苏 南京

Email: sjsyxx@163.com

收稿日期: 2017年1月27日; 录用日期: 2017年2月11日; 发布日期: 2017年2月16日

## 摘要

动态环境下的局部避碰是自主地面车的一项基本功能, 在自主导航中占有重要作用。由于遗传算法具有早熟收敛、局部最优解和占据较大的存储空间等缺陷, 为了提高自主地面车的避碰能力, 本文提出了一种基于学习分类器的自主地面车在狭隘环境中路径避碰规划方法, 设计改进了特殊的遗传算子。不同环境的仿真实验结果表明LS-SVM和学习分类器结合用于自主地面车的路径规划是收敛的, 提高了ALV在狭隘环境中快速发现安全路径的能力。

## 关键词

避碰规划, 最小二乘支持向量机, 自主地面车, 学习分类器, 遗传算法

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

狭隘环境的路径避碰规划是地面自主车(Autonomous Land Vehicle, ALV)自主导航的关键技术之一, 是指在各种复杂的地面环境中, 无需人工干预就可以自主完成各种工作[1] [2]。自主地面车 ALV 最初由马丁·玛丽埃塔公司研制, 该车的主要部件有视觉系统、导航系统和装载平台。由于 ALV 工作环境复杂, 本文仅就导航方面[3] [4] [5] [6], 对自主地面车在狭隘环境下的避碰能力进行了研究, 主要解决 ALV 在狭隘环境下的早熟收敛、局部最优解、占据存储空间较大、收敛数度慢等问题。

学习分类器[7] [8]是基于信用分配的强化学习机制和基于遗传算法(GA)的规则发现机制。本文提出融合 LCS 的衍生系统 XCS 和 LS-SVM 算法, 通过对环境的学习和反馈来解决 ALV 的路径避碰规划问题, 设计了在狭隘环境下的集成适应度函数, 改进设计了新的遗传算子。仿真实验结果也表明 XCS 和 LS-SVM 结合用于 ALV 的路径避碰是高效的。

## 2. 相关技术原理

### 2.1. 学习分类器(Accuracy-Based Learning Classifier System, XCS)

XCS [9]作为 LCS 的衍生系统是由 Wilson 提出的基于精度的学习分类器, 集成了强化学习、遗传算法和覆盖算法。XCS 是由条件、动作和预测回报三个主要的参数组成: 条件指明 XCS 可以匹配的环境输入状态; 动作表示在该 XCS 匹配的环境状态下 XCS 的建议动作; 预测回报  $P$  表示该 XCS 被匹配且执行了该 XCS 所建议的动作后系统预期从环境获得的回报。预测误差  $\varepsilon$  表示预测回报  $P$  的误差; 适应强度  $P$  表示预测回报  $P$  的精度。

#### 2.1.1. 执行机制

初始分类器集合是随机生成的, 每个分类器表示为状态、动作对。根据环境输入, 产生一个匹配环境输入的匹配规则集, 然后依据每个分类器的回报预测和适应强度值, 生成整个系统的预测列表和动作集, 最后依据分类器的动作概率选择一个作用于环境的动作。

$$P(a_i) = \frac{\sum_{cl_k \in |M| a_i P_k \times F_k} cl_k}{\sum_{cl_k \in |M| a_i F_k} cl_k} \quad (1)$$

### 2.1.2. 强化机制

分类器从环境获得一个回报  $R$ ，回报预测、分类器适应值强度更新如下：

$$P \leftarrow R + \gamma \max_a P(a) \quad (2)$$

其中  $0 \leq \gamma \leq 1$  是学习效率。

$$P_j \leftarrow P_j + \beta(P - P_j) \quad (3)$$

$$\varepsilon_j \leftarrow \varepsilon_j + \beta(|P - \varepsilon_j| - \varepsilon_j) \quad (4)$$

$$k_j = \begin{cases} 1 & \text{if } \varepsilon_j \leq \varepsilon_0 \\ \alpha(\varepsilon_j / \varepsilon_0)^{-\nu} & \text{otherwise} \end{cases} \quad (5)$$

$$k'_j = \frac{(k_j \times \text{num}_j)}{\sum_{cl_k \in |A|_{-1}} (k_k \times \text{num}_k)} \quad (6)$$

其中参数  $\varepsilon_0$  ( $\varepsilon_0 > 0$ ) 控制预测误差  $\varepsilon$  的冗余度；参数  $\alpha$  ( $0 < \alpha < 1$ ) 和  $\nu$  ( $\nu > 0$ ) 是常量，控制当  $\varepsilon_0$  被超过时精确度  $k$  的下降速度。动作集中的绝对精确度值  $k$  被转换为相对精确度值  $k'$ ，XCS 的适应度值  $F$  是依据相对精确度值进行更新：

$$F_j \leftarrow F_j + \beta(k'_j - F_j) \quad (7)$$

### 2.1.3. 规则发现机制

规则发现机制是从现有的较优规则中派生出新的规则。规则集中的所有规则作为遗传算法的个体，每一避碰规则的强度值作为遗传算法的适应度函数值。采用遗传算子如交叉、变异等进行遗传操作，通过进化得到新的避碰规则。

当匹配规则集中没有分类器与环境输入信息相匹配时，启动规则发现机制的覆盖算法，随机产生一个分类器。覆盖算法产生的新规则的条件部分与环境信息匹配，该规则淘汰规则集中强度最小的规则，存入规则集中。

## 2.2. 最小二乘支持向量机的学习(Least Squares Support Vector Machine, LS-SVM)

典型强化学习系统的映射关系包括：状态空间到动作空间的映射( $S \rightarrow A$ )、状态空间到回报函数的映射( $S \rightarrow r$ )、〈状态, 动作〉对到值函数的映射( $S \times A \rightarrow Q$ )以及〈状态, 动作〉对到转移函数的映射( $S \times A \rightarrow P$ )。王雪松等[10]提出了协同最小二乘支持向量机的  $Q$  学习算法。该学习系统由一个最小二乘支持向量回归机(Least squares support vector regression machine, LS-SVRM)和一个最小二乘支持向量分类机(Least squares support vector classification machine, LS-SVCM), 分别采用 LS-SVRM 和 LS-SVCM 实现对  $S \times A \rightarrow Q$  和  $S \rightarrow A$  映射的学习估计。

协同最小二乘支持向量机的  $Q$  学习过程如下[10]:

**Step 1:** 在初始学习阶段当前状态  $s_t$  下，系统直接选择由 LS-SVCM 给出的建议动作值  $a_t$ ，该动作作用于环境后获得回报信号  $r_t$ ，更新该时间步的  $Q$  值  $Q_t$ 。

**Step 2:**  $Q_t$  与系统的  $(s_t, a_t)$  组合后形成 LS-SVRM 训练样本集合 D1，对 LS-SVRM 模型进行在线训练； $(s_t, a_t)$  则送入 LS-SVCM 的训练样本集 D2，对其进行在线训练。

Step 3: 在 $((s_t, a_t), Q_t)$ 和 $(s_t, a_t)$ 加入样本集 D1 和 D2 时, 需要进行样本有效性判断, 当实际动作执行前后系统的变化满足下列条件(8)时,  $((s_t, a_t), Q_t)$ 加入 D1, 满足条件(9)时,  $(s_t, a_t)$ 加入集合 D2。

$$\|s_{t+1} - s_t\| > \theta \quad (8)$$

$$\|s_{t+1} - s_d\| < \|s_t - s_d\| \quad (9)$$

其中,  $\theta$  为状态阈值,  $s_d$  表示目标状态。

Step 4: 当由 LS-SVRM 给出的动作与 LS-SVCM 给出的动作相同时, 学习进入深层次学习阶段, 这时 LS-SVCM 失去作用, LS-SVRM 对 Q 值进行在线估计, 学习系统根据得到的数据 $((s_t, a_t), Q_t)$ , 依据动作  $a_t$  被选择的概率  $p$ , 及时对 LS-SVRM 模型作出调整, 不断重复上述过程 Step 1、2、3, 直至可以准确地学习到最优避碰策略。

$$p(\tilde{a} = a_t | s_t) = \frac{\exp\left(\frac{Q(s_t, a_t)}{T_t}\right)}{\sum_a \exp\left(\frac{Q(s_t, a)}{T_t}\right)} \quad (10)$$

其中:  $T_t > 0$  为温度系数, 控制动作选择的随机程度, 一般在学习初期选择较高的温度, 在学习的过程中逐渐降低温度, 以保证较好的学习效果。

$$\begin{cases} T_0 = T_{\max} \\ T_{t+1} = T_{\min} + \beta(T_t - T_{\min}) \end{cases} \quad (11)$$

其中:  $0 \leq \beta \leq 1$  为退火因子,  $T_{\max}$  和  $T_{\min}$  为设定的最大和最小温度。

### 3. 基于 XCS 的 ALV 路径避碰规划设计

#### 3.1. ALV 避碰规划的集成适应度函数设计

狭隘环境下的避碰规划是在确保 ALV 安全的情况下, ALV 能够避开静态或动态运行的障碍物。因此我们把 ALV 是否在安全范围内和 ALV 本身的适应度值两方面来确定 ALV 的集成适应度函数。

1) 是否在安全之内

我们把所有的动态障碍物或 ALV 都视为质点, 每个障碍物或 ALV 都有一个安全半径, 若机器人与障碍物或机器人间的距离大于安全半径, 则认为是安全的; 若小于安全半径, 则认为是不安全。障碍物与机器人每一个路点 $(x_i, y_i)$ 之间的距离  $d$  与安全半径  $R$  之间的关系如下:

$$fit1 = \begin{cases} 1 & d > R \\ 0 & d \leq R \end{cases} \quad (12)$$

其中  $d = \sqrt{(x_0 - x_r)^2 + (y_0 - y_r)^2}$ , 障碍物坐标 $(x_0, y_0)$ , ALV 坐标为 $(x_r, y_r)$ 。fit1 的值表明只要机器人质点与障碍物的距离均在安全半径外, 则其适应度为 1, 若有一个路径点与障碍物的距离在安全半径之内, 则适应度为 0。

2) 基于 XCS 的适应度函数:

$$fit2 = F_j \quad (13)$$

结合(12) (13), ALV 的集成适应度函数为:

$$f = fit1 * fit2 \quad (14)$$

### 3.2. 遗传算子的设计与改进

1) 选择算子。首先选择规则集中权值强度大的规则向下一代繁殖。然后选择那些因提供错误动作而受到惩罚的规则进行遗传操作。新规则的条件部分为机器人声纳当前的输入, 遗传算子仅对规则的动作部分操作。新规则准确体现环境的实时信息和依此而产生新的避碰规则。

2) 交叉算子(覆盖算子)。当匹配规则集中没有分类器与环境输入信息相匹配时, 启动规则发现机制的覆盖算法, 随机产生一个分类器。覆盖算法产生的新规则的条件部分与环境信息匹配。

3) 变异算子(合并算子)。两个规则动作编码相同, 均属于权值较高的规则, 且条件编码只有一位不同, 合并形成一个新的规则个体, 新规则条件编码是原两规则的条件编码, 将不同位改为#, 其权值为两规则权值的较大值。定期对规则集中的规则进行合并算子运算, 滤除相似的个体, 减少基因的单一性, 加快了学习收敛速度。

### 3.3. 基于 XCS 和 LS-SVM 的 ALV 避碰规划策略

LS-SVM 获得的最优学习策略作为 XCS 的初始规则集。XCS 通过与环境的交互, 可以发现一组用于指导 ALV 避碰的学习规则, 为 ALV 选择动作提供实时、动态的反馈, 使 ALV 自主地学习到最优避碰规划策略。

Step 1: 最小二乘支持向量机产生的最优学习策略作为 XCS 的初始规则集, 并随机执行其中的学习策略。

Step 2: 执行后将当前避碰规则的条件部分与消息列表中的消息进行匹配, 匹配的避碰规则送入匹配规则集。

Step 3: 从匹配规则集中选择合适的避碰编码规则送往效应器, 引导 ALV 产生相应的避碰动作, 并依据反馈值来判断避碰效果。

Step 4: 重复 Step2-Step3 过程, 不断从匹配规则集中选择合适的避碰动作编码规则送往效应器, 引导 ALV 产生相应的动作, 直到匹配规则集为空或定时启动避碰规则发现机制。

Step 5: 规则发现系统采用遗传算法和规则覆盖算法构造新的避碰规则。

Step 6: 对产生的新的避碰规则进行规则合并, 使其能概括以前的两个样本, 滤除相似的个体, 减少基因的单一性, 加快了学习收敛速度。产生的新学习个体规则送入规则集, 经过 XCS 发送到其公共规则集中, 供其它机器人 XCS 产生新避碰规则时共享。

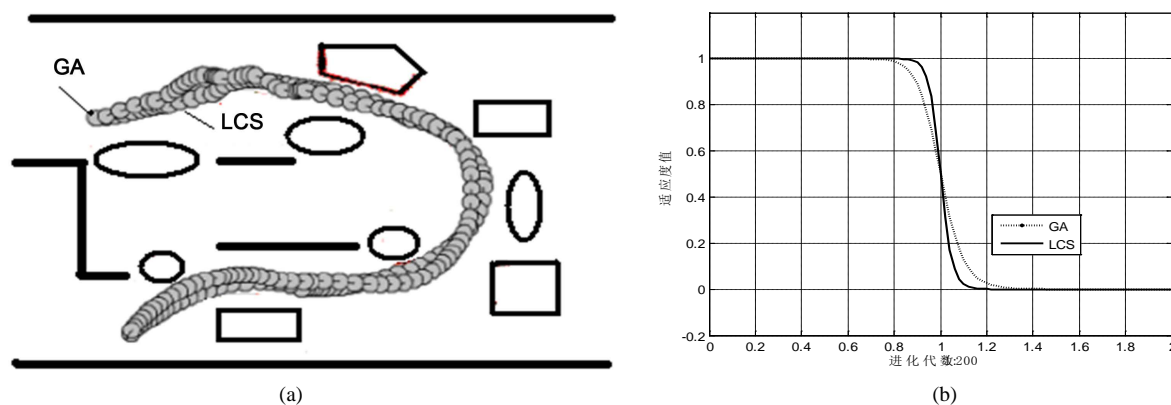
Step 7: 如果 ALV 未达到避碰效果, 则返回 Step 2。

## 4. 仿真实验和分析

实验环境是用 Matlab 开发的, 运行于 PC 机, 分别实验了在狭隘环境下的单机器人和多机器人进行静态和动态仿真, 同时进行了迭代分析, 最终给出了基于 LCS 和 XCS 的仿真比较结果, 验证了文中所提出的路径规划算法。

**实验场景一** 在狭隘环境下的单机器人仿真

假设机器人的活动区域为一狭隘环境, 诸如易燃、易爆和易碎等危险障碍物的分布情况等如图 1 所示。为了对基于遗传算法(GA) [5]和基于强度学习(LCS)的方法进行全方位的比较, 实验在不同的场景中。图 1(a)和图 1(b)分别为静态和动态场景,  $S$  和  $G$  分别为机器人初始和目标位置,  $W$  为易爆、易燃、易碎障碍物, 编号 1~3 为动态障碍物, 其它均为静态障碍物。在动态或静态狭隘环境中, 分别使用 GA 和 LCS 算法对机器人进行导航仿真实验。



**Figure 1.** (a) ALV trajectories in static narrow environment; (b) The average number of iteration steps in static narrow environment

**图 1.** (a) 静态场景下机器人运行轨迹; (b) 静态环境下平均迭代步数

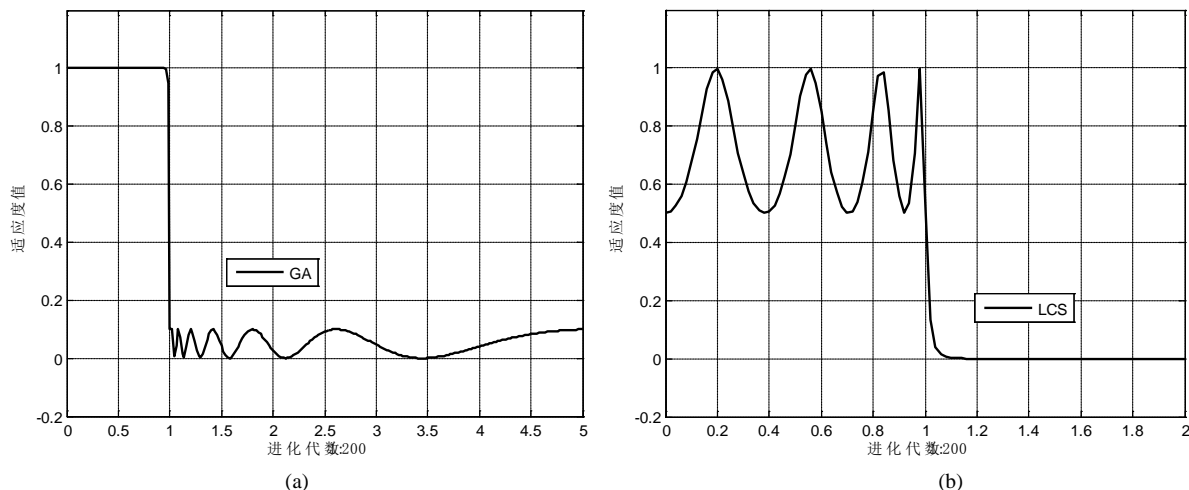
图 2(a)分别给出了 GA 和 LCS 方法在静态狭缝场景下的运行轨迹,由于这两种方法均以 GA 为基础,所以都能产生平滑的轨迹。比较这 2 种方法所用的时间,基于 GA 算法的机器人平均耗时 49.85 s,基于 LCS 算法的机器人平均耗时 43.65 s,由此可见,基于 LCS 算法的机器人导航效率要优于 GA 算法。图 2(b)为在静态狭缝 a 环境下,由于 LCS 方法以 GA 为基础,所以都能产生平滑的轨迹。基于 GA 和 LCS 算法的机器人在迭代过程中均取得较满意的收敛曲线。比较 2 种方法的迭代次数,GA 算法平均迭代次数 260, LCS 算法平均迭代次数 240,由此可见, LCS 算法的导航效率均优于 GA 算法。

图 2 分别为基于 GA 的算法和 LCS 的算法在动态环境下平均迭代步数图,由于 GA 算法的避障能力较弱,且与环境的交互能力较差,不可避免的存在一些局部特性,较难选择较优角度,容易在多个障碍物中迷失方向。最终导致机器人不能及时预测动态障碍物的位置,仿真曲线最后出现跳跃现象,机器人未能取得理想的规划收敛效果。LCS 集成算法在动态狭缝环境下,虽经过前期剧烈振荡,由于该算法具有较强的环境学习能力和行为调节控制能力,能够预测下几个周期机器人的位置和周围环境的关系,在几个可能角度之间进行优化选择,使得机器人能在局部范围找到一个较优运动方向,避免机器人陷入局部陷阱,机器人最终取得满意的规划收敛效果。具体实现过程:机器人在避过障碍物 03 之后根据目标函数选择与当前运动方向最近的区域前进,即障碍物 02 左侧的区域作为通道,由于该路径与 02 的运行路径交叉,不可避免地造成机器人运行轨迹不平滑,在成功避开障碍物 02 后,在门口附近的狭窄区域,观察障碍物 01 的运行轨迹,最终机器人成功避开障碍物 01,仿真曲线体现了机器人能够在快速性和安全性之间进行权衡,保证任务的顺利完成。

从上述任务完成平均时间可以知道,在静态环境下, LCS 继承了 GA 的快速性的优点,而在动态环境中,则体现了 LCS 安全性的优点。由于把碰撞预测模型与强化学习相结合, LCS 表现出良好的动态避障性能。在安全性方面, LCS 在 40 次试验中成功 39 次,仅有的一次失败发生在门口的狭窄区域,当机器人发现障碍物目标 02 时,障碍物 02 已经离机器人太近而无法避开,从而造成失败。

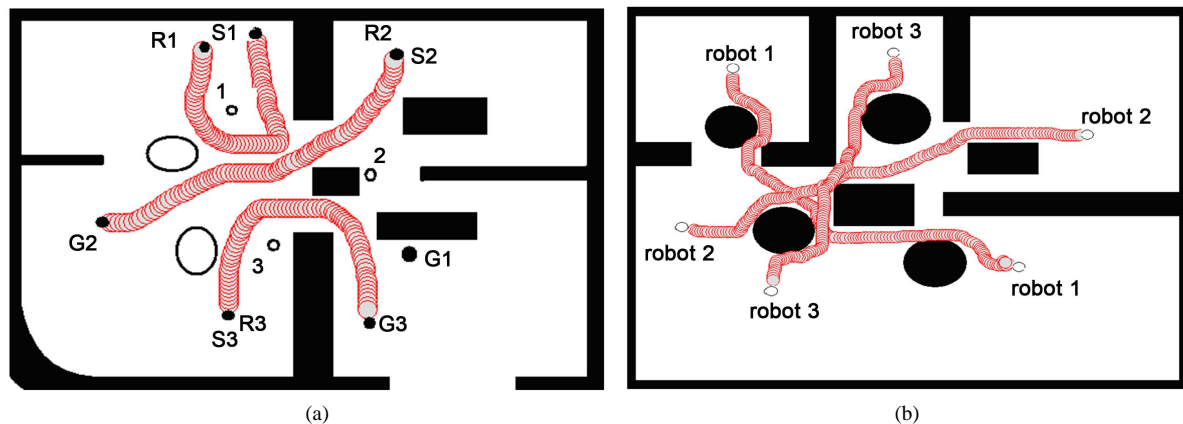
### 实验场景二 在狭缝环境下的多机器人仿真

在图 3(a)中, S1~S3 是三个机器人的初始位置, G1~G3 是目标位置,图标 1~3 的虚心圆为三个动态障碍物,其它为静态障碍物。由于 GA 算法无法预测排除有可能发生碰撞的区域,机器人 R1 在在避过障碍物 01 之后,根据目标函数选择与当前运动方向,由于选择的路径与机器人 R2 的路径交叉,为避免碰撞发生,机器人 R1 折路返回,没能到达其目标位置 G1。对于机器人 R2,由于 02 在运动过程中上侧区域的宽度增加,机器人 R2 和 R3 均顺利到达其目标位置 G2、G3,但由于机器人 R1 的无功而返,造成整个多机器人系统未能实现路径规划。



**Figure 2.** (a) The average number of iterations in unknown narrow dynamic environment; (b) The average number of iterations in unknown narrow dynamic environment

**图 2.** (a) 未知动态环境下平均迭代步数; (b) 未知动态环境下平均迭代步数



**Figure 3.** (a) Trajectory of the ALV in dynamic narrow environment based on GA; (b) Trajectory of the ALV in dynamic narrow environment based on LCS

**图 3.** (a) 动态狭隘环境下基于 GA 的多机器人运动轨迹; (b) 动态狭隘环境下基于 LCS 的多机器人运动轨迹

在图 4(b)中, LCS 方法能够利用强化学习机制, 预测排除有可能发生碰撞的区域, 从而在确保 robot1、robot2、robot3 三个机器人在动态狭隘环境下安全的情况下, 确保多机器人系统任务的顺利完成, 同时通过引入规则合并技术和冲突消解功能, 有效地提高了机器人的收敛速度, 较好地实现了多机器人系统的路径规划。从 robot1、robot2、robot3 三个机器人的运动轨迹来看, robot3 获得优化的路径长度最短, robot2 次之, robot1 路径长度最长。

图 4(a)为基于 GA 算法的多机器人系统在动态环境下的收敛曲线, 在开始的增强学习阶段, 由于学习空间较小, robot1、robot2、robot3 三个机器人的初始仿真曲线比较平稳。但随着学习环境空间的增大, 状态、动作对的大量增加, 造成计算复杂度剧增, 机器人不能及时得到精确的环境信息。最后结果是: robot1、robot2、robot3 三个机器人均没达到收敛, 且出现局部跳跃现象。

图 4(b)为基于 LCS 算法的多机器人在动态环境下的收敛曲线, 由于在学习的初始阶段, robot1、robot2、robot3 三个机器人获得的环境信息、较少, 不可避免出现随机行走现象, 在仿真曲线上体出现剧烈震荡。但随着迭代次数的增加, robot1、robot2、robot3 三个机器人通过 LCS 较强的强化学习能力和与外部环境

的融合能力, 不仅都取得了良好的收敛性, 且收敛时间与图 4(b)相比大为减少, 同时也没有在仿真曲线上出现图 5(b)类似的收敛局部最小或局部跳跃现象, 三个机器人均能以最优的路径到达各自的终点。从 robot1、robot2、robot3 三个机器人的迭代曲线来看, robot3 获得优化路径的迭代时间最短, robot2 次之, robot1 的迭代时间最长。

### 实验场景三 多机器人的 LCS 与 XCS 迭代曲线与收敛仿真结果

一般情况下, 一个路径规划算法的性能好坏需要从路径算法的收敛性和收敛速度两方面来判定[6]。XCS 作为 LCS 的衍生系统是目前较热的学习分类器, 由 Wnson 改进引入[7], 对基本的 LCS 系统结构进行了简化, 取消了系统内部消息列表, 采用强化学习中著名的 Q 学习算法, 其适应度值是建立在分类器回报预测的精度上, 用于生成高效、精简的问题解决方案。从图 5(a)和图 5(b)可以看出, 基于 XCS 的路径规划效果较好。

## 5. 结束语

在狭隘环境下的路径规划是机器人自主导航的一项重要基本功能。本文仅就导航方面, 对 ALV 在狭隘环境下的路径避碰规划进行了研究, 设计了集成适应度函数和改进遗传操作算子, 提高了 ALV 在

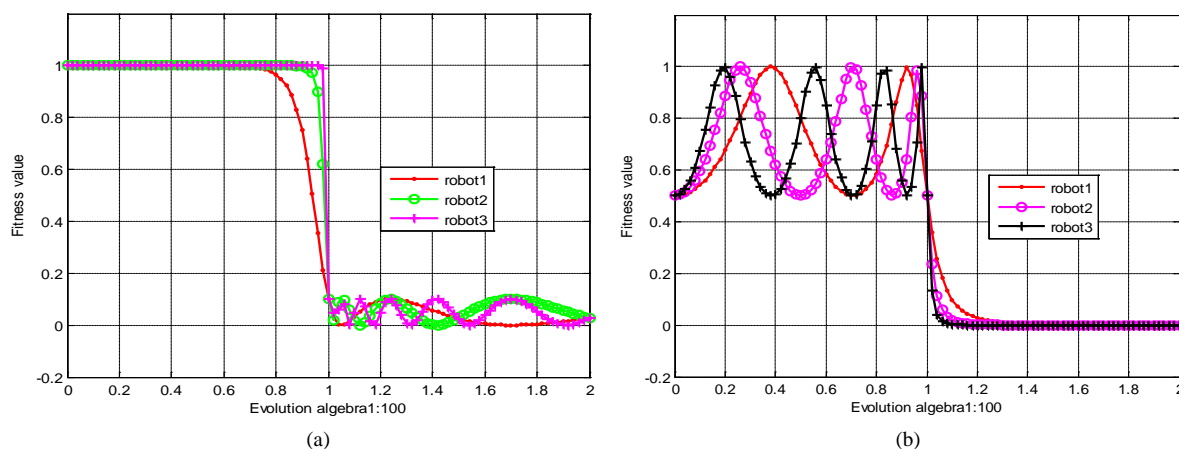


Figure 4. (a) Convergence curve of multi robot in dynamic environment based on GA; (b) Convergence curve of multi robot in dynamic environment based on LCS

图 4. (a) 基于 GA 算法的多机器人在动态环境下的收敛曲线; (b) 基于 LCS 算法的多机器人在动态环境下收敛曲线

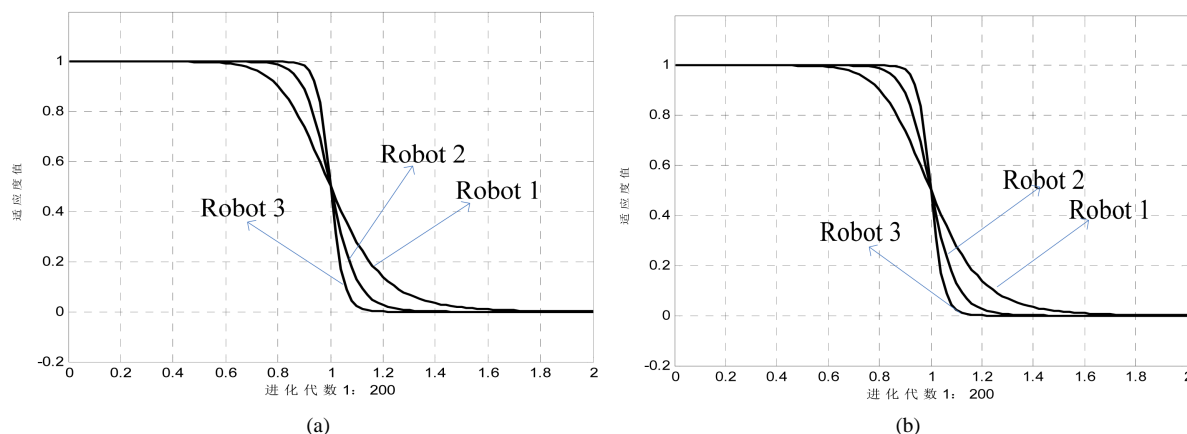


Figure 5. (a) Evolutionary algebra and fitness curve; (b) Evolutionary algebra and fitness curve

图 5. (a) 基于 LCS 的进化代数与适应度曲线图; (b) 基于 XCS 的进化代数与适应度曲线图



复杂环境下的路径避碰能力, 仿真实验也验证了本文所提算法的有效性。

目前, 虽然在分类器理论和应用领域开发出了各种算法, 但将 XCS 用于多机器人定位、路径规划和地图创建等关键技术方面的理论、算法及应用都非常少。随着如高速计算机计算结构、并行计算技术应用、图像识别和处理技术、人工智能和专家系统等技术的逐渐成熟, XCS 系统在机器人领域的应用开发前景是非常广阔的。尽管 ALV 自主能力还难以适应环境的复杂状况, 但随着各种机器人技术的日益成熟, 自主地面车辆在未来的地位将日益受到重视, 其作用会更大。

## 基金项目

国家自然科学基金资助项目(61462064)郑州市特聘高层次人才基金资助。

## 参考文献 (References)

- [1] Jie, S. and Yang, J.Y. (2010) Multi-Robot Reinforcement Learning Based on Learning Classifier System with Gradient Descent Methods. *Journal of Computational Information Systems*, **6**, 2449-2455.
- [2] 朱大奇, 颜明重. 移动机器人路径规划技术综述[J]. 控制与决策, 2010, 25(7): 961-967.
- [3] 邵杰, 杨静宇, 等. 基于学习分类器的多机器人路径规划收敛性研究[J]. 计算机研究与发展, 2010, 47(5): 948-955.
- [4] 焦殿科, 石川. 共享经验的多主体强化学习研究[J]. 计算机工程, 2008, 34(11): 219-221.
- [5] 尚艳玲, 肖文雅. 多 Agent 系统的 Q 值强化学习算法[J]. 河南师范大学学报(自然科学版), 2013, 41(2): 158-160.
- [6] 杨献峰, 付俊辉. 移动机器人路径规划的仿真研究[J]. 计算机仿真, 2012, 29(7): 223-226.
- [7] Musilek, P. (2005) Enhanced Learning Classifier System for Robot Navigation. 2005 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2-6 August 2005, 3390-3395. <https://doi.org/10.1109/iros.2005.1545150>
- [8] 邵杰, 杜丽娟, 等. XCSG 在多机器人强化学习中的应用[J]. 计算机科学, 2013, 8(2): 249-252.
- [9] Bernad'o-Mansilla, E. and Garrell, J. (2003) Accuracy-Based Learning Classifier Systems: Models, Analysis and Applications to Classification Tasks. *Evolutionary Computation*, **11**, 209-238. <https://doi.org/10.1162/106365603322365289>
- [10] 王雪松, 田西兰, 程玉虎, 易建强. 基于协同最小二乘支持向量机的 Q 学习[J]. 自动化学报, 2009, 35(2): 214-219.

### 期刊投稿者将享受如下服务:

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [airr@hanspub.org](mailto:airr@hanspub.org)