

菠萝果园多田块全覆盖作业路径规划方法研究

刘天湖 陈嘉鹏 孙伟龙 梁兆正 宋英楠
(华南农业大学工程学院, 广州 510642)

摘要: 为了优化菠萝老苗清理农机在菠萝果园中的作业路线,减少非工作区域消耗,本研究设计了混合改进的遗传粒子群优化(Genetic particle swarm optimization, GPSO)算法和Q-Learning(QL)算法全覆盖作业路径规划方法。利用无人机航拍获取目标菠萝果园田块图并转化为轮廓坐标地图;对各田块从0°到355°每间隔5°遍历覆盖路径寻找最优行进角度;针对传统GPSO算法收敛速度慢、易陷入局部最优等问题,提出改进遗传粒子群优化(Improved hybrid genetic particle swarm optimization, IHGPSO)算法获得菠萝果园多田块作业最佳遍历顺序;针对传统QL算法探索效率低、收敛慢的问题,提出改进的QL算法(Order initialization Q-Learning, OI-QL)完成田块之间的路径连接。仿真结果表明,在子区点数为30求解最佳遍历顺序时,IHGPSO算法适应度小于GPSO算法。OI-QL算法相比QL算法路径长度缩短约28.1%,平均收敛迭代次数减少约37.3%,平均收敛时间减少约25.7%,表明本文方法能有效地实现菠萝果园多田块全覆盖作业路径规划。

关键词: 菠萝果园; 多田块; 全覆盖作业路径规划; 遗传粒子群优化算法; Q-Learning 算法

中图分类号: TP18; S24 **文献标识码:** A **文章编号:** 1000-1298(2026)08-0001-12 **OSID:** 

Multi-field Complete Coverage Operation Path Planning Method of Pineapple Orchard

LIU Tianhu CHEN Jiapeng SUN Weilong LIANG Zhaozheng SONG Yingnan
(College of Engineering, South China Agricultural University, Guangzhou 510642, China)

Abstract: In order to optimize the operation routes of agricultural machinery for cleaning old pineapple seedlings in pineapple orchards and reduce consumption in non-working areas, a path planning method integrating improved genetic particle swarm optimization (GPSO) and Q-Learning(QL) algorithms was proposed to optimize the complete coverage operation path of unmanned machines in multi-field pineapple orchards. Images of the target farmland were obtained by unmanned aerial vehicle and converted into a coordinate map. Each field was traversed at 5° intervals from 0° to 355° to find the optimal travel angle. To address the inherent problem such as slow convergence and getting stuck into a local optimum in traditional GPSO algorithms, an improved GPSO algorithm, namely IHGPSO, was proposed to improve the generation and selection method of initial populations, a multi-objective weighted fitness function was designed, and a dynamic adjustment probability was added to particle exchange mechanism, thereby obtaining the optimal traversal order for multi-field complete coverage operation of pineapple orchard. To address the inherent problem in original QL algorithm, such as low exploration efficiency and slow convergence, an improved QL algorithm, namely order initialization Q-Learning (OI-QL) was proposed, which improved initialization of Q-table, proposed learning rate, and designed a reward function, thereby connecting multi-field complete coverage operation path of pineapple orchard. Simulation results showed that when the sub-area point groups contained 30 points, the IHGPSO algorithm outperformed the GPSO algorithm in solving the problem of the optimal traversal order. The average path length planned by OI-QL algorithm was 28.1% shorter than that planned by QL algorithm, and the average number of convergence iterations and the average convergence time by using OI-QL algorithm was 37.3% and 25.7% less than those by using QL algorithm, respectively. These results indicated that the method can effectively complete the complete coverage operation path planning of pineapple orchard.

Key words: pineapple orchard; multi-field; complete coverage operation path planning; genetic particle swarm optimization; Q-Learning algorithm

0 引言

菠萝通常种植在伞形缓坡地带的果园,田块难以整齐划一^[1]。菠萝苗首次结果后,老苗结出的二代果食用口感明显变差。为保证菠萝果实品质,采摘完后需将老苗处理掉,例如喷洒除草剂、铲除等,再重新种植。在清理作业场景中,对菠萝老苗清理农机需实现全覆盖路径规划。由于已采摘的老苗均为废株,农机行进角度可灵活调整,无需担心伤苗问题。解决多田块全覆盖路径规划问题,优化农机清理菠萝老苗作业路线,减少非工作区域消耗,是机械化、智能化作业的关键所在。全覆盖路径规划是农机路径规划的重要组成部分,可用于引导农机执行喷药、播种、收割和清理等特种作业^[2-3]。全覆盖路径规划包括单田块全覆盖路径规划与多田块路径连接2部分。当前研究多聚焦于单田块覆盖路径,针对多田块路径连接研究较少^[4-5]。

多田块路径连接分为多田块作业遍历顺序规划和田块之间具体路径规划。规划多田块作业遍历顺序可转化为广义旅行商问题,经典算法面对旅行商问题时,在适应环境和效率方面都表现出一定的局限性^[6]。越来越多研究采用生物群体智能算法解决此问题,如遗传算法^[7-8]、蚁群算法^[9]、粒子群算法^[10]等。

当前多数全覆盖路径规划研究在确定田块间的作业遍历顺序后,往往仅以直线距离作为代价指标连接具体的路径。然而菠萝果园中存在田垄等特殊区域,农机无法直接跨越,同时农机在作业区域内的移动与普通道路上的移动相比其代价也存在显著差异,如果单纯以距离作为衡量标准会影响田块之间路径规划的实际效果。传统路径规划方法,如A*算法^[11]、RRT算法^[12]、人工势场法^[13]、动态窗口法^[14]等,计算量大、灵活性不足并且对先验知识依赖程度高。强化学习法是一种借助与环境的持续交互实现自我迭代优化方法,能够灵活适应复杂多变的环境,追求长期累积回报,具备出色的可扩展性与广泛的通用性^[15]。Q-Learning(QL)是强化学习的一种经典算法^[16]。传统QL算法在初始化过程将Q值设置为随机值,探索概率与学习率固定,在学习过程收敛缓慢且容易陷入局部最优解^[17]。为了解决上述问题,众多学者开展了广泛研究。李潇宇等^[18]提出一种自适应步长和变 ϵ -greedy融合的QL算法。ZHOU等^[19]采用RMSprop方法调整传统QL算法的探索率,以加快学习速度并提高路径规划效率。LOW等^[20]则给智能体添加了距离指标、并引入虚拟目标优化QL算法。

本研究通过对比采用不同行进角度生成的单田块覆盖路径质量,选取最佳行进角度,以确立各田块出入口,提出改进遗传粒子群优化算法求解多个菠萝田块间的作业遍历顺序,采用改进QL算法规划田块之间的连接路径,以期实现菠萝果园多田块全覆盖路径规划,完成菠萝果园采摘后清理老苗的任务。

1 单田块全覆盖路径规划

1.1 覆盖与转弯方式

目前主要的局部覆盖路径规划模式有往复式覆盖和螺旋式覆盖,如图1所示。由于菠萝果园地形和种植模式的特殊性,往复式覆盖比螺旋式覆盖在起步阶段更为便捷且转弯次数较少。另外,螺旋式行走方式会导致与下一田块的路径连接性较差。因此本研究选用往复式覆盖模式。

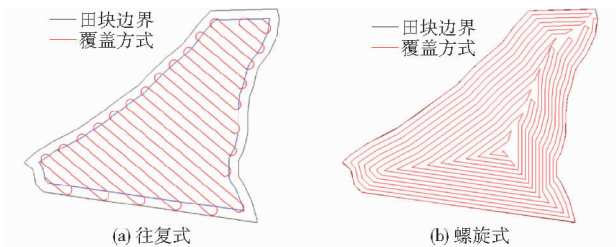


图1 局部覆盖路径规划

Fig. 1 Diagrams of local complete coverage path planning

不同作业幅宽 W 和最小转弯半径 R 有4种常见的农机地头转弯方式,半圆形、灯泡形、交叉形和弓形。对菠萝老苗清理农机通常具备较小的转弯半径以在狭窄的田间灵活转向。半圆形转弯时, $W=2R$,是直线行进之后的对称回转,路径相对简单,对农机的最小转弯半径要求较低,能有效避免因转向空间不足而影响作业流畅性,如图2所示。因此本研究农机地头转弯方式选择半圆形转弯。

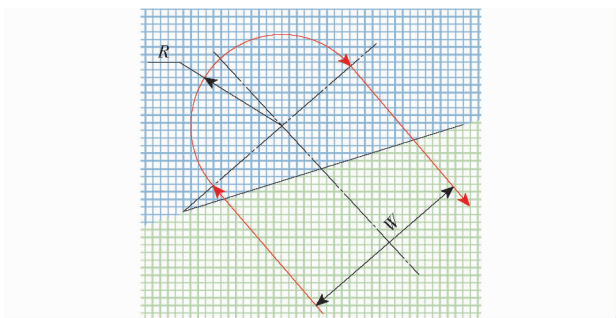


图2 半圆形

Fig. 2 Semi-circular turn

1.2 半圆形转弯路径作业消耗时间

菠萝老苗清理时,在农机全覆盖作业路径规划中,农机在田块内部作业区域的行驶路径称为直线路径,在转弯预留区域的行驶路径则称为转弯路径。

由于农机在转弯预留区域完成转弯,会消耗额外燃料与时间,因此在选取最优行进角度时,必须纳入转弯路径的作业消耗时间这一关键因素。图3为半圆形转弯示意图。

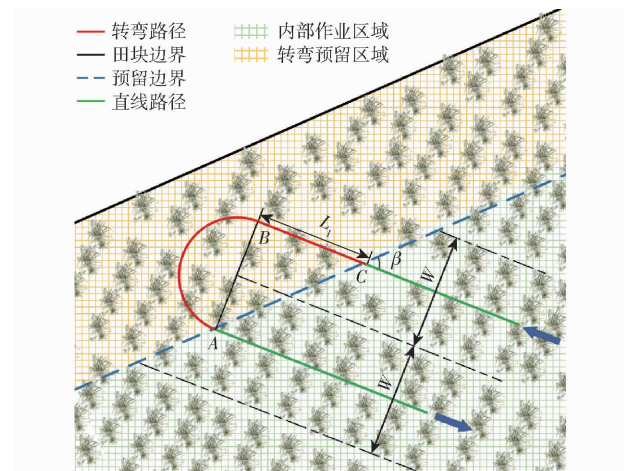


图3 半圆形转弯示意图

Fig. 3 Schematic of semi-circular turn

农机在田块转弯预留区域的单次转弯时作业消耗分为2部分。菠萝老苗清理农机在BC段消耗,为农机在转弯前行驶的直线路段,长度为 L_1 ;农机在红色半圆边路段的消耗,为农机在半圆形转弯路段,长度为 L_2 。计算式分别为

$$L_1 = \frac{W}{\tan\beta} \quad (1)$$

$$L_2 = \frac{1}{2} \pi W \quad (2)$$

其中

$$\tan\beta = \left| \frac{l_\alpha - l_k}{1 + l_\alpha l_k} \right| \quad (3)$$

式中 β ——作业方向与预留边界夹角, ($^\circ$)

l_α ——农机行进方向斜率

l_k ——预留边界所在直线斜率

设农机直线与转弯行驶速度恒定,分别为 v_1 和 v_2 ,行驶时间分别为 t_1 和 t_2 。农机在单个半圆形转弯中作业消耗时间为 T_a ,计算式为

$$T_a = t_1 + t_2 = \frac{L_1}{v_1} + \frac{L_2}{v_2} = \frac{W}{v_1 \tan\beta} + \frac{\pi W}{2v_2} \quad (4)$$

2 改进遗传粒子群算法求解遍历顺序

2.1 遗传粒子群优化算法

在实际清理采摘后的果苗作业中,需要对多个田块进行遍历。多田块最佳遍历顺序问题本质上是一种广义的旅行商问题。本研究采用改进的遗传粒子群优化算法(Improved hybrid genetic particle swarm optimization, IHGPSO)求解多田块最佳作业顺序,以便规划出更高效的全覆盖路径。

遗传粒子群优化算法(Genetic particle swarm

optimization, GPSO)是结合遗传算法(GA)和粒子群优化算法(PSO)的混合算法,粒子相当于路径,基因相当于路径中节点。通过结合GA算法选择、交叉和变异机制所具有的全局搜索能力和PSO算法局部快速收敛能力提升优化性能,但仍存在若干局限性。例如,GPSO算法的初始种群通常是随机生成,若分布不均可能导致收敛速度慢或陷入局部最优。此外GPSO算法适应度函数设计单一,仅基于目标函数,忽略多样性,可能引发早熟收敛。GPSO算法采用静态的GA与PSO协作机制,无法根据优化阶段自适应调整,间接影响了全局搜索与局部搜索之间的平衡。

2.2 改进遗传粒子群优化算法

本研究设计的IHGPSO算法相较传统GPSO算法改进了3个方面,改进初始种群生成与选择方法、设计新的适应度函数、采用动态概率调整粒子的交换机制。

(1)改进初始种群生成与选择

传统GPSO算法初始种群是随机生成,若分布不均匀可能缺乏多样性,且易陷入局部最优解,导致算法早熟收敛。同时随机生成无法利用先验知识,可能包含大量低质量解,增加无效搜索计算成本。

通过k-means聚类算法对可行解分簇,再结合最近邻贪婪法和2-opt局部优化策略初始化种群。给定 n 个基因集合 $G = \{g_1, g_2, \dots, g_n\}$,其中每个基因 g_i 在二维空间中坐标为 (x_i, y_i) 。算法有5个步骤:

①通过k-means聚类把 G 划分为 k 个簇, $C = \{C_1, C_2, \dots, C_k\}$,以最小化簇内平方误差 S_E 为优化目标, S_E 计算式为

$$S_E = \argmin \sum_{j=1}^k \sum \|g_i - \mu_j\|^2 \quad (5)$$

式中 μ_j ——簇 C_j 的质心

②对每个簇 C_j 采用最近邻贪婪法构造初始路径 π_j 。从随机选择的簇内起始基因 $\zeta_{j,1} \in C_j$ 开始,迭代选择当前未访问的最近邻基因 $\zeta_{j,t+1}$ 构造路径,计算式为

$$\zeta_{j,t+1} = \arg \min_{\zeta \in C_j, \{\pi_j\}} \|\zeta_{j,t} - \zeta\| \quad (6)$$

式中 ζ ——簇内基因

直至遍历整个簇。

③对 π_j 执行2-opt局部优化,即通过交换基因对 $(\zeta_{j,p}, \zeta_{j,q})$ 减少路径长度,其中交换条件满足

$$\|\zeta_{j,p} - \zeta_{j,q}\| + \|\zeta_{j,p+1} - \zeta_{j,q+1}\| < \|\zeta_{j,p} - \zeta_{j,p+1}\| + \|\zeta_{j,q} - \zeta_{j,q+1}\| \quad (7)$$

④通过贪心策略拼接各子路径,初始全局路径

Π 为随机选择的路径 π_a , 迭代将剩余子路径中与当前路径末端质心 μ_a 最近的簇 C_b 对应的路径 π_b 接入 Π 。 C_b 计算式为

$$C_b = \arg \min_{C_j \in C'} \|\mu_a - \mu_j\| \quad (8)$$

式中 C' ——未拼接簇的集合

μ_j ——候选簇 C_j 质心

⑤ 拼接最后生成总路径, 长度为 $L(\Pi)$, 计算式为

$$L(\Pi) = \sum_{i=1}^{n-1} \|g_{\Pi(i)} - g_{\Pi(i+1)}\| + \|g_{\Pi(n)} - g_{\Pi(1)}\| \quad (9)$$

最终选择生成的全局路径 Π 为初始种群个体。

(2) 新的适应度函数

适应度函数是算法与实际问题之间的桥梁, 合理适应度函数应能准确反映解的质量, 平衡多样性与收敛性。在旅行商问题的实际应用中, 传统 GPSO 算法以路径总长度为唯一的优化目标, 往往难以满足实际场景对路径平滑性、均匀性及局部跳跃等多方面的需求。本研究设计一种综合考虑路径总长度以及路径分布均匀性为指标的适应度函数, 从而提升路径的整体质量和实际可行性, 计算式为

$$F(P) = f_1(P) + f_2(P) \quad (10)$$

$$\text{其中 } f_1(P) = \frac{D(P)}{D_{\max}} \quad (11)$$

$$D(P) = \sum_{k=1}^{n-1} d(p_k, p_{k+1}) + d(p_n, p_1) \quad (12)$$

$$f_2(P) = \frac{\sigma^2(d(p_1, p_2), \dots, d(p_n, p_1))}{\sigma_{\max}^2} \quad (13)$$

式中 $F(P)$ ——适应度函数

P ——包含 n 个节点 $p_1, p_2, \dots, p_n$ 的路径

$f_1(P)$ ——路径总长度

$f_2(P)$ ——路径分布均匀性

$D(P)$ ——当前路径长度, m

$D_{\max}$ ——当前种群中最长路径长度, m

d ——两点之间距离, m

p ——路径节点

σ^2 ——节点段长距离方差, m^2

$\sigma_{\max}^2$ ——当前种群最大方差, m^2

(3) 动态概率调整粒子交换机制

传统 GPSO 算法的粒子群采用速度和位置更新, 计算式为

$$v_i^{t+1} = 0.5v_i^t + c_1r_1(p_{best_i} - x_i^t) + c_2r_2(g_{best} - x_i^t) \quad (14)$$

$$\text{其中 } x_i^{t+1} = x_i^t + v_i^{t+1} \quad (15)$$

式中 t ——当前迭代数

c_1, c_2 ——加速常数

r_1, r_2 ——随机数, 取 $0 \sim 1$

x_i^t —— t 时刻粒子位置

v_i^t —— t 时刻粒子速度

p_{best_i} ——粒子 i 历史最优位置

g_{best} ——粒子群历史最优位置

本研究要解决旅行商问题的解是访问排列, 直接使用传统的更新公式会导致非法解。IHGPSO 算法粒子不再更新速度和位置, 而是通过交换操作调整路径。计算每个粒子 X 适应度与个体最优 p_{best_i} 和全局最优 g_{best} 差异, 计算式为

$$diff_p = \frac{|F(X) - F(p_{best_i})|}{|F(X)| + \varepsilon} \quad (16)$$

$$diff_g = \frac{|F(X) - F(g_{best})|}{|F(X)| + \varepsilon} \quad (17)$$

式中 $diff_p$ ——当前粒子 X 适应度与个体最优差异

$diff_g$ ——当前粒子 X 适应度与全局最优差异

$F(X)$ ——当前粒子 X 适应度

$F(p_{best_i})$ ——个体最优适应度

$F(g_{best})$ ——全局最优适应度

ε ——极小值, 防止除零

根据个体和全局最优适应度差异计算相应交换概率 $prob_p$ 和 $prob_g$, 计算式为

$$prob_p = 0.5 + 0.5diff_p \quad (18)$$

$$prob_g = 0.5 + 0.5diff_g \quad (19)$$

对粒子 X 的每个基因 x_i 位置执行交换操作, 过程为: 以概率 $prob_p$ 交换 p_{best_i} 中与 x_i 对应位置基因, 以概率 $prob_g$ 交换 g_{best} 中与 x_i 对应位置基因。以动态概率交换粒子内部基因, 使适应度较差的粒子更倾向于向 p_{best_i} 和 g_{best} 学习, 而适应度较好的粒子则保持一定的探索能力。该操作让个体在进化时向局部最优和全局最优靠拢。

2.3 IHGPSO 算法复杂度

设子区数量为 n_s , 种群规模为 N , 最大迭代次数为 $I_{\max}$, k -means 聚类数 k 通常为 $n_s^{1/2}$, k -means 迭代次数为 i_k , 为常数。

IHGPSO 算法的时间复杂度由初始化和迭代 2 部分组成。初始化阶段包括 k -means 聚类复杂度 $O(n_s k i_k)$ 、最近邻贪婪法复杂度 $O(n_s^2/k)$ 、2-opt 优化复杂度 $O(n_s^2/k)$ 和路径拼接复杂度 $O(k^2)$ 。初始化复杂度 T_{init} 计算式为

$$T_{init} = O(n_s k i_k + n_s^2/k + n_s^2/k + k^2) \approx O(n_s^{1.5}) \quad (20)$$

迭代阶段每次包括适应度计算、遗传操作和粒子交换, 单次迭代复杂度为 $O(N n_s)$ 。IHGPSO 总时间复杂度 T_{IHGPSO} 计算式为

$$T_{IHGPSO} = O(n_s^{1.5}) + O(I_{\max} N n_s) \quad (21)$$

传统 GPSO 采用随机初始化 $O(N n_s)$, GPSO 总时间复杂度 T_{GPSO} 计算式为

$$T_{GPSO} = O(I_{\max} N n_s) \quad (22)$$

虽然 IHGPSO 初始化复杂度为 $O(n_s^{1.5})$ 高于 GPSO 的 $O(N n_s)$, 但该操作仅执行一次。种群规模 N 为 100, 最大迭代次数 $I_{\max}$ 为 200, 当 $n_s < 100$ 时, T_{init} 占总时间 T_{total} 比例为

$$\frac{T_{\text{init}}}{T_{\text{total}}} = \frac{O(n_s^{1.5})}{O(n_s^{1.5}) + O(I_{\max} N n_s)} < 10\% \quad (23)$$

与此同时高质量初始种群可减少收敛所需的实际迭代次数, 因此 IHGPSO 算法即使计算复杂度比 GPSO 高, 并不会降低模型计算的实时性和计算效率。图 4 为 IHGPSO 算法流程图。

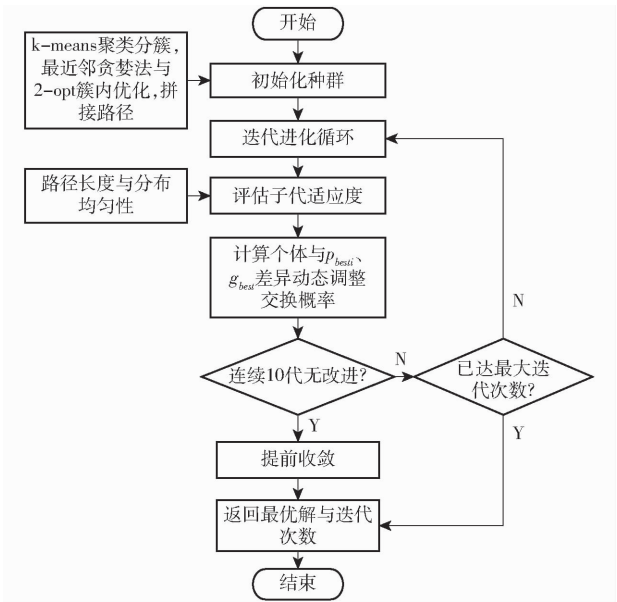


图 4 IHGPSO 算法流程图
Fig.4 Flowchart of IHGPSO

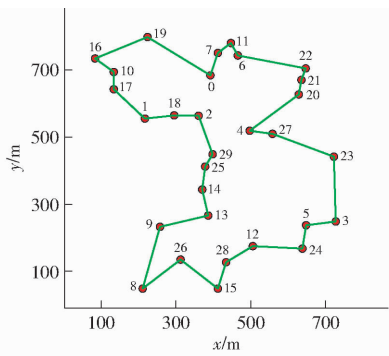
2.4 IHGPSO 算法解决广义旅行商问题

分别用 IHGPSO 和 GPSO 算法求解含有 30 坐标点的子区遍历顺序测试解决广义旅行商问题的能力。种群规模 N 设置为 100, 最大迭代次数 $I_{\max}$ 设置为 200。图 5 为 IHGPSO 和 GPSO 求得的含有 30 个坐标点的子区的最佳遍历顺序路径。图 6 为适应度迭代曲线。

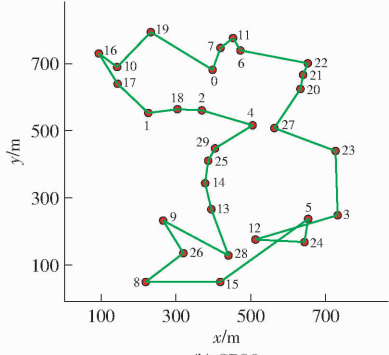
由图 6 可知, 在迭代前期, IHGPSO 算法收敛速度稍慢于 GPSO 算法, 但中后期收敛速度明显更快。IHGPSO 在第 93 次收敛, 适应度为 153.26。GPSO 在第 141 次收敛, 适应度为 262.04。仿真结果表明, IHGPSO 算法解决广义旅行商问题的效果优于 GPSO 算法。

3 改进 QL 算法优化路径连接

IHGPSO 算法得到田块的最佳遍历顺序后, 由



(a) IHGPSO



(b) GPSO

图 5 最佳遍历顺序路径

Fig.5 Optimal traversal order path

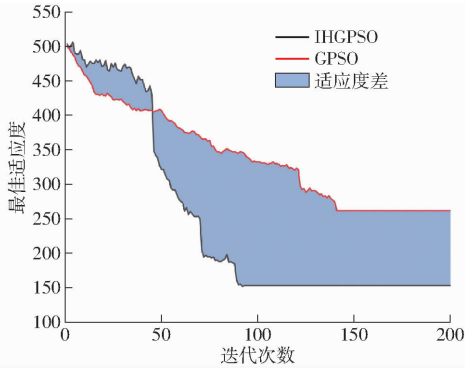


图 6 适应度迭代收敛曲线

Fig.6 Iterative convergence curves of fitness value

于菠萝果园田块之间存在田垄, 实际连接路径无法简单采用直线距离。如何高效地规划上一田块出口至下一田块入口间的路径成为全覆盖作业的关键环节。本研究通过改进的 QL 算法 (Order initialization Q-Learning, OI-QL) 完成田块之间的路径连接。

3.1 传统 QL 算法

QL 算法是一种经典的无模型离线强化学习算法, 其核心思想是通过智能体与环境的交互学习每个状态-动作对的价值即 Q 值, 使得在每个状态下选择奖励最大化的动作。在强化学习过程中, 智能体的行为遵循马尔可夫决策过程 (Markov decision process, MDP)。对于 MDP, 下一时刻的状态和奖励只与上一时刻所采取的动作有关, 而与前面历史动作信息无关。

QL 算法初始时会建立函数 $Q(s, a)$, 即 Q 表。

对于本研究连接菠萝田块的规划问题, s 表示农机所在地图像素坐标位置, a 表示前进的上下左右 4 个方向之一, 图 7 为 QL 算法流程图。学习初期 Q 值通常为 0, 学习过程 QL 算法通过不断更新函数 $Q(s, a)$ 逼近最优策略。更新公式为

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \max_{a'} Q(s', a') - Q(s, a)) \quad (24)$$

式中 α ——学习率, 控制更新步长

γ ——折扣因子, 平衡当前与未来的奖励

r ——当前动作直接反馈

$\max Q(s', a')$ ——下一状态 Q 表最大值

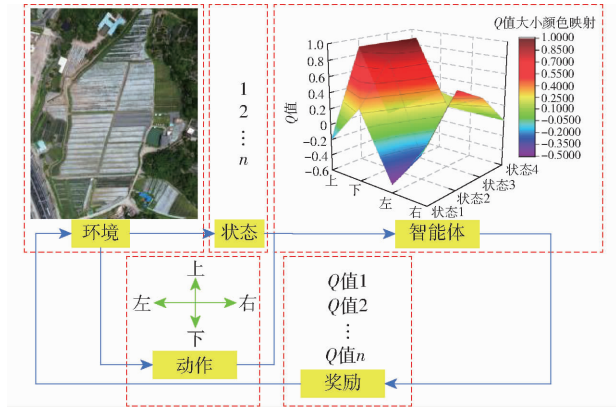


图 7 QL 算法流程图

Fig. 7 Flowchart of QL algorithm

传统 QL 算法由于 Q 表初始化为零, 早期探索效率低且易陷入局部最优; 同时, 固定的探索率和学习率可能造成收敛速度慢或策略振荡。

3.2 OI-QL 算法改进

本研究设计的 OI-QL 算法相较传统 QL 算法改进了 Q 表初始化、改进了学习率、设计新的奖励函数, 以缩短连接果园田块的路径长度, 提高路径质量。

(1) 改进 Q 表初始化

把全局遍历顺序 $U = \{u_1, u_2, \dots, u_K\}$ 融合进 QL 算法中, 将全局顺序信息编码为局部动作偏好对 Q 表进行初始化。对于状态 $s = (x, y)$ 和动作 a , 新的 Q 表初始化方法计算式为

$$Q_{\text{init}}(s, a) = -\|s + a - g_c\|_2 + 0.5 \text{Order}(s, a, U) \quad (25)$$

其中

$$\text{Order}(s, a, U) =$$

$$\begin{cases} \sum_{k=1}^{N_f} (\text{Ex}(s, a, K) + \text{En}(s, a, K)) & (\|s - \text{exit}_k\| < \tau) \\ 0 & (\text{其他}) \end{cases} \quad (26)$$

$$\text{Ex}(s, a, K) = \frac{a(\text{entry}_{K+1} - \text{exit}_K)}{\|\text{entry}_{K+1} - \text{exit}_K\|_2} \quad (27)$$

$$\text{En}(s, a, K) = \frac{a(\text{entry}_K - \text{exit}_{K-1})}{\|\text{entry}_K - \text{exit}_{K-1}\|_2} \quad (28)$$

式中 K ——田块序号

N_f ——田块总数

$\text{Ex}(s, a, K)$ ——田块出口对齐度函数

$\text{En}(s, a, K)$ ——田块入口对齐度函数

exit_K ——第 K 田块出口坐标

entry_K ——第 K 田块入口坐标

$\text{Order}(s, a, U)$ ——全局遍历顺序对齐函数

$\|s - \text{exit}_k\|$ ——状态 s 到出口的欧氏距离, m

τ ——距离阈值, 控制对齐度的激活范围, m

$Q_{\text{init}}(s, a)$ ——改进 Q 表的初始化值

$s + a$ ——执行动作后状态坐标

g_c ——目标状态坐标

OI-QL 算法通过融合全局遍历顺序与局部几何信息初始化 Q 表, 提升初始策略的质量。核心流程为得到田块的最优作业顺序 $U = \{u_1, u_1, \dots, u_K\}$ 后, 提取每个田块的出入口坐标, 然后用范数欧氏距离公式计算当前状态到目标点的路程。改进点为引入顺序对齐函数, 当检测到状态 s 进入出入口某个阈值的邻域时, 通过向量点积求得当前动作方向与全局最优方向 ($\text{exit}_k \rightarrow \text{entry}_{K+1}$) 的夹角余弦值, 加权后注入 Q 表初始化值。上述过程形成了具有拓扑意识的初始 Q 值分布, 强化了在路径关键节点出入口处方向的一致性, 在普通区域保留几何距离引导, 最终生成兼顾全局顺序与局部平滑性的初始化策略, 以避免传统 QL 冷启动问题。

(2) 学习率

传统学习率无论是固定值还是简单递减都无法有效地抑制偶发的大误差或噪声对学习率的影响, 易导致学习率剧烈波动, 进而影响 Q 值的稳定更新。针对此问题本研究提出基于误差动量反馈的学习率, 计算式为

$$\alpha_t(s, a) = \frac{\eta}{1 + N_t(s, a)^{0.5} m_t(s, a)^{0.5}} \quad (26)$$

其中

$$m_t(s, a) = 0.7 m_{t-1}(s, a) + (1 - \beta) |\delta_t(s, a)| \quad (27)$$

$$\delta_t(s, a) = r + \gamma \max_{a'} Q(s', a') - Q(s, a) \quad (28)$$

式中 $m_t(s, a)$ ——误差动量

$\alpha_t(s, a)$ ——动态学习率

η ——基础学习率

$N_t(s, a)$ ——状态-动作对的访问次数

$\delta_t(s, a)$ ——TD (Temporal-difference) 误差, 表示当前 Q 值与目标 Q 值的差

基于误差动量反馈的学习率在每一次 Q 值更新时,首先计算当前的 TD 误差 $\delta_t(s, a)$,通过动量机制计算其加权指数平均值,得到误差动量 $m_t(s, a)$ 。随着学习过程的推进, $N_t(s, a)$ 不断增加,动量项 $m_t(s, a)$ 也将根据误差的变化动态调整,最终,学习率 $\alpha_t(s, a)$ 由初始值、访问次数和误差动量项共同决定。在学习初期或误差较大时,学习率较高,有利于快速吸收新信息,而在学习后期或误差趋于平稳时,学习率会自动减小,自适应地调整智能体的学习步长,从而保证 Q 值更新的平滑性和稳定性。

(3) 奖励函数的设计

QL 算法的核心是设置合理的奖励值,算法需要通过定义奖励函数量化其在每个状态下的效用值,为策略迭代提供明确的优化方向。传统 QL 算法将靠近田块入口作为唯一的奖励来源。改进 QL 算法设计 3 个奖励函数对生成连接田块之间的路径进行正负奖励导向。

第 1 个奖励函数计算靠近下 1 个田块入口的奖励 $R_{\text{goal}}(S')$,计算式为

$$R_{\text{goal}}(S') = \begin{cases} 200 + \frac{\psi_1}{1 + d_{\text{goal}}(S')} & (S' \text{ 为入口}) \\ \frac{\psi_1}{1 + d_{\text{goal}}(S')} & (S' \text{ 为其他}) \end{cases} \quad (29)$$

式中 $d_{\text{goal}}(S')$ ——状态 S' 到终点的欧氏距离, m
 ψ_1 ——靠近终点距离的权重

第 2 个奖励函数计算靠近田块边缘的奖励 $R_{\text{edge}}(S')$,计算式为

$$R_{\text{edge}}(S') = \exp\left(-\frac{(d_{\text{edge}}(S') - d_{\text{ideal}})^2}{2\rho^2}\right) \quad (30)$$

式中 $d_{\text{edge}}(S')$ ——状态 S' 到田块边缘的距离, m
 d_{ideal} ——理想的边缘距离, m
 ρ ——控制奖励分布宽度的参数

第 3 个奖励函数计算直线运动奖励 $R_{\text{straight}}(S')$,计算式为

$$R_{\text{straight}}(S') = \begin{cases} \psi_2 & (\text{直行}) \\ \psi_2 \cos \iota & (\text{转弯}) \end{cases} \quad (31)$$

式中 ι ——前后前行方向的夹角, ($^\circ$)

ψ_2 ——直线距离权重

总奖励函数 $R(S')$ 为

$$R(S') = \begin{cases} -500 & (S' \text{ 进入田块}) \\ R_{\text{goal}}(S') + R_{\text{edge}}(S') + R_{\text{straight}}(S') & (\text{正常路径}) \end{cases} \quad (32)$$

总奖励函数能通过参数 ψ_1 、 ψ_2 、 d_{ideal} 和 ρ 灵活调

节行为偏好引导智能体靠近终点从而达到完成路径连接的目标。图 8 为 OI-QL 算法流程图。

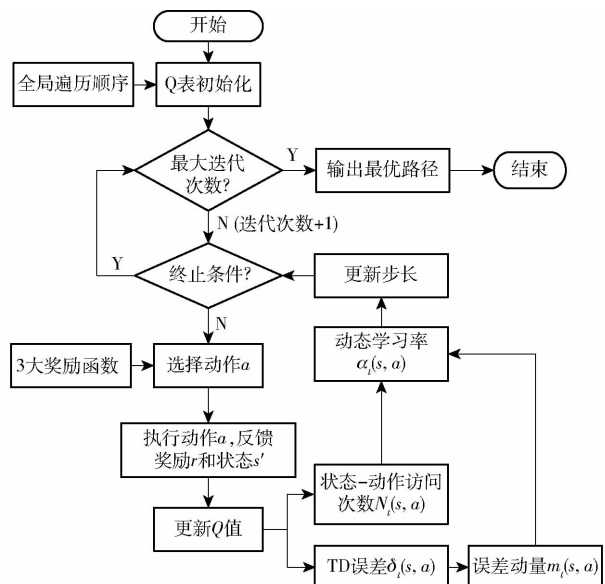


图 8 OI-QL 算法流程图

Fig. 8 Flowchart of OI-QL

4 仿真试验

4.1 目标地图

选取位于广东省中山市神湾镇 2 个不同特征的菠萝果园进行仿真试验。果园 A 共 12 块田, 规模较大, 形状不规则性更强; 果园 B 共 15 块田, 规模较小, 田块形状较为规则, 多为四边形。首先利用无人机对 2 个菠萝果园进行拍照采集完整的初始地图, 分别如图 9a、9b 所示。再把初始地图导入 Matlab 软件转换为田块轮廓坐标图。果园 A 田块按顺序从 0 到 11 进行编号, 果园 B 田块按顺序从 0 到 14 进行编号, 分别如图 9c、9d 所示。

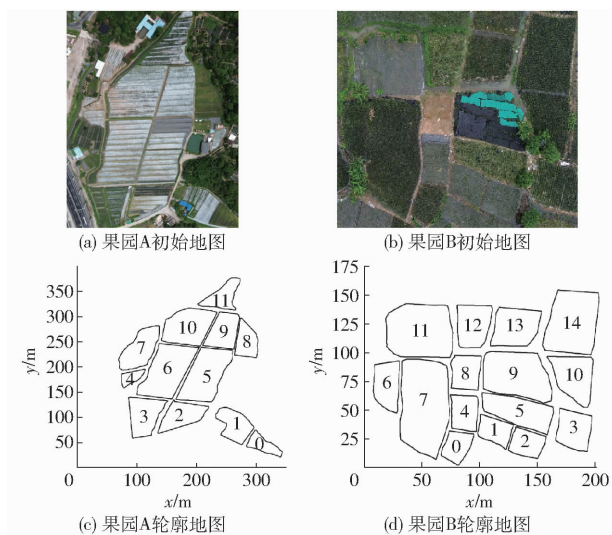


图 9 目标菠萝果园地图

Fig. 9 Map of target pineapple orchard

4.2 仿真方法

试验硬件配置为 NVIDIA GeForce RTX4070 8 GB 显存,搭载 AMD Ryzen 9 7940HX CPU,运行内存为 16 GB,试验平台为 Visual Studio Code。目标田块边界已知,不考虑动态障碍。设转弯预留区域宽度为 4 m,菠萝老苗清理机具搭载在宽度为 1.2 m 的龙门式履带机上,履带轴距为 1.15 m,可以原地掉头,符合最小转弯半径的要求。机具宽度略大于履带机宽度,设定作业幅宽 W 为 2.5 m。农机在预留区域进行半圆形转弯时,单次转弯作业消耗分为 2 个部分路段,转弯前的直线路段,设其速度 v_1 为 1.2 m/s;实际转弯的半圆形路段,设其速度 v_2 为 1 m/s。

行进角度定义为农机在田块内作业时的行进方向与水平方向(X 轴正方向)的夹角,范围为 $0^\circ \sim 360^\circ$ 。不同的行进角度会导致不同的转弯消耗时间,进而影响作业效率,因此本研究使各田块都按每隔 5° 的作业角度遍历覆盖路径,进行对比从而求得最优行进角度并以此确立各个田块的出入口坐标。分别用 QL 算法、OI-QL 算法和启发式 A^* 算法按照 IHGPSO、GPSO 算法得出的遍历顺序连接各田块出入口。

4.3 评价指标

对于最优行进角度评价指标,以转弯路径消耗总时间最小为标准。对于连接路径评价指标,除常规的路径长度、路径重复率、路径转折数指标外,增加路径平均弧度 $\bar{\theta}$ 作为评价指标体系分别对传统 QL 算法、OI-QL 算法和启发式 A^* 算法在目标菠萝果园中生成的连接路径进行对比分析。路径平均弧度 $\bar{\theta}$ 计算式为

$$\bar{\theta} = \frac{1}{N_t} \sum_{i=1}^{n-2} |\theta_i| \tag{33}$$

其中 $\theta = \arccos \frac{\mathbf{v}_f \mathbf{v}_b}{\|\mathbf{v}_f\| \|\mathbf{v}_b\|} \tag{34}$

式中 N_t ——有效转折数量

$\mathbf{v}_f$ 、 $\mathbf{v}_b$ ——前后速度向量

5 仿真结果与分析

5.1 最优作业行进角度

图 10 为以果园 A 田块 11 为例用 4 种不同行进角度规划的全覆盖路径,图 11 为以果园 A 中 12 个田块为例行进角度从 0° 到 355° 遍历的转弯路径消耗时间结果。表 1 为 2 个果园不同田块遍历所得最佳行进角度。果园 A、果园 B 遍历所得的最优全覆盖路径如图 12 所示。

5.2 路径连接仿真试验结果

分别用 GPSO 算法和 IHGPSO 算法求解 2 个果

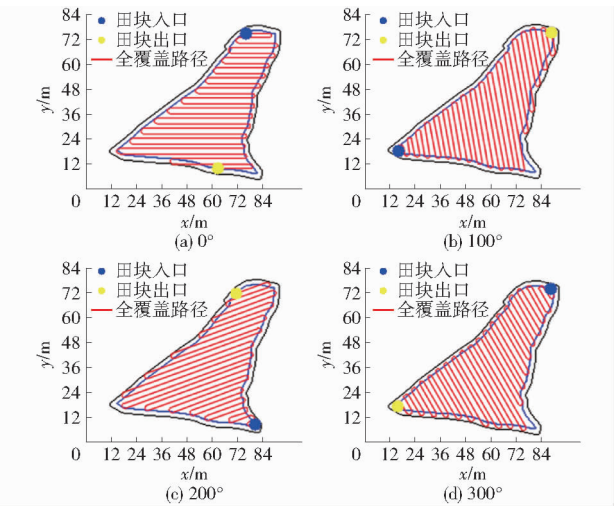


图 10 果园 A 中田块 11 不同行进角度的全覆盖路径
Fig. 10 Complete coverage path of field No. 11 at different travel angles in orchard A

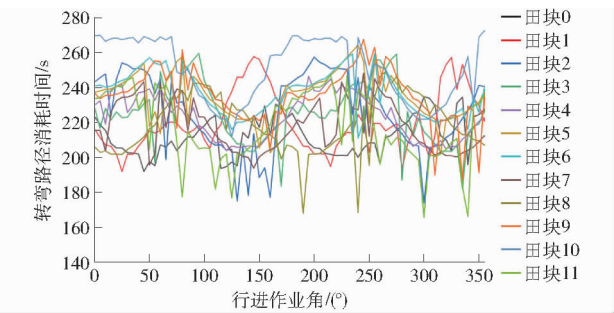


图 11 果园 A 中田块 0~11 的转弯路径消耗时间
Fig. 11 Time consumption result of turning path from plot 0 to 11 in orchard A

表 1 田块的最优行进角度

Tab.1 Optimal travel angle of field

果园 A		果园 B	
田块序号	最佳行进角度/ (°)	田块序号	最佳行进角度/ (°)
0	45	0	120
1	25	1	230
2	300	2	300
3	335	3	30
4	120	4	220
5	315	5	250
6	125	6	5
7	145	7	15
8	190	8	120
9	310	9	35
10	125	10	20
11	300	11	50
		12	135
		13	315
		14	310

园田块的最佳遍历顺序,再分别用 QL 算法、OI-QL 算法和启发式 A^* 算法按照最佳遍历顺序连接出入口,并进行分析比较。

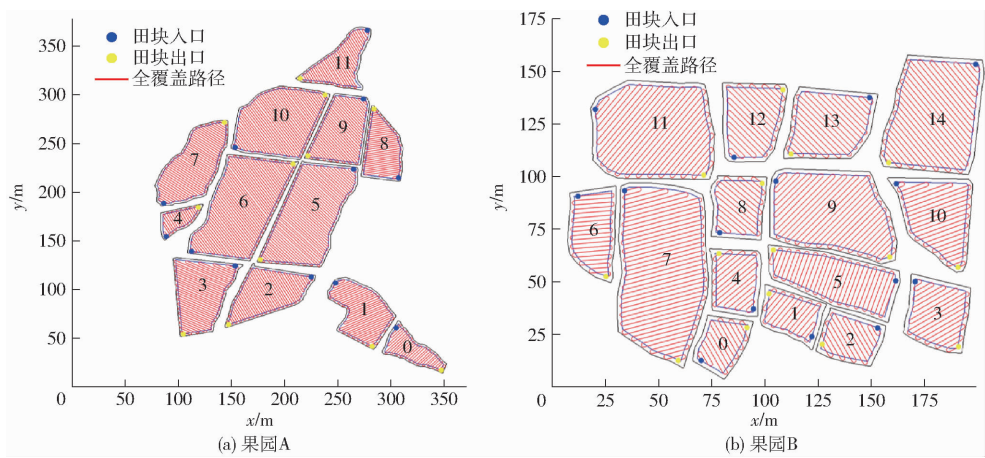


图 12 目标果园最优全覆盖路径
Fig. 12 Optimal complete coverage path for target orchard

GPSO 算法得出果园 A 最佳遍历顺序为 6-3-5-2-1-0-8-9-11-10-7-4, 果园 B 最佳遍历顺序为 12-14-7-11-9-13-6-3-1-5-0-2-8-10-4。IHGPSO 算法得出的果园 A 最佳遍历顺序为 3-6-4-7-10-11-9-8-5-2-1-0, 果

园 B 最佳遍历顺序为 6-7-11-8-12-13-14-10-3-9-5-2-1-4-0。

A* 算法采用欧氏距离作为启发函数, OI-QL 算法与 QL 算法基本参数如表 2 所示。图 13、14 为果园 A 3 种路径规划算法仿真路径。图 15、16 为

表 2 QL 和 OI-QL 算法的初始参数
Tab. 2 Initial parameters of QL algorithm and OI-QL algorithm

算法	迭代次数	Q 表初始化	学习率	折扣因子	探索概率	ψ_1	ψ_2	d_{ideal}	ρ
QL	500	0	0.7	0.9	0.2	150			
OI-QL	500	全局顺序	误差动量	0.9	0.2	150	200	5	0.2

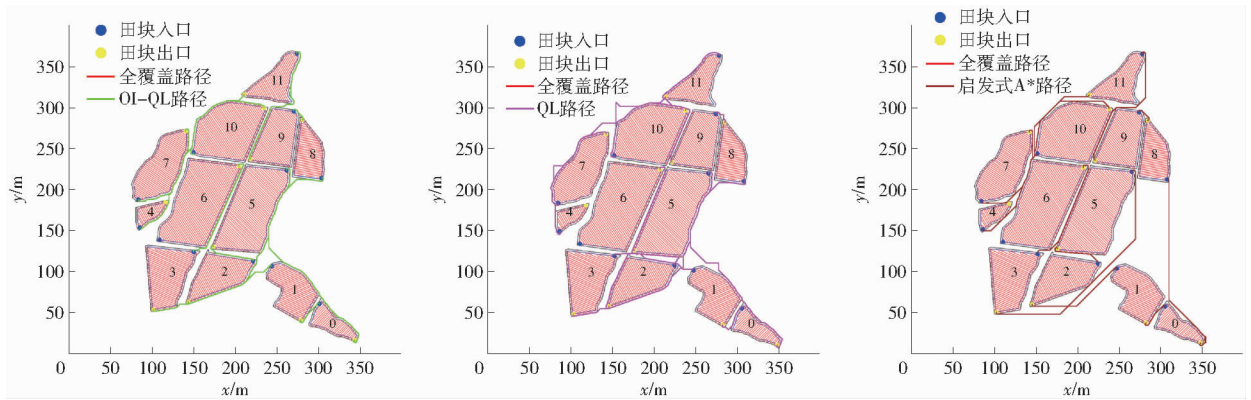


图 13 果园 A 3 种算法按 GPSO 顺序连接路径
Fig. 13 Path diagrams of three algorithms under GPSO order in orchard A

果园 B 3 种路径规划算法仿真的路径。

5.2.1 路径质量

图 17、18 为果园 A、B 路径质量对比。2 个果园仿真结果表明 3 种路径规划算法根据 IHGPSO 算法遍历所得最佳顺序连接的路径质量均优于 GPSO 算法; 在同一遍历顺序下 OI-QL 算法路径规划质量都优于 QL 算法和启发式 A* 算法。

5.2.2 迭代曲线

图 19a 为 QL 和 OI-QL 算法在果园 A 的连接路径迭代收敛曲线。根据 GPSO 遍历所得最佳顺序

进行迭代的过程中, QL 算法在第 305 次迭代时收敛, 收敛时路径长度为 2 014 m; OI-QL 算法在第 199 次迭代时收敛, 收敛时路径长度为 1 505 m。根据 IHGPSO 遍历所得最佳顺序进行迭代过程中, QL 算法在第 216 次迭代时收敛, 收敛时路径长度为 1 689 m; OI-QL 算法在第 138 次迭代时收敛, 收敛时路径长度为 1 131 m。

图 19b 为 QL 和 OI-QL 算法在果园 B 的连接路径迭代收敛曲线。根据 GPSO 遍历所得最佳顺序进行迭代的过程中, QL 算法在第 201 次迭代时收

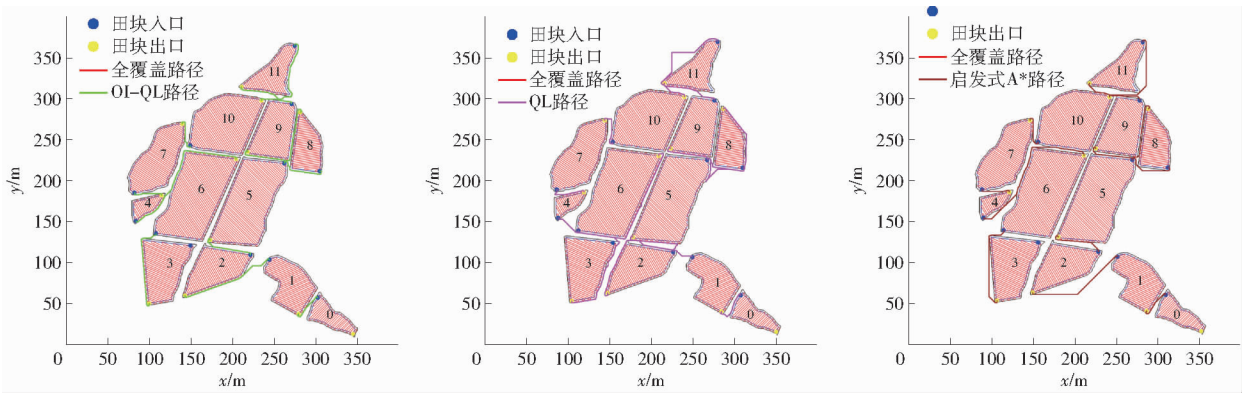


图 14 果园 A 3 种算法按 IHGPSO 顺序连接路径

Fig. 14 Path diagrams of three algorithms under IHGPSO order in orchard A

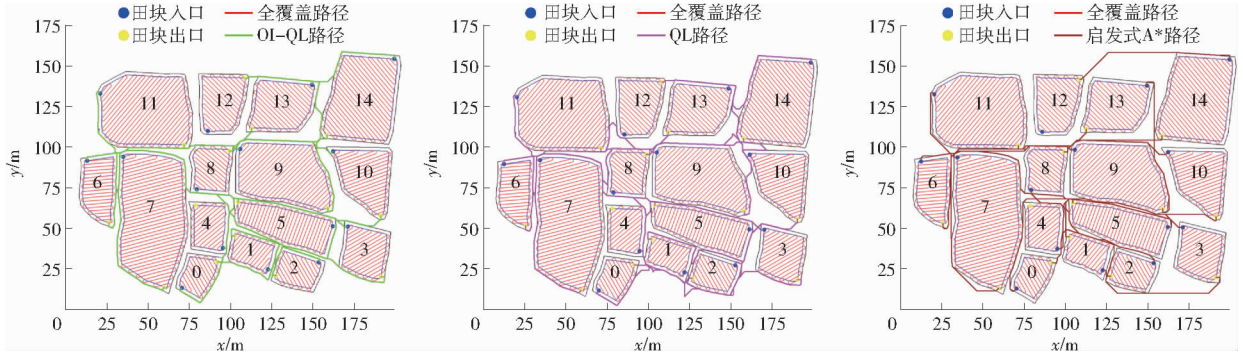


图 15 果园 B 3 种算法按 GPSO 顺序连接路径

Fig. 15 Path diagrams of three algorithms under GPSO order in orchard B

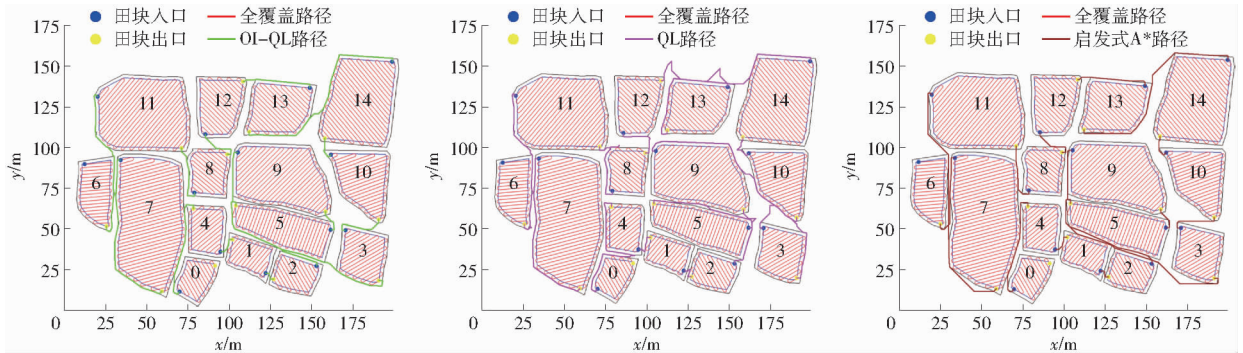


图 16 果园 B 3 种算法按 IHGPSO 顺序连接路径

Fig. 16 Path diagrams of three algorithms under IHGPSO order in orchard B

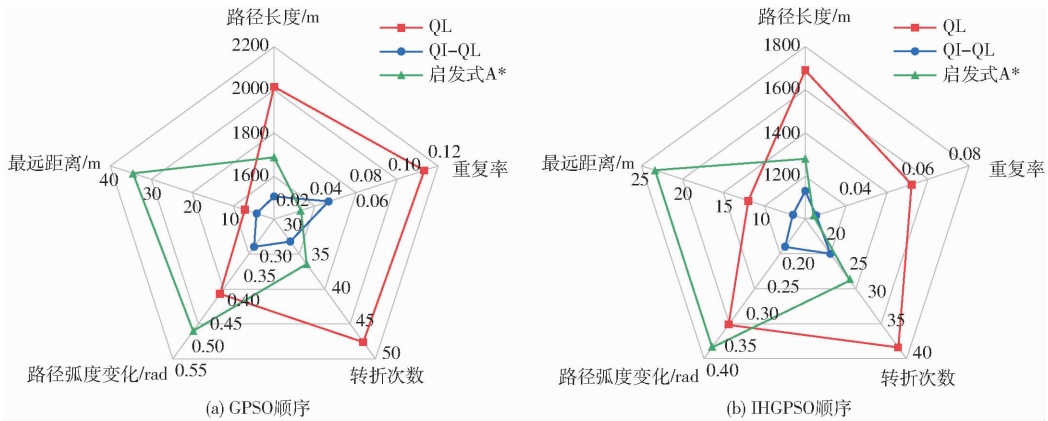


图 17 果园 A 路径质量对比

Fig. 17 Path quality of orchard A comparison chart

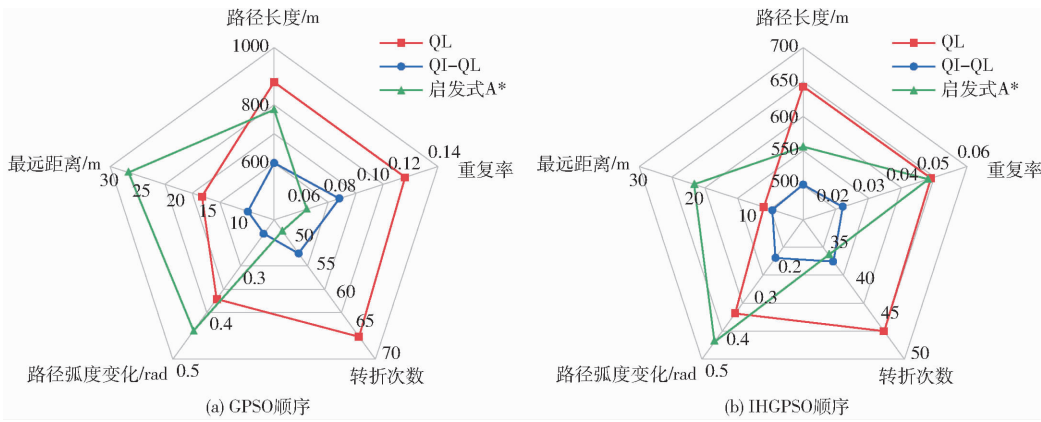


图 18 果园 B 路径质量对比

Fig. 18 Path quality of orchard B comparison chart

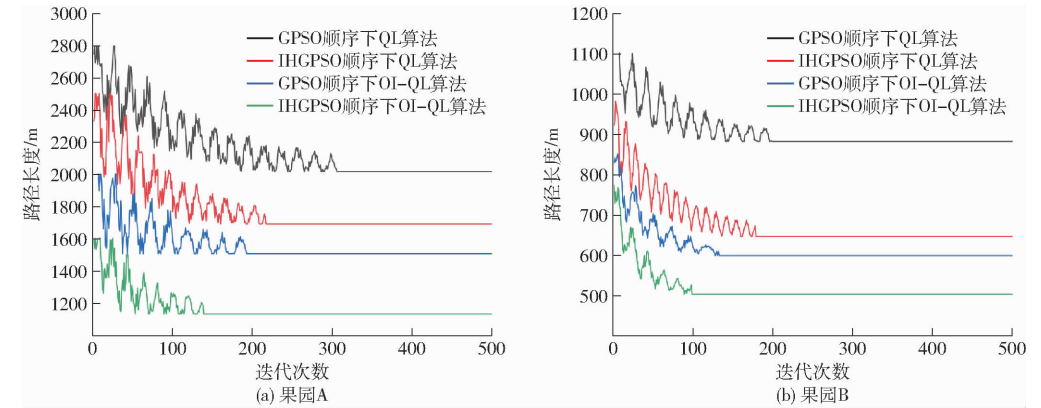


图 19 路径迭代曲线

Fig. 19 Path iterative curves

敛,收敛时路径长度为 880 m;OI-QL 算法在第 132 次迭代时收敛,收敛时路径长度为 597 m。根据 IHGPSO 遍历所得最佳顺序进行迭代过程中,QL 算法在第 178 次迭代时收敛,收敛时路径长度为 644 m;OI-QL 算法在第 98 次迭代时收敛,收敛时

路径长度为 501 m。

5.2.3 不同算法性能比较

为了进一步验证 IHGPSO 算法以及 OI-QL 算法性能,使用改进前后的 GPSO 算法与 QL 算法进行 20 次仿真试验取平均值以及标准差,结果如表 3 所示。

表 3 不同算法在目标菠萝果田的仿真结果							
Tab.3 Simulation results of different algorithms in target pineapple orchard							
目标果园	最佳遍历顺序	平均路径长度与标准差/m		平均收敛迭代次数与标准差		平均收敛时间与标准差/s	
		QL	OI-QL	QL	OI-QL	QL	OI-QL
果园 A	GPSO	2 014.5 ± 16.1	1 505.3 ± 6.4	305.3 ± 9.7	199.6 ± 10.1	424.7 ± 11.9	316.3 ± 9.2
	IHGPSO	1 689.6 ± 15.5	1 130.9 ± 4.8	216.1 ± 11.2	138.3 ± 8.9	351.9 ± 10.4	212.8 ± 5.9
果园 B	GPSO	880.1 ± 16.3	597.7 ± 7.1	201.4 ± 13.4	132.7 ± 9.1	258.6 ± 8.0	228.4 ± 9.8
	IHGPSO	644.5 ± 13.6	501.2 ± 3.8	177.9 ± 10.2	98.2 ± 4.4	184.7 ± 5.6	135.6 ± 5.5

通过对 2 个果园的数据进行统计分析,使用 IHGPSO 算法遍历所得最佳顺序生成的路径相比使用 GPSO 算法遍历所得最佳顺序生成的路径,平均长度缩短约 20.9%;平均收敛迭代次数减少约 24.3%;平均迭代时间减少约 29.7%,说明 IHGPSO 算法相比 GPSO 算法求解得的最佳遍历顺序的更优。在同一遍历顺序下,OI-QL 算法相比 QL 算法,平均长度缩短约 28.1%;平均收敛迭代次数减少约 37.3%;平均迭代时间减少约 25.7%,说明

OI-QL 算法相比 QL 算法在路径连接表现更好。

6 结论

(1)分析了不同行进角度对单田块全覆盖路径的影响,通过仿真求得各田块的最佳行进角度,并规划出各田块的最佳全覆盖路径。

(2)针对 GPSO 算法易陷入局部最优、无法根据优化阶段自适应调整等问题,IHGPSO 算法提出先分簇后使用最近邻算法来初始化种群,设计了新的

适应度函数以及动态调整粒子的交换机制。仿真结果表明,使用 IHGPSO 算法遍历所得最佳顺序生成的路径相比使用 GPSO 算法遍历所得最佳顺序生成的路径,平均长度缩短约 20.9%;平均收敛迭代次数减少约 24.3%;平均迭代时间减少约 29.7%。

(3) 针对 QL 算法探索效率低、迭代时间长的问

题,OI-QL 算法融合全局遍历顺序与局部几何信息初始化 Q 表,融合误差动量反馈动态学习率,设计了新的奖励函数。仿真试验表明,在同一遍历顺序,OI-QL 算法相比 QL 算法平均路径长度缩短约 28.1%,平均迭代次数减少约 37.3%,平均收敛时间减少约 25.7%。

参 考 文 献

[1] 刘天湖,赖嘉上,孙伟龙,等. 基于跳点优化蚁群算法的菠萝田间导航路径规划[J]. 农业机械学报, 2025, 56(4): 387-396.
LIU Tianhu, LAI Jiashang, SUN Weilong, et al. Navigation path planning of pineapple planting field based on jump point optimized ant colony algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(4): 387-396. (in Chinese)

[2] 陈凯,解印山,李彦明,等. 多约束情形下的农机全覆盖路径规划方法[J]. 农业机械学报, 2022, 53(5): 17-26, 43. CHEN Kai, XIE Yinshan, LI Yanming, et al. Full coverage path planning method of agricultural machinery under multiple constraints[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 17-26, 43. (in Chinese)

[3] 罗承铭,熊陈文,黄小毛,等. 四边形田块下油菜联合收获机全覆盖作业路径规划算法[J]. 农业工程学报, 2021, 37(9): 140-148.
LUO Chengming, XIONG Chenwen, HUANG Xiaomao, et al. Coverage operation path planning algorithms for the rape combine harvester in quadrilateral fields[J]. Transactions of the CSAE, 2021, 37(9): 140-148. (in Chinese)

[4] 王宁,韩雨晓,王雅萱,等. 农业机器人全覆盖作业规划研究进展[J]. 农业机械学报, 2022, 53(增刊1): 1-19. WANG Ning, HAN Yuxiao, WANG Yaxuan, et al. Research progress of agricultural robot full coverage operation planning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp.1): 1-19. (in Chinese)

[5] 周俊,何永强. 农业机械导航路径规划研究进展[J]. 农业机械学报, 2021, 52(9): 1-14.
ZHOU Jun, HE Yongqiang. Research progress on navigation path planning of agricultural machinery[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 1-14. (in Chinese)

[6] KATO E R R, INOUE R S, FRANCO L S. A method for planning multirotor unmanned aerial vehicle flight paths to cover areas using the ant colony optimization metaheuristic[J]. Computers and Electronics in Agriculture, 2025, 231: 109983.

[7] 李涛,邱权,赵春江,等. 矮化密植果园多臂采摘机器人任务规划[J]. 农业工程学报, 2021, 37(2): 1-10. LI Tao, QIU Quan, ZHAO Chunjiang, et al. Task planning of multi-arm harvesting robots for high-density dwarf orchards[J]. Transactions of the CSAE, 2021, 37(2): 1-10. (in Chinese)

[8] 吴越安,杜昌平,杨睿,等. 基于改进遗传算法的倾转旋翼无人机区域覆盖路径规划[J]. 浙江大学学报(工学版), 2024, 58(10): 2031-2039.
WU Yuean, DU Changping, YANG Rui, et al. Area coverage path planning for tilt-rotor unmanned aerial vehicle based on enhanced genetic algorithm[J]. Journal of Zhejiang University (Engineering Science), 2024, 58(10): 2031-2039. (in Chinese)

[9] 李文峰,徐蕾,杨琳琳,等. 基于改进蚁群算法的农业机器人多田块路径规划方法与试验[J]. 南京农业大学学报, 2024, 47(4): 823-834.
LI Wenfeng, XU Lei, YANG Linlin, et al. Multi field path planning method and experiment of agricultural robot based on improved ant colony algorithm[J]. Journal of Nanjing Agricultural University, 2024, 47(4): 823-834. (in Chinese)

[10] 武晓雯,郭孟营,胡阿建,等. 基于改进遗传粒子群算法的无人机路径规划[J]. 仪器仪表学报, 2025, 46(4): 315-325. WU Xiaowen, GUO Mengying, HU Ajian, et al. Path planning for UAV based on improved hybrid genetic particle swarm algorithm[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 315-325. (in Chinese)

[11] 张万枝,赵威,李玉华,等. 基于改进 A* 算法 + LM-BZS 算法的农业机器人路径规划[J]. 农业机械学报, 2024, 55(8): 81-92.
ZHANG Wanzhi, ZHAO Wei, LI Yuhua, et al. Path planning of agricultural robot based on improved A* and LM-BZS algorithms[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 81-92. (in Chinese)

[12] 刘慧,张世义,段云鹏,等. 基于改进双向 RRT* 的果园机器人运动规划算法[J]. 农业机械学报, 2022, 53(11): 31-39.
LIU Hui, ZHANG Shiyi, DUAN Yunpeng, et al. Orchard robot motion planning algorithm based on improved bidirectional RRT* [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 31-39. (in Chinese)

[13] 万俊,孙薇,葛敏,等. 基于含避障角人工势场法的机器人路径规划[J]. 农业机械学报, 2024, 55(1): 409-418. WAN Jun, SUN Wei, GE Min, et al. Robot path planning based on artificial potential field method with obstacle avoidance angles[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 409-418. (in Chinese)

processing[J]. Computers and Electronics in Agriculture, 2023, 210: 107923.

[18] KOIRALA A, WALSH K B, WANG Z L, et al. Deep learning-method overview and review of use for fruit detection and yield estimation[J]. Computers and Electronics in Agriculture, 2019, 162: 219 – 234.

[19] MINAEE S, BOYKOV Y Y, PORIKLI F, et al. Image segmentation using deep learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3523 – 3542.

[20] 孙红, 李松, 李民赞, 等. 农业信息成像感知与深度学习应用研究进展[J]. 农业机械学报, 2020, 51(5): 1 – 17.
SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 1 – 17. (in Chinese)

[21] ZHU X, CHEN F, ZHANG X, et al. Detection the maturity of multi-cultivar olive fruit in orchard environments based on Olive-EfficientDet[J]. Scientia Horticulturae, 2024, 324: 112607.

[22] ZHENG C, CHEN P, PANG J, et al. A mango picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard[J]. Biosystems Engineering, 2021, 206(6): 32 – 54.

[23] 周馨翌, 吴烽云, 邹湘军, 等. 基于多目标识别的葡萄果串采摘点定位方法[J]. 农业工程学报, 2023, 39(22): 166 – 177.
ZHOU Xinzhaoy, WU Fengyun, ZOU Xiangjun, et al. Method for locating picking points of grape clusters using multi-object recognition[J]. Transactions of the CSAE, 2023, 39(22): 166 – 177. (in Chinese)

[24] WANG C, WANG C, WANG L, et al. A lightweight cherry tomato maturity real-time detection algorithm based on improved YOLOV5n[J]. Agronomy, 2023, 13(8): 2106.

[25] MA D M, LI P H, HUANG X Y, et al. Efficient semantic segmentation based on improved DeepLabV3[J]. Computer Engineering & Science, 2022, 44(4): 737 – 745.

[26] WANG Q, WU B, ZHU P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.

[27] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236 – 239.

(上接第 12 页)

[14] 劳彩莲, 李鹏, 冯宇. 基于改进 A* 与 DWA 算法融合的温室机器人路径规划[J]. 农业机械学报, 2021, 52(1): 14 – 22.
LAO Cailian, LI Peng, FENG Yu. Path planning of greenhouse robot based on fusion of improved A* algorithm and dynamic window approach[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 14 – 22. (in Chinese)

[15] 熊俊涛, 李中行, 陈淑绵, 等. 基于深度强化学习的虚拟机器人采摘路径避障规划[J]. 农业机械学报, 2020, 51(增刊 2): 1 – 10.
XIONG Juntao, LI Zhongxing, CHEN Shumian, et al. Obstacle avoidance planning of virtual robot picking path based on deep reinforcement learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp. 2): 1 – 10. (in Chinese)

[16] WANG Y, LU C, WU P, et al. Path planning for unmanned surface vehicle based on improved Q – Learning algorithm[J]. Ocean Engineering, 2024, 292: 116510.

[17] MAOUDJ A, HENTOUT A. Optimal path planning approach based on Q-Learning algorithm for mobile robots[J]. Applied Soft Computing, 2020, 97: 106796.

[18] 李潇宇, 张君华, 郭晓光, 等. 基于强化学习的机器人底盘能量管理与路径规划优化算法[J]. 农业工程学报, 2024, 40(21): 175 – 183.
LI Xiaoyu, ZHANG Junhua, GUO Xiaoguang, et al. Reinforcement learning-based optimization algorithm for energy management and path planning of robot chassis[J]. Transactions of the CSAE, 2024, 40(21): 175 – 183. (in Chinese)

[19] ZHOU Q, LIAN Y, WU J, et al. An optimized Q – Learning algorithm for mobile robot local path planning[J]. Knowledge-Based Systems, 2024, 286: 111400.

[20] LOW E S, ONG P, LOW C Y, et al. Modified Q-Learning with distance metric and virtual target on path planning of mobile robot[J]. Expert Systems with Applications, 2022, 199: 117191.