

文章编号:1673-1522(2019)03-0297-06

DOI: 10.7682/j.issn.1673-1522.2019.03.007

基于认知结构SOAR的机器人路径规划

穆 杨

(海军航空大学,辽宁 葫芦岛 125000)

摘 要:针对环境未知、目标位置已知情况下机器人路径规划问题,提出了一种基于认知结构SOAR的探索方法。首先,结合箱子移动问题介绍了SOAR的基本概念及求解机制;其次,在讨论了人的行为模型和SOAR的决策过程基础上,设计了基于SOAR-Agent的行为建模方法;最后,利用本文提出的方法,设计长期知识,使机器人能够模拟人在未知动态环境下向目标搜索的行为。通过仿真实验,对比分析了长期知识的完备程度对机器人顺利完成任务的影响,同时也验证了该方法的有效性。

关键词:认知结构;SOAR;机器人;未知环境

中图分类号:TP242

文献标志码:A

实时避障及路径规划是自主移动机器人的核心技术,随着机器人应用环境的复杂化,对这2项核心技术的要求也越来越高。根据环境及任务的复杂性,可以将研究背景划分为以下4个等级:

- 1)环境已知、静态、目标位置已知;
- 2)环境未知、静态、目标位置已知;
- 3)环境未知、动态、目标位置已知;
- 4)环境未知、动态、目标位置未知。

当前,国内研究主要集中在第一等级的全局规划中^[1-3],而随着环境的逐步恶劣,预载地图由于精度及范围等原因并不能够提供足够的信息^[4-5],而且当环境变化时,原规划路径很可能被障碍物截断,导致采用原来规划的路径移动时,可能陷入陷阱或者发生无法到达目标位置的问题。所以,在其他3个等级的态势下,机器人还须要具有局部规划能力,使用传感器探测到的信息控制机器人避免碰撞障碍物^[6-7]。本文引入认知结构SOAR,通过Visual SOAR开发环境设计SOAR-Agent控制机器人,模仿人在未知、动态环境下向已知的目标位置搜索路径的行为,实现了在第三等级态势下的路径规划能力。

1 SOAR的概念及基本原理

随着人工智能技术的发展,为了弥补智能设计只能应用于设计好的领域的限制^[8],国外诸多学者于90年代开展了认知结构的研究^[9]。认知结构旨在开发一种通用的架构^[10],通过合理的设计,使开发的系统具有人的相同的认知能力,实现人级智能体(Human-level Agents)^[11]。在之后的20多年的研究中,先后涌现出了

诸如SOAR^[12]、ACT-R^[13]、EPIC^[14]、Clarion^[15]、SAL^[16]、Icarus^[17]等认知结构。

SOAR是由Newell等人于1983年开发出来的一种认知结构,它是State, Operator, Act and Result的缩写,译为状态、操作、行动和结果。简单地说,就是选择操作,作用环境,改变状态,从而产生相应的结果。经过20多年的发展,SOAR在计算机生成兵力^[18],无人机任务规划^[19]等方面取得了诸多成果。以下结合图1所示的箱子移动问题介绍SOAR的基本概念及其求解机制。

箱子移动问题,即将随机放置的3个方块按照期望状态顺序排列,每次只能移动1个方块,期望状态如图1所示。

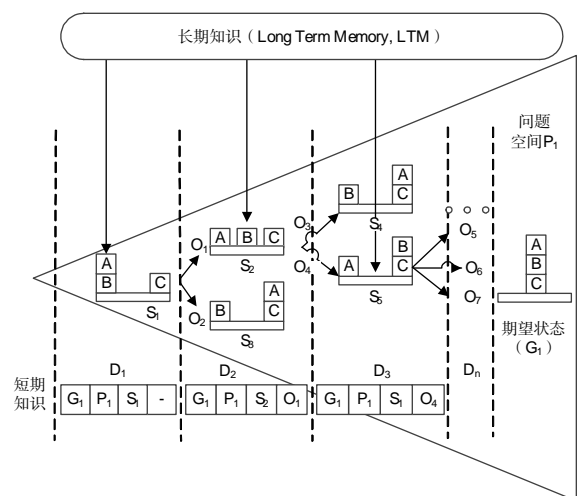


图1 箱子移动的问题空间

Fig.1 Problem space of box moving

收稿日期:2019-04-28; 修回日期:2019-05-24

作者简介:穆 杨(1989-),女,硕士。

1.1 基于问题空间的求解机制

从箱子的初始状态转换到箱子的期望状态过程中,存在许多其他状态,这些状态构成了箱子移动问题的“问题空间”。

1)问题空间^[13](Problem Space):当前状态到目标状态转换过程中,所有可能出现状态的集合以及使状态改变对应的操作。对于箱子移动问题,问题空间包括了箱子所有位置状态的组合,如图1中的 $S_1 \sim S_3$ 、 G_1 、 $O_1 \sim O_7$ 。

2)状态(State)。当前世界信息:对于箱子移动问题,状态包含了箱子当前的位置信息。如图1中 D_1 下的 S_1 、 D_2 下的 S_2 等。

3)操作(Operator):控制状态向目标转化。操作包括前提条件及相应的行为,可对操作进行评估,设置优先度(Preference),控制操作的选择。对于箱子移动问题,操作包含了对箱子的移动。如图1中的 $O_1 \sim O_7$ 。

4)短期知识(Working Memory, WM):用于存储各种临时信息,包括当前的环境信息、当前的操作、当前的目标。

5)长期知识(Long Term Memory, LTM):用于存储规则等信息,长期知识只有通过操作才可以更改。

6)困境(Impasse):基于问题空间的求解机制在求解过程中会出现以下4种困境,如表1所示。

表1 困境类型

Tab.1 Impasse types

困境类型	困境描述
无合适操作	当前状态下,没有可选操作
多个操作可选	当前状态下,多个操作优先度相同,缺乏足够知识选择最优操作
操作冲突	当前状态下,操作优先权存在冲突
操作无变化	操作不存在动作,无法改变状态 操作对外界的改变需要一定时间才能完成

针对以上4种困境,SOAR通过在短期知识中创建子状态(Substate),并在子状态中创建新的问题空间去解决困境。具体解决方法如表2所示。

表2 困境解决方案

Tab.2 Solution of impasse

困境类型	解决方案
无合适操作	创建新的问题空间,以设计新的操作为目标求解
多个操作可选	复制上一层状态,使用向前搜索策略,对操作进行评估,从而选出最优操作
操作冲突	操作冲突意味着长期知识中存在冲突规则,该问题解决方案与多个操作可选的解决方案相同
操作无变化	创建新的问题空间,以设计对应的动作为目标求解 在子状态中重复选择该操作直到状态产生变化

1.2 基于经验的试错学习机制——经验学习

多个操作可选以及操作存在冲突的解决方案是通过采用复制上一层状态,在子状态的问题空间中使用向前搜索策略,重新对操作进行评估,对比人的行为方式,当已有规则无法判断哪一种决策更合适时,人通常是对决策的下一状态进行评估,最优解即使状态更接近目标的决策。通过采用向前搜索策略,人可以对决策进行评估,将评估结果及环境信息存储在记忆中,在将来碰到相同情况下,直接使用,避免二次推理。Newell等学者通过引入实践法则^[20]设计了基于经验的试错学习机制——经验学习,实现了相同的能力,从而使在子状态的问题空间中推理得到的规则能够作为长期知识直接使用。

1.3 基于强化学习的学习机制

在加入了 chunking 功能之后,利用 SOAR 开发的 agent 具备一定的学习能力,但是,该学习能力是建立在外部环境不变的前提下,当外部环境出现变化, chunking 习得的知识以及已有的规则知识显然不再正确。为了能够调整已有的知识以应对变化的态势,SOAR 加入了如图2所示的基于强化学习(Reinforcement Learning, RL)的学习机制^[21],即通过对能够到达目标的一系列操作给予奖励,并将奖励信息作为知识的一部分,从而使 SOAR 开发的 agent 能够在行动—评价的模式中改变操作优先度以适应充满噪声、变化的动态环境。

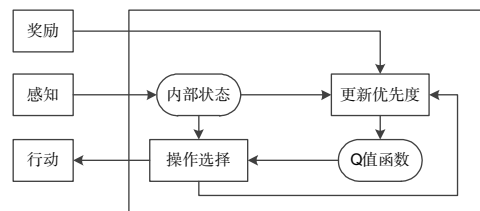


图2 基于强化学习的学习机制

Fig.2 Learning mechanism based on reinforcement learning

SOAR 支持 ϵ -贪婪算法以及玻尔兹曼算法。在 ϵ -贪婪算法下,agent在固定的 Q 值下依概率选择操作。在玻尔兹曼算法下,agent在当前状态下,对每一个操作 O_1, O_2, \dots, O_n 分配一个预期 Q 值,分别为 $Q(s, Q_1), Q(s, Q_2), \dots, Q(s, Q_n)$,对于操作 O_i 的选择概率为 $\frac{\exp(Q(s, O_i)/\tau)}{\sum_{j=1}^n \exp(Q(s, O_j)/\tau)}$,其中 τ 是阈值,控制操作被选择的概率。当 $\tau=0$ 时,选择 Q 值最大的操作;当 $\tau=1$ 时,依 Q 值概率选取操作,当 $\tau \rightarrow \infty$ 时,随机选取操作。

2 基于 SOAR-Agent 的行为建模方法

由于在未知的、动态环境下无法使用任务先验信息,从而无法使用任何算法对全局路径进行规划。为了保证机器人在不触碰障碍物的前提下,同时实现到达目标位置的目的,本文通过使用认知结构 SOAR 开发 agent,模拟人在此环境下的行为方式,使机器人具有人的思考能力,从而能够应对动态的未知环境。首先对人的行为模型进行分析。

2.1 人的行为模型

Rasmussen 等人在分析了人类行为模型当中不同元素之间的区别及联系基础上,总结出行为的 3 个层级^[22],如图 3 所示。

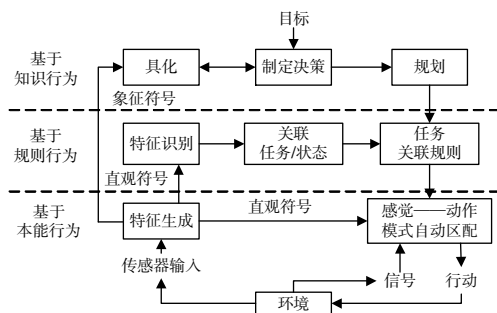


图 3 人的行为模型

Fig.3 Model of human behavior

1) 基于本能层级。该层级,人是基于本能对外界做出反应,是未经思考的。例如,开车时,车前突然出现一只动物,没有经验的驾驶员通常会立即打方向盘以趋避(结果非常危险)。

2) 基于规则层级。人通过感知器官获取外界信息,部分信息能够直接转换成可以识别的符号,被识别之后,结合当前任务或是状态,将之与记忆当中类似问题相关联,确定是否有明确的行动规则,最后寻找到相应的反应行为。

3) 基于知识层级。人在面对未知事物时,外界信息只是具有象征意义的符号,记忆当中不包含相应规则。此时,需要利用已有知识结合所要达到的目标进行推理并找到正确的答案。这一层级是人类行为的最高层级。

2.2 SOAR 的决策过程

SOAR 的决策过程如图 4 所示。具体分析如下:

1) 信息输入。使用 SOAR 设计的 agent 可以通过输入链接(input-link)动态的获取诸如传感器探测到的外部信息,将这部分信息存储在 WM 中。

2) 状态细化、操作建议、操作评估。状态细化

(Elaborate State)复制外部信息,通过与 LTM 中的长期知识比对产生环境的理解知识;操作建议根据当前的环境知识,给出符合条件的操作,并将这些操作存储在 WM 中;操作评估通过比较备选操作,设置优先级。

3) 决策。根据备选操作的优先级决定操作的选择。此时,可能会出现困境,产生子状态,重新执行步骤 2)。

4) 操作细化。在步骤 2) 中只给出了符合条件的操作,如需对其动作进行修改,可以在操作细化阶段重新设计。

5) 操作执行。操作指令写入 WM 中,将操作对应的动作作用于内部状态,如果存在对外部环境的作用,则将操作指令传至输出链接(output-link)中。由于环境的改变可能需要时间,从而需要重复执行操作。如果当前给出的操作过于抽象,需要重新执行步骤 4),将操作分解成外部执行机构能够执行的动作。

6) 输出。执行输出指令,作用于外部环境。

7) 目标状态识别。目标状态知识隐含在长期知识中,所以目标状态识别需要通过与环境知识匹配,使用相应的操作判断,然后终止整个过程。

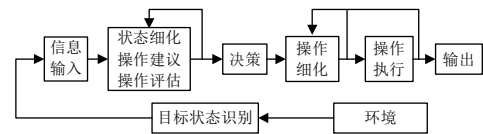


图 4 SOAR 的决策过程

Fig.4 Decision process of SOAR

2.3 基于 SOAR-Agent 的行为建模方法

对比人的行为模型和 SOAR 的决策过程可以发现,通过将感知到的外界信息以及内部产生的操作信息存储在 agent 的短期知识中,设计长期知识库(特征知识、任务知识、目标知识、规划知识),利用 SOAR 的基于问题空间的求解机制,对相关任务及目标进行推理,从而实现模拟人的行为能力,进而应用到机器人的路径规划当中,具体设计如图 5 所示。

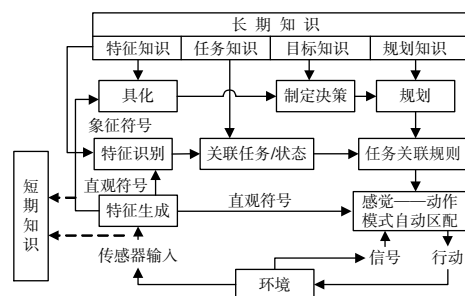


图 5 基于 SOAR-Agent 的行为建模方法

Fig.5 Behavior modeling based on SOAR-Agent

3 仿真实验分析

3.1 环境描述

为了验证本方法的有效性,采用文献[23]中的栅格模型地图,同时增加地图难度,其中一个障碍物能够向左移动2个单位距离,从而影响机器人的行进路线,机器人能够感知周围一个单位距离的信息,如图6所示。

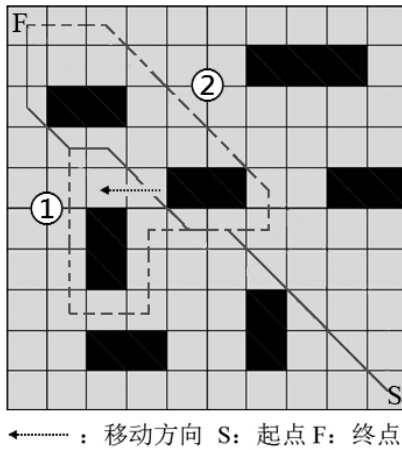


图6 避障地图
Fig.6 Avoidance map

3.2 长期知识设计

长期知识设计体现了任务的描述程度以及人的行为方式、人对外界环境的认知。本文通过设计2组长期知识以体现不同长期知识的设计对机器人的影响。

第1组:

- 1)特征知识。黑色物体是障碍物。
- 2)任务知识、目标知识。本实验任务知识和目标知识一致,在不触碰障碍物的前提下到达F点。
- 3)规划知识。规划知识控制着操作的选择,具体设计如表3所示。

表3 规划知识(1)

Tab.3 Planning knowledge(1)

优先度	操作
1	选择无障碍物方向移动
2	选择最近路径向F点移动
3	随机选择几种方向: ↑、↓、←、→、↖、↗、↘、↙

第2组:

特征知识、任务知识、目标知识不变,规划知识设计如表4所示。

表4 规划知识(2)

Tab.4 Planning knowledge(2)

优先度	操作
1	选择无障碍物方向移动
2	选择未探索区域方向移动
3	选择最近路径向F点移动
4	随机选择几种方向: ↑、↓、←、→、↖、↗、↘、↙

3.3 仿真实验分析

使用 Visual SOAR 开发环境设计 SOAR-Agent, 写入上述长期知识, 在 Visual C++ 环境下设计地图, 作为外部环境, 并使用 SOAR Debugger 软件观察 SOAR-Agent 的决策过程以及其与外界环境的交互过程。第1组长期知识设计下的运动轨迹如图6中的实线所示, 但是, 当移动障碍向左运动一个单位之后, 形成了U型陷阱, 此时, 机器人无法逃离, 当障碍物向右移动至原位时, 机器人才能继续前进。第2组长期知识设计下的运动轨迹如图6中的虚线①所示, 实验结果表明, 机器人可以应对动态的未知环境, 并从U型陷阱中逃离, 但是, 该路线并不是全局最优的, 最优路径如图6中的虚线②所示。

从实验结果可以看出, 通过预先载入的认知模型(长期知识)进行路径探索, 符合人在未知环境下探索路径的认知规律。在未知且变化的环境中, 依然可以找到通往目标的路径。与传统路径规划研究方法(基于A*[23]搜索算法、遗传算法[24]、蚁群算法[25]、深度学习[26]等算法的路径规划)不同的是, 该方法是将人思考问题的机制引入到路径规划中, 即根据当前状态(环境), 在可选方案中选择最佳方案行动, 行动作用于环境, 改变当前状态, 在未到达目的地之前, 按照长期知识继续行动。而传统研究方法的重点是设计模型及相关函数, 且需要进行多次训练, 而且当环境变化时, 需要对模型进行重新训练, 如果环境持续变化, 则不能采用传统算法, 而本文提出的方法能够弥补传统方法的缺点, 实现在环境持续变化情况下的路径规划。

4 结论

本文提出了一种基于认知结构 SOAR 的机器人路径规划方法。与传统方法相比, 该方法的优点在于能够让机器人模拟人的行为, 通过预先载入的认知模型(长期知识)进行路径探索, 在改进规则之后, 能够处理变化环境下的路径选择。从仿真实验结果可以看出, 由于没有任何先验信息, 机器人无法找到最优路径。但是, 随着地图信息的逐渐完备, 可以先采用传统的全局规划方法, 在找到全局最优路径, 并将其

离散化为多个航路点的基础上,结合本文提出的方法,实现机器人向下一航路点移动同时避开动态的障碍物。同时,可以加入 chunking 功能,设计强化学习机制,提高机器人探索环境的速度,这也是下一步工作的重点。

参考文献:

- [1] LIANG XD, LI LY, WU JG, et al. Mobile robot path planning based on adaptive bacterial foraging algorithm[J]. Journal of Central South University, 2013, 20(12): 3391-3400.
- [2] ZHANG X. Path planning for unmanned aerial vehicles in surveillance tasks under wind fields[J]. Journal of Central South University, 2014, 21(8): 3079-3091.
- [3] YAN X F, GU F, SONG C, et al. Dynamic formation control for autonomous underwater vehicles[J]. Journal of Central South University, 2014, 21(1): 113-123.
- [4] OH C, CHOI B S, KIM J J, et al. Modified A* algorithm for outdoor environments with risk and velocity map[C]//IEEE International Symposium on Industrial Electronics. Seoul, South Korea: IEEE, 2009: 1587-1591.
- [5] SAUL SIMHON G D. A global topological map formed by local metric maps[C]//IEEE/RSJ International Conference on Intelligent Robots & Systems. Victoria, BC, Canada: IEEE, 1998: 1708-1714.
- [6] YAHJA A, SINGH S, STENTZ A. An efficient on-line path planner for mobile robots operating in vast environments[J]. Robotics & Autonomous Systems, 2000, 32(2): 129-143.
- [7] TOUCHTON R, KENT D, GALLUZZO T, et al. Planning and modeling extensions to the joint architecture for unmanned systems for application to unmanned ground vehicles[C]//Spie Defense & Security Symposium. 2005: 138-144.
- [8] TROY D K, LYLE N L. Deep blue cannot play checkers: the need for generalized intelligence for mobile robots[J]. Journal of Robotics, 2010, 2(1): 1-8.
- [9] NEWELL A. Soar: a cognitive architecture in perspective [M]. Cambridge, Mass, USA: Harvard University Press, 1990: 8-25.
- [10] HUNT R B E. GPS: a case study in generality and problem solving by George W Ernst[J]. The American Mathematical Monthly, 1971, 78(8): 923-924.
- [11] LANGLEY P, LAIRD J E, ROGERS S. Cognitive architectures: research issues and challenges[J]. Cognitive Systems Research, 2009; 10(2), 141-160.
- [12] ANDERSON J, LEBIERE C. The atomic components of thought[J]. Journal of Mathematical Psychology, 1995, 45(6): 917-923.
- [13] MEYER D E, KIERAS D E. A computational theory of executive cognitive processes and multiple-task performance: part 1[J]. Psychological Review, 1997, 104(1): 3-65.
- [14] SUN R. The CLARION cognitive architecture: extending cognitive modeling to social simulation[M]. New York: Cambridge University Press, 2006: 120-122.
- [15] JILK D J, LEBIERE C, O'REILLY R C, et al. SAL: an explicitly pluralistic cognitive architecture[J]. Journal of Experimental & Theoretical Artificial Intelligence, 2008: 20(1), 197-218.
- [16] LANGLEY P, MCKUSICK K, ALLEN J. A design for the icarus architecture[J]. ACM SIGART Bulletin, 1991, 2(4): 104-109.
- [17] NEWELL A. Reasoning, problem solving, and decision processes: the problem space as a fundamental category [C]//International Conference on Computer Graphics. ACM, 1980: 25-30.
- [18] CHIEN W C. SOAR CGFS that learn inductively: a hybrid autonomous approach based on a modified naïve bayes learning algorithm[D]. Orlando, Florida: College of Engineering and Computer Science, University of Central Florida, 2003: 23-25.
- [19] PAOLO G, HAYDN T, TONY D. Autonomous mission management for UAVs using soar intelligent agents[J]. International Journal of Systems Science, 2013, 44(5): 831-852.
- [20] LAIRD J, ROSENBLOOM P, NEWELL A. Universal subgoaling and chunking: the automatic generation and learning of goal hierarchies[M]. New York: Kluwer Academic Publishers, 1986: 26-30.
- [21] SUTTON R, BARTO A. Reinforcement learning: an introduction[M]. Cambridge: MIT Press, 1998: 30-35.
- [22] RASMUSSEN J. Skills, rules and knowledge, signals, signs and symbols, and other distinctions in human performance models[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1983, SMC-13(3): 257-266.
- [23] LIAO WEIQIANG. Genetic algorithm based robot path planning[C]//Proceedings of 2008 International Confer-

- ence on Intelligent Computation Technology and Automation. Changsha, 2008: 56-59.
- [24] 侯翔. 基于人工势场和量子遗传算法的移动机器人路径规划方法[J]. 计算机应用与软件, 2018, 35(6): 263-266, 333.
HOU XIANG. Mobile robot path planning method based on artificial potential field and quantum genetic algorithm [J]. Computer Applications and Software, 2018, 35(6): 263-266, 333. (in Chinese)
- [25] 王辉, 王景良, 朱龙彪, 等. 基于改进蚁群算法的泊车系统路径规划[J]. 控制工程, 2018, 25(2): 253-258.
WANG HUI, WANG JINGLIANG, ZHU LONGBIAO, et al. Path planning system based on improved ant colony algorithm[J]. Control Engineering, 2018, 25(2): 253-258. (in Chinese)
- [26] 马琼雄, 余润笙, 石振宇, 等. 基于深度强化学习的水下机器人最优轨迹控制[J]. 华南师范大学学报: 自然科学版, 2018, 50(1): 128-132.
MA QIONGXIONG, YU RUNSHENG, SHI ZHENYU, et al. Autonomous underwater vehicles optimal trajectory control based on deep reinforcement learning[J]. Journal of South China University: Natural Science Edition, 2018, 50(1): 128-132. (in Chinese)

Path Planning of Robot Based on Cognitive Structure SOAR

MU Yang

(Naval Aviation University, Huludao Liaoning 125000, China)

Abstract: In a situation that environment was unknown and target position was already known, a search method based on cognitive structure SOAR was proposed. The theory and problem solving mechanism of SOAR were discussed with the box moving problem, behavior model of human and decision process of SOAR were analyzed, and behavior modeling method was designed based on SOAR-Agent. By applying the proposed method, robot could simulate the searching pattern of human in a unknown and dynamic environment, the effects of completeness of long term memory on the task of navigating a robot to the destination were analyzed through simulating experiment, and the effectiveness of this method was also verified.

Key words: cognitive structure; SOAR; robot; unknown environment

(上接第265页)

A New Method of Channel Partition on Parallel Detection

ZHANG Haibo¹, TIAN Wenbiao¹, ZANG Zhihe², LIU Youming³, LI Tingjun¹

(1. Naval Aviation University, Yantai Shandong 264001, China;

2. Yantai Huale Real Estate Co., Ltd., Yantai Shandong 264001, China;

3. The 92863rd Unit of PLA, Qingdao Shandong 266001, China)

Abstract: In MIMO system with N_t transmitting antennas and N_r receiving antennas, the parallel detection algorithms can achieve optimal detection performance. However, the channel partition method used in the prior art is to calculate the N_t times of the pseudo-inverse of the channel matrix to complete the channel partition. With the increasing number of antennas, the complexity of the pseudo-inverse of the channel matrix will increase rapidly, which further aggravates the computational complexity of parallel detection algorithms. For this problem, in this paper it's proved that in parallel detection algorithms, only one calculation of the pseudo-inverse of the channel matrix was needed. To complete channel partition. Simulation results show that the performance of the lattice reduction parallel detection algorithm applying the new channel partition method is unchanged.

Key words: MIMO system; parallel detection; channel partition; lattice reduction