arXiv:1905.00532v1 [cs.RO] 1 May 2019

一种基于高效可达性框架的、可证明安全的未知环境自主导航方法

Andrea Bajcsy*、Somil Bansal*、Eli Bronstein、Varun Tolani、Claire J. Tomlin

摘要——现实世界中的自动驾驶车辆通常运行在先验未知的环境中。由于其中大多数系统都属于安全关键系统，因此，确保它们在面对环境不确定性（例如未观测障碍物）时安全运行至关重要。当前的安全分析工具能够在先验掌握环境状态全部信息的情况下，使自主系统对安全性进行推理。然而，当环境通过实时传感进行感知时，例如执行导航任务期间，这些工具难以很好地扩展到相应场景。在本文中，我们提出了一种基于 Hamilton-Jacobi 可达性的全新实时安全分析方法，即使存在环境不确定性，也能提供强有力的安全保证。我们的方法与规划器无关，并且能够针对各种建图传感器提供保证。我们在仿真和硬件平台上验证了该方法，为一种当前先进的、基于视觉的基于学习的规划器提供安全保证。有关我们方法和实验的视频可在项目网站上获取[1]。

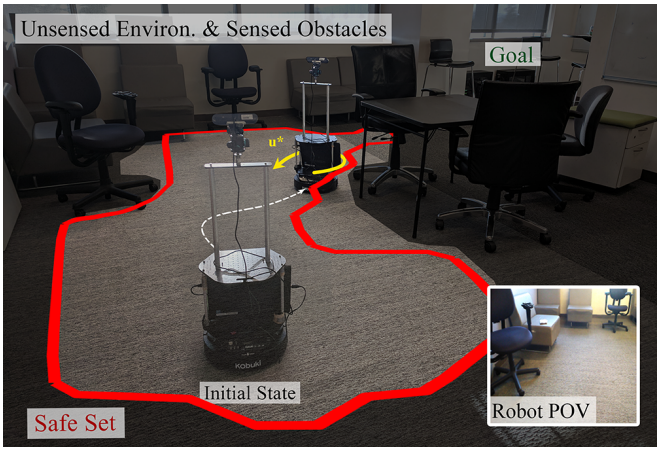


图 1. **概述：**我们研究在先验未知环境中，从初始状态安全导航至目标状态的问题。我们将未感知的环境视为障碍物，并采用 HJ 可达性框架为车辆计算安全控制器；随着车辆探索环境，该控制器会实时更新。我们展示了将该方法应用于搭载基于视觉的规划器的 Turtlebot 的实例。当机器人面临碰撞风险时，安全控制器（u^*）可确保系统安全。

I. 引言

在现实世界中运行的自动驾驶车辆必须使用车载的有限范围传感器，在先验未知环境中进行导航。随着车辆朝着目标前进并接收到有关环境的新传感器信息，严谨的安全性分析对于确保系统行为不会导致危险碰撞至关重要。为了为真实车辆提供此类安全保证，任何分析都应考虑多种不确定性来源，例如建模误差、外部扰动以及环境中未知的部分。

人们提出了各种机制，以确保对建模误差和外部扰动具有鲁棒性 [24][16], [34]。此外，针对使用有限范围传感器在未知环境中运行的系统，其安全保证也已成为近期研究的主题 [21], [22], [32], [19]。尽管这些研究取得了一些有趣的成果，但其安全保证依赖于对传感器和/或规划器施加特定假设，而这些假设对于众多现实世界中的自动驾驶系统以及用于导航的传感器而言相当严格。相比之下，我们并不提出一种能够保证安全的特定规划和感知范式，而是旨在设计一种能够兼容广泛传感器、规划器和动力学的安全框架。

构建此类框架面临两项主要挑战。第一项挑战是，在尽量不干扰以目标为导向的行为的同时，确保车辆在面对环境未知部分和外部扰动时的安全性。第二，随着获取新的环境信息，需要提供实时安全保障，这就要求采用计算效率高且不过度保守的近似方法。此外，这种安全性分析还应适用于各种真实世界中的传感器、规划器和车辆。

本文提出了一种安全框架，可在假设传感器在其量程内正常工作的前提下，帮助在先验未知的静态环境中运行的自动驾驶车辆克服这些挑战。传感器存在错误和噪声会使安全性分析变得更加复杂，这一问题将留待未来研究。我们的框架基于 Hamilton-Jacobi（HJ）可达性分析[25], [27], ，这是一种用于验证具有一般非线性动力学和扰动的动力系统的安全性与性能的方法。该方法不仅能够提供动力系统始终满足状态约束（即保持无碰撞）的状态集合，还能提供一个保证系统永不违反状态约束的最优控制器。具体而言，我们将在任意时刻的未知环境视为障碍物，并利用 HJ 可达性计算后向可达集（*BRS*），即车辆即使采取最佳控制仍可能进入环境中未知且潜在不安全部分的状态集合。BRS 的补集

*贡献相同。所有作者均就职于加州大学伯克利分校电气工程与计算机科学系。[1]项目网站：https://abajcsy.github.io/safe navigation/

因此表示车辆的安全集。通过这一计算，我们还获得了相应的限制最少的安全控制器；只有当车辆安全面临风险时，该控制器才会干预规划器。因此，在我们的框架中使用 HJ可达性分析使我们能够克服第一个挑战——我们的框架可应用于一般的非线性车辆、传感器和规划器。

一般而言，由于 HJ 计算的计算成本高昂，这种方法尚未用于环境事先未知、而是在运行时通过传感器感知的场景。为克服这一挑战，我们提出了一种用于计算后向可达集（BRS）的新型实时算法。该算法根据新获取的环境信息，仅在局部更新后向可达集，从而显著减轻 HJ 可达性的计算负担，同时始终维持安全保证。总而言之，我们的主要贡献如下：

1）一种可证明安全的静态未知环境导航框架，适用于广泛类型的传感器、规划器和动力学；2）一种在机器人导航过程中根据新的传感器测量在线更新安全集的算法；3）在存在外部扰动的情况下，展示我们的方法在具有非线性动力学的车辆上与不同传感器和规划器结合使用的效果；4）通过硬件演示展示我们的方法如何为一种最先进的基于学习的规划器提供安全保障，该规划器仅使用单目 RGB 图像进行规划。

II. 相关工作

大量研究探讨了未知环境中机器人的运动规划和安全探索，其中一些研究重点关注尽管存在建模误差和外部扰动，仍能提供安全保证。我们无法在此总结所有相关工作，但将尝试讨论几种最密切相关的方法。

A. 安全运动规划

尽管存在建模误差和扰动仍能确保安全性的方法，很大程度上源于实时规划中安全性与效率之间的权衡。一种常用做法是进行离线 计算，以量化扰动和建模误差，并在在线阶段利用这些结果确定无碰撞轨迹 [24], [16], [34]。或者， [2], [3] 使用控制障碍函数来设计可证明稳定的控制器，同时满足给定的状态空间约束。然而，这些方法假设即使在未知环境中也能获得递归可行的无碰撞路径，而这在真实环境中未必可能。若干工作在模型预测控制框架 [29], [31], 内针对单智能体场景解决了这一问题，也有工作针对多辆车采用序贯轨迹规划 [33], [5]。然而，这些工作假设对所有障碍物具有先验 知识，而我们的框架则针对可能的高阶非线性动力学，保证在先验 未知环境中的安全性。

同时针对建模误差和有限感知视界确保安全性的问题，已借助平方和 [21]、线性时序逻辑 [22], 、反应式综合方法 [32]、基于图的运动动力学规划器 [19] 等方法得到研究。这些工作通常对传感器或规划器施加限制，以确保相对于未知环境的安全性。相比之下，所提出的框架与传感器和规划器无关，前提是传感器能够准确识别其感知区域内的障碍物。

B. 安全探索

在机器人探索文献中，研究者也针对采用简化运动学运动模型的机器人，使用前沿探索方法 [36] 和 D* [20], 研究了在未知环境中寻找通往指定目标的可行轨迹这一问题。其他研究还包括针对无漂移动力学的基于采样的运动规划器 [8], 以及针对具有有限停止时间的车辆的动态探索方法 [17]。机器人探索也在完全可观测和部分可观测的马尔可夫决策过程 [30]、[28] 以及强化学习 [1], [18]的背景下得到研究，以降低碰撞概率；然而，这些研究通常并未提供理论上的安全保证。

安全探索也曾根据李雅普诺夫稳定性 [9], [11]进行研究。尽管稳定性通常是理想的，但它不足以保证碰撞规避。相比之下，我们的形式化方法采用了更强的安全性定义，并且更接近于 [15], [13], ，后者使用可达集来刻画安全性。

III. 问题陈述

在这项工作中，我们研究在先验未知的静态环境中的自主导航问题。考虑车辆的一个稳定、确定性、非线性动力学模型

$$\dot{x} = f(x, u, d), \tag{1}$$

其中， $x \in \mathbb{R}^n$、 $u \in \mathcal{U}$和 $d \in \mathcal{D}$ 分别表示车辆的状态、控制和所受扰动。这里，d 可以包含外部扰动或动力学不匹配的影响。为方便起见，我们将状态x 划分为位置分量 $p \in \mathbb{R}^{n_p}$ 和非位置分量 $h \in \mathbb{R}^{n-n_p}$： $x = (p, h)$。我们假设流场 f： $\mathbb{R}^n \times \mathcal{U} \times \mathcal{D} \to \mathbb{R}^n$ 关于时间一致连续，并且在固定 u 和 d时关于 x 利普希茨连续。在这一假设下，给定 $u(\cdot) \in \mathbb{U}$， $d(\cdot) \in \mathbb{D}$，存在唯一的轨迹满足式（1） [12]。我们还假设始终能够准确感知车辆状态 x 。

令 x_0 和 x^* 分别表示车辆的起点和目标状态。车辆旨在从 x_0 导航至 x^*，穿越先验未知环境 $\mathcal{E}$；机器人无法获取该环境的地图或拓扑结构。在任意时刻 t 和状态 $x(t)$下，车辆都有一个规划器 $\Pi(x(t), x^*, \mathcal{E})$，用于输出要在时刻$t$施加的控制命令 $u(t)$。车辆还配备了一个传感器，该传感器在任意给定时刻都能探测状态空间中的一个区域 $\mathcal{S}_t \subset \mathbb{R}^n$，并对 $\mathcal{S}_t$内的占据情况提供保守估计。例如，如果车辆配备了摄像头传感器， $\mathcal{S}_t$ 将是

一个三角形区域（在三维中为棱柱形），表示相机的视场。我们假设在此有限传感器范围内能够实现完美感知。处理错误感知、传感器噪声和动态环境本身就是各自独立的问题，我们将留待未来工作研究。最后，我们假设 x_0 周围存在一个已知的初始无障碍区域，由 $\mathcal{X}_{\text{init}} \subset \mathbb{R}^n$给出；例如，车辆从静止状态起步且其初始状态无碰撞时，就是这种情况。

给定 x_0、 x^*、 $\mathcal{X}_{\text{init}}$、规划器 Π 和传感器 $\mathcal{S}$，本文旨在设计一种最小限制控制机制，使车辆能够驶向目标状态，同时保持安全，即始终避开障碍物。由于环境 $\mathcal{E}$ 未知，因此必须在仅能通过传感器获得环境部分观测的情况下确保安全，而这通常颇具挑战。我们采用 HJ可达性框架，确保即使对环境仅有部分了解，也能保障安全性。

IV. 运行示例

为阐释我们的方法并帮助理解所提出框架背后的直觉，我们引入一个简单的运行示例：一个速度受到扰动的三维 Dubins汽车系统。该系统的动力学如下：

$$\dot{p}_x = v\cos\phi + d_x, \quad \dot{p}_y = v\sin\phi + d_y, \quad \dot{\phi} = \omega, \\ \underline{v} \le v \le \bar{v}, \quad |\omega| \le \bar{\omega}, \quad |d_x|, |d_y| \le d_r \tag{2}$$

其中， $x := (p_x,\ p_y,\ \phi)$ 为状态， $p = (p_x,\ p_y)$ 为位置， ϕ 为航向， $d = (d_x,\ d_y)$ 为车辆所受的扰动。车辆的控制量为$u := (v,\ \omega)$，其中 v 为速度， ω 为转向速率。两种控制量均有上下界，本例中取值为 $\underline{v} = 0.1m/s$、 $\bar{v} = 1m/s$、 $\bar{\omega} = 1rad/s$。扰动界取为 $d_r = 0.1m/s$。

环境设置如图2所示。车辆的起点和目标状态分别为 $x_0 = [2, 2.5, \pi/2]$（黑色标记）和 $x^* = [8.5, 3, -\pi/2]$（绿色区域的中心）。目标是在 $0.3m$ 范围内到达 x^*（浅绿色区域）。然而，环境中存在一个车辆事先未知的障碍物（灰色标记）。在运行示例导航任务开始时，我们假设 x_0的 $1.5m$ 范围内没有障碍物，并得到初始无障碍区域 $\mathcal{X}_{\text{init}} := \{x : \|p - p_0\| \le 1.5\}$（虚线黑线以内的区域）。

为展示我们方法不依赖特定传感器的特性，我们使用两种不同的传感器对 Dubins汽车进行仿真：激光雷达和相机。对于激光雷达，感知区域 $\mathcal{S}_t$是以当前位置 $p(t)$为圆心、半径为 R 的圆，其中 $R = 3m$ （如图 2a 所示）。对于相机，感知区域 $\mathcal{S}_t$ 由一个三角形区域确定，其立体角为 F （也称为视场），顶点位于当前车辆航向处，最大范围为 R。在仿真中，我们采用 $F = \pi/3$ 和 $R = 20m$ （如图 2b 所示）。不过，部分 $\mathcal{S}_t$ 区域可能被障碍物遮挡，现实世界中的传感器也会出现这种情况。

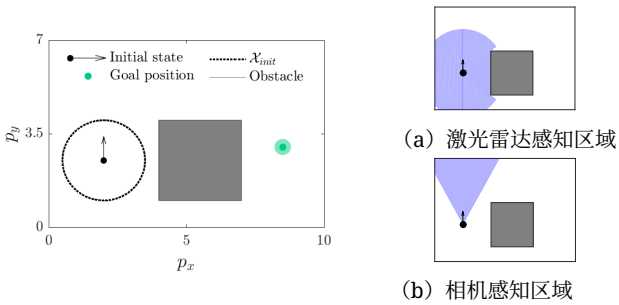


图2。运行示例的初始设置。目标是在存在未知障碍物（灰色方块）的情况下，从初始位置（黑色标记）安全到达目标（绿色区域的中心）。我们还展示了激光雷达和摄像头传感器的初始感知区域。

此外，对于每种传感器，我们使用两种不同的规划器 Π展示该方法：基于采样的规划器和基于模型的规划器。对于基于采样的规划器，我们采用快速扩展随机树（RRT）[23]，；对于基于模型的规划器，我们采用基于样条的规划器 [35]。我们的目标是在存在未知障碍物的情况下，仍能安全地导航至目标位置。

V. HAMILTON-JACOBI 可达性

我们的框架基于 Hamilton-Jacobi (HJ) 可达性分析 [25], [27]。该分析已成功应用于多个领域，例如飞机自动着陆和安全多车辆路径规划 [5], [6]。在本研究中，我们将使用可达性分析，在给定一组不安全状态 $\mathcal{L}$ 的情况下，计算后向可达集（BRS） $\mathcal{V}(\tau)$。直观而言， $\mathcal{V}(\tau)$是这样一组状态：从该集合出发的系统轨迹，在某种扰动作用下，即使控制已尽最大努力，仍能在 τ 的时间范围内进入 $\mathcal{L}$ 。相反，对于任何从 $\mathcal{V}^c(\tau)$出发的轨迹，都存在一种控制，使系统轨迹即使面对最坏情况扰动，也永远不会进入 $\mathcal{L}$。这里，$\mathcal{V}^c(\tau)$表示集合 $\mathcal{V}(\tau)$的补集。

BRS 的计算可以表述为控制与扰动之间的微分博弈，并可利用动态规划原理求解。该微分博弈对应的代价泛函为：

$$J(x, \tau, u(\cdot), d(\cdot)) = \inf_{s\in[\tau,0]} l(\xi^{u,d}_{x,\tau}(s)), \tag{3}$$

其中， $\xi^{u,\ d}_{x,\ \tau,}(s)$ 表示从时间 τ 的状态 x 出发，在扰动 $d(\cdot)$ 下施加控制 $u(\cdot)$ 后，系统在时间 s 的状态。在式 (3) 中，函数 $l(x)$ 是表示不安全集 $\mathcal{L} = \{x$ 的隐式曲面函数： $l(x) \le 0\}$。直观而言， J 用于记录系统轨迹在时间范围 $[\tau, 0]$ 内是否曾进入不安全集；如果进入了，则与该轨迹对应的代价为负。

与式 (3) 中代价泛函对应的值函数为：

$$V(\tau, x) = \min_{d(\cdot)\in\Gamma[u]} \max_{u(\cdot)} J(x, \tau, u(\cdot), d(\cdot)), \tag{4}$$

其中， Γ 表示非预见性策略 [27]的集合。如果给定状态的值函数为负，则从该状态开始

系统最终无法避免进入不安全集。因此，式（4）中的值函数跟踪系统的所有不安全轨迹，进而可用于计算系统的安全轨迹。有关此公式的更多详细信息，请参阅 [27], [6]。

式（4）中的值函数可以通过动态规划获得，从而得到哈密顿–雅可比–艾萨克斯变分不等式（HJI-VI） [14], [25]。最终，可以通过求解以下终值HJI-VI来计算后向可达集：

$$\begin{aligned}\min\{D_\tau V(\tau,x)+H(\tau,x,\nabla V(\tau,x)),l(x)-V(\tau,x)\}=0\\ V(0,x)=l(x),\quad \tau\le 0.\end{aligned}\tag{5}$$

这里， $D_\tau V(\tau,x)$和 $\nabla V(\tau,x)$分别表示值函数 $V(\tau,x)$的时间导数和空间导数。哈密顿量 $H(\tau,x,\nabla V(\tau,x))$编码了系统动力学、控制和扰动的作用，其定义为

$$H(\tau,x,\nabla V(\tau,x))=\max_{u\in\mathcal{U}_i}\min_{d\in\mathcal{D}}\nabla V(\tau,x)\cdot f(x,u,d).\tag{6}$$

一旦计算出值函数 $V(\tau,x)$，后向可达集以及由此得到的安全状态集合为

$$\mathcal{V}(\tau)=\{x:V(\tau,x)\le 0\},\tag{7}$$

$$\mathcal{W}(\tau)=\mathcal{V}^c(\tau)=\{x:V(\tau,x)>0\}.\tag{8}$$

HJI可达性还提供了使系统保持在安全集合中的最优控制，其表示为

$$u^*(\tau,x)=\arg\max_{u\in\mathcal{U}}\min_{d\in\mathcal{D}}\nabla V(\tau,x)\cdot f(x,u,d).\tag{9}$$

事实上，只要系统不处于不安全区域的边界上，就可以安全地施加任意控制。如果系统到达 $\mathcal{V}(\tau)$的边界，式（9）中的控制会将系统引导离开不安全状态。HJI可达性提供的这一最小限制控制器，也是我们框架以最小限制方式确保任意规划器安全性的基础。

VI. 我们的方法

我们提出一种基于HJ可达性的框架，以确保在先验未知环境中的安全性。我们的框架还采用一种新颖的实时计算方法，根据车辆行驶过程中对环境的新观测，计算安全集的保守近似。我们首先介绍该框架，然后给出一种实时更新安全集的算法。

A. 确保未知环境中的安全性

我们的框架将任意时刻尚未感知的环境视为障碍物。利用未感知环境和已感知障碍物，通过HJI可达性为车辆计算安全区域。即使存在最坏情况扰动，这也能确保车辆不会进入环境中未知且可能不安全的区域。

令 $\mathcal{F}_t$ 表示环境在任意时刻 t感知到的无障碍区域。给定初始无障碍区域 $\mathcal{F}_0=\mathcal{X}_{\text{init}}$，我们通过计算相应的安全集 $\mathcal{W}_0$ 来求解式 (5) 中的 HJI-VI，并假设 $\mathcal{F}_0$ 之外的所有区域均为障碍物。对于此计算，式 (5) 中的初始值函数定义为 $V_0(0,\ x):=l_0(x)$，其中 $l_0(x)$ 在 $\mathcal{F}_0$ 内部为正，在外部为负。一个这样的函数是到 $\mathcal{F}_0^c$的有符号距离。从 $l_0(x)$ 开始，求解 HJI-VI，得到值函数 $V_0(x):=\lim_{\tau\to-\infty}V_0(\tau,\ x)$。其中， τ是式 (5) 中的虚拟计算变量。随后使用 $V_0(x)$ 计算安全区域 $\mathcal{W}_0$ （见式 (8)）。只要车辆处于 $\mathcal{W}_0$内部，就存在一个控制器，能够确保车辆不会与已知或未知障碍物发生碰撞。

接下来，我们根据以下控制律，在系统上执行控制器，控制时域为 $t\in[0,T]$：

$$u(t)=\begin{cases}\Pi(x(t),x^*,\mathcal{E}), & \text{if } x(t)\in\mathcal{W}_t\\ u^*(t,x(t)), & \text{otherwise}\end{cases}\tag{10}$$

其中， $u^*(t,\ x(t))$ 是与 $\mathcal{W}_t$ 对应的最优安全控制器，其定义见式 (9)。此外，在安全集更新之前，我们使用最近一次计算得到的安全集来寻找最优安全控制器，即 $\mathcal{W}_t=\mathcal{W}_0\ \forall t\in[0,T]$。式 (10) 中的控制机制限制最少，因为它允许规划器在系统上执行期望控制，除非系统面临违反安全性的风险。请注意，系统设计者可以任意选择我们框架中的控制时域 T，同时仍能确保安全性。

系统执行式 (10) 中的控制律时，将在每个时刻 t 获取新的传感器测量值 $\mathcal{S}_t$ ，并据此得到 $\mathcal{F}_t$，即该时刻传感器探测到的自由空间。如果传感器在任一时刻被障碍物完全遮挡，则对应的自由空间为空集。因此，时刻 t 的整体已知自由空间为：

$$\mathcal{M}_t=\bigcup_{s\in[0,t]}\mathcal{F}_\tau.\tag{11}$$

在控制时间范围结束时，我们假设 $\mathcal{M}_T$ 之外的一切都是障碍物，并计算另一个安全区域 $\mathcal{W}_T$ 。该安全区域通过求解 HJI-VI 直至收敛而得到。随后，我们执行类似于式 (10) 的控制律，不同之处在于，只有当系统位于 $\mathcal{W}_T$ 的边界时，安全控制器才会介入。重复整个过程，直至系统到达目标状态。

由于安全控制器不允许系统轨迹离开已知自由空间，因此所提出的框架能够保证始终避免碰撞。然而，安全集可能相当保守，尤其是在早期环境仍大多未被探索时；这是我们为确保能够抵御所有意外障碍物而作出的权衡。如果已知环境中障碍物的其他信息，也可以将其纳入考虑，从而只会降低安全集的保守性。

请注意，安全集不一定需要每隔 T 秒更新一次。它可以比控制时域更新得更快、更慢，也可以以相同的频率更新。从本质上说，可以在式（10）的控制律中使用最新的安全集，同时仍能确保始终满足安全性。这是因为任意时刻 t_1 的安全集都只会小于时刻 t_2的安全集

当 $t_1 < t_2$时。不过，应尽可能快速地更新安全集，以尽量减少对规划器的干扰。

B. 后向可达集的高效更新

我们的框架要求在车辆穿行环境时实时计算安全集。一般而言，由于 HJI可达性的计算复杂度会随状态维度呈指数级增长，因此这项任务颇具挑战性 [6]。为缓解部分计算难题，我们提出了两种计算后向可达集的新方法：热启动和局部值函数更新。

1) 热启动方法：在任意给定时刻，车辆只能感知环境中一小部分新增区域。因此，自由空间地图 $\mathcal{M}$ 只会在较短的时间范围内发生少量变化。直观而言，这应该只会导致安全区域发生微小变化。我们利用这一直觉，提出一种更新可达集的新方法，使其速度更快。为简洁起见，我们假设安全集每隔 T 秒更新一次来介绍该方法，但当安全集以非固定频率更新时，结论同样成立。

给定时间 t_{last} 计算得到的最后一个安全区域，以及 t_{last} 处的地图和当前时间 t，我们将式（5）中的值函数“热启动”，以便在时间 t 进行后向可达集计算，具体如下：

$$V_t(0,x) = \begin{cases} l_t(x), & \text{if } x \in \mathcal{M}_t \cap \mathcal{M}^c_{t_{\text{last}}} \\ V_{t_{\text{last}}}(x), & \text{otherwise} \end{cases} \tag{12}$$

其中，$l_t(x)$ 的定义与之前相同，即在 $\mathcal{M}_t$ 内部为正，在外部为负。直观而言，与其在整个状态空间中都使用 $l_t(x)$ 初始化值函数，式（12）不如在自上次计算以来未获得新信息的状态处使用上一次计算得到的值函数进行初始化，而仅在先前被假定为已占用、但实际上无障碍物的状态处使用 $l_t(x)$ 进行初始化。这样可以大幅加快后向可达集的计算，因为只需更新数量少得多的、刚刚发现为空闲的状态处的值函数。在其他所有状态处，值函数已经几乎准确，只需进行小幅调整。有趣的是，该过程还保持了安全区域的保守性，这足以确保始终实现碰撞规避。

~引理*1*：设 $V_t(\tau,x)$ 为以下热启动 HJI-VI 的解：

$$\min\{D_\tau \tilde{V}_t(\tau,x) + H(\tau,x,\nabla\tilde{V}_t(\tau,x)), l_t(x) - \tilde{V}_t(\tau,x)\} = 0,$$

~其中 $V_t(0,x)$ 按 (12) 定义。设 $V_t(\tau,x)$ 为 (5) 中 HJI-VI 在 $V_t(0,x) = l_t(x)$ 下的解。则 ~$V_t(\tau,x) \le V_t(\tau,$~$x$ $\tau \le 0$ $\mathcal{V}_t$ $-\infty$ $\supseteq$) 对所有 。特别地，( ) $\mathcal{V}_t(-\infty)$ 且 $\mathcal{W}_t(-\infty) \subseteq \mathcal{W}_t(-\infty)$。

证明：由于 $V_{t_{\text{last}}}(x)$ 表示时刻 t_{last} 收敛后的值函数，因此有 $V_{t_{\text{last}}}(x) \le l_t(x)\ \forall x \in \mathbb{R}^n$. ~因此，根据 (12)， $V_t(0,x) \le l_t(x)\ \forall x \in \mathbb{R}^n$. 现在我们可以证明我们的结论了。

算法1： 利用 HJ 可达性的安全导航

x_0, x^*：起点和目标状态
$\mathcal{F}_0 \longleftarrow \mathcal{X}_{\text{init}}$：初始无障碍区域
$\mathcal{E}$：未知环境
$\Pi(\cdot,\ x^*,\ \mathcal{E})$：车辆的规划器
T：控制时域
$\mathcal{W}_0$：通过求解 HJI-VI 得到的初始安全区域 在 (5) 中
$\mathcal{W}_{\text{last}} \longleftarrow \mathcal{W}_0$； $\mathcal{V}_{\text{last}} \longleftarrow \mathcal{W}_0^c$； $t_{\text{last}} \longleftarrow 0$：最后计算出的安全集、后向可达集及对应时间
当车辆未到达目标**时**
  获取当前传感器观测 $\mathcal{S}_t$ 和自由空间 $\mathcal{F}_t$
  应用式 (10) 给出的限制最少的控制 $u(t)$
  每隔 T 秒**执行**
    获取当前环境地图 $\mathcal{M}_t$ 使用 (11)
    使用 $\mathcal{V}_{\text{last}}$热启动 BRS 计算，$\mathcal{M}$，以及 (12)
    通过求解来获得新的安全区域， $\mathcal{W}_t$，使用热启动后的值函数求解 HJI-VI
    $\mathcal{W}_{\text{last}} \longleftarrow \mathcal{W}_t$； $\mathcal{V}_{\text{last}} \longleftarrow \mathcal{W}_t^c$； $t_{\text{last}} \longleftarrow t$

对于 $\tau \le 0$，根据动态规划（详情见 [6] ），我们知道：

$$\begin{aligned} \tilde{V}_t(\tau,x(\tau)) &= \max_u \min_d \min\left\{ \inf_{s\in[\tau,0)} l_t(x(s)), \tilde{V}_t(0,x(0)) \right\}, \\ &\le \max_u \min_d \min\left\{ \inf_{s\in[\tau,0)} l_t(x(s)), l_t(x(0)) \right\}, \\ &= V_t(\tau,x(\tau)) \end{aligned}$$

其中，第二个不等式根据以下事实得出：~$V_t(0,\ x) \le l_t(x)\ \forall x \in \mathbb{R}^n$。因此，由 (7) 可知，对于所有 $\tau \le 0$，都有 ~$\mathcal{V}_t(\tau) \supseteq \mathcal{V}_t(\tau)$。特别地，这意味着 ~ ~$\mathcal{V}_t(-\infty) \supseteq \mathcal{V}_t($ $-\infty$)。最后，由于安全集是后向可达集的补集，(8) 意味着 $\mathcal{W}_t(-\infty) \subseteq \mathcal{W}_t(-\infty)$。

■

直观而言，引理1表明，使用热启动方法得到的安全集是通过求解完整 HJI-VI 得到的实际安全集的下近似。因此，该方法可用于确保车辆安全，同时具有较高的计算效率。实际上，对于本文中的传感器和导航问题，热启动所引入的保守性非常小，我们将在第七节中对此进行展示。我们采用热启动更新安全集的整体方法总结于算法1中。我们从初始已知自由空间 $\mathcal{X}_{\text{init}}$ 开始，并使用 HJI-VI 计算初始安全集 $\mathcal{W}_0$ （第 6 行）。此次计算的值函数以到 $\mathcal{X}_{\text{init}}$ 的有符号距离进行初始化。我们还维护最近一次计算得到的后向可达集 $\mathcal{V}_{\text{last}}$、安全集 $\mathcal{W}_{\text{last}}$ 及相应时间 t_{last} （第 7 行）。在每个状态下，车辆获取当前传感器观测，并提取感知到的自由空间（第 9 行和第 10 行）。接下来，向车辆施加控制指令（第 11 行）。如果车辆位于 $\mathcal{W}_{\text{last}}$ 内，则使用规划器获取控制指令；否则，

算法 2： 后向可达集的局部更新

1 $\mathcal{Q} \longleftarrow \mathcal{M}_t \cap \mathcal{M}^c_{t_{\text{last}}}$：初始化状态列表，以便对其值函数进行更新
2 $\mathcal{Q} \longleftarrow \mathcal{Q} \cup \mathcal{N}(\mathcal{Q})$：将邻近状态添加到 $\mathcal{Q}$
3 对 $\mathcal{Q}$ 中的状态进行热启动，$V_t(0,\ \mathcal{Q})$，使用 (12)
4 $V_{\text{old}} \longleftarrow V_t(0,\ \mathcal{Q})$：上一次计算得到的值函数，用于状态中 $\mathcal{Q}$
5 **当** $\mathcal{Q}$ 不为空 **时执行**
6 $V_{\text{updated}} \longleftarrow$ 更新值函数 V_{old}，针对一个时间步长 ΔT
7 $\Delta V = \|V_{\text{updated}} - V_{\text{old}}\|$：值函数的变化函数
8 $\mathcal{Q}_{\text{remove}} \longleftarrow \{x \in \mathcal{Q}:\ \Delta V = 0\}$：具有未改变值的状态
9 $\mathcal{Q} \longleftarrow \mathcal{Q} - \mathcal{Q}_{\text{remove}}$：移除值未改变的状态值
10 $\mathcal{Q} \longleftarrow \mathcal{Q} \cup \mathcal{N}(\mathcal{Q})$：将邻近状态添加到 $\mathcal{Q}$
11 $V_{\text{old}} \longleftarrow V_{\text{updated}}$

应用安全控制器。每隔 T 秒，根据车辆迄今感知到的自由空间，使用 HJI-VI 更新安全集和控制器（第 15 行）。该计算的值函数使用 V_{last} 进行热启动，但自 t_{last} 以来被发现为无障碍的状态除外，如式（12）所述（第 14 行）。重复整个过程，直至车辆到达目标。

2)后向可达集的局部更新：在上一节中，我们讨论了对值函数计算进行热启动可能会使值函数更快收敛；然而，值函数仍然是在整个状态空间上计算的。本节将介绍一种更实用的算法，该算法利用热启动的优势，仅在自上次值函数计算以来获得新信息的状态处，对值函数进行局部计算和更新。我们的安全框架与算法1中所述的相同——只是修改了安全集计算的计算过程（算法1中的第15行），以便对值函数进行局部更新。我们在算法2中概述这一过程。

在算法 2 中，我们维护一个状态列表 $\mathcal{Q}$，其中的值函数需要根据新的环境观测进行更新。$\mathcal{Q}$ 初始化为自 t_{last}以来新发现的自由状态集合，即$\mathcal{Q} = \mathcal{M}_t \cap \mathcal{M}^c_{t_{\text{last}}}$（第 1 行）。由于 $\mathcal{Q}$ 中的状态值相较于 $V_{t_{\text{last}}}(x)$ 发生变化，也会导致邻近状态的值发生变化，因此我们也将这些状态加入 $\mathcal{Q}$（第 2 行）。因此，$\mathcal{Q} = \mathcal{Q} \cup \mathcal{N}(\mathcal{Q})$。通常，HJI-VI 中的值函数通过将状态空间离散化为网格，并在该网格上求解 VI 来计算。在这种情况下，值函数的空间导数（用于计算式（5）中 HJI-VI 的哈密顿量）通过数值方法，利用邻近网格点计算得出。正是这个空间导数使一个状态上的值函数变化传播到其邻近状态。在这种情况下，$\mathcal{N}(\mathcal{Q})$ 可能表示用于计算 $\mathcal{Q}$ 中各状态的值函数空间导数的邻近网格点；不过，也可以采用其他邻近判定标准。

将邻居添加到 $\mathcal{Q}$后，按照式（12）初始化 $\mathcal{Q}$ 中所有状态的值（第 3 行），并在一定时间步长 ΔT 内使用式（5）中的 HJI-VI 更新这些状态的值（第 6 行）。由于通常只需对少得多的状态执行此计算，因此它比经典 HJI-VI计算快得多。接下来，我们从 $\mathcal{Q}$ 中移除所有在此 ΔT 期间值函数变化不显著的状态（第 8 行和第 9 行），因为这些状态不会导致其他状态的值函数发生变化。随后，将剩余状态的邻居添加到 $\mathcal{Q}$（第 10 行），并重复整个过程，直到所有状态的值函数都收敛。请注意，算法 2 仍保持了安全集的保守性，因为它只是计算热启动值函数的一种不同计算过程，而该值函数仍用于算法 1 的安全框架中。

VII. 仿真

A. 运行示例再探

我们现在回到运行示例，并通过仿真展示所提出的方法（详见第四节）。我们采用三种不同的方法来更新后向可达集，从而实现安全框架：使用完整的 HJI-VI、热启动方法（第六节 B.1）以及局部更新方法（第六节 B.2）。图3展示了这三种方法在不同规划器和传感器组合下所对应的系统轨迹。对于所有规划器和传感器组合，尽管存在外部扰动，且对障碍物没有先验 知识，所提出的框架仍能安全地将车辆导航至其目标位置（没有轨迹穿过障碍物）。车辆在环境中行驶时，规划器会在若干状态下做出可能导致碰撞的乐观决策；然而，安全控制器会介入以确保安全性。应用安全控制器的状态以红色标出。请注意，与激光雷达相比，使用相机传感器时安全控制器介入得更频繁。这是因为相机的视场角通常比激光雷达窄得多（激光雷达可在一定范围内感知所有方向上的障碍物）。由于视场角有限，安全控制器需要考虑范围大得多的未探索环境，进而采取更谨慎的控制。

我们比较了三种方法分别为相机和激光雷达传感器计算后向可达集所需的计算时间，结果见表I。所有计算均在配备 Core i7 5820K 处理器的台式计算机上，使用水平集工具箱 [26] 的 MATLAB 实现完成。正如预期，在所有场景中，对 BRS 计算的值函数进行热启动，相比完整的 HJI-VI 能显著缩短计算时间；然而，该计算时间对大多数实际应用而言可能仍不实用。仅在热启动之外再局部更新值函数，才能使计算时间进一步显著改善，而后向可达集

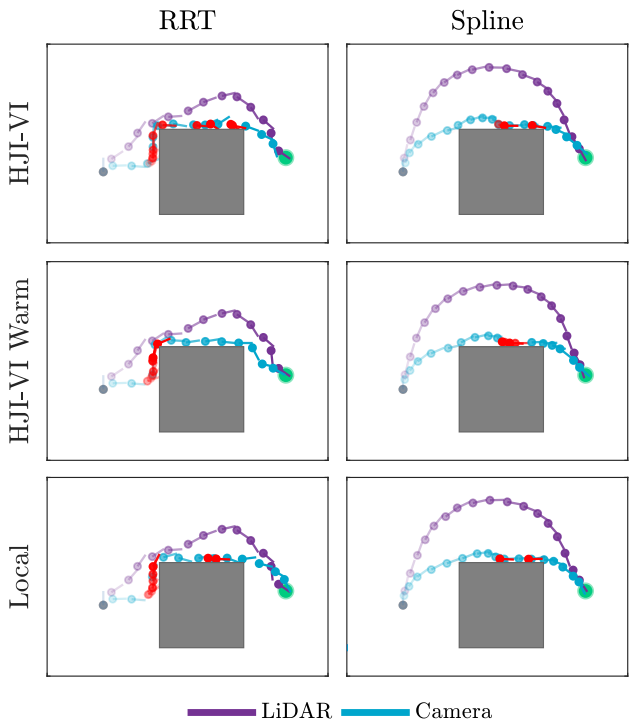


图3. 图2所示问题设定下，两种规划器（RRT和样条规划器）及两种传感器（激光雷达和摄像头传感器）对应的车辆轨迹；安全控制器由三种候选安全方法分别计算得出。在所有情况下，所提出的框架都能安全地将车辆导航至目标点。当规划器做出不安全决策时，安全控制器会介入（以红色标记的状态），以确保安全性。

在所有传感器和规划器上平均约在 1 秒内完成更新。考虑到该计算是在 MATLAB 中完成的，且未使用任何已知可将计算时间降低 100 [10] 倍的并行化，这一改进令人印象深刻。

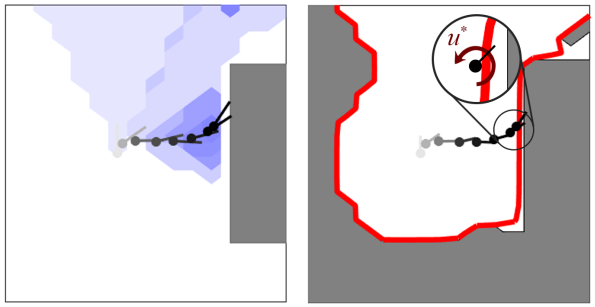
引理1表明，通过热启动值函数得到的安全集相比通过完整HJI-VI得到的安全集更为保守。因此，我们还比较了安全集较为保守的状态所占的体积百分比。这部分过度保守的体积通常仅为0.5%，表明热启动方法能够很好地近似真实值函数。

最后，我们更仔细地考察当系统使用距离受限传感器运行时，安全控制如何发挥作用。图4a展示了一辆配备摄像头传感器和RRT规划器的Dubins汽车，其中当前机器人状态以黑色显示，相应的感知区域为深蓝色，轨迹及相应的感知区域分别以灰色和浅蓝色显示。由于在当前状态下相机的视场被障碍物遮挡，它无法感知障碍物以外的环境。图4b展示了相应的当前环境信念地图，即车辆迄今所感知自由空间的并集（以白色显示）。由于当前感知区域包含在前一状态的感知区域内，因此未获得新的环境信息，后向可达集也未更新。当前车辆航向下安全集的切片如图4b所示（

模拟摄像头结果				
指标	规划器	HJI-VI	Warm	局部
平均计算时间（秒）	RRT	45.688	26.290	0.596
	样条	51.723	12.489	0.898
% 过度保守状态	RRT	0.0	1.112	0.517
	样条	0.0	0.474	0.506

模拟激光雷达结果				
指标	规划器	HJI-VI	Warm	局部
平均计算时间（秒）	RRT	21.145	6.075	1.108
	样条	25.318	3.789	1.158
% 过度保守状态	RRT	0.0	0.032	0.290
	样条	0.0	0.024	0.240

表I. 不同 BRS 更新方法下，各规划器和传感器对应的过度保守状态的平均计算时间和相对体积的数值比较。局部更新可在 ≈ 1 秒内计算出几乎精确的 BRS，速度显著快于 HJI-VI 和热启动。



(a) (b)

图4. (a) 摄像头传感器在不同时刻车辆处于不同状态时的感知区域。(b) 车辆感知到的整体自由空间及对应的安全集（红色边界内部）。由于车辆位于安全集的边界上，因此应用安全控制器 u^*，使机器人驶入安全集并确保碰撞规避。

红色边界内的区域）。由于车辆位于安全集的边界，安全控制器介入并施加控制 u^*，使系统朝向安全集内部（红色箭头），以确保碰撞规避。

B. 基于学习的规划器的安全性

由于所提出的安全框架与规划器无关，我们可以用它来确保即使存在基于学习的规划器时也能实现安全导航。具体而言，我们使用 [7],中提出的基于视觉的规划器，该规划器以RGB 相机图像和目标位置作为输入，并使用基于卷积神经网络的感知模块来生成期望的下一状态，使机器人朝其目标移动，同时尽量避开途中的障碍物。该期望的下一状态由基于模型的低层规划器用于生成从车辆当前状态到期望状态的平滑轨迹。作者表明，所提出的规划器能够利用机器人的先验经验，在新颖的室内杂乱环境中高效导航；然而，在若干真实场景中它仍会导致碰撞，例如当车辆需要穿过狭窄空间时。我们使用所提出的安全框架来确保既安全又

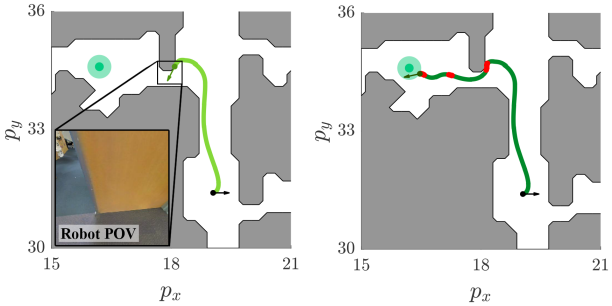


图 5. 所提出的框架可用于为将学习纳入闭环的基于视觉的规划器提供安全保证。基于视觉的规划器规划了一条穿过门口的路径。不使用安全控制时（a），这会导致碰撞；而采用安全控制时（b），机器人能够避免碰撞并到达目标。

在此类复杂导航场景中实现高效规划。

仿真的任务设置如图 5a 所示。机器人需要穿过一条非常狭窄的走廊，再经过一道门进入房间，以到达其目标（绿色圆圈）；机器人从初始状态（黑色箭头）出发。起初，机器人对障碍物（深灰色区域所示）一无所知。我们使用 S3DIS 数据集来仿真这一场景，该数据集包含斯坦福多栋建筑的网格扫描 [4]。通过在任意状态下渲染该网格，我们可以获得相机观测到的图像（供规划器使用），以及机器人视场内的占据信息（供安全计算使用）。对于机器人动力学，我们采用四维 Dubins 小汽车模型：

$$\dot{p}_x = v\cos\phi, \quad \dot{p}_y = v\sin\phi, \quad \dot{v} = a, \quad \dot{\phi} = \omega \tag{13}$$

其中，$p = (p_x,\ p_y)$ 表示位置，ϕ 表示航向，v表示车辆的速度。控制量为 $u := (a,\ \omega)$，其中$|a| \le 0.4$ 表示加速度，$|\omega| \le 1.1$ 表示转向速率。

在没有安全模块的情况下，基于学习的规划器所采用的轨迹如图 5a 所示。尽管车辆能够通过狭窄的走廊，但最终还是与门发生了碰撞。规划器与所提出的安全框架结合时，车辆所采用的轨迹如图 5b 所示。当规划器在门附近采取不安全动作时，安全控制器会介入（以红色标记），引导机器人安全地通过门口。我们还展示了机器人在门口附近观察到的图像，如图 5a 所示。尽管门挡住了机器人的大部分视野，规划器仍做出了相当乐观的决策，选择向前移动，结果导致碰撞。相比之下，安全控制器采取了更保守的决策：原地旋转，以便在向前移动前进一步探索环境，最终通过门口到达目标。我们的框架不依赖于特定规划器，因此也能为基于学习的规划器提供安全保证。

VIII. 实验

我们使用搭载立体 RGB 相机的 TurtleBot 2，在硬件上测试所提出的方法。对于车辆状态测量，我们使用 TurtleBot 上的机载里程计传感器。在实验中，车辆需要穿过未知且杂乱的室内环境，到达其目标位置（如图 1 所示）。对于 BRS 计算，我们使用式（13）中的动力学模型。我们使用开源的同步定位与地图构建（SLAM）算法和机载立体相机对环境进行预先建图。此预先建图步骤用于避免实时 SLAM 地图更新中的显著延迟和不准确性。然而，在部署期间，不会向机器人提供完整地图。相反，为了在任意给定时刻计算安全集，会将相机当前的视场投影到 SLAM 地图上，并仅使用视场内的信息更新安全集。这模拟了 Turtlebot 相机传感器范围有限的情况。尽管如此，这也指向了一个重要且有趣的未来研究方向：即使存在传感器噪声，也要确保安全性。

对于规划，我们使用第 VII-B 节中描述的基于学习的规划器，该规划器利用当前 RGB 图像确定候选下一状态。图 6 展示了实验场景的俯视图。车辆的起始位置和航向用黑色标示，目标区域用绿色标示，障碍物（车辆事先未知）用灰色标示。我们分别在启用和未启用安全控制器的情况下进行了实验，并在图 6 中展示了相应的轨迹。未启用安全控制器时，基于学习的规划器难以在拐角附近完成急转弯，最终撞上障碍物（本例中为椅子）。作为参考，我们还展示了规划器在拐角附近接收到的 RGB 观测。尽管机器人离椅子非常近，规划器仍做出了继续向前移动这一不安全的决策。然而，当基于学习的规划器在所提出的安全框架中使用时，安全控制器能够应对这种不安全情况，并安全地引导车辆避开障碍物。我们展示了车辆处于障碍物边界时对应的安全集，以及使用安全控制器得到的相应车辆轨迹。之后，规划器接管控制，并引导车辆驶向目标。

IX. 实际考量

在真实系统上实现安全框架时，需要考虑许多实际因素。下面我们将讨论在仿真和硬件实验中使用所提出的框架时遇到的一些主要实际问题：

- 由于值函数是在离散化的状态空间上计算的，因此可能会产生一些数值误差。采用更精细的离散化以及更高阶的空间导数近似，通常有助于缓解这些问题；不过，计算时间也会相应增加。根据我们的经验，三阶或更高阶的近似方法效果相当不错。

- 由于现实世界障碍物的复杂几何形状，任意时刻的感知地图都可能呈现出高度不规则的形态。这类不规则地图会引发不规则

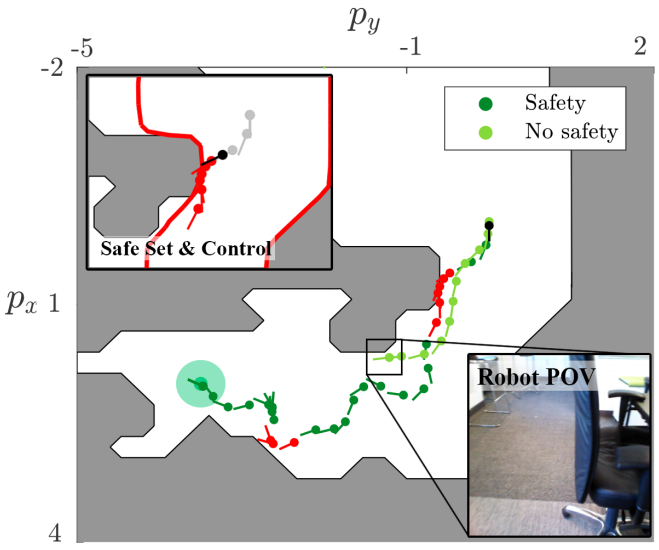


图 6。我们展示了实验场景的俯视图，以及采用和未采用所提出安全框架时相应的系统轨迹。未采用安全框架时，机器人会与椅子发生碰撞。相比之下，我们的安全框架能够在车辆过于接近障碍物时进行干预，从而安全地引导机器人到达其目标。

隐式曲面函数，这可能会严重阻碍值函数的计算。因此，通常可能需要将占据地图转换为规则且性质良好的函数，例如有符号距离地图，并使用该函数进行值函数计算。

- 从理论上讲，只有当车辆处于安全集的边界时，才需要应用安全控制器。然而，在实际应用中，由于数值误差，可能需要在值函数为正值的情况下应用安全控制器。

X. 结论与未来工作

为在先验未知环境中运行的真实世界自主系统提供安全保证，是一项具有挑战性但十分重要的问题。本文提出了一种基于 HJ 可达性的安全框架，用于未知环境中的导航，该框架适用于多种规划器和传感器。为克服经典 HJ 可达性分析的计算复杂性，我们提出了一种新颖的实时算法，可在车辆穿越环境的过程中更新可达集。我们在多种传感器和规划器上验证了所提出的方法，其中包括一种基于学习的规划器，并分别在仿真环境和硬件测试平台上进行了实验。

这项工作展现出若干有趣的未来研究方向。首先，所提出的框架目前假设状态估计和传感器测量都是完美的。然而，这些假设在实践中可能并不成立，因此需要在框架中予以适当考虑。此外，当前的安全保证适用于静态环境——将其扩展到动态、多智能体环境，是另一个有趣且有价值的研究方向。

致谢

本研究部分得到DARPA Assured Autonomy计划（协议编号FA8750-18- C-0101）以及NSF CPS Frontier项目VeHICaL的支持项目（1545126），以及SRC下属CONIX中心和Berkeley Deep Drive的支持。

作者还要感谢 Kene Akametalu、Ellis Ratner 和 Anca Dragan，就本文所探讨的各种技术问题提供了有益建议。

参考文献

[1] J. Achiam、D. Held、A. Tamar 和 P. Abbeel。约束策略优化。载于 ICML，2017。[2] A. Agrawal 和 K. Sreenath。用于离散系统安全关键控制的离散控制障碍函数及其在双足机器人导航中的应用。载于 RSS，2017。[3] A. D. Ames 、 X. Xu 、 J. W. Grizzle 和 P. Tabuada。基于控制障碍函数的安全关键系统二次规划。IEEE 自动控制汇刊，62(8):3861–3876，2017。[4] I. Armeni、A. Sax、A. R. Zamir 和 S. Savarese。用于室内场景理解的联合二维-三维语义数据。arXiv 预印本，2017。[5] S. Bansal、M. Chen、J. Fisac 和 C. J. Tomlin。扰动和信息不完备条件下的安全序贯路径规划。载于 ACC，2017。[6] S. Bansal、M. Chen、S. Herbert 和 C. J. Tomlin。Hamilton-Jacobi 可达性：简要概述与最新进展。载于 CDC，2017。[7] S. Bansal、V. Tolani、S. Gupta、J. Malik 和 C. Tomlin。结合最优控制与学习，在新颖环境中进行视觉导航。arXiv 预印本，2019。[8] K. E. Bekris 和 L. E. Kavraki。运动学与动力学约束下贪心但安全的重新规划。载于 ICRA，2007。[9] F. Berkenkamp、M. Turchetta、A. Schoellig 和 A. Krause。具有稳定性保证的安全模型驱动强化学习。NIPS '17。[10] M. Chen、S. Bansal、K. Tanabe 和 C. J. Tomlin。通过序贯路径规划实现可证明安全且鲁棒的无人机航线规划：旧金山及湾区案例研究。arXiv 预印本，2017。[11] Y. Chow、O. Nachum、E. Duenez-Guzman 和 M. Ghavamzadeh。基于李雅普诺夫函数的安全强化学习方法。载于 NIPS '18。[12] E. A. Coddington 和 N. Levinson。常微分方程理论。McGraw-Hill，纽约，1955。第 1–42 页。[13] J. Fisac、A. Akametalu、M. Zeilinger、S. Kaynama、J. Gillula 和 C. J. Tomlin。面向不确定机器人系统中基于学习的控制的通用安全框架。IEEE 自动控制汇刊，2018。[14] J. Fisac、M. Chen、C. J. Tomlin 和 S. Sastry。具有时变动力学、目标和约束的可达-规避问题。载于 HSCC，2015。[15] T. Fraichard 和 H. Asama。不可避免的碰撞状态——迈向更安全的机器人？Advanced Robotics，18(10):1001–1024，2004。[16] S. Herbert、M. Chen、S. Han、S. Bansal、J. Fisac 和 C. J. Tomlin。Fastrack：用于快速且有安全保证的运动规划的模块化框架。载于 CDC，2017。[17] L. Janson、T. Hu 和 M. Pavone。未知环境中的安全运动规划：最优性基准与易处理策略。arXiv 预印本，2018。[18] G. Kahn、A. Villaflor、V. Pong、P. Abbeel 和 S. Levine。用于碰撞规避的感知不确定性的强化学习。arXiv 预印本，2017。[19] D.F. Keil、J. Fisac 和 C. J. Tomlin。未知环境中的安全、完备实时规划与探索。arXiv 预印本，2018。[20] S. Koenig 和 M. Likhachev。用于未知地形导航的快速重新规划。IEEE 机器人学汇刊，21(3):354–363，2005。[21] S. Kousik、S. Vaskov、F. Bu、M. J. Roberson 和 R. Vasudevan。弥合移动机器人滚动时域轨迹设计中安全性与实时性能之间的差距。arXiv 预印本，2018。[22] M. Lahijanian、M. R. Maly、D. Fried、L. E. Kavraki、H. Kress-Gazit 和 M. Vardi。具有部分满足保证的不确定环境迭代时序规划。IEEE 机器人学汇刊，32:583–599，2016。[23] Steven M LaValle。快速扩展随机树：一种新的路径规划工具。1998。[24] A. Majumdar 和 R. Tedrake。用于实时鲁棒反馈运动规划的漏斗库。The International Journal of Robotics Research，36(8):947–982，2017。[25] K. Margellos 和 J. Lygeros。可达-规避微分博弈的 Hamilton-Jacobi 公式。IEEE 自动控制汇刊，56(8):1849–1861，2011。

[26] I. Mitchell。水平集方法工具箱。*http://www. cs. ubc. ca/mitchell/ToolboxLS/toolboxLS. pdf,* 技术报告 *TR-2004-09*，2004。[27] I. Mitchell、A. Bayen 和 C. J. Tomlin。连续动态博弈中可达集的时变 Hamilton-Jacobi 公式。*IEEE* 自动控制汇刊，50(7):947–957，2005。[28] T. Moldovan 和 P. Abbeel。马尔可夫决策过程中的安全探索。*arXiv* 预印本，2012。[29] A. Richards 和 J. How。具有完工时间保证和鲁棒可行性的车辆机动模型预测控制。*ACC '03*。[30] C. Richter、W. Vega-Brown 和 N. Roy。用于未知环境中安全高速导航的贝叶斯学习。载于 机器人研究，第 325–341 页。2018。[31] U. Rosolia 和 F. Borrelli。面向迭代任务的学习型模型预测控制：一种数据驱动的控制框架。*IEEE* 自动控制汇刊，63(7):1883–1896，2018。[32] S. Sarid、Bingxin X. 和 H. Kress-gazit。在探索部分已知地图时保证高层行为。载于 *RSS*，2012。[33] T. Schouwenaars、J. How 和 E. Feron。具有严格安全保证的多架飞机分布式协同轨迹规划。载于 *AIAA* 制导、导航与控制会议，2004。[34] S. Singh、A. Majumdar、J. Slotine 和 M. Pavone。基于收缩理论和凸优化的鲁棒在线运动规划。载于 *ICRA '17*。[35] R. Walambe、N. Agarwal、S. Kale 和 V. Joshi。基于样条插值的汽车型移动机器人的最优轨迹生成。*IFAC* 论文在线，49(1):601–606，2016。[36] L. Yoder 和 S. Scherer。利用微型飞行器自主探索并构建基础设施模型。载于 野外与服务机器人，2016。