

基于非合作模型预测控制的人机共驾策略

刘 瑞, 朱西产, 刘 霖, 马志雄

(同济大学 汽车学院, 上海 201804)

摘要: 提出一种基于非合作模型预测控制(model predictive control, MPC)的智能汽车人机共驾策略. 首先,建立了驾驶员和控制系统两者共同控制车辆的人机共驾系统模型. 接着,得到了驾驶员和控制系统的代价函数. 然后,求解了非合作 MPC 人机共驾策略的纳什均衡解. 最后,通过仿真验证了非合作 MPC 人机共驾策略的优点和有效性. 证明了非合作 MPC 的纳什均衡解可以通过非迭代的方法求解,并通过驾驶员和控制系统置信度矩阵的更新实现了驾驶权的逐渐交接. Matlab 仿真表明,非合作 MPC 人机共驾策略可以在智能车辆遇到危险时将驾驶权逐渐从驾驶员转交给控制系统,同时保证驾驶员实时在环.

关键词: 人机共驾; 博弈论; 模型预测控制; 驾驶权

中图分类号: U461.6

文献标志码: A

Cooperative Driving Strategy Based on Non-cooperative Model Predictive Control

LIU Rui, ZHU Xichan, LIU Lin, MA Zhixiong

(School of Automotive Studies, Tongji University, Shanghai 201804, China)

Abstract: An intelligent vehicle cooperative driving strategy based on non-cooperative model predictive control (MPC) was proposed. Firstly, the cooperative driving model was presented, in which the shared control of the vehicle was realized. Next, the cost functions of the driver and the control system were obtained. Then, the Nash equilibrium solution of the non-cooperative MPC was achieved. At last, simulations were used to verify the advantages and effectiveness of the strategy. It is shown that the Nash equilibrium solution of the non-cooperative MPC can be achieved by a non-iterative method. And gradual handover of the driving privilege is realized by using the updated confidence matrixes of the driver and the control system. Simulations based on Matlab show that the non-cooperative MPC cooperative driving strategy can deliver the driving

privilege from the driver to the control system gradually when the intelligent vehicle encounters danger. And this strategy can guarantee that the driver is in the control loop all the time.

Key words: cooperative driving; game theory; model predictive control; driving privilege

智能化是减少交通事故,降低驾驶员操作负荷,提高交通效率的重要途径之一. 目前,以 ADAS (advanced driver assistance systems)为代表的驾驶辅助系统已经在量产车上有很多应用. 但完全自动驾驶或高等级自动驾驶实现起来仍有较大困难. 因此人机共驾成为了近些年的一个研究热点^[1-5].

广义上的人机共驾指所有驾驶员和智能系统共同驾驶车辆的系统. 从这个意义上讲,人机共驾可以根据控制模式分为单驾双控、双驾单控和双驾双控. 单驾双控是指驾驶指令只来源于驾驶员或控制系统其中之一,而实际车辆控制由驾驶员和控制系统共同完成. ABS (anti-lock brake system) 和 ESP (electronic stability program) 都是典型的单驾双控系统. 在 ABS 中所有的制动指令都来源于驾驶员踩制动踏板的行为,即单驾;ABS 系统根据车轮角加速度传感器获得的轮胎滑移率信息和驾驶员的制动行为进行制动操作,即双控. 单驾双控可以较好的补偿驾驶员在控制层^[6]的不足. 但单驾双控仍然将驾驶员作为理想驾驶员来考虑,即驾驶员的所有操作都是正确的. 统计表明^[7],93%的事故是由于驾驶员和驾驶环境之间的信息交互错误和驾驶员的误操作导致的. 双驾单控和双驾双控系统可以在一定程度上对驾驶员予以预先纠正,因此可以获得更大的安全收益.

双驾单控是指驾驶指令可以来源于驾驶员和控

收稿日期: 2018-08-06

基金项目: 国家重点研发计划(2016YFB0100904-2)

第一作者: 刘 瑞(1989—),博士生,主要研究方向为汽车主动安全、智能车控制、自动驾驶研究和智能车评价与测试.

E-mail: liurui@163.com

通信作者: 朱西产(1962—),教授,博士生导师,工学博士,主要研究方向为汽车主、被动安全. E-mail: xczhu@163.com

制系统,但同一时刻只执行驾驶员和控制系统其中之一的驾驶指令. ACC(adaptive cruise control)和 AEB(autonomous emergency braking)都是典型的双驾单控系统. 关于智能汽车事故的研究表明,在自动驾驶系统完全成熟之前保持驾驶员时刻在环是非常重要的^[8-9]. 双驾单控不能保证驾驶员始终在环,并且双驾单控系统的驾驶权是在驾驶员和控制系统之间无过度转换的. 这些都带来一定的安全隐患. 双驾双控是指驾驶指令来源于驾驶员和控制系统,且同一时刻按照某种策略同时执行驾驶员和控制系统驾驶指令. 双驾双控可以保证驾驶员始终对车辆保持控制,同时又可以在驾驶员操作失误时在一定程度上予以纠正. 相比于在驾驶员和控制系统之间无过度的转换,双驾双控是一种更好的模式.

由于人机共驾的双驾双控策略中的驾驶指令来源于驾驶员和智能车的控制系统,如何根据两者的指令控制车辆成为一个难题. 触觉共享控制是一种较早的双驾双控模式. 当触觉共享控制的控制系统期望接管驾驶权时,会在方向盘上作用一个附加力矩来控制车辆^[10]. 在触觉共享控制中,从方向盘到前轮的传力路径仍然是机械连接. 文献[11]基于触觉共享控制提出一种人机共驾策略,虚拟驾驶员根据驾驶员作用在方向盘上的力矩来感知驾驶员的驾驶意图,并且通过方向盘力矩来辅助驾驶员. 文献[12]使用驾驶模拟器研究了触觉共享控制对驾驶员弯道通过的辅助效果. 最近发展起来的线控转向(steer-by-wire)技术为驾驶权分配提供了新的可能. 文献[13]提出一种使用驾驶员和控制系统操作的加权和来得到智能车的实际控制输入的方法. 文献[14]基于加权和与 MPC(model predictive control)来实现人机共驾的双驾双控.

在双驾双控的驾驶权分配中,驾驶员和控制系统都希望智能车沿自己的期望轨迹行驶. 人机共驾双控双驾策略是要在两者期望轨迹之间根据某一规则找到一个最优解. 非合作动态博弈主要用来解决多个决策者共同作用于同一个动力系统的问题^[15],因而双驾双控中的驾驶权分配问题可以使用动态博弈理论来描述. 动态博弈在车辆系统中的应用仍相当有限. 文献[16]和文献[17]使用动态博弈研究了在恶劣工况下车辆评价方法,并使用卡车侧翻和折叠这两个极端工况验证了其评价方法. 文献[18]和文献[19]使用纳什均衡策略研究了将驾驶员行为考虑在内的车辆控制策略. 文献[20]使用动态博弈对驾驶员和前轮主动转向系统共同进行轨迹跟踪时的

控制行为进行了建模.

本文基于非合作 MPC(model predictive control)提出了一种人机共驾双控双驾策略. 本文的主要创新点包括:① 证明了非合作 MPC 存在唯一的纳什均衡解的条件,并表明在求解非合作 MPC 时文献[20]中的迭代是没有必要的. ② 使用驾驶员和控制系统的置信度矩阵更新实现了驾驶员与控制系统之间驾驶权的逐渐交接. 本文提出的非合作 MPC 人机共驾策略可以在智能车遇到危险时实现智能汽车由驾驶员驾驶到系统控制驾驶的平稳过渡,从而实现在提高车辆安全性的同时保持驾驶员时刻在环.

1 人机共驾系统模型

基于车辆单轨模型构建一个人机共驾车辆侧向控制模型,以实现驾驶员和控制系统同时对车辆转向的控制. 记系统状态变量为 $\mathbf{x}=[y, v_y, \phi, \omega]^T$. 其中, y 为车辆侧向位移; v_y 为车辆侧向速度; ϕ 为车辆朝向角; ω 为车辆横摆角速度. 系统状态方程可以表示为

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}_c \mathbf{x} + \mathbf{B}_{1,c} \mathbf{u}_D + \mathbf{B}_{2,c} \mathbf{u}_A \\ \mathbf{z} &= \mathbf{C} \mathbf{x} \end{aligned} \quad (1)$$

$$\text{式中, } \mathbf{A}_c = \begin{bmatrix} 0 & 1 & u & 0 \\ 0 & a_{11} & 0 & a_{12} \\ 0 & 0 & 0 & 1 \\ 0 & a_{21} & 0 & a_{22} \end{bmatrix}, \mathbf{B}_{1,c} = \mathbf{B}_{2,c} = \begin{bmatrix} 0 \\ b_1 \\ 0 \\ b_2 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, a_{11} = -\frac{2C_f + 2C_r}{mv_x},$$

$$a_{12} = -v_x - \frac{2aC_f - 2bC_r}{mv_x}, a_{21} = -\frac{2aC_f - 2bC_r}{I_z v_x},$$

$$a_{22} = -\frac{2a^2 C_f + 2b^2 C_r}{I_z v_x}, b_1 = \frac{2C_f}{m}, b_2 = \frac{2aC_f}{I_z}.$$

式中: v_x 为车辆纵向速度; C_f 为车辆前轮侧偏刚度; C_r 为车辆后轮侧偏刚度; m 为车辆质量; a 为车辆前轴中心到质心距离; b 为车辆后轴中心到质心距离; I_z 为车辆绕 z 轴转动惯量; \mathbf{u}_D 为非合作 MPC 人机共驾策略中驾驶员的输入; \mathbf{u}_A 为非合作 MPC 人机共驾策略中控制系统的输入; \mathbf{z} 为系统可量测状态.

将连续系统离散化可以得到

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A} \mathbf{x}(k) + \mathbf{B}_1 \mathbf{u}_D(k) + \mathbf{B}_2 \mathbf{u}_A(k) \\ \mathbf{z}(k) &= \mathbf{C} \mathbf{x}(k) \end{aligned} \quad (2)$$

式中: \mathbf{A} 为系统离散化之后 \mathbf{A}_c 对应的矩阵; \mathbf{B}_1 和 \mathbf{B}_2 为系统离散化之后 $\mathbf{B}_{1,c}$ 和 $\mathbf{B}_{2,c}$ 对应的矩阵.

通过离散系统模型的连续迭代可以得到

$$\mathbf{Z}(k) = \Psi \mathbf{x}(k) + \Theta_1 \mathbf{U}_1(k) + \Theta_2 \mathbf{U}_2(k) \quad (3)$$

$$\text{式中, } \mathbf{Z}(k) = \begin{bmatrix} \mathbf{z}(k+1) \\ \mathbf{z}(k+2) \\ \vdots \\ \mathbf{z}(k+N_p-1) \\ \mathbf{z}(k+N_p) \end{bmatrix},$$

$$\mathbf{U}_1 = \begin{bmatrix} \mathbf{u}_D(k) \\ \mathbf{u}_D(k+1) \\ \vdots \\ \mathbf{u}_D(k+N_u-2) \\ \mathbf{u}_D(k+N_u-1) \end{bmatrix}, \mathbf{U}_2 = \begin{bmatrix} \mathbf{u}_A(k) \\ \mathbf{u}_A(k+1) \\ \vdots \\ \mathbf{u}_A(k+N_u-2) \\ \mathbf{u}_A(k+N_u-1) \end{bmatrix},$$

$$\Psi = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{N_p-1} \\ \mathbf{CA}^{N_p} \end{bmatrix},$$

$$\Theta_1 = \begin{bmatrix} \mathbf{CB}_1 & 0 & \cdots & 0 \\ \mathbf{CAB}_1 & \mathbf{CB}_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{N_u-1} \mathbf{B}_1 & \mathbf{CA}^{N_u-2} \mathbf{B}_1 & \cdots & \mathbf{CB}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{N_p-1} \mathbf{B}_1 & \mathbf{CA}^{N_p-2} \mathbf{B}_1 & \cdots & \mathbf{CA}^{N_p-N_u} \mathbf{B}_1 \end{bmatrix},$$

$$\Theta_2 = \begin{bmatrix} \mathbf{CB}_2 & 0 & \cdots & 0 \\ \mathbf{CAB}_2 & \mathbf{CB}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{N_u-1} \mathbf{B}_2 & \mathbf{CA}^{N_u-2} \mathbf{B}_2 & \cdots & \mathbf{CB}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{N_p-1} \mathbf{B}_2 & \mathbf{CA}^{N_p-2} \mathbf{B}_2 & \cdots & \mathbf{CA}^{N_p-N_u} \mathbf{B}_2 \end{bmatrix}.$$

式中: N_p 为所谓的优化域 (preview horizon), N_u 为所谓的控制域 (control horizon). 系统根据 N_p 步的信息来求解局部最优解, \mathbf{U}_1 和 \mathbf{U}_2 可以在 N_u 步内调节, 因而有 $N_u \leq N_p$.

2 非合作 MPC 人机共驾策略

基于非合作 MPC 提出一种人机共驾策略, 实现人机共驾中的双驾双控. 在人机共驾策略中, 智能车根据驾驶员的操作输入、驾驶员模型、和车辆模型得到驾驶员的期望轨迹. 同时, 智能车根据环境感知系统构建的环境模型得到车辆可行域. 通过比较车辆可行域和驾驶员期望轨迹, 智能车可以判定当前车辆处于安全域、过渡域、或危险域. 过渡域通常指虽然不危险但安全裕量已经较小.

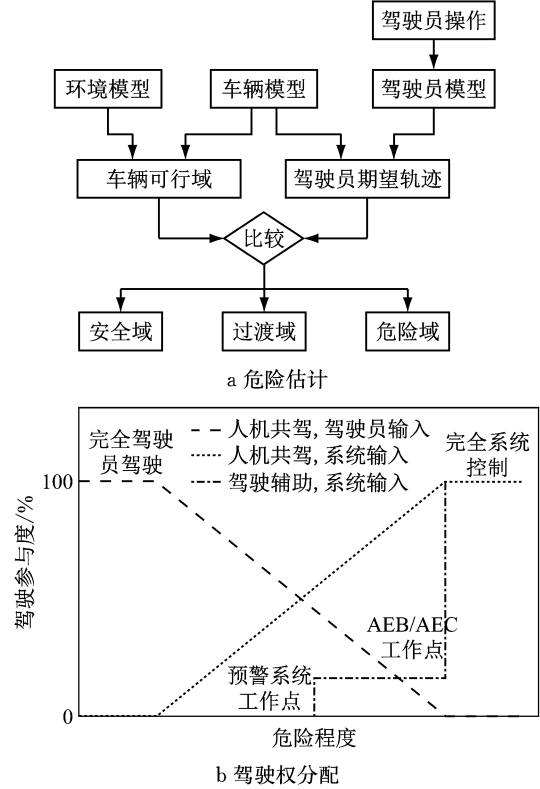


图1 人机共驾策略原理

Fig.1 Schematic diagram of cooperative driving

对于驾驶辅助系统而言, 当车辆从安全域到过渡域再到危险域的过程中通常有两个工作点. 即当车辆从安全域进入过渡域时, 包括 FCW (forward collision warning) 和 LDW (lane departure warning) 在内的预警系统开始工作; 当车辆在危险域进入预碰撞状态时, AEB 或 AEC (autonomous emergency control) 系统开始工作. 驾驶辅助系统可以在一定程度上降低事故风险. 但驾驶辅助系统的这种在某一工作点突然介入的模式让很多驾驶员感到不适应或不习惯. 同时, 如何准确确定预警系统和辅助系统的工作点时刻或位置成为一个难题. 这些都限制了驾驶辅助系统的接受程度和安全收益.

在非合作 MPC 人机共驾策略中, 智能车在行驶过程中驾驶指令同时来源于驾驶员和控制系统. 智能车根据两者的驾驶指令来规划运动轨迹. 当智能车处于安全驾驶状态时, 智能车根据驾驶员的操作来控制车辆. 当由于驾驶员操作失误或分心导致车辆进入过度域或危险域时, 非合作 MPC 人机共驾策略可以将驾驶权从驾驶员逐渐交接给控制系统, 来避免危险. 这种逐渐过渡的方式可以较好地兼顾舒适性和安全性, 同时保证驾驶员时刻在环.

非合作 MPC 人机共驾策略中的两个参与者 (驾驶员和控制系统) 都期望使关于自身目标的代价函

数尽可能小,即

$$\min_{U_1} \mathbf{V}_1(k), \quad \min_{U_2} \mathbf{V}_2(k)$$

$$\text{s. t. } \mathbf{Z}(k) = \Psi \mathbf{x}(k) + \Theta_1 \mathbf{U}_1(k) + \Theta_2 \mathbf{U}_2(k) \quad (4)$$

式中, $\mathbf{V}_1(k)$ 为驾驶员的代价函数, $\mathbf{V}_2(k)$ 为控制系统的代价函数。

两个参与者的代价函数定义为

$$\mathbf{V}_1(k) = \|\mathbf{Z}(k) - \mathbf{T}_1(k)\|_{Q_1(k)}^2 + \|\mathbf{U}_1(k)\|_{R_1}^2$$

$$\mathbf{V}_2(k) = \|\mathbf{Z}(k) - \mathbf{T}_2(k)\|_{Q_2(k)}^2 + \|\mathbf{U}_2(k)\|_{R_2}^2 \quad (5)$$

$$\text{式中, } \mathbf{T}_1(k) = \begin{bmatrix} t_1(k-N_p+1) \\ t_1(k-N_p+2) \\ \vdots \\ t_1(k-1) \\ t_1(k) \end{bmatrix},$$

$$\mathbf{T}_2(k) = \begin{bmatrix} t_2(k-N_p+1) \\ t_2(k-N_p+2) \\ \vdots \\ t_2(k-1) \\ t_2(k) \end{bmatrix}, \mathbf{R}_1 = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_1 \end{bmatrix},$$

$$\mathbf{R}_2 = \begin{bmatrix} r_2 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_2 \end{bmatrix},$$

$$\mathbf{Q}_1(k) = \begin{bmatrix} q_1(k+1) & 0 & \cdots & 0 \\ 0 & q_1(k+2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_1(k+N_p) \end{bmatrix},$$

$$\mathbf{Q}_2(k) = \begin{bmatrix} q_2(k+1) & 0 & \cdots & 0 \\ 0 & q_2(k+2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_2(k+N_p) \end{bmatrix},$$

$$q_1(k) = \begin{bmatrix} \kappa_1(k) & 0 \\ 0 & \lambda_1(k) \end{bmatrix}, q_2(k) = \begin{bmatrix} \kappa_2(k) & 0 \\ 0 & \lambda_2(k) \end{bmatrix},$$

$$t_1(k) = \begin{bmatrix} y_D(k) \\ \psi_D(k) \end{bmatrix}, t_2(k) = \begin{bmatrix} y_A(k) \\ \psi_A(k) \end{bmatrix}.$$

在人机共驾非合作 MPC 模型中,有以下两个要点:

(1) $\mathbf{Q}_1(k)$ 和 $\mathbf{Q}_2(k)$ 均为时变矩阵. $q_1(k)$ 为驾驶员置信度矩阵, $q_2(k)$ 为智能车控制系统置信度矩阵. 通过 $\mathbf{Q}_1(k)$ 和 $\mathbf{Q}_2(k)$ 中的两个置信度矩阵随时间的变化可以实现驾驶员与智能车控制系统之间驾驶权的转换. 通过后文仿真分析可知, $\kappa_1(k)$ 和 $\kappa_2(k)$ 是与驾驶权分配相关的参数, 因此称其为分配系数; $\lambda_1(k)$ 和 $\lambda_2(k)$ 是与动态特性相关的参数, 因此称其为

动态调整系数. 当智能车感知到车辆进入危险状态时, 可以通过逐渐调低 $\kappa_1(k)$ 并逐渐调高 $\kappa_2(k)$, 使车辆沿着智能车控制系统的规划轨迹行驶以躲避危险. 而当车辆躲避危险逐渐进入正常行驶状态时, 可以通过逐渐调高 $\kappa_1(k)$ 并逐渐调低 $\kappa_2(k)$, 将驾驶权逐渐交还给驾驶员. 这样就实现了驾驶员与控制系统之间驾驶权的平缓交接.

(2) MPC 需要根据未来 N_p 步内的预测信息来求局部最优解. 但在非合作 MPC 人机共驾策略中, 期望使用驾驶员实时转向操作实现双驾双控. 因此, 使用一种预测域提前的方法, 即将预测域取为当前时刻向前 N_p 步的区间. 这样可以较好地解决非合作 MPC 的实时求解, 但会产生 N_p 步的延迟. 因此预测域 N_p 不能取得过大. 本文选取 $N_p=10$, 而每一步的时长为 0.01 s. 这样产生的 0.1 s 延迟仍在可接受范围之内.

$\mathbf{T}_1(k)$ 和 $\mathbf{T}_2(k)$ 是两个参与者的局部目标轨迹, 在每一次优化前都需要滚动更新. 其更新方程为

$$\mathbf{T}_1(k+1) = \mathbf{G}\mathbf{T}_1(k) + \mathbf{H}t_1(k+1)$$

$$\mathbf{T}_2(k+1) = \mathbf{G}\mathbf{T}_2(k) + \mathbf{H}t_2(k+1) \quad (6)$$

$$\text{式中, } \mathbf{G} = \begin{bmatrix} 0 & \mathbf{I}_m & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{I}_m & 0 & \cdots & 0 \\ 0 & 0 & 0 & \mathbf{I}_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mathbf{I}_m \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \mathbf{H} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \mathbf{I}_m \end{bmatrix},$$

\mathbf{I}_m 为 m 维单位矩阵, m 为状态变量的个数.

3 人机共驾中的纳什均衡解

定义两个误差变量 $\epsilon_1(k)$ 和 $\epsilon_2(k)$ 为

$$\epsilon_1(k) = \mathbf{T}_1(k) - \Psi \mathbf{x}(k) - \Theta_2 \mathbf{U}_2(k)$$

$$\epsilon_2(k) = \mathbf{T}_2(k) - \Psi \mathbf{x}(k) - \Theta_1 \mathbf{U}_1(k) \quad (7)$$

则有

$$\mathbf{V}_i(k) = \|\Theta_i \mathbf{U}_i(k) - \epsilon_i(k)\|_{Q_i(k)}^2 + \|\mathbf{U}_i(k)\|_{R_i}^2, \quad i=1,2 \quad (8)$$

$\mathbf{V}_i(k)$ 对 $\mathbf{U}_i(k)$ 的偏导为

$$\frac{\partial \mathbf{V}_i(k)}{\partial \mathbf{U}_i(k)} = -2\Theta_i^T \mathbf{Q}_i(k) \epsilon_i(k) + 2[\Theta_i^T \mathbf{Q}_i(k) \Theta_i + \mathbf{R}_i] \mathbf{U}_i(k) \quad (9)$$

由于 $\mathbf{Q}_i(k)$ 都是半正定矩阵且 \mathbf{R}_i 都是正定矩阵, 因此 $\mathbf{V}_i(k)$ 对 $\mathbf{U}_i(k)$ 的二阶偏导始终大于 0. 则 $\mathbf{V}_i(k)$ 对 $\mathbf{U}_i(k)$ 的偏导等于 0 的解即为代价函数最小的最优控制序列. 因此有

$$\begin{aligned} U_i^0(k) &= F_i(k)\varepsilon_i(k) \\ U_i^0(k) &\in \Pi_i \end{aligned} \quad (10)$$

其中

$$F_i(k) = [\Theta_i^T Q_i(k) \Theta_i + R_i]^{-1} \Theta_i^T Q_i(k) \quad (11)$$

$U_i^0(k)$ 表示第*i*个($i=1,2$)控制输入 $U_i(k)$ 的最优控制序列。 Π_i 表示 $U_i(k)$ 的所有可能最优控制序列的集合。通过式(7)可知,式(10)中一个参与者的最优控制序列总与另一个参与者的控制决策有关。即 $U_1^0(k)$ 是与 $U_2(k)$ 相关的, $U_2^0(k)$ 是与 $U_1(k)$ 相关的。因此式(10)表示了一簇最优控制序列,这一簇最优控制序列组成的集合可以表示为 $\Pi_1 \times \Pi_2$ 。符号 \times 表示集合的直积。

非合作博弈问题可以通过所谓的纳什均衡(Nash equilibrium)^[21]来求解。不失一般性的,我们以两个参与者的动态博弈来说明这个问题。记博弈中第*i*个参与者的控制为 U_i ($i=1,2$)。其中 $U_i \in \Pi_i$, Π_i 为第*i*个参与者全部可行控制的集合。若存在一组控制 $\{U_1^*, U_2^*\}$ ($U_i^* \in \Pi_i, i=1,2$)使得关于控制的代价函数式(12)成立,那么 $\{U_1^*, U_2^*\}$ 就称为一组纳什均衡解^[15]。

$$\begin{aligned} V_1(U_1^*, U_2^*) &\leq V_1(U_1, U_2^*), \forall U_1 \in \Pi_1 \\ V_2(U_1^*, U_2^*) &\leq V_2(U_1^*, U_2), \forall U_2 \in \Pi_2 \end{aligned} \quad (12)$$

纳什均衡表明,当一个参与者执行纳什均衡策略时,其他参与者无法通过选择非纳什均衡的其他策略来增加自己的收益。因此对每一个参与者来说,纳什均衡解是当前博弈条件下的最优解。

对于非合作博弈MPC人机共驾策略,其纳什均衡解可以由定理1给出。

定理1 对于如式(2)所描述的博弈系统和式(4)所描述的代价函数,当且仅当 $I-L(k)$ 可逆时系统具有唯一的纳什均衡解。且该纳什均衡解为

$$\begin{bmatrix} U_1^*(k) \\ U_2^*(k) \end{bmatrix} = K(k) \left\{ \begin{bmatrix} T_1(k) \\ T_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} x(k) \right\}$$

式中, $K(k)=[I-L(k)]^{-1}M(k)$,

$$M(k) = \begin{bmatrix} F_1(k) & 0 \\ 0 & F_2(k) \end{bmatrix},$$

$$L(k) = \begin{bmatrix} 0 & -F_1(k)\Theta_2 \\ -F_2(k)\Theta_1 & 0 \end{bmatrix}.$$

证明:

(1) 存在性与唯一性。 $V_i(k)$ ($i=1,2$)为一个二次型函数的博弈通常也被称为二次博弈(quadratic game)。二次博弈是否存在唯一纳什均衡解可以通过矩阵的可逆性来判别。文献[15]表明,存在一个矩阵 $P(k)$,当且仅当 $P(k)$ 可逆时,二次博弈存在唯一的

纳什均衡解。其中

$$P(k) = \begin{bmatrix} \Theta_1^T Q_1(k) \Theta_1 + R_1 & \Theta_1^T Q_1(k) \Theta_2 \\ \Theta_2^T Q_2(k) \Theta_1 & \Theta_2^T Q_2(k) \Theta_2 + R_2 \end{bmatrix}$$

对 $P(k)$ 进行变换可得

$$P(k) = \begin{bmatrix} \Theta_1^T Q_1(k) \Theta_1 + R_1 & 0 \\ 0 & \Theta_2^T Q_2(k) \Theta_2 + R_2 \end{bmatrix} \cdot [I-L(k)]$$

因为 $Q_i(k)$ 都是半正定矩阵且 R_i 都是正定矩阵,所以 $P(k)$ 的可逆性等价于 $I-L(k)$ 的可逆性。因此当且仅当 $I-L(k)$ 的可逆性,非合作MPC具有唯一的纳什均衡解。

(2) 构造性。当 $I-L(k)$ 可逆时,非合作MPC的唯一纳什均衡解可以直接求得闭式解析表达而不需要迭代。下面来构造非合作MPC的纳什均衡解。

将 $U_1^0(k)$ 与 $U_2^0(k)$ 表示为矩阵形式可得

$$\begin{bmatrix} U_1^0(k) \\ U_2^0(k) \end{bmatrix} = M(k) \left\{ \begin{bmatrix} T_1(k) \\ T_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} x(k) \right\} + L(k) \begin{bmatrix} U_1(k) \\ U_2(k) \end{bmatrix} \quad (13)$$

通过式(13)得到的 $U_1(k)$ 和 $U_2(k)$ 的一对最优控制序列满足

$$\begin{aligned} V_1(U_1^0, U_2) &\leq V_1(U_1, U_2), \forall U_1 \in \Pi_1, \forall U_2 \in \Pi_2 \\ V_2(U_1, U_2^0) &\leq V_2(U_1, U_2), \forall U_1 \in \Pi_1, \forall U_2 \in \Pi_2 \end{aligned} \quad (14)$$

当且仅当 $I-L(k)$ 可逆时,非合作MPC有唯一的纳什均衡解。记系统的纳什均衡解为 $(U_1^*(k), U_2^*(k))$ 。系统的纳什均衡解应满足最优反应函数式(13),即 $\{U_1^*(k), U_2^*(k)\} \in \Pi_1 \times \Pi_2$ 。因此有

$$\begin{bmatrix} U_1^*(k) \\ U_2^*(k) \end{bmatrix} = M(k) \left\{ \begin{bmatrix} T_1(k) \\ T_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} x(k) \right\} + L(k) \begin{bmatrix} U_1^*(k) \\ U_2^*(k) \end{bmatrix} \quad (15)$$

式(15)得到的 $U_1^*(k)$ 和 $U_2^*(k)$ 满足

$$\begin{aligned} V_1(U_1^*, U_2^*) &\leq V_1(U_1, U_2^*), \forall U_1 \in \Pi_1 \\ V_2(U_1^*, U_2^*) &\leq V_2(U_1^*, U_2), \forall U_2 \in \Pi_2 \end{aligned} \quad (16)$$

因此,通过式(15)得到的最优控制序列对 $\{U_1^*(k), U_2^*(k)\}$ 即为系统的一组纳什均衡解。当 $I-L(k)$ 可逆时,式(15)有解。非合作博弈MPC的纳什均衡解为

$$\begin{bmatrix} U_1^*(k) \\ U_2^*(k) \end{bmatrix} = [I-L(k)]^{-1} M(k) \cdot \left\{ \begin{bmatrix} T_1(k) \\ T_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} x(k) \right\} \quad (17)$$

证毕.

在 MPC 中,通常采取一种域后退的策略.即在每一个优化域中求解局部最优 MPC 控制策略,但优化域随时间一直向后推移进行滚动优化,因此只有控制序列的第一个控制输入起作用.两个参与者的反馈增益 $\mathbf{K}_1(k)$ 和 $\mathbf{K}_2(k)$ 为

$$\begin{bmatrix} \mathbf{K}_1(k) \\ \mathbf{K}_2(k) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I}_l & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{N_p} \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ \mathbf{I}_l & 0 & \cdots & 0 \end{bmatrix}}_{N_p} \mathbf{K}(k) \quad (18)$$

式中: \mathbf{I}_l 为 l 维单位矩阵, l 为每个参与者控制输入的个数.

人机共驾中两个参与者的非合作博弈 MPC 控制输入可以表示为

$$\begin{aligned} \mathbf{u}_D^*(k) &= \mathbf{K}_1(k) \left\{ \begin{bmatrix} \mathbf{T}_1(k) \\ \mathbf{T}_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} \mathbf{x}(k) \right\} \\ \mathbf{u}_A^*(k) &= \mathbf{K}_2(k) \left\{ \begin{bmatrix} \mathbf{T}_1(k) \\ \mathbf{T}_2(k) \end{bmatrix} - \begin{bmatrix} \Psi \\ \Psi \end{bmatrix} \mathbf{x}(k) \right\} \quad (19) \end{aligned}$$

非合作 MPC 人机共驾策略的控制输入是综合考虑驾驶员操作,控制系统规划轨迹,和当前危险状态后对车辆的控制.即根据危险程度在驾驶员和控制系统之间动态的分配驾驶权,因此可以兼顾舒适性和安全性.

文献[22]表明,虽然通过式(10)推导 MPC 的最优控制序列比较直观方便,但使用式(10)计算矩阵 $\mathbf{F}_i(k)$ 的方法通常数值不稳定.文献[22]提供了一种较好的解决方法,本文使用该方法来计算 $\mathbf{F}_i(k)$. 则 $\mathbf{F}_1(k)$ 和 $\mathbf{F}_2(k)$ 可以表示为

$$\begin{aligned} \mathbf{F}_1(k) &= \begin{bmatrix} \mathbf{S}_{Q_1}(k) \Theta_1 \\ \mathbf{S}_{R_1} \end{bmatrix}^+ \begin{bmatrix} \mathbf{S}_{Q_1}(k) \\ 0 \end{bmatrix} \\ \mathbf{F}_2(k) &= \begin{bmatrix} \mathbf{S}_{Q_2}(k) \Theta_2 \\ \mathbf{S}_{R_2} \end{bmatrix}^+ \begin{bmatrix} \mathbf{S}_{Q_2}(k) \\ 0 \end{bmatrix} \\ \begin{cases} \mathbf{S}_{Q_1}(k)^T \mathbf{S}_{Q_1}(k) = \mathbf{Q}_1(k) \\ \mathbf{S}_{R_1}^T \mathbf{S}_{R_1} = \mathbf{R}_1 \end{cases} \\ \begin{cases} \mathbf{S}_{Q_2}(k)^T \mathbf{S}_{Q_2}(k) = \mathbf{Q}_2(k) \\ \mathbf{S}_{R_2}^T \mathbf{S}_{R_2} = \mathbf{R}_2 \end{cases} \end{aligned} \quad (20)$$

式中: \mathbf{A}^+ 表示矩阵 \mathbf{A} 的广义逆.

4 仿真结果及分析

使用 Matlab 软件对非合作 MPC 人机共驾策略进行仿真验证. 根据实际车辆参数选取车辆参数如表 1 所示.

表 1 车辆参数

参数	数值
质心到前轴中心距离 a/m	1.0
质心到后轴中心距离 b/m	1.5
整备质量 m/kg	1 270
转动惯量 $I_z/\text{kg} \cdot \text{m}^2$	1 443.1
前轮侧偏刚度 $C_f/\text{kN} \cdot \text{rad}^{-1}$	30
后轮侧偏刚度 $C_r/\text{kN} \cdot \text{rad}^{-1}$	30

使用驾驶员期望向左侧变道而智能车控制系统期望直行的场景验证非合作 MPC 人机共驾策略的效果. 驾驶员变道轨迹使用 5 次多项式变道轨迹拟合^[23]. 变道轨迹长度为 50 m, 宽度为 3.5 m. 仿真中车辆行驶速度为 $20 \text{ m} \cdot \text{s}^{-1}$. 预测域 $N_p=10$, 控制域 $N_u=10$, 仿真步长 $T=0.01 \text{ s}$. 与最优控制类似, 在非合作 MPC 中, 只有 $\mathbf{q}_1(k)$ 和 $\mathbf{q}_2(k)$ 与 \mathbf{r}_1 和 \mathbf{r}_2 的相对值会对控制结果产生影响. 本文中每个参与者只有一个输入, \mathbf{r}_1 和 \mathbf{r}_2 均为标量. 因此在仿真中均设定 $\mathbf{r}_1=1, \mathbf{r}_2=1$.

第 1 组仿真验证分配系数 $\kappa_1(k)$ 和 $\kappa_2(k)$ 对控制结果的影响. 这一组 5 个工况的参数设置如下:

(1) 仿真工况 1.1: $\kappa_1(k)=0.1, \lambda_1(k)=10$; $\kappa_2(k)=0.1, \lambda_2(k)=10$.

(2) 仿真工况 1.2: $\kappa_1(k)=0.4, \lambda_1(k)=40$; $\kappa_2(k)=0.1, \lambda_2(k)=10$.

(3) 仿真工况 1.3: $\kappa_1(k)=0.1, \lambda_1(k)=10$; $\kappa_2(k)=0.3, \lambda_2(k)=30$.

(4) 仿真工况 1.4: $\kappa_1(k)=0, \lambda_1(k)=0$; $\kappa_2(k)=0.1, \lambda_2(k)=10$.

(5) 仿真工况 1.5: $\kappa_1(k)=0.1, \lambda_1(k)=10$; $\kappa_2(k)=0, \lambda_2(k)=0$.

第 1 组仿真中, $\mathbf{q}_1(k)$ 和 $\mathbf{q}_2(k)$ 均为常数, 仿真结果如图 2 所示. 通过图 2 可以看出, 改变 $\mathbf{q}_1(k)$ 和 $\mathbf{q}_2(k)$ 的相对比例(即等比例的缩放 $\kappa_i(k)$ 和 $\lambda_i(k)$, ($i=1,2$)), 会使最终的规划轨迹处于驾驶员期望轨迹和控制系统期望轨迹之间. 当 $\mathbf{q}_1(k)$ 为 0 时, 最终规划轨迹会与控制系统期望轨迹完全相同, 此时智能车完全受系统控制; 当 $\mathbf{q}_2(k)$ 为 0 时, 最终规划轨迹会与驾驶员期望轨迹完全相同, 此时智能车完全受驾驶员控制.

第 2 组仿真验证动态调整系数 $\lambda_1(k)$ 和 $\lambda_2(k)$ 对控制结果的影响. 这一组 5 个工况的参数设置如下:

(1) 仿真工况 2.1: $\kappa_1(k)=0.1, \lambda_1(k)=10$; $\kappa_2(k)=0.1, \lambda_2(k)=10$.

(2) 仿真工况 2.2: $\kappa_1(k)=0.1, \lambda_1(k)=2$; $\kappa_2(k)=0.1, \lambda_2(k)=10$.

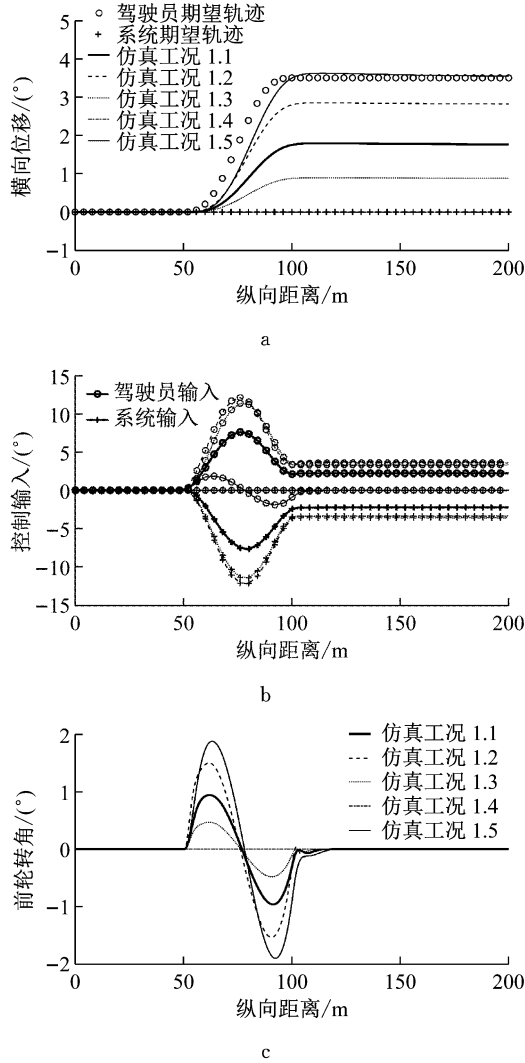


图2 第1组仿真工况结果

Fig.2 Results of the first group of simulations

(3) 仿真工况 2.3: $\kappa_1(k) = 0.1, \lambda_1(k) = 6;$

$\kappa_2(k) = 0.1, \lambda_2(k) = 10.$

(4) 仿真工况 2.4: $\kappa_1(k) = 0.1, \lambda_1(k) = 10;$

$\kappa_2(k) = 0.1, \lambda_2(k) = 2.$

(5) 仿真工况 2.5: $\kappa_1(k) = 0.1, \lambda_1(k) = 10;$

$\kappa_2(k) = 0.1, \lambda_2(k) = 6.$

第2组仿真中, $q_1(k)$ 和 $q_2(k)$ 同样均为常数, 仿真结果如图3所示. 通过图3可以看出, 在 $\kappa_1(k)$ 和 $\kappa_2(k)$ 保持不变的情况下, 改变 $\lambda_1(k)$ 和 $\lambda_2(k)$ 对规划轨迹的最终结果不产生影响. 即5个工况的规划轨迹最终收敛到同样的位置. 改变 $\lambda_1(k)$ 和 $\lambda_2(k)$ 主要影响了规划轨迹的动态特性. 当 $\lambda_1(k)$ 比 $\lambda_2(k)$ 大时, 规划轨迹表现出一种超调特性, 并且 $\lambda_1(k)$ 与 $\lambda_2(k)$ 的差值越大这种超调特性越明显; 当 $\lambda_1(k)$ 比 $\lambda_2(k)$ 小时, 规划轨迹表现出一种过阻尼特性, 并且 $\lambda_1(k)$ 与 $\lambda_2(k)$ 的差值越大这种过阻尼特性越明显.

第1组和第2组仿真使我们对 $\kappa_i(k)$ 和 $\lambda_i(k)$

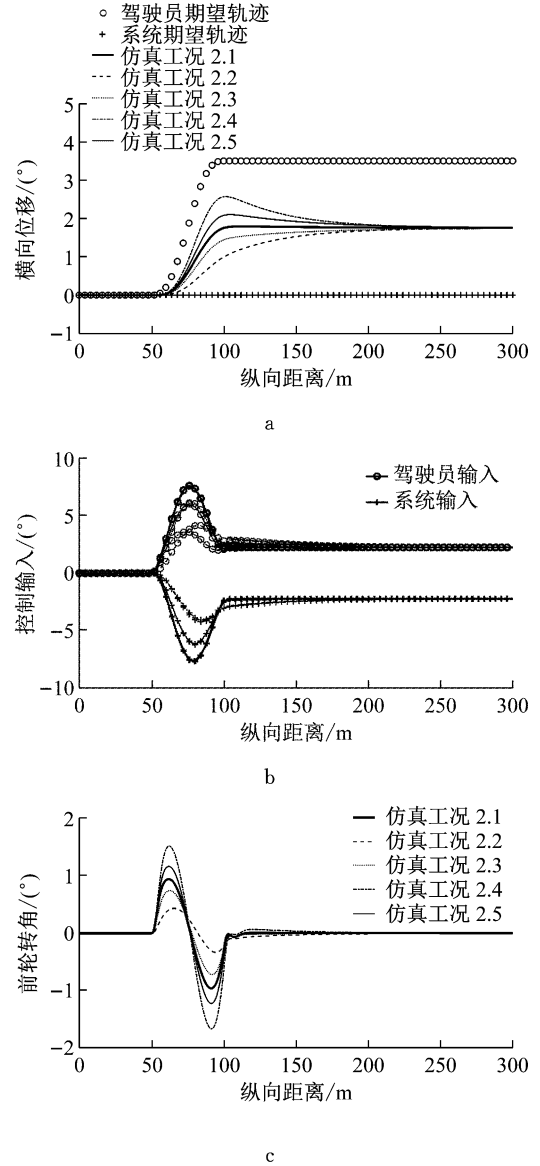


图3 第2组仿真工况结果

Fig.3 Results of second group of simulations

($i=1,2$)对规划轨迹的影响有了较为清晰的了解. 接下来的第3组仿真将表示非合作MPC人机共驾策略最为明显的优点, 即实现驾驶员和控制系统之间驾驶权的逐渐交接. 这主要通过 $\kappa_1(k)$ 和 $\kappa_2(k)$ 的逐渐变化来实现.

第3组仿真中, $\lambda_1(k)$ 和 $\lambda_2(k)$ 设为常数, 即 $\lambda_1(k) = 2, \lambda_2(k) = 2$. 3个工况中开始均设定 $\kappa_1(k) = 0.1, \kappa_2(k) = 0$. 在某一时刻, $\kappa_1(k)$ 逐渐线性的减小到0, 同时 $\kappa_2(k)$ 逐渐线性的增大到0.1. 第3组仿真3个工况 $\kappa_1(k)$ 和 $\kappa_2(k)$ 的变化如图4所示. 仿真工况3.1是在3s时(60m处)开始驾驶权交接, 并在1s内(20m内)完成从驾驶员驾驶到系统驾驶的转换. 注意到仿真工况设定是在50m处驾驶员期望开始变道, 并在100m处变道结束. 因此仿真工况3.1表

示了在变道过程中进行驾驶权交接的结果. 仿真工况 3.2 是在 9 s 时(180 m 处)开始驾驶权交接,并在 6 s 内(120 m 内)完成从驾驶员驾驶到系统驾驶的转换. 仿真工况 3.3 是在 9 s 时(180 m 处)开始驾驶权交接,并在 1 s 内(20 m 内)完成从驾驶员驾驶到系统驾驶的转换. 仿真工况 3.2 和 3.3 表示了当驾驶员和控制系统存在分歧的稳定状态时进行驾驶权转换的结果.

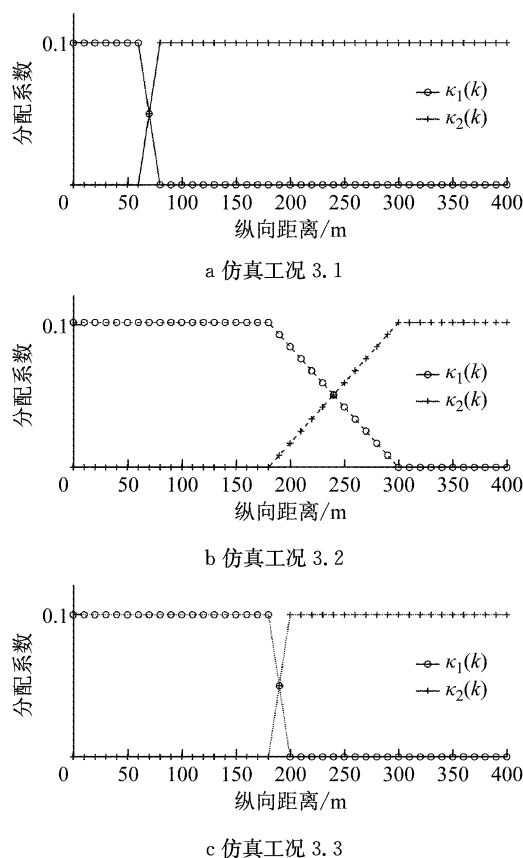


图 4 分配系数的变化

Fig.4 Assignment coefficients in the third group of simulations

第 3 组仿真 3 个工况结果如图 5 所示. 通过图 5a 可以看出,当驾驶员与控制系统发生分歧时进行驾驶权的转换,3 个工况都可以规划出一条非常符合车辆动力学的稳定轨迹. 并且 3 个工况都实现了平稳的从驾驶员控制转换到系统控制. 在仿真工况 3.1 中,变道进行到一半时进行驾驶权转换,车辆可以平顺的由向左变道逐渐回到本车道. 在仿真工况 3.2 和 3.3 中,当车辆已经按照驾驶员的操作进行左变道后进行驾驶权交接,车辆可以平顺的右变道回到本车道. 且驾驶权交接时间越短,变道轨迹越紧急. 通过图 5c 可以看出,3 个工况在驾驶权交接的过程中,车辆前轮转角输入始终保持在较小范围内,没

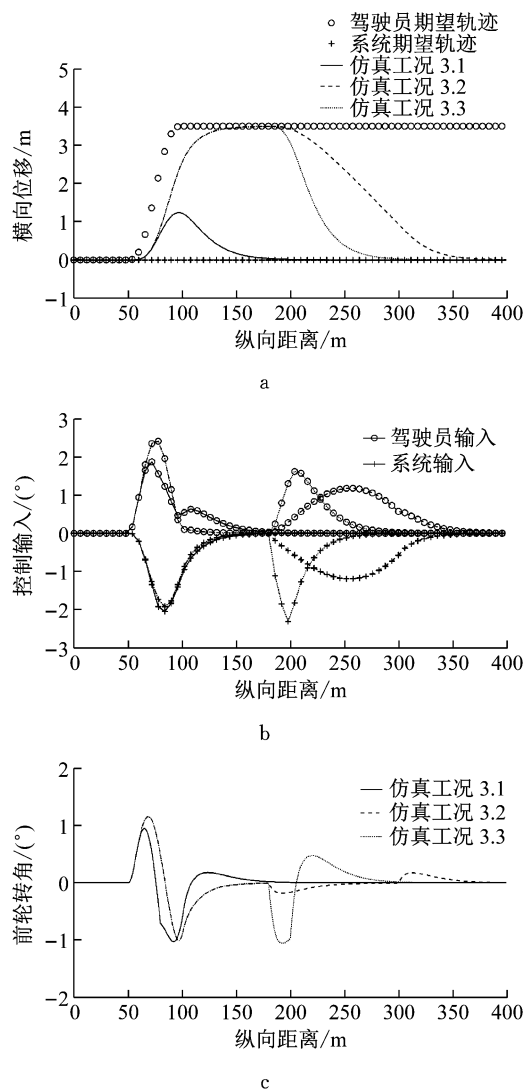


图 5 第三组仿真工况结果

Fig.5 Results of the third group of simulations

有较大范围的突然剧烈变化. 这对保持车辆稳定是非常有利的.

第 3 组仿真工况是有其实际意义的. 当驾驶员左转方向盘期望左变道时,智能车在之前可以由于感知系统的遮挡等没有检测到左侧车道的障碍物. 在车辆左变道过程中,智能车在检测到左侧车道障碍物的危险后在变道过程中马上进行驾驶权交接,使本车回到原车道以避免危险.

5 结语

使用非合作 MPC 实现了一种人机共驾的双驾双控策略. 即驾驶员和智能车控制系统同时发出驾驶指令,智能车根据车辆当前状态和两者的控制指令依某一规则规划车辆运动轨迹. 之前关于非合作 MPC 求解的文献中均使用了迭代法,本文表明非合

作 MPC 可以通过非迭代的方法求解. 非合作 MPC 人机共驾策略的主要优点是可以实现驾驶员和控制系统之间驾驶权的逐渐交接. 这样,既能保证驾驶员实时在环,又不会在智能车控制系统接管车辆时过于突兀,给驾驶员带来不适感. 非合作 MPC 的这种驾驶权逐渐交接是通过驾驶员和控制系统置信度矩阵的实时更新实现的. Matlab 仿真验证表明在危险工况时,非合作 MPC 人机共驾策略可以完成驾驶权从驾驶员到智能车控制系统的平稳交接.

参考文献:

- [1] Da LIO M, BIRAL F, BERTOLAZZI E, *et al.* Artificial co-drivers as a universal enabling technology for future intelligent vehicles and transportation systems [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(1): 244.
- [2] ROSENSTATTER T, ENGLUND C. Modelling the level of trust in a cooperative automated vehicle control system [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(4): 1237.
- [3] SAWADE O, SCHULZE M, RADUSCH I. Robust communication for cooperative driving maneuvers [J]. *IEEE Intelligent Transportation Systems Magazine*, 2018, 10(3): 159.
- [4] WU Jian, CHENG Shuo, LIU Binhao, *et al.* A human-machine-cooperative-driving controller based on AFS and DYC for vehicle dynamic stability [J]. *Energies*, 2017, 10(11): 1737.
- [5] MENG Yue, LI Li, WANG Feiyue, *et al.* Analysis of cooperative driving strategies for nonsignalized intersections [J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(4): 2900.
- [6] MICHON J A. A critical view of driver behavior models: what do we know, what should we do? [C]//*Human Behavior and Traffic Safety*. New York: Springer, 1985: 485-520.
- [7] RUMAR K. The role of perceptual and cognitive filters in observed behavior[M]. *Human Behavior and Traffic Safety*. New York: Springer, 1985.
- [8] KHOSRAVANI S, KHAJEPOUR A, FIDAN B, *et al.* Development of a robust vehicle control with driver in the loop [C]//*American Control Conference*. Portland: IEEE, 2014: 3482-3487.
- [9] LI Renjie, LI Shengbo, GAO Hongbo, *et al.* Effects of human adaptation and trust on shared control for driver-automation cooperative driving [C]//*SAE Technical Paper Series*. Detroit: SAE Publication Group, 2017: 2017-01-1987.
- [10] ABBINK D A, MULDER M, BOER E R. Haptic shared control: smoothly shifting control authority? [J]. *Cognition Technology & Work*, 2012, 14(1): 19.
- [11] SOUALMI B, SENTOUH C, POPIEUL J C, *et al.* Automation-driver cooperative driving in presence of undetected obstacles [J]. *Control Engineering Practice*, 2014, 24(1): 106.
- [12] MULDER M, ABBINK D A, BOER E R. sharing control with haptics: seamless driver support from manual to automatic control [J]. *Human Factors*, 2012, 54(5): 786.
- [13] OMAE M, FUJIOKA T, HASHIMOTO N, *et al.* The application of RTK-GPS and steer-by-wire technology to the automatic driving of vehicles and an evaluation of driver behavior [J]. *IATSS Research*, 2006, 30(2): 29.
- [14] LI Renjie, LI Yanan, LI Shengbo, *et al.* Driver- automation indirect shared control of highly automated vehicles with intention-aware authority transition [C]//*Intelligent Vehicles Symposium*. Los Angeles: IEEE, 2017: 26-32.
- [15] BASAR T, OLSDER G J. Dynamic noncooperative game theory [M]. Philadelphia: SIAM, 1999.
- [16] MA W H, PENG H. Worst- case vehicle evaluation methodology? examples on truck rollover/jackknifing and active yaw control systems [J]. *Vehicle System Dynamics*, 1999, 32(4/5): 389.
- [17] MA W H, PENG H. A worst-case evaluation method for dynamic systems [J]. *Journal of Dynamic Systems Measurement and Control*, 1999, 121(2): 191.
- [18] TAMADDONI S H, TAHERI S, AHMADIAN M. Optimal VSC design based on nash strategy for differential 2-player games [C]//*International Conference on Systems, Man and Cybernetics*. San Antonio: IEEE, 2009: 2415-2420.
- [19] TAMADDONI S H, TAHERI S, AHMADIAN M. Optimal preview game theory approach to vehicle stability controller design[J]. *Vehicle System Dynamics*, 2011, 49(12): 1967.
- [20] NA X, COLE D J. Linear quadratic game and non-cooperative predictive methods for potential application to modelling driver-AFS interactive steering control [J]. *Vehicle System Dynamics*, 2013, 51(2): 165.
- [21] ENGWERDA J. LQ dynamic optimization and differential games [M]. London: John Wiley & Sons Ltd, 2005.
- [22] MACIEJOWSKI J M. Predictive control with constraints[M]. London: Pearson Education, 2001.
- [23] SLEDGE N H, MARSHEK K M. Comparison of ideal vehicle lane-change trajectories [C]//*SAE Technical Paper Series*. Detroit: SAE Publication Group, 1997: 971062.