

第1章 引言

1.1 概述

近年来,随着汽车电控技术的发展,汽车的智能驾驶系统,例如自动紧急制动(autonomous emergence braking, AEB)、车道偏离预警(lane departure warning, LDW)、自适应巡航控制(adaptive cruise control, ACC)乃至无人驾驶被大量引入。这些系统的应用可以减少事故发生的概率。减小事故后果的安全措施(被动安全技术),例如安全带和安全气囊,已经在全球不同的新车评价规程(new car assessment programme, NCAP)的测试试验中展示出其性能。但是相比与被动安全技术,目前还没有针对这些智能驾驶系统的被广泛认可的测试方法。

从车辆使用者的角度考虑,如果相关部门设立了针对智能驾驶系统功能的有效性评价试验,那么车辆使用者也会逐渐意识到智能驾驶技术的重要性,并且可在购买车辆时对比不同品牌之间的智能驾驶性能。保险公司等投资机构也需要第三方对每个汽车公司的智能驾驶系统功能进行独立可靠的评价,以作为其投保的重要参考。更进一步讲,政府也需要根据技术预期的效果来制定汽车行业的发展导向。

从企业技术开发的角度出发,各大汽车公司投入大量资金和研发人员对汽车智能驾驶系统进行开发,其不得不面对这样一个事实:通常情况下,各种技术所能带来的好处只有在产品投入到市场一段时间后才能得到验证。几乎每一个汽车公司都希望在产品进入市场之前对安全技术效果进行验证,这样有助于公司投资及研发方向的确定,同样,若针对不同国家或地区的交通状况评价某些产品效果,则可根据其结果对该技术进行适应改进。

因上述智能驾驶系统评价的重要性,美国、欧盟、日本等发达国家或地区均给予了极大关注与项目经费支持。第三方机构与汽车企业纷纷开展了智能驾驶系统评价相关的理论与实验研究。由于智能驾驶系统集成了感知、决策和控制等多种功能,涵盖了交通、人工智能、机械等多学科领域,该

系统的评价方法,不能简单借鉴以往被动安全技术评价方法。两者的区别主要有以下几点:

1. 测试环境的未知性

被动安全技术的应用不会影响驾驶员的驾驶行为,那么,驾驶员所遇到的交通运行环境不会因其改变。因此基于目前已有的事故情况,提炼可表征交通环境的场景,即可对被动安全技术进行评价。相对而言,智能驾驶系统的功能则会与驾驶员行为产生未知的耦合作用,这一点在该系统应用前是不可预估的,造成其测试环境不能通过现有的交通环境分析提取得到。

2. 待评价对象的未知性

目前材料科学研究较为完善,且车身材料较为公开,能够方便地通过有限的碰撞数据来推断出未测试的碰撞情形下的结果。不同厂家实现智能驾驶系统功能所采用的硬件配置存在着较大差异,且其决策控制逻辑具有保密性。任一款智能驾驶系统对于第三方来说均是一个黑箱。

3. 评价指标的综合性

NCAP 仅依靠几个典型场景的综合评分即可对被动安全技术予以定量评价。而由于上述的测试环境与对象的未知性,评价指标仅能选取事故避免率、乘员损伤避免率、乘员死亡避免率等综合性指标。其中乘员损伤避免率因其能够综合反映由于技术应用带来的安全性的提高程度,而最为常用。

4. 评价方法的综合性

由于智能驾驶系统评价的特殊性(测试环境与对象的未知性、评价指标的综合性),该系统的评价方法应能够综合实现测试环境搭建、系统辨识、碰撞过程分析、乘员损伤计算等功能,并保证满足可扩展性、精度高、运算快速和结果可重复等要求。

智能驾驶系统评价方法由于其测试环境与对象的未知性、评价指标和评价方法的综合性等特点,已经成为国际研究热点。本书将从评价指标、系统辨识、有效性评价方法三个方面进行深入研究:

(1) 待评价对象辨识精度直接决定着有效性评价结果的准确性与合理性。合理地利用常见汽车运行数据,综合通用车辆动力学特点对待评价系统的关键参数和控制逻辑进行辨识,是实现有效性评价的关键基础工作。智能驾驶系统通常会将车辆控制处于稳态区域。这样轮胎滑移率或侧偏角均较小,参数间的相互影响较大。仅依靠时域信息的观测方法难以剥离其

他参数对待估计参数的影响。本书通过引入频域分析方法,实现该对象关键参数的估计,进一步采用神经网络学习智能驾驶系统的控制逻辑。

(2) 为计算乘员损伤风险,往往将其表述为车辆在碰撞前后运动状态变化量的函数,其中最为常见的是速度变化量。而该变化量的获得过程较为复杂,需要专业人员对事故情形进行深入调查。我国事故深入调查数据库的建立尚处于起步阶段,故拥有的样本量还不足以拟合出可信的以速度变化量为自变量的乘员损伤风险函数。那么,本书将基于事故数据,利用基本统计、回归分析、聚类分析等多种数据挖掘方法,论证车身变形深度作为评价乘员损伤风险指标的合理性,在此基础上,提出了基于变形深度的乘员损伤风险计算方法。

(3) 有效性评价方法应同时满足可扩展性、精度高、运算快速和结果可重复等要求。首先,不同智能驾驶系统的评价方法是相通的,考虑到所提出的评价方法将应用于不同智能驾驶系统评价,其应该采用模块化的方式进行搭建,通过替换待评价对象所对应模块,避免重复性开发工作。其次,该方法只有在具有较高精度的条件下,才能够应用于车辆智能驾驶系统的匹配验证和损伤风险估计,因此所提出的汽车智能驾驶系统有效性评价方法应该尽量准确反映车辆在不同场景下的运动状态与碰撞过程。再次,由于随机变量多样,运行场景数目繁多,评价过程中的运算快速性必须得到满足。最后,为了比较相同工况下不同算法的效果以便设计合理的控制参数,该方法需要满足可重复性。

1.2 汽车智能驾驶系统有效性评价方法研究现状

汽车智能驾驶系统有效性评价方法研究总体分为四大类:实车道路实验、测试矩阵、危险场景评价、蒙特卡罗仿真。

1.2.1 评价方法详细介绍

1.2.1.1 实车道路实验

自然驾驶实车道路实验(naturalistic field operational tests, FOT)^[1]已经被广泛采用,用以评价不同种类的智能驾驶系统。在自然驾驶实车道路实验中,由多名驾驶员驾驶着装配有该系统的车辆,进行长时间的数据收集工作^[2]。美国进行了大量的 FOT 项目,见表 1.1。

表 1.1 美国主要的 FOT 项目

项目名称	执行机构	时间	里程/ 1000 英里 ^①	传感器	车辆 数目	驾驶员	关注点
100 Car Naturalistic Driving Study ^[3]	弗吉尼亚理工大学	2001—2009 年	2000	摄像头	100	241	追尾碰撞
ACAS ^[4]	密歇根大学	2004—2005 年	137	摄像头、毫米波雷达	11	96	前撞预警
RDCW ^[5]	密歇根大学	2005—2006 年	83	摄像头、毫米波雷达	11	11	车道偏离预警
SeMiFOT ^[6]	密歇根大学	2008—2009 年	107	摄像头、毫米波雷达	14	39	多种驾驶辅助系统
IVBSS ^[7]	密歇根大学	2010—2011 年	815	摄像头、毫米波雷达	26	126	集成式预警
谷歌无人驾驶车	谷歌	2012 年至今	1300	激光雷达、摄像头、毫米波雷达	约 50	未知	全自动无人驾驶

① 1 英里=1.609344 千米。

其中,由弗吉尼亚理工大学主持的项目 100 Car Naturalistic Driving Study 关注于事故发生的原因。它的数据被用于研究在紧急事件、接近事故、事故发生等条件下的驾驶员行为、道路环境和其他影响因素^[3]。由密歇根大学主持的多个 FOT 项目(包括 the Automotive Collision Avoidance System Project(ACAS)、the Road Departure Crash Warning(RDCW)、the Sweden-Michigan Naturalistic Field Operational Test (SeMiFOT)、the Integrated Vehicle-Based Safety Systems(IVBSS))主要关注于评价多种驾驶辅助系统对驾驶员行为的影响。以谷歌^[8]、百度^[9]、奥迪为代表的无人车品牌纷纷在美国加利福尼亚州申请了无人驾驶测试牌照,被允许进行上路测试。

FOT 虽然能够全面分析一个系统的适用性,但具有明显的局限性。由于危险事件在自然驾驶过程中出现的概率极低^[10],因此 FOT 需要进行极长时间的实验和庞大的资金支持^[11,12]。这种方法可以用于整车厂进行技术的开发验证,不适用于第三方对于不同产品的快速评价过程。

1.2.1.2 测试矩阵

测试矩阵这种方法^[13]首先根据 FOT 或事故数据等进行分析,定义一系列特定场景,然后令待评价车辆在这些场景下进行测试,最后根据这些车辆的表现给出综合评分。这种方法与 EuroNCAP^[14]中碰撞实验的理论方法是一脉相承的。EuroNCAP 作为汽车安全性能评价的标杆,2011 年正式要求对汽车的主动安全系统进行评价,最初只是对车辆的电子稳定性控制(electronic stability control, ESC)性能进行评价,到 2013 年底,其又将 AEB 列入到评价对象中^[15],而这个测试标准就是一种测试矩阵方法。其他国家的 NCAP 组织^[16-18]的评价体系相对落后,但发展趋势与 EuroNCAP 保持一致,也正在针对 AEB 制定相应标准。

国际上有几个正在进行的或最近完成的性能测试方法,该方法关注纵向主动安全系统,如:前撞预警(forward collision warning, FCW)和自主紧急制动系统。参与研究的组织有 ADAC^[19]、AEB^[20]和 vFSS(BASt)^[21]。欧盟支持多个研究测试矩阵方法的项目,如 TRACE^[22]、APROSYS^[23]、interactIVe^[24]、ASSESS^[25]和 AsPeCSS^[26]。这些项目以国家事故统计数据 and 事故深度调查数据库(如 German In-Depth Accident Study 数据库,以下简称 GIDAS^[27])为基础,提炼出针对行人、骑车人、轿车、卡车等的 AEB 测试场景。

美国也有相应的研究。早在 Crash Avoidance Metrics Partnership(CAMP)项目^[28-30]中就已总结出了 26 个 FCW 系统的测试场景。虽然完成这些场景的测试仍需要两周时间^[31-33],但这已经比 FOT 要缩短许多。Development of Performance Evaluation Procedures for Active Safety Systems^[34]项目提出了 11 种场景来测试两种辅助制动系统的表现。

尽管各研究组织和项目所参照的数据库不同,但最后提出的场景^[35-38]都颇为相似。场景均包括了前方车辆制动、恒定的速度行驶和前车静止等工况。也有些组织还提出车辆切入和交叉口的场景,场景中包括易受伤害的道路使用者,如行人和摩托车。

由于我国事故深入调查数据库研究刚刚起步,没有前述的大型纵向研究项目支撑。相关学者在横向项目支持下开展了区域性的测试矩阵研究。同济大学朱西产^[39]基于自然驾驶数据,挑选出其中的危险情形数据,进一步提炼出可用于评价 AEB 的区域性测试情景。同济大学王宏雁^[40]基于所建立的事故深入调查数据库,得到了上海地区典型的事故场景。中国汽车

工程研究院股份有限公司则是通过对国外的 AEB 测试工况进行总结,并结合我国交通特点,提出了适应国内交通状况的 ADAS 测试场景和评价规程^[41]。

基于提炼出的场景,道路实验、硬件在环实验、驾驶模拟仪等手段均可以被用来测试待评价的系统。其中,计算机仿真方法可以进一步降低评价过程所需的时间与资金消耗,而且这些评价结果可以复现验证。然而,如果智能驾驶系统的功能与驾驶员行为产生未知的耦合作用,那么基于人驾驶状态下的事故数据分析得到的测试环境并不能表征该系统应用后所遇到的交通环境。

1.2.1.3 危险场景评价

危险场景评价方法基于事故再现数据库,仿真对比有无待评价智能驾驶系统间的乘员损伤差异。这种方法起源于对 ESC 的评价研究。有些学者利用事故再现软件仿真与 ESC 系统相关的事故场景,对比加载该系统 and 未加载该系统两种情况的事故结果,算出 ESC 系统在这种场景下的有效性^[42]。Badea-Romero^[43,44]将该方法应用于对辅助制动系统的评价。Jamson^[45]基于已有的事故场景,利用驾驶模拟仪实现了对 FCW 的评价。江丽君^[46]利用行车记录仪中的危险工况数据,进行统计分析,结合驾驶员反应时间,尝试得到适合我国城区道路的车对车的 AEB 测试情景。刘颖^[47]则采用同样的方法得到了 AEB 针对行人工况的测试场景。

这种方法与测试矩阵方法类似,其前提条件是待评价系统不可影响事故场景出现的可能性。与测试矩阵方法的区别在于,其可以详细分析待评价系统所带来的损伤变化^[48]。但这要求待评价系统是已知的,而整车厂往往将待评价系统作为一核心竞争力,并不会向第三方公开。而且若要得到该系统对整个交通环境的有效性,则需知道每个仿真的事故情形在整个交通环境所占的比重。

1.2.1.4 蒙特卡罗仿真

近来 Peng 基于 FOT 数据分析得到前车随机驾驶模型^[49]和有错误驾驶操作机制的跟车模型^[50,51]。Yang 基于这两种模型再现了事故发生的机制^[52],并应用于 FCW 的评价^[53]。Woodrooffe^[54]利用这种方法生成了 150 万种碰撞情景,并用以评价重型卡车应用碰撞预警和紧急制动系统的有效性。由于蒙特卡罗仿真过程中产生了大量无用的情景,Zhao^[55]提出了一种基于重要性采样理论的加速仿真方法,并成功实现了对于 ACC 的评价。

该方法的优势在于提炼的模型能够复现真实交通环境,这样就能够收获与实车道路实验相同的普遍适用性。而这也就带来了大量的仿真运算量,但总体较实验方法的时间仍短。由于其仿真建模的需求,待评价系统的被控对象和控制逻辑均需已知,这就极大限制了其应用的潜力。

1.2.1.5 综合评价方法

ASSESS 项目提出了一种综合评价方法(integrated assessment)^[25],总体思路是通过对事故数据分析提炼出典型场景,通过驾驶模拟台架和实车试验搭建驾驶员反应模型,结合反应模型设置机器人的操作,在试验场地做实验,测出汽车减速情况,并进一步进行驾驶员损伤程度的分析,根据损伤程度得出该系统的最后评分。

美国 NHTSA 所支持的项目提出了一种被称为 safety impact methodology (SIM)的评价体系^[56]。与综合评价方法相比,它采用的数据源更为丰富,将自然驾驶的数据、前期实车试验数据和事故数据库相结合来提炼评价场景,采用的评价指标也更为细化,加入了避免事故的数目、减少的致死率等。该方法不需要进行实车实验,仅利用仿真平台就可对系统进行评价,减少测试成本,适合对系统的前期开发。该方法得到了 Toyota、Volvo、GM^[57-59]等整车厂的普遍认可。

1.2.2 各方法归类对比分析

总的来说,从设备条件上,现有评价方法可分为两类:实验方法和仿真方法。实验方法又可分为两类:FOT 和标准测试(测试矩阵)。前者:对于不同智能驾驶系统的评价,具有普遍适用性;但花费大,耗时长,具有危险性。后者:标准统一,操作简单,可重复性好;工况单一,无法进行多工况下的统计分析,不具有普遍适用性,评价具有主观性。仿真方法也可分为两类:危险场景仿真和蒙特卡罗仿真。前者:可重复性好,无危险性,操作简单;但样本数量少,且反应真实情况的能力较差。后者:具有普遍适用性,但运算量较大。

就方法成立的前提条件来讲,测试矩阵和危险场景评价是相同的,均要求待评价系统的应用不影响正常的驾驶行为,这保证了从现有数据提炼的场景对该系统的适用性。而实车道路实验和蒙特卡罗仿真则是采用遍历法,对待评价系统没有特殊要求。

现有评价方法所采用的原始数据主要有两类:事故数据和 FOT 数据。

事故数据：能够提炼现有交通环境下的各种危险场景，能够建立乘员损伤模型；数据不完整，数据库不完善，需要大量的标准化整理分类工作。FOT数据：数据完整且数目多，能够建立随机仿真场景。

综合上述方法的优缺点，为适应未来更加复杂的智能驾驶系统评价，本书研究基于多源数据挖掘的蒙特卡罗与危险场景相结合的仿真评价方法。为实现该方法，需要解决如下两个难点：

(1) 待评价的被控对象与控制逻辑辨识。不同厂家实现智能驾驶系统功能所采用的硬件配置存在着较大差异，且其决策控制逻辑具有保密性。任一款智能驾驶系统对于第三方来说均是一个黑箱。合理地利用常见汽车运行数据，综合通用车辆动力学特点对待评价系统的关键参数和控制逻辑进行辨识，是实现对待评价系统建模仿真评价的前提条件。

(2) 评价指标计算方法。该方法往往是以车辆在碰撞前后运动状态变化量与乘员损伤风险的函数关系为基础的。其中最为常见的是速度变化量。而该变化量的获得过程较为复杂，需要专业人员对事故情形进行深入调查。我国事故深入调查数据库的建立尚处于起步阶段，故拥有的样本量还不足以拟合出可信的以速度变化量为自变量的乘员损伤风险函数。如何搭建适用于中国国情的全新乘员损伤风险模型是亟待解决的问题。

接下来的 1.3 节和 1.4 节将围绕这两个难点展开论述。

1.3 汽车智能驾驶系统辨识研究现状

智能驾驶系统包含了两个层次：被控对象和控制逻辑。针对系统辨识的研究现状，也需要从这两方面着手。

1.3.1 车辆关键参数估计

本书需要辨识出能够实现质点化车辆建模的关键参数，包括：反映轮胎运动状态与轮胎力之间关系的参数——轮胎刚度、表征控制输出期望值与实际车辆状态间差别的参数——时间延迟系数、反映车辆运动状态与轮胎力之间关系的参数——整车质量。下面分别介绍这三种参数的估计方法相关研究。

1.3.1.1 轮胎刚度估计

轮胎刚度与路面情况^[60]、胎压、轮胎型号等均有关。对于一个商品车

辆来讲,胎压监测设备已经成为标配,且胎压测量信号也作为车辆故障诊断所需实时获取存储的信号。轮胎型号亦为一固定可知的参数。故而,在文献中,轮胎刚度的估计往往与路面情况估计同时进行^[61]。

国内外对于轮胎刚度估计算法已经展开了深入研究。这些方法可以分为基于 Cause 和 Effect 两类^[62]。

基于 Cause 的方法利用光学或超声波传感器来检测路面的覆盖物(如水、冰、雪等)来估计路面种类。该方法需要外加昂贵的传感器^[63],并不具有对不同车辆和轮胎的普遍适用性。由于环境复杂且这些复杂因素与轮胎刚度的耦合机理并不明确,这种方法需要对不同路面进行大量的实验标定,因此其对于环境的鲁棒性也是不足的。

基于 Effect 的方法则是直接利用车辆与轮胎的动力学特性来估计轮胎刚度。由于这种方法仅采用车载传感器,能够在较少增加成本的前提下,实现刚度的估计,因此其研究和应用较为普遍。这些方法所考虑的动力学特性分别有大加速或制动工况下纵向滑移与纵向力的特性^[64]、转向工况下的侧偏角与回正力矩^[65]或侧向力间的特性^[66]、垂向运动时垂向位移与轮胎垂向力的特性^[67]。在不同的方向上,相关学者分别从观测原理^[68]、轮胎模型^[69,70]、观测方法^[71,72]等不同角度尝试进行轮胎刚度的估计。下面将从这些角度阐述该领域研究的现状与局限性。

观测原理:文献提出了利用纵向力-纵向滑移的曲线斜率,结合卡尔曼(Kalman)滤波器来估算轮胎刚度的方法。该方法适用于轮胎处于较大滑移且轮胎仍处在线性区的工况,即车轮轮速与车速存在一定差异,此时滑移率能够较为准确地得到,这也就显示出该方法的局限性,即不适用正常行驶工况(轮胎力很小,相应的滑移率难以观测得到),同时其不能兼顾到非线性区的工况。有些学者将该方法拓展到了存在侧向力的工况,利用侧向力-侧偏角曲线斜率来估算轮胎刚度。但是相比纵向力和滑移率,轮胎侧向力和轮胎的侧偏角并不能由传感器或相关公式直接计算得到。这两个量也是需要估计得到,且与轮胎刚度本身存在耦合关系。观测器需要同时观测多个状态量和参数,还需考虑它们的耦合关系,这样就带来了观测器模型复杂、所采用的矩阵维数高、求解速度慢等问题。进而针对其中的问题,有的学者采用双卡尔曼观测器^[73]解决矩阵维数高的问题,有的则采用电机转矩估计轮胎纵向力^[74]、侧向力传感器测量侧向力^[75,76]或 GPS 与 INS 相结合来估计质心侧偏角^[77],以降低观测器模型的复杂度。由于轮胎回正力矩或轮胎拖距与轮胎刚度的线性关系相比侧向力与它的关系更为显著,所以有些学

者利用力矩传感器^[78],通过转向系模型,来估计在一定侧偏角下的回正力矩,进而得到轮胎刚度,该方法相比前述方法能够提高估计准确度。

轮胎模型:观测原理部分讨论的方法大多采用的是线性化的简易轮胎模型,该模型能够表征一定滑移率或侧偏角下的轮胎力。当轮胎的状态处于接近抱死或滑转时,轮胎非线性特征明显,不能用上述模型表征,这就需要引入适合较宽滑移率或侧偏角范围的轮胎模型^[70]。如果考虑纵向力和侧向力同时存在工况下的轮胎刚度估计,则需要引入更为复杂的联合滑移轮胎模型^[79]。虽然复杂模型的引入能够提高估计准确度和适用范围,但也带来了更复杂的变量间的耦合特性,使得观测器的设计变得困难^[80]。

观测方法:当采用简单的车辆和轮胎模型时,则可采用线性观测器(如RLS^[81]、卡尔曼滤波器等)来实现参量的观测。但是如上所述,为了提高估计的准确度和适用范围,需要引入较为复杂的轮胎模型,这样也会导致整体车辆模型的复杂。如果仍然采用常规的线性观测器,不仅存在着矩阵维数灾的问题,还不能保证相应准确度的提高。现代观测理论提出了扩展卡尔曼滤波、无味卡尔曼滤波^[82]、粒子滤波^[83]等新型的观测方法。已经有一部分学者尝试用这些方法来估计轮胎刚度,验证了其对于估计准确度提高的有效性。

总结以上研究现状,存在的局限性有以下几点:小滑移率和小侧偏角下轮胎刚度不可观,这是现有方法均不能解决的难点问题;轮胎模型和观测方法复杂程度与估计准确性的矛盾关系;车辆模型和轮胎模型复杂造成观测器复杂、维数灾、求解速度慢;观测成本与观测器复杂程度的矛盾关系;大部分方法为单方向观测,适用范围局限。

本书所评价的智能驾驶系统在其绝大部分控制状态下,车辆处于稳态行驶,轮胎的滑移率或侧偏角较小,而且不同车辆级别所能拥有的车辆状态信号不同。为此,本书尝试解决在小滑移率和小侧偏角工况下,采用通用信号(轮速、方向盘转角等)实现轮胎刚度估计的问题。

1.3.1.2 时间延迟系数估计

车辆模型通常利用一阶延迟环节表述控制期望输出与真实车辆响应间的关系。

目前时间延迟系数估计的方法主要有以下两类:

(1) 采用标准阶跃信号输入,测量动态响应曲线,观察得到时间延迟。清华大学李升波^[84]为准确描述车辆执行器的动态特性,基于执行器的动态响应数据,结合面积法^[85]得到了执行器的延迟系数。而车辆系统时间延迟