

DOI:10.3969/j.issn.1003-5060.2025.11.002

基于多模态人机交互的汽车天窗控制系统设计

方博一¹, 汪春华², 耿展², 姚卫军², 江昊³, 陈奇¹

(1. 合肥工业大学 机械工程学院, 安徽 合肥 230009; 2. 黄山奥特斯电气股份有限公司, 安徽 黄山 245400; 3. 合肥工业大学 汽车与交通工程学院, 安徽 合肥 230009)

摘要:随着汽车科技的发展,智能座舱利用人机交互技术能够为驾乘人员提供更舒适的体验,然而目前的人机交互主要依赖基于摄像头的手势识别技术,存在成本高、算力需求大的问题。为此,文章设计了基于红外手势识别和语音识别的人机交互方案。红外手势识别方案以 STM32F103 为控制核心,设计 2 种实现方式,即采用 APDS-9960-3.3 RGB 传感器的集中式和使用通用红外传感器的分布式;语音识别方案采用 US516P6 纯离线语音识别芯片作为核心;通过搭建控制模块并与汽车天窗连接,对 2 种方案的性能进行测试。试验结果表明,所设计的方案在有效降低成本、减少算力消耗的同时,仍能保持较高的识别准确率。该研究为汽车智能座舱的发展提供了一种新的技术路径。

关键词:手势识别;红外传感器;语音识别;人机交互;智能座舱

中图分类号:U463.859 **文献标志码:**A **文章编号:**1003-5060(2025)11-1448-05

Design of vehicle sunroof control system based on multimodal human-machine interaction

FANG Boyi¹, WANG Chunhua², GENG Zhan², YAO Weijun², JIANG Hao³, CHEN Qi¹

(1. School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China; 2. Huangshan Automaster Electric Co., Ltd., Huangshan 245400, China; 3. School of Automobile and Traffic Engineering, Hefei University of Technology, Hefei 230009, China)

Abstract: With the development of automotive technology, intelligent cockpits using human-machine interaction techniques can provide a more comfortable experience for drivers and passengers. However, current mainstream human-machine interaction techniques rely on camera-based gesture recognition technology, which involves high costs and significant computational power requirements. Therefore, this paper proposes a human-machine interaction scheme based on infrared gesture recognition and voice recognition. The infrared gesture recognition scheme uses the STM32F103 as the control core, and two infrared gesture recognition methods, centralized and distributed, are designed using the APDS-9960-3.3 RGB sensor and general infrared sensors, respectively. The voice recognition scheme adopts the US516P6 pure offline voice recognition chip as the core. A control module is built and connected to the vehicle sunroof to test the performance of each scheme. It is verified that the designed system effectively reduces costs and computational power consumption while maintaining the accuracy of human-machine interaction recognition. This provides a new technical path for the development of intelligent cockpits in vehicles.

Key words: gesture recognition; infrared sensor; voice recognition; human-machine interaction; intelligent cockpit

收稿日期:2023-11-15; **修回日期:**2023-12-31

基金项目:安徽省重点研究与开发计划资助项目(2022a05020024);合肥工业大学校地合作产业创新引导资金资助项目(JZ2023YDZJ0011)

作者简介:方博一(2001—),男,安徽安庆人,合肥工业大学硕士生;

耿展(1983—),男,安徽黄山人,黄山奥特斯电气股份有限公司工程师,通信作者,E-mail:pthh230790@outlook.com;

陈奇(1979—),男,安徽肥东人,博士,合肥工业大学教授,博士生导师。

0 引言

传统汽车座舱的设计主要围绕驾驶体验和乘坐舒适性展开,而随着智能化技术的发展,汽车座舱的功能和交互方式得到了极大的拓展和提升。驾驶员能够通过智能座舱获取导航、路况等信息,座舱还能提供娱乐、语音助手等人性化服务,因此,实现汽车智能座舱的设计很有必要^[1]。此外通过智能座舱,车辆能感知道路环境、驾驶员状态等信息,为自动驾驶技术的发展提供了可能^[2]。

在汽车智能座舱中,人机交互的方式不再局限于传统的按键、触摸屏或语音交互,而是融入了更多如人体动作、表情、手势等自然交互元素,具备多种交互方式的人车设计意味着能够建立准确、高效的有效交互通道,对驾驶安全具有重要意义^[3]。此外,多模态人机交互也能提供更丰富、自然的交互体验,有助于提升驾驶安全性,并为驾驶员带来更加个性化的服务^[4]。例如:通过手势控制,驾驶员可以在不离开驾驶环境的情况下完成对车辆功能的选择和调整;智能座舱能综合分析驾驶员的语音、手势等多模态信息,从而更准确地理解其意图,提供更个性化的服务。

文献[5]提出红外结构光的交互手势跟踪识别方法,通过肤色像素分割图像,使用光流跟踪定位和跟踪手势,利用 HMM 模型跟踪手势运动的轨迹完成交互手势的识别;文献[6]开发了基于视觉手势识别的人机交互实验平台,采用传感器跟踪并采集手部运动数据信息,提取用户的手部特征并提前预测识别;文献[7]设计了基于深度学习的手势识别系统,通过摄像头获取用户手部图像数据,运用机器视觉库进行图像处理,再使用神经网络对处理后的图像进行手势识别模型训练。以上述基于视觉的手势识别为代表的人工智能识别方案具备识别精度高的优点,但其优势主要依赖于复杂算法和高成本传感器。

为解决现有手势识别方案成本高、算力需求大的问题,本文以汽车天窗控制系统为例,使用基于红外手势与语音识别的人机交互方案。通过设计分层控制系统架构及各模块控制算法,搭建测试平台,验证交互方案的效果与控制系统性能。

1 天窗控制系统架构

传统汽车的控制架构基于分布式方法,根据实现的功能采用不同的传感器和专用控制器,成本较高且资源浪费;而使用基于功能应用的划分

方法可有效解决此类问题^[8]。本文所设计的天窗控制系统作为智能座舱的一部分,其架构遵循基于功能应用的分层设计方法。该系统架构如图 1 所示。

1) 感知层:负责采集座舱乘员的手势或语音信息。通过红外传感器和语音识别系统,实时监测并采集座舱内的手势和语音信息。

2) 决策层:负责处理和识别感知层所采集的信息,并根据这些信息确定相应的动作。通过对手势和语音信息的处理和分析,识别出相应的指令,并根据这些指令生成相应的控制命令。

3) 执行层:根据决策层所确定的天窗控制命令来控制电机执行相应的动作。通过驱动电机的方式实现天窗的开启、关闭等动作,以满足座舱乘员的需求。

通过功能分层架构的设计,天窗控制系统可以更加高效、准确地采集和处理座舱乘员的手势或语音信息,从而实现更加智能化的天窗控制。

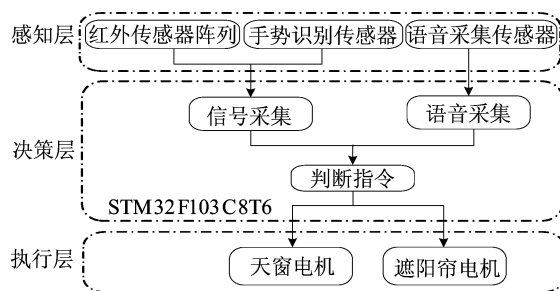


图 1 天窗控制系统架构图

2 控制原理设计

本文天窗控制系统按传感器类型可分为集中式红外手势识别模块、分布式红外手势识别模块和语音识别模块。

2.1 集中式红外手势识别模块

集中式红外手势识别模块以 APDS-9960-3.3 RGB 传感器为核心,该传感器集成光电二极管,用于探测红外 LED 发射并反射的红外光。

集中式红外手势识别模块的工作步骤如下:

1) 发射红外信号。集中式红外手势识别模块中的红外发射管发射出红外信号。

2) 反射和接收红外信号。当手部或其他物体在红外信号的范围移动时,部分红外信号会被反射回来,并被红外接收器所接收。

3) 信号转换。红外接收器接收到反射回来的红外信号后,将其转化为电信号,电信号的强弱

取决于反射回来的红外信号的强度。

4) 信号处理和识别。电信号通过算法进行处理和解析,通过计算手掌经过光电二极管的时间差,识别出手势的轨迹和动作。

集中式红外手势识别模块的算法流程如图 2 所示。

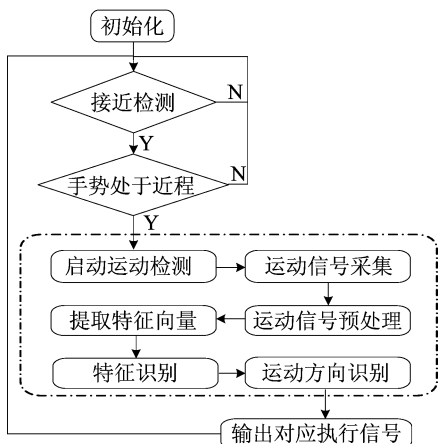


图 2 集中式红外手势识别模块算法流程

乘员在传感器前方做出手势时,红外传感器

识别动作方向,解析为特定指令。指令被传递给控制模块,控制模块根据信号控制天窗和遮阳帘的开启、关闭或暂停。

2.2 分布式红外手势识别模块

分布式红外手势识别模块由 4 个红外传感器组成阵列,红外阵列的上、下、左、右 4 个方向上分别分布着红外传感器,4 个红外传感器呈“十”字型分布。红外传感器的工作检测距离为 20~30 cm,检测范围为以传感器中轴线为中心的 35°锥形范围。

由于特殊的几何形状,传感器在检测接近中轴线的目标时具有较高的灵敏度,而在远离中轴线的区域时,检测灵敏度会逐渐降低,故使用 4 个红外传感器组成阵列,提高灵敏度。

当乘员在红外传感器阵列前使用手势挥动时,对应传感器前方有遮挡物,传感器响应。此时,红外传感器阵列根据响应的红外传感器随时间变化的分布和先后判断手势的类型,从而识别使用者的指令。分布式红外手势识别模块算法流程图如图 3 所示。

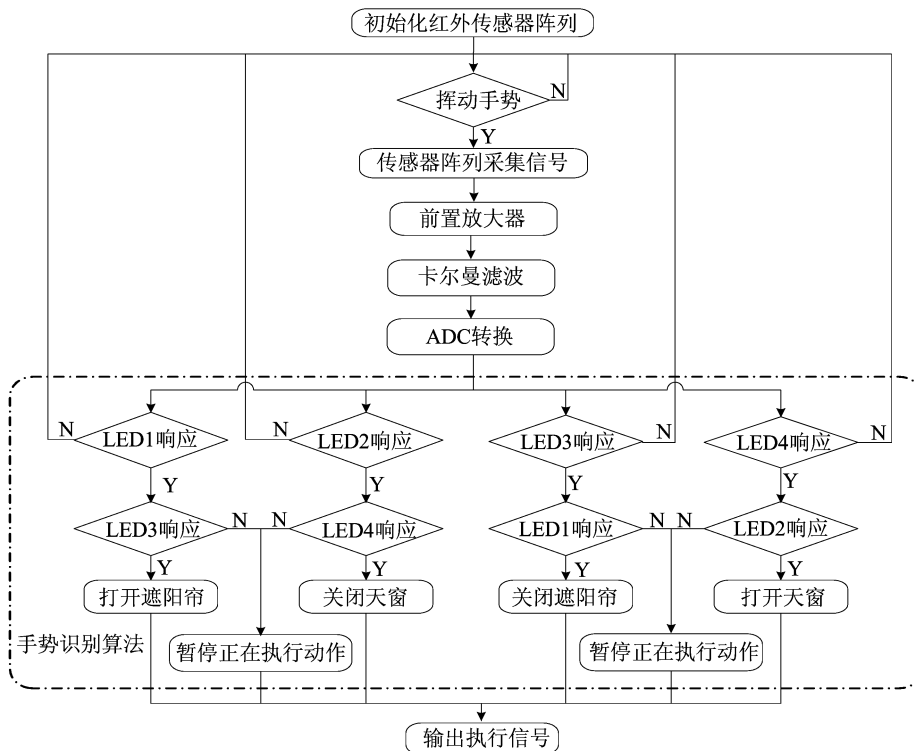


图 3 分布式红外手势识别模块算法流程图

2.3 语音识别模块

语音识别模块由控制芯片、麦克风和扬声器组成。麦克风采集声音,控制芯片识别声信号并发送指令,扬声器则播报所执行命令给用户形成

交互,其中控制芯片采用 US516P6 纯离线语音识别芯片。US516P6 芯片通过麦克风接收语音信号,再将信号进行转换和识别处理,最后发送控制信号,控制天窗和遮阳帘的开启、关闭和暂停。语

音识别模块算法流程如图 4 所示。

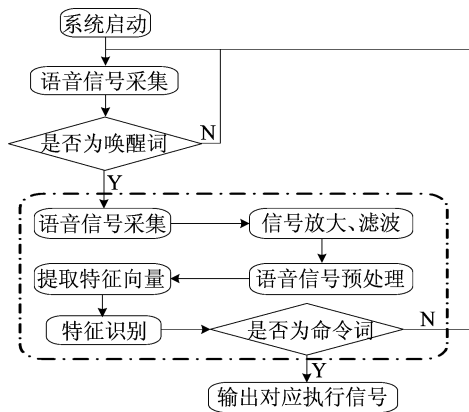


图 4 语音识别模算法流程图

3 测试平台搭建及试验

本文所设计的控制模块如图 5 所示。

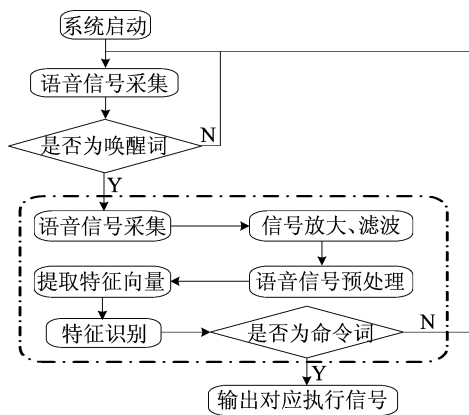


图 5 控制模块

天窗控制系统的分布式红外手势识别模块、集中式红外手势识别模块的手势定义见表 1 所列。

表 1 手势定义

指令	对应手势
打开天窗	手掌在传感器前从下向上挥动
关闭天窗	手掌在传感器前从上向下挥动
暂停天窗(分布式)	手掌在任意传感器前停顿
暂停天窗(集中式)	手掌与目前动作对应手势相反方向挥动
打开遮阳帘	手掌在传感器前从左向右挥动
关闭遮阳帘	手掌在传感器前从右向左挥动
暂停遮阳帘(分布式)	手掌在任意传感器前停顿
暂停遮阳帘(集中式)	手掌与目前动作对应手势相反方向挥动

本文搭建座舱模拟试验台,用于测试主驾驶和副驾驶位置使用所设计的人机交互模块的识别准确率。试验内容包括语音识别模块、分布式红

外手势识别模块和集中式红外手势识别模块的功能测试。试验以各模块的识别准确率为主要测试目标,试验人员分别在主、副驾驶位置,并分别使用语音识别方案、分布式红外手势识别方案、集中式红外手势识别方案模块控制天窗和遮阳帘,完成 50 次开启、关闭和暂停操作。各模块功能测试结果见表 2 所列。

表 2 各模块功能测试结果

控制方案	手势动作	准确识别次数	识别准确率/%
分布式红外 手势识别	打开天窗	44	88
	关闭天窗	46	92
	暂停天窗	45	90
	打开遮阳帘	47	94
	关闭遮阳帘	46	92
	暂停遮阳帘	45	92
集中式红外 手势识别	打开天窗	49	98
	关闭天窗	48	96
	暂停天窗	50	100
	打开遮阳帘	50	100
	关闭遮阳帘	47	94
	暂停遮阳帘	47	94
语音识别	打开天窗	50	100
	关闭天窗	50	100
	暂停天窗	49	98
	打开遮阳帘	50	100
	关闭遮阳帘	48	96
	暂停遮阳帘	50	100

由表 2 可知,分布式红外手势识别方案识别准确率最差,集中式红外手势识别方案居中,语音识别方案最好。

综合功能性、准确性、可靠性和成本因素对 3 种方案进行评价,结果见表 3 所列。

表 3 3 种方案的评价结果

控制方案	语音识别	分布式红外 手势识别	集中式红外 手势识别
功能性	合格	合格	合格
准确率/%	99	91	97
可靠性	优	良	良
成本	中	低	中

功能性方面,3 种方案的综合识别率均达到 90%以上,故功能性均为合格。可靠性方面,语音识别方案只受外界噪音的影响,可靠性为优;分布式红外手势识别方案和集中式红外手势识别方案均受乘员所处环境影响,若物品在传感器前挥动,则存在误识别的可能,故可靠性为良。成本方面,

(下转第 1458 页)

- the Acoustical Society of America, 1992, 92(2): 533-549.
- [12] BI C X, CHEN X Z, CHEN J, et al. Nearfield acoustic holography based on the equivalent source method[J]. Science in China (Series E), 2005, 48(3): 338-353.
- [13] ZHANG Y B, JACOBSEN F, BI C X, et al. Near field acoustic holography based on the equivalent source method and pressure-velocity transducers[J]. Journal of the Acoustical Society of America, 2009, 126(3): 1257-1263.
- [14] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [15] CHARDON G, DAUDET L, PEILLOT A, et al. Near-field acoustic holography using sparse regularization and compressive sampling principles[J]. Journal of the Acoustical Society of America, 2012, 132(3): 1521-1534.
- [16] FERNANDEZ-GRANDE E, XENAKI A, GERSTOFT P. A sparse equivalent source method for near-field acoustic holography[J]. Journal of the Acoustical Society of America, 2017, 141(1): 532-542.
- [17] BI C X, LIU Y, XU L, et al. Sound field reconstruction using compressed modal equivalent point source method[J]. Journal of the Acoustical Society of America, 2017, 141(1): 73-79.
- [18] HU D Y, LI H B, HU Y, et al. Sound field reconstruction with sparse sampling and the equivalent source method[J]. Mechanical Systems and Signal Processing, 2018, 108(9): 317-325.
- [19] HALD J. A comparison of compressive equivalent source methods for distributed sources[J]. Journal of the Acoustical Society of America, 2020, 147(4): 2211-2221.
- [20] PHOTIADS D M. The relationship of singular value decomposition to wave-vector filtering in sound radiation problems[J]. Journal of the Acoustical Society of America, 1990, 88(2): 1152-1159.
- [21] BORGOTTI G V. The power radiated by a vibrating body in an acoustic field and its determination from boundary measurements[J]. Journal of the Acoustical Society of America, 1990, 88(4): 1884-1893.

(责任编辑 吴亮)

(上接第 1451 页)

语音识别方案和集中式红外手势识别方案均使用专用传感器,增加了成本,成本为中;分布式红外手势识别方案使用通用红外传感器,有效降低成本,成本为低。

考虑可靠性、准确率和成本,使用集中式红外手势识别方案和语音识别方案效果最佳。

4 结 论

本文使用基于红外手势识别和语音识别作为智能座舱的人机交互方案,使用通用红外传感器、专用红外手势识别传感器和语音识别传感器的组合实现了多模态人机交互,实现了对智能座舱的天窗系统控制,经测试,可得结论如下。

1) 分布式红外手势识别模块识别准确率为 91%,可以正确识别大多数动作,且占用范围较大,需要专门设计空间容纳;在可靠性方面,基本可以识别所定义的手势;使用通用红外传感器,有效降低了成本。

2) 集中式红外手势识别模块识别准确率为 97%,体积较小,无需专门设计空间;在可靠性方面,多种环境下均可以识别所定义的手势;使用专用红外手势传感器,成本中等。

3) 语音识别模块在安静的座舱环境内可以识别所有指令;可靠性优秀,噪音信号基本无法被识别,有效避免了误识别情况;由于使用 US516P6 纯离线语音识别芯片,成本中等。

综上所述,集中式红外手势识别方案和语音识别方案效果最佳。本文设计的汽车天窗控制系统,能够在降低硬件成本、避免消耗大量算力的同时,为车辆乘员提供准确高效的交互通道,这对保障驾驶安全具有重要意义。作为智能座舱的组成部分,本文汽车天窗控制系统的设计可拓展至天窗以外的其他部分,对智能座舱领域的其他设计具有较好的借鉴意义。

[参 考 文 献]

- [1] 赵逸彬,刘晶,雷聪锐,等. 基于 MPV 车型的智能座舱设计[J]. 制造业自动化, 2023, 45(8): 164-166.
- [2] 孟健,王馨,李函遥. 自动驾驶与智能座舱的关系及技术难点[J]. 时代汽车, 2023(11): 132-134.
- [3] 孙博文,杨建明,孙远波. 汽车人机交互界面层级设计研究[J]. 机械设计, 2019, 36(2): 121-125.
- [4] 吉岩,晏江华,董玮,等. 智能座舱交互式功能测试技术研究[J]. 中国汽车, 2022(5): 39-4348.
- [5] 徐娜,郭广通. 基于红外结构光的交互手势变形跟踪识别方法[J]. 激光杂志, 2020, 41(8): 72-76.
- [6] 占宏,吴碧霄,杨辰光. 基于视觉手势识别的人机交互实验平台[J]. 实验室研究与探索, 2022, 41(7): 23-25, 29.
- [7] 李澍,蔡振雄,詹文杰,等. 基于深度学习的手势识别系统的设计与实现[J]. 现代信息科技, 2022, 6(11): 106-109.
- [8] 陈思宇,黄菊花,刘引引,等. 高度集成化的汽车智能座舱工程应用[J]. 农业装备与车辆工程, 2023, 61(8): 62-66.

(责任编辑 胡亚敏)