

本文引用格式: 彭龙江.融合大语言模型的智能音控大屏系统设计与实现[J].自动化与信息工程,2026,47(3):23-31.

PENG Longjiang. Design and implementation of an intelligent voice-controlled display system integrating large language models[J]. Automation & Information Engineering, 2026,47(3):23-31.

融合大语言模型的智能音控大屏系统设计与实现

彭龙江^{1,2,3}

(1.上海西派埃智能化系统有限公司, 上海 200233

2.重庆鲁班机器人技术研究院, 重庆 400714

3.重庆工程学院, 重庆 400056)

摘要: 为推进生产运营管理平台的“数智化”转型,提升大屏功能模块的智能化水平与人机交互能力,本文融合大语言模型(LLM)及相关技术,构建智能音控大屏系统。该系统的人机交互终端模块通过自动语音识别技术将移动端用户语音口令转为文本,并利用文语合成技术实现语音播报;以本地部署的 LLM 为基座,结合检索增强生成、提示词及领域知识库等技术,构建用于解析用户语音口令并生成控制大屏操作的文本指令的增强控制模块;采用 WebSocket 协议和 HTTP 协议实现各模块与大屏功能模块的数据通信。测试结果表明,该系统能有效识别和解析用户语音口令,并准确执行大屏功能模块的相关控制功能,其可靠性、准确性与鲁棒性均达到预期要求,能够为相关行业应用 LLM 及其相关技术进行系统智能化升级改造提供借鉴和参考。

关键词: 大语言模型;检索增强生成;提示词;知识库;音控

中图分类号: TP391.4;TP183

文献标志码: A

文章编号: 1674-2605(2026)03-0004-09

DOI: 10.12475/aie.20260304

开放获取

Design and Implementation of an Intelligent Voice-controlled Display System Integrating Large Language Models

PENG Longjiang^{1,2,3}

(1. Shanghai SIPAI Intelligent Systems Co., Ltd., Shanghai 200233, China

2.Chongqing Luban Robotics Technology Research Institute, Chongqing 400714, China

3.Chongqing Institute of Engineering, Chongqing 400056, China)

Abstract: To advance the “digital and intelligent” transformation of the production operations management platform and enhance the intelligence and human-machine interaction capabilities of the large-screen functional modules, this paper integrates large language models (LLMs) and related technologies to build an intelligent voice-controlled large-screen system. The system’s human-machine interaction terminal module uses automatic speech recognition (ASR) technology to convert voice commands from mobile users into text, and employs text-to-speech synthesis technology to generate voice announcements. Based on a locally deployed LLM, the system incorporates technologies such as retrieval-augmented generation, prompts, and domain-specific knowledge bases to build an enhanced control module capable of parsing user voice commands and generating text instructions for controlling the large-screen display. The WebSocket and HTTP protocols are used to facilitate data communication between the various modules and the large-screen functional modules. Test results indicate that the system can effectively recognize and parse user voice commands and accurately execute the relevant operational control functions of the large-screen modules. Its reliability, accuracy, and robustness all meet the expected requirements, providing valuable insights and references for related industries seeking to upgrade and modernize their systems through the application of LLMs and related technologies.

Keywords: large language models; retrieval-augmented generation; prompts; knowledge base; voice-control

0 引言

大语言模型（large language model, LLM）作为推动数字时代向智能时代转变的核心技术之一，具备知识融合^[1]、跨领域理解^[2]、通用任务处理^[3-4]、复杂推理与逻辑^[3,5-6]等方面的能力。当前，国内传统行业的大规模“数字化”平台基础建设已基本完成，正向基于 LLM 的“数智化”深度赋能阶段转型。相关行业的海量历史数据、算力基础设施^[7]以及政策标准^[8-9]等为 LLM 的落地应用提供了良好的基础条件和强劲的驱动力。同时，模型量化、知识蒸馏以及模型剪枝等^[10]技术可对高速迭代的 LLM 瘦身，构建适合端侧设备部署的轻量化模型，并借助微调技术^[13]优化 LLM 在特定任务中的能力、调整回答问题风格或增强推理能力，特别是融合产业知识的各种工业垂直领域模型^[4]，正为电力能源^[11-13]、冶金化工^[14]、生命科学^[15]、智能制造^[16]、具身智能^[17-18]等领域的高端产业应用场景升级赋能。文献[11]系统梳理了 LLM 在电力系统智能应用中的逻辑推理与程序理解进展，展望了其在负荷预测、运行规划、故障诊断与电力市场等场景的应用前景，并指出数据质量、可解释性与隐私保护仍是关键挑战。文献[13]提出的面向测井领域的专用模型“盖亚”采用了微调技术，在提升训练效率的同时，增强了对测井曲线序列模式与空间特征的感知能力。文献[14]面向冶金工业领域，通过集成多源异构数据构建的智能化多模态大模型，有效解决了复杂炉况解析、冶炼终点预测与过程控制等核心问题。文献[15]探讨了 LLM 凭借迁移学习能力处理多模态生物数据的潜力，并介绍了其在 DNA、RNA 及蛋白质序列分析等任务中的典型应用。文献[16]提出一种融合 LLM 与机器视觉的智能制造人机协同框架，通过融合视觉感知与语义理解，提升人机协作的智能化水平。

本文基于某集团的发展需求，借助 LLM 对现有生产运营管理平台的大屏功能模块进行音控智能化升级改造，以提升系统操作效率与用户体验。然而，LLM 在落地应用中仍面临若干挑战：1）生成内容可能包含不实信息的“幻觉”现象；2）对特定业务场

景适配能力弱、数值计算易产生误差；3）受限于训练数据的时效性，存在知识滞后等问题。

为解决上述问题，本文通过集成检索增强生成（retrieval-augmented generation, RAG）、程序辅助语言（program-aided language, PAL）模型、推理行动结合（reasoning and acting, ReAct）、自动语音识别（automatic speech recognition, ASR）与文语合成（text-to-speech, TTS）等技术对 LLM 进行针对性增强，并将其应用于大屏功能模块的操作控制优化，构建适应于业务需求的智能音控大屏系统。

1 基于 LLM 的智能控制技术框架

本文基于 LLM 的智能控制技术框架，对现有生产运营管理平台大屏功能模块进行音控智能化升级改造，如图 1 所示。

针对 LLM 存在的“幻觉”、知识滞后等问题，本文构建了语音输入输出交互模式与融合 RAG 技术、PAL 模型以及 ReAct 框架的 LLM 增强体系，并集成 ASR 与 TTS 技术实现语音输入输出交互模式，实现了对大屏功能模块的音控智能化升级改造。

1.1 LLM

LLM 通过大规模的数据、参数和算力进行预训练，可具备类似人类的语言文本搜索、分类、聚类、提取、生成、总结以及重写等能力。自然语言处理模型的架构经历了从循环神经网络（recurrent neural networks, RNN）^[19]到长短期记忆网络（long short-term memory, LSTM）^[20]再到基于 Transformer 架构的 LLM^[21]的演进，三种模型的核心结构差异如图 2 所示。

RNN 模型、LSTM 模型虽在序列数据建模领域占据主导地位，但存在如下缺陷：1）时序依赖的计算机制导致训练无法并行，制约了训练效率；2）随着序列增长，远程依赖的捕捉能力急剧下降，出现“长程遗忘”现象。LSTM 模型虽通过门控机制（输入门、遗忘门、输出门）增强了长期依赖的建模能力，但仍未突破串行计算的瓶颈。而基于 Transformer 架构的 LLM 采用编码器-解码器结构，其自注意力机制能够实现全局上下文建模，不仅能高效捕捉长距离语义关

联，还支持并行化计算。其核心原理为：通过位置编码对输入词嵌入的序列顺序信息进行并行化处理；利用自注意力机制（包括多头注意力）全局建模词元间的依赖关系；借助前馈神经网络进一步提取特征并捕

捉长程依赖；通过残差连接与层归一化稳定训练过程，缓解梯度问题；最终由线性层与 Softmax 层完成特征变换与概率化输出预测。

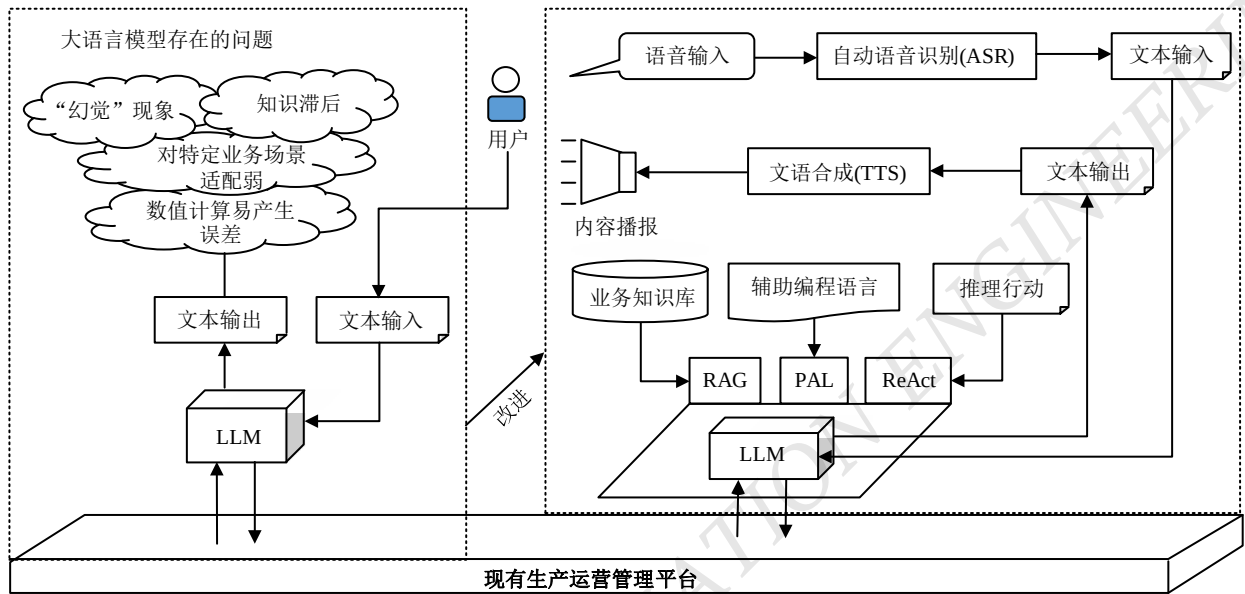


图1 基于 LLM 的智能控制技术框架

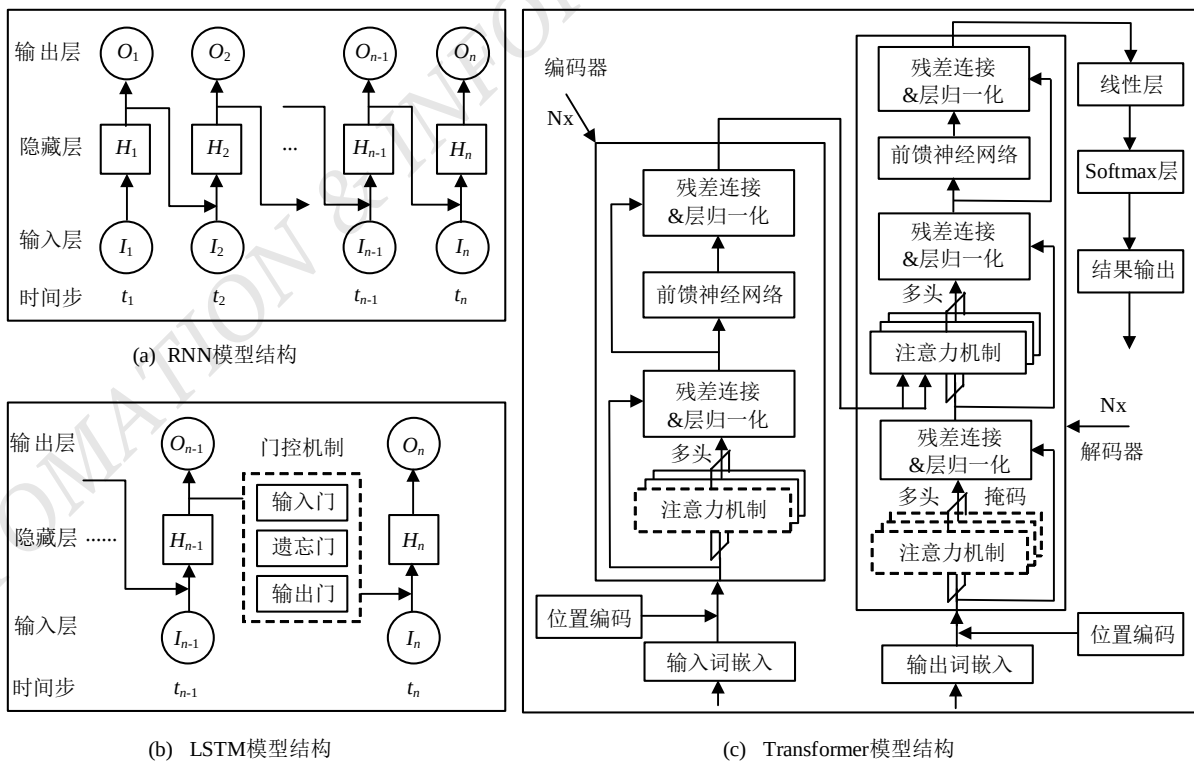


图2 RNN、LSTM 及 Transformer 模型结构

该架构凭借多头注意力机制实现并行训练和全局上下文建模，提升了 LLM 对序列语义的理解与生成能力及训练效率。因此，Transformer 已成为当前主流大模型统一的技术基座。

1.2 RAG 技术

基于对企业数据、商业秘密、个人隐私等信息的安全考虑，私域数据通常不作为 LLM 的预训练数据集，导致其在回答问题时易出现“答非所问、编造事实”的“幻觉”现象。本文采用 RAG 技术来解决该问题，其实现原理示意图如图 3 所示。

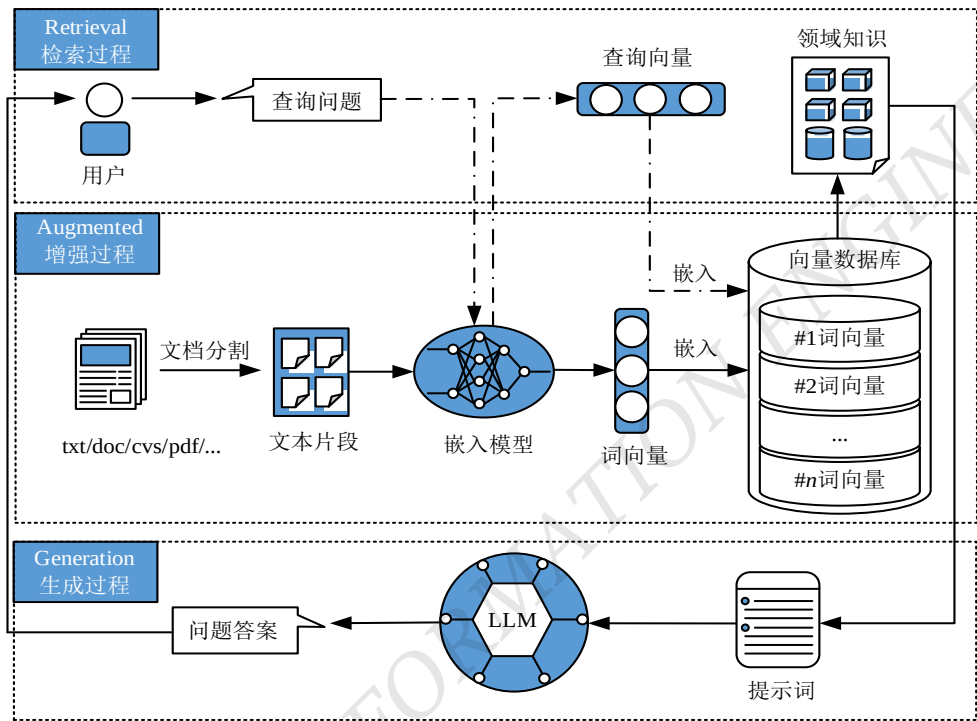


图 3 RAG 技术实现原理示意图

RAG 技术的实现过程如下：

1) 检索过程 (Retrieval)，用户输入查询问题后，系统采用与构建向量数据库完全一致的嵌入模型，将查询文本转化为查询向量，然后通过向量相似度计算检索出向量数据库中 with 查询向量语义相关的词向量对应的内容。

2) 增强过程 (Augmented)，主要通过文档分割将多源异构数据拆分为粒度统一的文本片段，再利用嵌入模型对其进行语义编码，生成高维词向量，最终将高维词向量与对应的原始文本片段一同存入结构化向量数据库，构建向量索引。从而实现对非结构化数据的语义化索引和高效检索。

3) 生成过程 (Generation)，系统将检索到的领域知识与查询问题融合，通过提示词构建标准化输入，

再输入到 LLM，生成精准、可靠的问题答案。

为进一步优化 LLM 的输出质量，本文采用小样本提示提供示例参考，通过约束系统限定生成范围并引入思维链技术引导多步推理，提升 LLM 在复杂任务中的准确性与逻辑连贯性。RAG 技术整合了知识表征、语义检索与提示优化，增强了 LLM 的知识利用与推理能力。

1.3 PAL 模型

针对 LLM 在处理计算任务时，因依赖概率性文本生成机制而出现的数值误差问题，本文利用 PAL 模型将逻辑推理与数值计算过程分离，提升计算可靠性。其主要实现原理如下：

1) 结合小样本提示与思维链技术，引导 LLM 将计算任务分解为结构化步骤，并生成可执行代码（如

Python 脚本)；

2) 外部解释器执行精确运算，并将计算结果返回 LLM 以整合输出。

1.4 ReAct 框架

针对 LLM 因预训练数据时效性限制而引发的知识滞后问题——即 LLM 无法掌握预训练截止日期后出现的新信息，导致生成内容过时或虚构，本文采用基于 ReAct 的框架解决方案。该方案通过融合小样本提示与思维链技术，引导 LLM 对问题进行动态推理与分步式外部交互，主要包括推理、行动与观察三个阶段：推理阶段解析问题或前序观察结果；行动阶段调用工具与外部环境（如实时文档、数据库或应用程序）进行交互；观察阶段获取外部反馈。该方案使 LLM 能够通过结构化工具交互突破自身知识局限，模拟人类任务分解与解决路径，不仅提升了答案的准确性与时效性，还增强了输出结果的可验证性与透明度，同

时避免了成本高昂的模型全量重训练。

1.5 ASR 与 TTS 技术

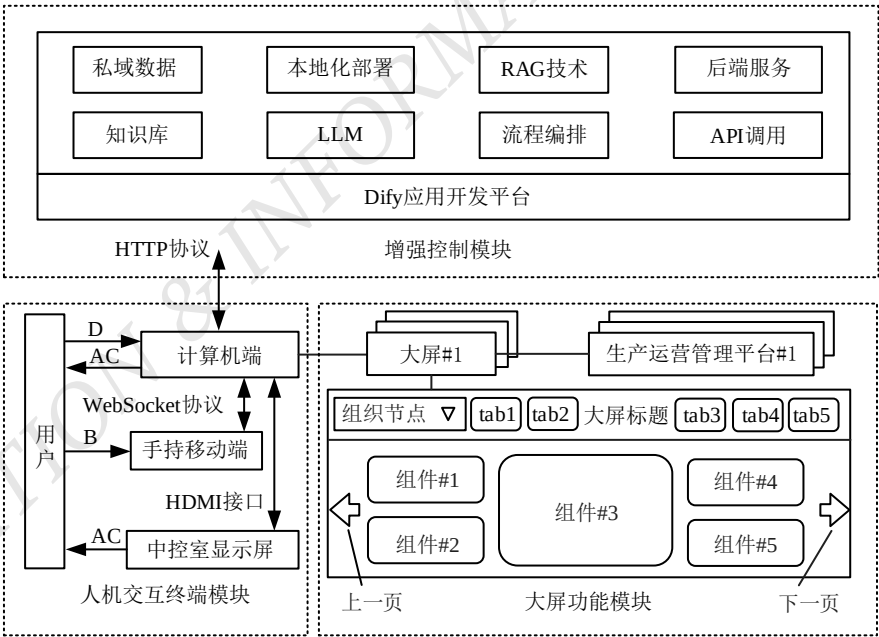
ASR 与 TTS 技术共同实现人机交互过程中语音与文本的双向转换。

ASR 技术通过将语音信号采样、量化并编码为数字信号；经特征提取（如 MFCC）得到频谱图，并切分为特征向量；利用声学模型匹配实现语音到文本的转换。

TTS 技术基于文本规范化、语言分析与语义理解，通过音素单元选择与连接生成语音特征，并利用信号处理算法合成波形，完成文本到语音的映射，实现语音播报功能。

2 系统设计

智能音控大屏系统主要由人机交互终端模块、增强控制模块和大屏功能模块三部分组成，整体架构如图 4 所示。



人机交互方式说明：A—眼看；B—口说；C—耳听；D—手动（鼠标、键盘操作）。

图 4 智能音控大屏系统整体架构图

2.1 人机交互终端模块

人机交互终端模块主要由计算机端、手持移动端和中控室显示屏等组成，负责完成用户与大屏功能模

块之间的交互。其中，计算机端中的 Web 页面为现有生产运营管理平台大屏功能模块的相关页面，负责文本化指令接收、大模型语义解析、控制逻辑调度等任

务；手持移动端负责语音口令采集、通过“长按采集、松开结束”的交互逻辑完成语音录制与 ASR 转写，将文本化指令通过 WebSocket 协议实时推送至计算机端的 Web 页面，实现大屏功能的远程无线操控；中控室显示屏负责大屏可视化内容的实时展示与动态更新，为生产运营决策提供直观的可视化交互界面，其采用高清多媒体接口（high definition multimedia interface, HDMI）与计算机端连接。

用户与相关终端交互的方式包括：

1) 眼看，用户通过计算机端、中控室显示屏查看大屏功能模块的相关 Web 页面；

2) 口说，用户通过手持移动端打开语音口令接收功能页面（采用跨平台多端应用开发框架 uni-app 实现），利用手持移动端麦克风发送动作指令，对大屏功能模块进行操作控制；

3) 耳听，用户通过计算机端、中控室显示屏配套的扬声器，与大屏功能模块进行语音交互；

4) 手动，用户通过计算机端的鼠标、键盘与大屏功能模块进行输入交互操作。

2.2 增强控制模块

增强控制模块基于 Dify 应用开发平台实现，主要由私域数据、本地化部署、RAG 技术、后端服务、知识库、LLM、流程编排、API 调用等部分组成，负责接收计算机端的 Web 页面、大屏功能模块的语音控制指令。实现将用户对大屏操作控制的语音口令进行转换、分析和识别，输出现有生产运营管理平台中大屏功能模块的计算程序代码可识别和执行的控制方法。其中，私域数据包括但不限于 JSON 或 TXT 格式的文件、组织架构数据文档以及方案组件数据文档等；本地化部署采用基于 Docker 容器的轻量容器化部署方案；采用 RAG 技术，使用不同功能节点完成文本指令的语义校验、逻辑推理、问题分类、输出格式标准化等核心任务的流程编排；后端服务提供私域数据的同步更新及高效检索、权限安全监控、指令处理操作日志等服务；知识库收录大屏功能模块的相关业务数据操作规范、大屏语音操作中指令的标准和

同义异构表述以及对代码映射关系、指令解析失败执行异常时的应对策略等核心信息；LLM 借助 Ollama 完成开源大模型的部署；流程编排完成各功能节点的逻辑串联，实现语音指令从接收、转换、解析、识别到生成可执行控制方法的全流程自动化；API 调用则为大屏功能模块提供便于集成和调用的相关控制指令。

2.3 大屏功能模块

大屏功能模块主要负责对生产运营数据进行可视化展示。为便于理解，从以下三个方面对该模块进行阐述：

1) 组织结构层级关系，组织节点自上而下主要包括集团、区域中心、项目公司、镇厂等层级，每个层级可根据需要配置不同的大屏页面展示内容；切换组织节点时，同层级节点的大屏页面展示内容不变，仅将数据同步切换到对应组织节点的生产和管理内容；

2) 大屏页面结构组成，一个大屏页面对应一个方案，一个方案由若干个组件组成，一个组件由若干个元素组成；用户可根据需求在默认方案基础上进行灵活配置；也可自定义创建方案，方案中配置组件库中显示不同内容元素的业务组件；

3) 大屏内容功能划分，大屏页面顶部为导航栏组件，主要包括组织节点、大屏标题、功能切换 tab 按钮（如企业介绍、数字运营、区域治理、科研创新、视频监控）等。

3 系统操作流程

智能音控大屏系统操作流程如图 5 所示。

首先，用户通过长按移动端语音口令发送页面的语音按钮获取语音口令，用户松开语音按钮后语音口令转换为文本指令后经 WebSocket 推送至计算机端，然后计算机端获取文本指令，通过 HTTP 协议调用 LLM 接口解析文本指令，获得指令类型和大屏控制参数，驱动大屏控制方法执行对应操作并更新展示。最终实现大屏的自然语音口令远程智能控制。

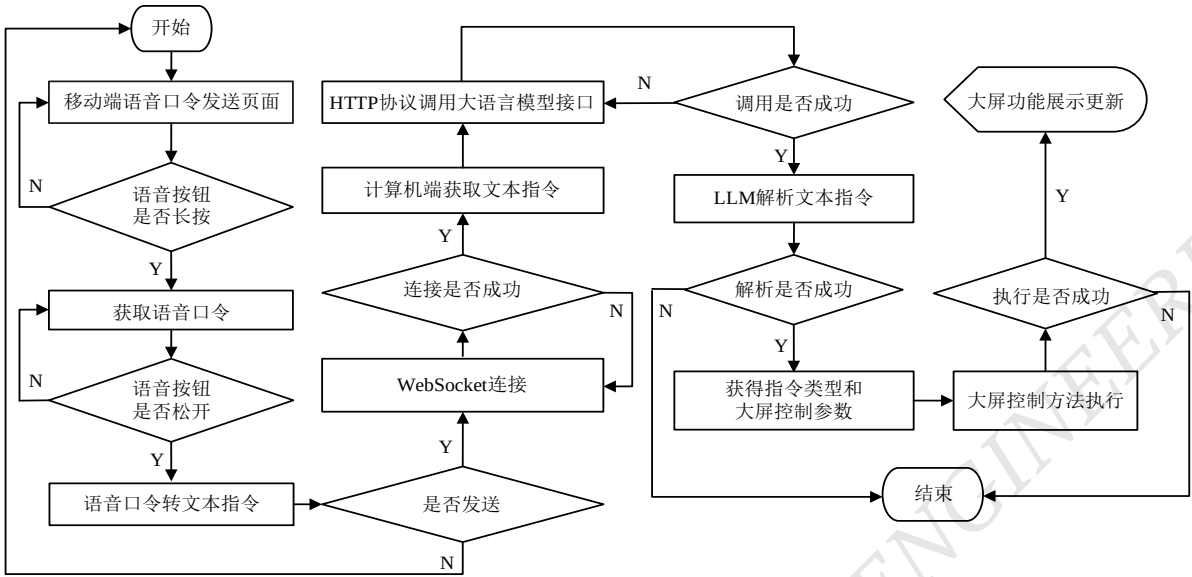


图5 智能音控大屏系统操作流程

智能音控大屏系统控制功能的实现过程如下：

- 1) 提供领域知识库文档给 LLM，该文档由不同指令类型及整个集团管理和业务类型的组织节点 ID、组件 ID 以及文本参数组成；
- 2) 采用 RAG 技术解析指令类型参数 `command-Type` 和大屏控制参数 `params`；
- 3) 根据指令类型参数及大屏控制参数分别与现有系统的不同大屏控制方法执行操作动作。

指令类型参数、大屏控制参数及执行操作动作之间的映射关系如表 1 所示。

表 1 系统功能控制指令映射集

指令类型参数 commandType	大屏控制参数 params	执行操作动作
prevPage	无	切换上一页
nextPage	无	切换下一页
switchCompany	组织节点 id:String	切换组织节点
locationCompany	组织节点 id:String	定位组织节点
switchTab	标签名称 tab:String	切换标签
componentFullScreen	组件 id:String	组件全屏
componentExitFullScreen	无	退出全屏
pageAnalyze	输入文本 txt:String	页面分析

4 系统部署、测试与优化

为提升智能音控大屏系统的可靠性、准确性与鲁

棒性，针对不同功能进行 20 轮测试并优化。

4.1 系统部署

1) 人机交互终端模块部署：操作系统 iOS（版本号 18.5）、处理器（A16，6 核）、存储容量 256 GB。在安装有微信（版本号 8.0.60）的手机上使用通过 uni-app 移动端应用开发框架开发的用户语音指令接收程序，将经过 ASR 技术处理的用户自然语言语音指令转换为文本指令。采用 WebSocket 协议将这些文本指令传输到智能音控大屏系统的 Web 页面。

2) 大屏功能模块部署：操作系统 Windows Server 2016（版本号 1607，OS 版本号 14393.4350）、处理器 Intel Xeon Processor (Skylake, IBRS) 2.29 GHz，8 核、内存 32 GB 系统盘 40 GB、数据盘 350 GB。使用 Nodejs（版本号 v16.17.0）完成前端应用的部署；使用 Nginx（版本号 1.22.0）作为 Web 应用服务器，代理不同的后端服务 Apache Tomcat（版本号 9.0.91）；结合 SSL 证书将大屏系统的 IP 地址映射为域名访问方式。

3) 增强控制模块部署：操作系统 Centos（版本号 6.8，64 bit）、处理器 16 核、内存 32 GB、系统盘 40 GB，数据盘 1 TB。通过 Docker（版本号:28.0.2）容器化技术，安装 Dify（版本号 v1.0.0）进行 AI 应

用平台的相关业务功能应用部署、安装 Ollama（版本号 v0.5.12）进行 LLM 的本地部署。

4.2 系统测试

4.2.1 测试方法

通过统计 LLM 对大屏控制文本指令解析为指令类型参数和大屏控制参数的准确率，评估智能音控大屏系统的准确性；通过 20 轮测试评估智能音控大屏系统的可靠性；通过模糊指令、同音词及口音人员测试，评估系统的鲁棒性。

4.2.2 测试结果与分析

用户语音口令（包括普通话口音和粤语口音）转换为文本指令后，经 LLM 解析得到 JSON 格式的指令类型参数（commandType: string）和大屏控制参数（params: []）如图 6 所示。



图 6 文本指令解析测试效果图

由图 6 可知，LLM 准确识别了“放大地图”和“企业介绍组件全屏”这两句文本指令对应的指令类型（组件全屏：componentFullScreen），params 并分别准确地给出了地图组件和企业介绍组件的唯一 ID。“切换到龙田镇”“切换 tab 为科研创新”以及“下一页”这三句文本指令，LLM 能够分别准确地识别出指令类型（切换到组织节点：switchCompany、切换标签：switchTab 以及下一页：nextPage），准确给出参

数类型和内容，为下一步驱动大屏切换动作的调用方法提供了准确的参数。

根据智能音控大屏系统的业务功能（如厂区数、大屏的 tab 标签数），确定各项测试内容的测试次数。系统测试的具体情况如表 2 所示。

表 2 系统测试的具体情况

测试内容	测试方法 (口令输入)	测试 次数	成功 次数	准确 率/%
切换组织节点	切换厂区为 XX 切换到 XX 等	228	226	99.12
定位组织节点	定位到 XX 等	162	160	98.77
切换 tab 标签	切换 tab 为 XX 等	5	5	100
切换到上一页	切换到上一页 上一页等	20	20	100
切换到下一页	切换到下一页 下一页等	20	20	100
组件全屏	放大 XX XX 组件全屏等	20	20	100
退出全屏	(退出 取消 恢复 缩小 还原)等	20	20	100
页面分析	帮我分析当前页面 介绍 XX 对 XX 部分进行分析	10	10	100

由表 2 可知：切换组织节点和定位组织节点的语音口令测试准确率分别为 99.12%、98.77%，其原因为组织节点名称存在同音词（如“霞山”和“峡山”）。后续针对该问题进行代码适配优化，其准确率提升至 100%。其他语音口令测试准确率均为 100%，表明系统的准确率、可靠性以及鲁棒性均达到预期目标。

4.3 系统优化

根据系统测试结果以及实际使用场景需求，通过以下三种方式对系统进行优化：

- 1) 记录 LLM 错误解析的组织节点名称，采用名称与唯一 ID 进行强关联匹配的方式，对组织节点名称进行代码适配优化。
- 2) 通过在领域知识库文档中增加与指令类型识别有语义关联且经过测试的可靠语料内容，增强 LLM 对多样化文本指令的泛化能力，进而提升系统对文本口令理解的鲁棒性。
- 3) 通过优化选择 ASR 与 TTS 技术，实现对普通话口音和粤语口音都能很好地进行语音和文本之

间的识别转换,赋予系统更具地域方言特色的语音交互能力,提升系统的鲁棒性和用户使用体验。

5 结论

依托现有生产运营管理平台中的大屏功能模块,结合由手持移动端设备构成的人机交互终端模块、集成 LLM 及相关技术的增强控制模块,构建了智能音控大屏系统。用户通过语音口令,将原有大屏功能模块的传统键盘、鼠标的人机交互方式升级为基于自然语音口令与 LLM 意图理解的智能交互模式。该系统构建了非接触式、口语化的交互入口,实现了用户操作意图的自动识别、理解与执行,拓宽了人机交互渠道,降低了交互操作的复杂度,提升了人机协同效率与系统易用性,进而提升了现有生产运营管理平台的整体智能化水平,较好地满足了平台升级改造的实际需求。未来,一方面可在现有语音交互及文本处理能力的基础上结合图像识别技术,进一步提升平台的多模态融合处理能力;另一方面可更深层次地挖掘平台业务需求功能和多源异构数据资源,借助现有集成的技术基础,推动系统由被动响应指令向主动推荐处置方案的智能认知决策能力跃升。

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献

- [1] 张鹤译,王鑫,韩立帆,等.大语言模型融合知识图谱的问答系统研究[J].计算机科学与探索,2023,17(10):2377-2388.
- [2] 许娜,陈曦,杨建平,等.大语言模型在土木工程领域中的研究现状、挑战及趋势[J].科学技术与工程,2025,25(21):8773-8783.
- [3] 徐月梅,胡玲,赵佳艺,等.大语言模型的技术应用前景与风险挑战[J].计算机应用,2024,44(6):1655-1662.
- [4] 籍欣萌,咎红英,崔婷婷,等.大模型在垂直领域应用的现状与挑战[J].计算机工程与应用,2025,61(12):1-11.

作者简介:

彭龙江,男,1989年生,硕士研究生,高级工程师,主要研究方向:智能控制系统集成、大语言模型应用开发以及智能化系统建设。E-mail: 450294634@qq.com

- [5] 文森,钱力,胡懋地,等.基于大语言模型的问答技术研究进展综述[J].数据分析与知识发现,2024,8(6):16-29.
- [6] 李荣涵,浦荣成,沈佳楠,等.基于思维链的大语言模型知识蒸馏[J].数据采集与处理,2024,39(3):547-558.
- [7] 钱德沛,栾钟治,刘轶.从网格到“东数西算”:构建国家算力基础设施[J].北京航空航天大学学报,2022,48(9):1561-1574.
- [8] 范德志,于水.生成式人工智能大模型助推实体经济高质量发展:理论机理、实践基础与政策路径[J].云南民族大学学报(哲学社会科学版),2024,41(1):152-160.
- [9] 陈鸿应.2025“工赋上海”创新大会在沪举行深度融合,打造工业智能新高地[J].上海化工,2025,50(1):2.
- [10] 王永威,沈弢,张圣宇,等.大小模型云端协同进化技术进展[J].中国图象图形学报,2024,29(6):1510-1534.
- [11] 赵俊华,文福拴,黄建伟,等.基于大语言模型的电力系统通用人工智能展望:理论与应用[J].电力系统自动化,2024,48(6):13-28.
- [12] 黄成燕,查小云,丁群晏,等.数据增强与规则引导的大语言模型电网法律答辩文书生成[J].国防科技大学学报,2025,47(4):180-188.
- [13] 陈掌星,张永安,李健,等.测井曲线重构的人工智能大模型[J].石油勘探与开发,2025,52(3):744-756.
- [14] 韩德浩,彭一帆,杨志浩,等.大模型驱动冶金行业智能化转型:技术路径与应用前瞻—以转炉炼钢过程为例[J].冶金自动化,2025,49(4):36-48.
- [15] 李晓晨,宋敏芳,刘扶苒.基于 Transformer 架构的大语言模型在生命科学中的应用与挑战[J].人工智能,2024(5):44-52.
- [16] 黄思翰,陈建鹏,徐哲,等.基于大语言模型和机器视觉的智能制造系统人机自主协同作业方法[J].机械工程学报,2025,61(3):130-141.
- [17] 王文晟,谭宁,黄凯,等.基于大模型的具身智能系统综述[J].自动化学报,2025,51(1):1-19.
- [18] 吴振宇,王子为,许修为,等.面向居家服务机器人任务规划的具身大模型[J].机器人,2025,47(4):517-525.
- [19] 刘建伟,宋志妍.循环神经网络研究综述[J].控制与决策,2022,37(11):2753-2768.
- [20] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation,1997,9(8):1735-1780.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. Advances in Neural Information Processing Systems. California, 2017,30.