

硕士论文中期检查报告

论文名称：智能家居语音控制系统的设计与实现

姓 名：张加杰

学 号：ZF1821334

专业方向：人工智能

学院指导教师：王丽华

企业指导教师：王宁

企业导师单位：百度时代网络技术（北京）有限公司

论文起止时间：2020 年 11月 ~ 2021 年 12月

2021 年 10月

**目 录**

[1. 课题简介 3](#_Toc25829)

[1.1 课题背景 3](#_Toc17342)

[1.2 主要研究内容和目标 3](#_Toc21364)

[2. 论文工作进展情况 3](#_Toc28160)

[2.1 开题报告工作计划 3](#_Toc5636)

[2.2 实际进展情况 4](#_Toc30961)

[3. 论文工作成果介绍 4](#_Toc14703)

[3.1 课题所实施的解决方案介绍 4](#_Toc3475)

[3.2 开题报告中所列关键问题的解决情况 5](#_Toc7164)

[3.3 创新性的方法、技术、成果 5](#_Toc11578)

[4. 论文后期工作及进度安排 5](#_Toc27974)

[5. 尚存的问题及措施 5](#_Toc5695)

[5.1 论文后期工作存在的困难和问题 5](#_Toc21164)

[5.2如期完成全部论文工作的可能性 5](#_Toc24392)

[参考文献 5](#_Toc31737)

# 课题简介

## 课题背景

论文研究课题来源于公司部门项目，我相继在百度智能生活事业群组小度智能音箱业务部和语音技术部门工作，部门主要产品是小度智能音箱及相关音频类产品，我的主要工作内容是小度智能音箱的系统开发，先后开发了小度智能音箱，小度智能车载支架，小度真无线蓝牙耳机等产品，主要功能模块包括语音识别SDK，OTA（Over the Air）空中升级，蓝牙配网，底层Linux系统和驱动等功能。受当前流行的智能语音控制能力启发，结合我当前从事的公司项目项目，我决定进行深入研究。

当前智能音箱普遍重云轻端，控制功能薄弱，强依赖网络，将主要算法如ASR（Automatic Speech Recognition），NLP（Natural Language Processing）等放在云端，音箱端仅负责数据采集和TTS（Text To Speech）播放，端云之间通过一套自研协议进行通信，基于这种架构形式可以有效削减硬件成本，相应的音箱端可实现的功能受到限制，无法发挥嵌入式系统控制优势，音箱端可新增功能受限，降低了灵活性。基于此，设计一款基于深度学习的嵌入式高性能语音控制系统非常有必要，应用场景可涵盖手机，车载，智能家居，行业智能终端等，该系统可将唤醒，识别和控制功能全部放在音箱本地系统中，可实现离线语音唤醒和识别功能，并能极大程度增加系统的控制功能，对该控制系统进行有效设计和实现是本课题的主要内容。

## 主要研究内容和目标

我的研究内容主要是在国内外研究现状的基础上，根据起优点和不足定制化一个智能家居语音控制系统，主要方面有：

1. 嵌入式开发板和系统选型。
2. 语音唤醒功能。
3. 语音识别。
4. 系统控制。

# 2. 论文工作进展情况

目前论文总体进度为80%，框架已基本完成，绪论和系统需求分析部分已完成，系统总体设计完成85%，需要进一步深化细节，增加流程图等。系统详细设计与实现部分实现70%，其中离线识别部分还需要进一步整理，蓝牙和WiFi通信部分缺少必要的结构图。系统测试分析还需要调整格式。

系统控制部分包含蓝牙，WiFi，红外控制，内容比较多，完成80%左右。

，还有较多需要补充和调整的地方。

## 2.1 开题报告工作计划

|  |  |  |  |
| --- | --- | --- | --- |
| 个人任务计划 | 时间：2020.11—2021.7 | | |
| 时间 | 主要任务目标 | 论文阶段 |
| 2020年11月 | 整体系统方案设计 | 熟悉毕设内容，完成开题报告 |
| 2020年11月-2020年12月 | 开发板选型+系统搭建 |  |
| 2020年12月-2021年1月 | 语音降噪算法和语音唤醒调研与实现 |  |
| 2021年1月-2021年3月 | 语音识别算法调研及实现 | 撰写论文 |
| 2021年3月-2021年5月 | TTS算法实现和无线控制模块 | 中期答辩 |
| 2021年5月-2021年6月 | 系统测试完善与整体满足度评估 | 撰写论文 |
| 2021年7月 | 项目完成 | 毕业答辩 |

## 2.2 实际进展情况

开题报告中时间规划不够成熟，忽略了可能的影响因素，比如工作变动和个人突发事情等因素，实际进展如下。

|  |  |  |  |
| --- | --- | --- | --- |
| 个人任务计划 | 时间：2020.11—2021.12 | | |
| 时间 | 主要任务目标 | 论文阶段 |
| 2020年11月 | 整体系统方案设计 | 熟悉毕设内容，完成开题报告 |
| 2020年12月-2021年1月 | 开发板选型+系统搭建 |  |
| 2021年1月-2021年3月 | 语音降噪算法和语音唤醒调研与实现 |  |
| 2021年3月-2021年6月 | 语音识别算法调研及实现 | 撰写论文 |
| 2021年6月-2021年8月 | TTS算法实现和无线控制模块实现 | 中期答辩 |
| 2021年8月-2021年10月 | 系统测试完善与整体满足度评估 | 撰写论文 |
| 2021年12月 | 项目完成 | 毕业答辩 |

# 3. 论文工作成果介绍

## 3.1 课题所实施的解决方案介绍

1）Jetson Xavier NX，NVIDIA开发套件，达21 TOPS的加速计算能力可并行运行现代神经网络并处理来自多个高分辨率传感器的数据，这是完整AI系统的要求。

2）唤醒模块，采用了snowboy，一个开源的、轻量级语音唤醒引擎，可定制唤醒词。

3）识别模块，离线模式采用了kaldi，kaldi是一个语音识别开发包，里面含有很多语音识别所需要的模型，比如：gmm，sgmm，dnn和hmm。你可以通过自己的数据在这平台上训练得到自己的模型，然后用于识别。你需要做的也就是改些脚本，选择适合自己的语音库。在线模式采用了百度AI开发平台语音识别能力。

4）语音合成模块，采用了百度AI开放平台语音合成能力。

## 3.2 开题报告中所列关键问题的解决情况

关键问题是离线识别，采用了kaldi软件开发包，数据集是thchs30，[清华大学30小时的中文语料库](http://www.baidu.com/link?url=m81r365Eij0hh5Q55miUibs6WYbUhGZBcKRirovUz5XRVc6QRGNOdsBu89r9K02jDm3LTZvqMmWXXCCduK6seP7_XBy8SP9OIQffK9ojzZq" \t "https://www.baidu.com/_blank)。目前已经训练出识别模型，可识别出该系统所需要的语音指令。打开空调，关闭空调。打开电视，关闭电视。灯光调亮，灯光调暗。

该功能从学习调研到方案对比耗时较长，虽然整体实现过程并不复杂，但涉及知识点非常多，目前该关键问题已基本完成。

## 3.3 创新性的方法、技术、成果

1. 系统支持离线运行，系统的核心功能均支持离线运行。语音唤醒模块是离线的，语音识别模块可同时支持离线和在线。

2. 同时支持红外，WiFi，蓝牙，可控制绝大多数的家居设备，WiFi可支持控制大部分智能设备，红外可支持带红外接收器的家居设备，如电视，空调等。蓝牙配合蓝牙智能开关可控制一切用电设备的开关操作。

# 4. 论文后期工作及进度安排

论文当前主体基本完成，但从论文格式和结构来看还存在诸多问题，后面的时间会及时调整和纠正，另外内容有如下几个方面需要修改：

1. 系统架构部分，需要进一步说明各模块细节，辅以流程图时序图等。
2. 系统详细设计与实现章节，目前结构还不够清晰，需要继续梳理，kaldi部分不必要的细节要去掉，尽量将重要内容描写清楚。
3. 系统测试部分还没有对识别率进行合理量化，目前只能靠句准和识别结果进行判定。
4. 一些自己无法发现的问题还需要依靠老师的指导。

# 5. 尚存的问题及措施

## 5.1 论文后期工作存在的困难和问题

目前论文结构不够系统化，有些细节还没有联系起来，比如识别结果的处理与外设的交互部分。

系统测试部分测试结果不好评判，如唤醒率，识别率，可接受程度等。

## 5.2如期完成全部论文工作的可能性

论文结构不够系统化，需要静下心来慢慢打磨，对于系统测试部分，可参考市面上相关产品的评价和用户感受，梳理得到一个评判标准。

# 参考文献