2026届课题

开题报告

**课题名称 低成本普惠型智能家居方言**

**指令转换器的研究**

**姓 名 陶理**

**指导教师 张继禄；殷立峰**

**日 期**

|  |  |  |
| --- | --- | --- |
| **课题研究领域** | | **工学** |
| **一、课题背景及意义** | **（课题的立题依据及研究价值和社会意义）**  随着普通话的推广，为了更好地与各族人民交流，方言的盛行注定会成为过去时。虽然方言不再常用，但是大地域的方言都还以各种各样的方式流传甚广，如上海话的沪剧、四川话的川剧、各歌手使用粤语所演唱的粤语歌曲等。如今讯飞科技等公司、组织都研发了此类大方言转换为人工智能的算法，并已经付诸实践，在微信等社交平台的语音识别中运用。而小众方言在使用人群数量上无法与大方言相比，自然在被关注度上也无法与大方言争锋。但是在这些小众方言流行的地域上，新一代的年轻人同样大多都说普通话，若是没有进一步的保护措施，这类的小众方言将很快走下历史的舞台。  各地迥异的方言是中华民族传承千年的瑰宝。以吴语为例：现如今普通话已经只剩下四个声调，即所谓的阴平、阳平、上声、去声，但是在古汉语中还有一个声调为入声，在古诗文的韵律中时常出现；如在浙江义乌话中，声调达到了10个。因为有入声，所以使用方言阅读古诗文能够接近古人在创作时的读音，相较于普通话更有韵味。若此般瑰宝失传，则是中华民族乃至世界的重大文化损失。若我们这一代人不及时研究，那么小众的方言就会更加濒危，增加研究的难度以及造成在细节上不可逆的损失。  虽然方言的小众化是不可逆的，但是及时的研究能够更多地保留这一非物质文化遗产。因此需要尽快进行研究。  如今语音识别的发展随着计算机算力的提升、算法的优化而使用人群增加，在各种应用、设施中都有语音识别技术的身影，就比如说智能家居。  智能家居是以语音识别为基础，智能家具通过语音模型理解人的指令进而执行指令。但是语音识别模型的缺陷在于大多模型是基于普通话进行训练的，所以导致对于方言的支持很差。而对于方言识别来说，如今只有大方言的语音识别模型，小众方言并没有对应的语音识别模型。这导致说小众方言的智能家居使用者只能使用普通话进行指令的描述。而小众方言的使用者占我国总人数的很大比重。虽然现在普通话推广范围很广，但是还是有很多人（以老年人为主）对于普通话的运用并不熟练。若一致使用普通话，会降低使用方言的频率，进而会使得方言的消失速度更快，特别对于小众方言。因此针对小众方言研究方言指令转换器是十分有必要的。 | |
| **二、课题研究综述** | **（课题研究领域内的代表作品与观点及可能的发展方向）**   1. 代表作品与观点   吴永焕[1]就方言的保护在必要性和紧迫性的方面进行了论证，强调了要抢记方言资料，尽可能延缓方言特征消失速度；黄涛[2]强调了方言在文化价值上的重要性；武瑞丰[3]提出了人工智能在方言建档中有很大的作用和优点；石佳影、黄威[4]利用梅尔倒谱系数（MFCCs）对四川方言进行特征提取，并使用了基于Kaldi平台的深度神经网络（DNN），构建了基于语音与普通话的四川方言语料库；杨波[5]构建了循环神经网络（RNN）为基础的声学模型，并以此搭建了桂柳方言的语音识别系统；张宇聪[6]分别运用了隐马尔科夫模型（HMM）和长短时记忆模型（LSTM）构建了语音识别模型和藏语声学特征提取器，并以此完成了藏语语音识别系统；余陆峰[7]使用多种深度学习算法，并利用TensorFlow框架进行实现，最终对比得出了效果最佳的关于客家方言的语音识别系统；彭煦潭等[8]利用“无监督跨语言词向量”的算法，分别研究了12种不同的语言并构建了语音识别系统。  科大讯飞所推出的讯飞输入法目前支持23种方言输入（截止至2023年10月9日），并推出了方言保护计划，通过征集各语言的语音数据进行翻译的训练。彭煦潭与团队伙伴司靖辉、陈冠毅等人利用百度飞桨PaddlePaddle框架，并运用“跨语言词向量”的算法，使用AI自动生成了独龙族语言和普通话的词典，其中独龙族语言的使用者不足万人。  河北工业大学刘阳、王巍等人[9]曾提出一种智能家居方言语音识别的装置，通过用户录音提升产品的识别准确率。  珠海格力电器股份有限公司和珠海联云科技有限公司提出了一种基于方言交互来控制智能家居的专利，通过获取用户的控制指令，直接控制对应的智能设备执行相关的指令。   1. 可能的发展方向 2. 落实于各项语音识别的应用：现如今各项需要语音识别技术的应用如：智能家居、语音对话问答等都难以做到对于方言的精准识别，从而提高了各种便民的软硬件应用的使用门槛。方言语音识别能够使得语音识别的使用门槛降低，使得更多的人受益于技术的发展。 3. 双向翻译：石佳影、杨波等人[4~7]的研究均局限于将各族方言翻译为普通话，而忽略了将普通话转化为各组方言亦是十分重要和有意义的。若掌握了双向翻译，则不懂得普通话，只了解该地域方言的人也可以较为流畅地与外人交流。 4. 提升易拓展性：目前的各种方言语音识别依赖于大量的双向语音数据，并对于语音数据的质量有极高的要求，这对于一些方言来说极难获得。对于特定的研究方向：如方言指令的语音识别，只需要相较于语音库总体来说极少的一部分数据。所以小数据量的语音识别以完成特定的识别任务是可能的发展方向，这也使得不同领域的研究者可以更深入地研究各方面不同的问题。 | |
| **三、课题研究内容** | **（研究的理论依据与运用研究方法，并阐明要解决的主要问题）**  **理论依据：**  通过傅里叶变换将语音数据看作时域信号，转化为频域数据，使用如梅尔倒谱系数（MFCCs）等特征值算法获取语音的特征值。并通过分析不同语音得出可鉴别不同语音的特征值。若干特征值组成多维向量，通过如余弦定理计算各语音数据特征值的多维向量的夹角或simhash算法[9]等相似度算法比较语音数据的相似度。  **研究方法：**   1. 调查法和文献研究法   通过调查、搜索、阅读文献，选择算法去实现本论文中需要的语音特征值提取及语音数据相似度的比较。   1. 数学建模法   通过算法提取语音数据的特征值，并建立有关模型描述语音数据。   1. 实验法   通过程序设计、产品设计，采用实验法验证和测试本研究的成果，并重复进行该进程对成果进行修正和优化。 | |
| **三、课题研究内容** | **（研究的理论依据与运用研究方法，并阐明要解决的主要问题）**  **主要问题：**   1. 受使用人数所限，小众方言没有足够的研究基础，并且小众方言无法获得足够的关注度。因此存在方言语音数据不足的问题。旨在通过少数据量完成智能家居方言语音指令的识别。 2. 由于少数据量，无法使用基于大数据的各种深度学习算法搭建的语音识别模型。目标为通过特征提取的方式达到方言语音指令的初步的、可接受的识别。但是特征提取的方式有一定概率无法达到在可接受范围内的准确率。 | |
| **四、预计的研究难点** | **（课题研究过程中可能遇到的理论难题或技术难点）**   1. 小众方言的语音数据难以获取，而更细致的方言智能家居指令数据更难获取。 2. 数据量缺少使得有效特征提取变得困难。 3. 由于数学学科的知识限制导致特征提取的算法选择空间不宽裕。 4. 若将研究成果产品化，所使用的单片机有可能算力不够。 | |
| **五、创新点** | **（选题、观点、理论、材料、方法等创新点）**  选题：目前对于小众方言的研究较少，并且并不聚焦于智能家居领域。而小众方言智能家居语音指令的识别具有研究的必要性。  观点：在智能家居方言语音指令识别的领域中，不应该盲目追随潮流，追求大模型带来的识别精确度，而是应该发展对于特定指令（词汇）的识别。目前的语音识别大多基于大数据为基础的大模型，从而限制了模型对于小众方言的适配性。  方法：启发自计算文本相似度的simhash算法[10]，比较两段语音的相似度 | |
| **六、进度计划** | **（根据研究内容拟定的研究思路和主要技术路线的进度安排）**   1. 收集小众方言智能语音指令的语音数据（通过录音）。 2. 通过特征提取的算法提取收集到的小众方言语音指令的特征，并通过筛选获取有效的特征值用于识别。 3. 设计并完成基于特征值的语音识别（语音比对）。 4. 初步目标：在电脑上成功运行有效的方言指令识别的程序。 5. 进一步目标：将程序烧录到单片机中，并完善成一个产品。 | |
| **七、资料来源** | **（指能够支持本表格中“一 ~ 六”的主要文献资料**）  [1] 吴永焕.汉语方言文化遗产保护的意义与对策[J].中国人民大学学报,2008(04):39-43.  [2] 黄涛.语言文化遗产的特性、价值与保护策略[J].中国人民大学学报,2008(04):27-33.  [3] 武瑞丰.人工智能视域下的林州方言建档[J].山西档案,2018(01):62-64.  [4] 石佳影,黄威.基于深度神经网络的四川方言识别[J].现代计算机(专业版),2016(13):3-6+10.  [5] 杨波.基于RNN的桂柳方言语音识别系统研究[J].现代计算机,2019(31):6-9+14.  [6] 张宇聪. 基于深度学习的藏语拉萨方言语音识别的研究[D].西北师范大学,2016.汉语方言文化遗产保护的意义与对策  [7] 余陆峰. 基于深度学习的客家方言语音识别[D].华南理工大学,2019.DOI:10.27151/d.cnki.ghnlu.2019.002580.  [8] Peng X , Stevenson M , Lin C ,et al.Understanding Linearity of Cross-Lingual Word Embedding Mappings[J]. 2020.DOI:10.48550/arXiv.2004.01079.  [9]刘阳,王巍,程佳慧等. 一种智能家居方言语音识别装置[P]. 河北省：CN211669966U,2020-10-13.  [10] Manku G S , Jain A , Sarma A D .Detecting near-duplicates for web crawling[C]//The Web Conference.ACM, 2007.DOI:10.1145/1242572.1242592. | |
| **指导教师意见：**  陶理同学的课题来源于自己的思考，机器学习对语言的特征识别已经成熟，但对小众方言还有待提高，因此陶理同学希望通过特定算法提取小众方言语音指令的特征，并用于智能家居的控制，该课题对高中生而言具有一定的挑战性。  **指导教师签字:**  **年 月 日** | | |