

第十二课 语音识别与翻译



学习目标

- 知道语音识别的概念。
- 了解语音识别系统及其原理。
- 了解语音识别的应用领域。

知识导图



随着科技的发展，人类陆续发明和制造了各种机器协助工作。尽管生产力得到大大提高，但在这个阶段，人类仍然需要动手操作机器。为了真正“解放双手”，科学家致力于研究怎样使机器更加智能化，变得能够听懂人类的语言，可以按照人类的指令工作，甚至可以与人对话。而语音识别与翻译技术，就是其中一项关键技术。

一、语音识别与翻译

语音识别是一门涵盖声学、语音学、语言学、信息理论、模式识别理论和神经生物学等的综合性学科。语音识别技术就是让机器通过识别和翻译把语音信号转变为相应的文本或指令的技术，并逐渐成为信息技术中人机交互的关键技术。



自主探究

如图12-1所示，利用语音识别软件，用语音输入的方法，输入诗人李白的《望庐山瀑布》：“日照香炉生紫烟，遥看瀑布挂前川。飞流直下三千尺，疑是银河落九天。”由此体验语音输入的便捷。也可以尝试转换

所使用的语言（如普通话、粤语、英语等），输入不同的内容，探究语音识别的特点。

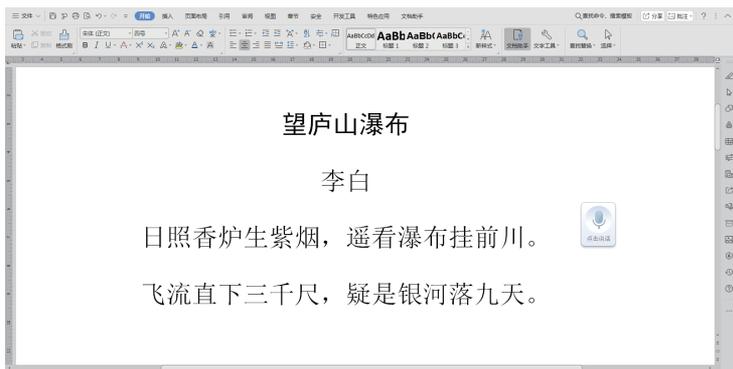


图12-1 语音识别输入文字

二、语音识别系统及其原理

1. 声音的本质

声音是由物体振动产生，并通过介质（空气或固体、液体）传播的一种波，如图12-2所示。因此，对声音的识别就是对波信号的处理。

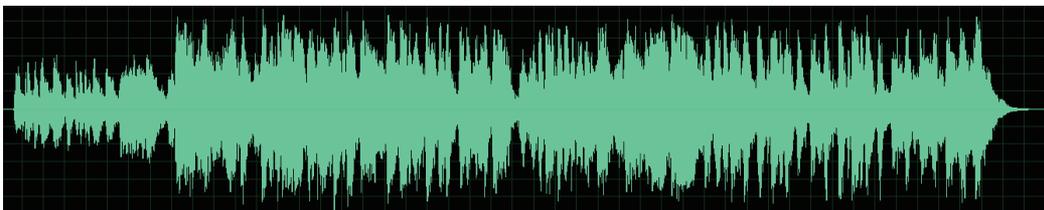


图12-2 声音波形

2. 语音识别系统

语音识别系统模型通常由声学模型和语言模型两部分组成，分别对应语音到音节概率的计算和音节到字概率的计算。

(1) 声学特征。

声学特征的提取与选择是语音识别的一个重要环节。声学特征的提取既是信息大幅度压缩的过程，也是信号分析和划分的过程。

（2）语言模型。

语言模型主要分为规则模型和统计模型两种。其中，统计模型是用概率统计的方法来揭示语言单位的内在统计规律。

（3）系统实现。

语音识别系统的实现，建立在大量数据训练的基础上，使系统生成能准确地定义语音片段的识别基元。系统所需的训练数据规模与模型复杂程度有关。

（4）稳健性与自适应。

语音识别系统的性能受许多因素的影响，包括不同的说话个体、说话方式、环境噪音、传输信道等。要提高系统稳健性，就要提高系统克服这些因素影响的能力，使系统在不同的应用环境、条件下保持性能稳定。自适应的目的是根据不同的影响来源，自动地、有针对性地对系统进行调整，并在运作过程中逐步提高性能。

三、语音识别的应用

语音识别能把语言转换成计算机可以处理和传播的数据，使计算机做出人们所期望的反应。

1. 语音与文字转换

语音转换成文字的功能除了使人们日常交流、听写数据录入更加便捷，还能帮助人们快速准确地记录发生的事情，如：法院庭审的发言笔录，访谈、报告和会议的即时字幕，手术情况记录等。

2. 机器控制

语音识别技术能使机器直接识别语言的内容和语义，并给予相应的反馈，实现相应的功能，如：自动导航、设备操控、文档检索等。

某品牌的智能音箱内置了语音助手，如图12-3所示。用户的所有操作都可通过语音控制完成。它内置Wi-Fi和蓝牙模块，能和计算机、手机等电子设备互联。用户唤醒它后，就可以向它询问天气，让它播放音乐，或进行闹钟设置等操作。



图12-3 内置语音助手的智能音箱

3. 信息挖掘

语音识别技术还能帮助我们进行信息挖掘。如通过智能语音识别技术，对客服的交谈内容进行全方位分析，记录通话双方的语速、情绪和交谈内容，从而快速发现问题，提升企业客服质量，提升客户满意度。此外，语音识别技术能有效检测音频和视频中的不良信息与敏感词，与传统人工审核方式相比，其成本更低、效率更高。



如图12-4所示，利用语音翻译软件，用普通话说：“早上好，欢迎来到广州。”观察中英文翻译的准确性，体验语音翻译的便捷。



图12-4 语音翻译



小组讨论

结合语音识别的体验过程，填写表12-1。

表12-1 学习讨论表

序号	讨论内容	讨论结果
1	语音识别是什么	
2	语音识别系统及其原理是什么	
3	语音识别的应用领域有哪些	



反思评价

学而不思则罔，思而不学则殆。对自己在本课中的学习进行评价，对照表12-2，在相应的○里画“√”，在“其他收获”栏里留言。

表12-2 学习评价表

序号	学习内容	能否掌握
1	知道语音识别的概念	能○ 不能○
2	了解语音识别系统及其原理	能○ 不能○
3	了解语音识别的应用领域	能○ 不能○
其他收获：		