

探索语音识别技术的前世今生

自动语音识别技术(Automatic Speech Recognition-ASR)是一种实现从“声音”到“文字”转换的技术,通过将人的语音直接转换成相应的文本以便计算机进行理解和产生相应的操作,并最终实现人与机器之间的自然语音交互。语音识别就好比“机器的听觉系统”,让机器通过识别和理解,能够把语音信号转变为相应的文本或命令,下面简要回顾语音识别技术的发展历程。

1. 最早的语音识别技术可以追溯到20世纪50年代,贝尔研究所 Davis 等研究成功了世界上第一个能识别10个英文数字发音的实验系统;1960年英国 Denes 等研究成功第一个计算机语音识别系统。



2. 大规模的语音识别研究始于20世纪70年代以后,在小词汇量、孤立词的识别方面取得了实质性的进展。



3. 20世纪80年代以后,语音识别研究的重点逐渐转向大词汇量、非特定人连续语音识别。同时,语音识别在研究思路上也发生了重大变化,由传统的基于标准模板匹配的技术思路开始转向基于统计模型的技术思路。此外,业内有专家再次提出了将神经网络技术引入语音识别问题的技术思路。



4. 1987年12月,李开复开发出世界上第一个“非特定人连续语音识别系统”,用统计方法提升了语音识别率。



5. 20世纪90年代以后,大词汇量连续语音识别得到优化,在语音识别技术的应用及产品化方面出现了很大的进展。1997年,IBM Via-voice 首个语音听写产品问世。

6. 2001年, Intel 的创始人之一戈登·摩尔 (Gordon Moore) 曾预言语音识别技术将大大改变未来科技的发展,之后的发展也印证了这一点。

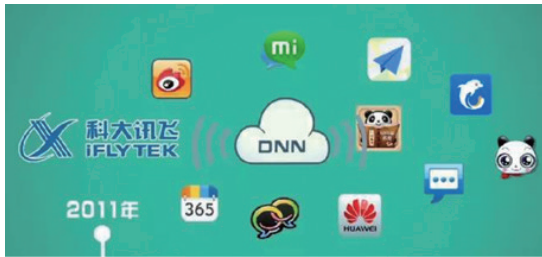


7. 自2009年以来,借助机器学习领域深度学习研究的发展以及大数据语料的积累,语音识别技术得到突飞猛进的发展。2010年,谷歌发布VoiceAction支持语音操作与搜索。

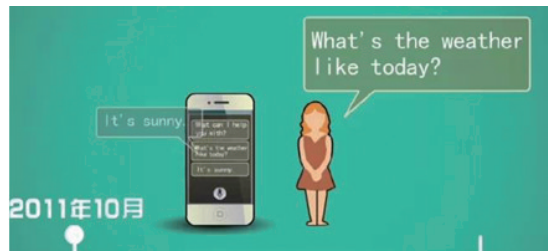
8. 2011年初,微软的深度学习神经网络(DNN)模型在语音搜索任务上获得成功。



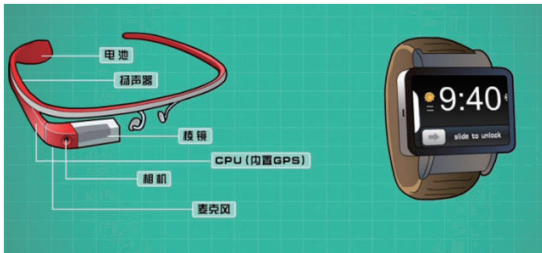
9. 2011年,科大讯飞在国内首次将DNN技术运用到语音云平台,并提供给开发者使用。



10. 2011年10月,苹果手机助理Siri首次亮相,人机交互掀开了新的篇章。



11. 2013年,谷歌发布Glass,使用语音交互,穿戴式语音交互设备成为新热点。



12. 2014年8月,科大讯飞发布讯飞语音云3.0,独家具备中文方言语音识别、高抗噪语音识别、个性化识别等功能、未来必能为用户带来更为智能、便捷的交互体验。



13. 2014年9月9日,苹果公司正式发布旗下第一款智能手表Apple Watch。该产品集成了语音功能,让大家对穿戴式语音交互设备的未来更加充满期待。



14. 2015年12月21日,科大讯飞在以“AI复始,万物更新”为主题的年度发布会上,提出了以前馈型序列记忆网络(FSMN, Feed-forward Sequential Memory Network)为代表的新一代语音识别系统。通过进一步的研究,在FSMN的基础之上,再次推出全新的语音识别框架,将语音识别问题创新性的重新定义为“看语谱图”的问题,并通过引入图像识别中主流的深度卷积神经网络实现了对语谱图的全新解析,同时打破了传统深度语音识别系统对DNN和RNN等网络结构的依赖,最终将识别准确度提高到了新的高度。

现今,语音识别技术已经实现了自由说识别,从算法到模型都有了质的发展,语音技术陆续进入工业、家电、通信、车载导航、医疗、家庭服务、消费电子产品等各个领域。类人机器人自从拥有语音识别技术,就可以与用户拟人化、趣味的对话,拥有一定程度的情感智商,与用户互动,甚至成为家庭一员!当有一天,机器能够真正“理解”人类语言,并作出回应,那时必将迎来一个崭新的时代。

文/科大讯飞
(责任编辑 刘志远)