

活动一 走近人工智能

今天，人工智能已融入了人们的日常生活。例如，智能音箱会识别人们的语音命令，播报天气情况或播放歌曲；智能手机能识别用户的脸或指纹，完成自动解锁；购物网站会根据用户近期的浏览或购买记录，自动向用户推荐可能感兴趣的商品……

在本活动中，将尝试使用智能工具处理声音、图像等，了解人工智能及其常见应用，畅想未来人工智能发展的趋势。

一、用智能工具处理声音

1. 使用语音输入法

以小组为单位，用语音输入法先以普通话朗读一段文字，比较语音转换为文字的准确率；再将语音识别模式修改为方言，用方言朗读该段文字，比较普通话与方言的识别效果。



图 1-1-1 使用语音输入法



提示板

在语音输入法的设置选项中，预先选好识别的语言，如普通话、英语，然后按住语音输入键，输入一段语音信息，语音输入法就会将这段语音自动转换成文字。

说说做做

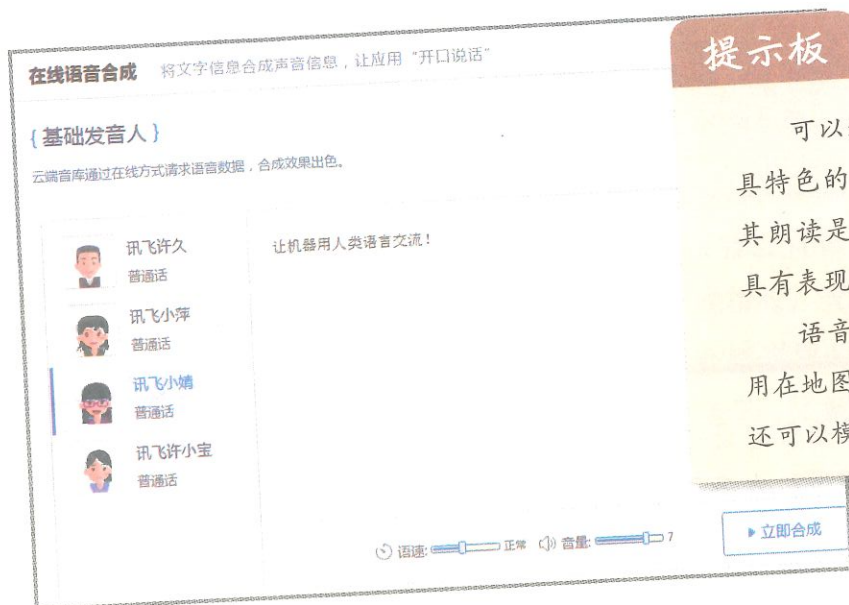
使用计算机、手机等智能设备时，你和家人一般使用的是哪几种输入法？语音输入法的使用频率高吗？大多用在哪些场合？比较各种输入法的优势和劣势，并完成表 1-1-1。

表 1-1-1 输入法的比较

输入法	优势	劣势
拼音输入法		
语音输入法		

2. 体验语音合成技术

访问一个开放的人工智能平台，选择平台上的“语音合成”功能，输入一段文字，让计算机将其读出来，感受计算机的合成音在音色、自然度等方面的表现，说说其是否接近人声。



提示板

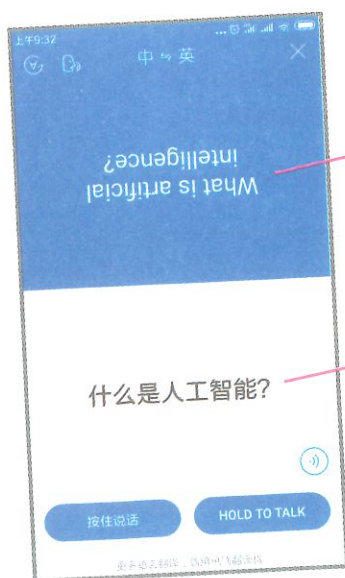
可以选择平台提供的各种极具特色的发音人(音库), 比较其朗读是否流畅、清晰、自然和具有表现力。

语音合成技术已经被广泛应用于地图导航、语言翻译等方面, 还可以模仿某个人的独特嗓音。

图 1-1-2 在线语音合成

说说做做

查找并试用一款人工智能语音翻译 App, 尝试朗读一些中文或英文句子(如图 1-1-3 所示), 了解其语音翻译功能和准确率, 说说其使用场合。



② 对文本进行理解后将其翻译为另一种语言(此例中英文倒立显示, 方便另一人浏览), 再由手机读出来

① 识别人的语音后将其转化为文本

图 1-1-3 语音翻译 App



3. 了解语音交互技术及其应用

上述案例中所使用的这些智能工具背后，都隐藏着同一类人工智能技术——语音交互技术。

信息时代，人们已能利用“语音”这一自然、便捷的手段进行人机交互。看似简单的人机语音交互，背靠的是让机器“听懂”“理解”和“说话”的众多技术支撑，如语音识别技术、语义理解技术和语音合成技术等。

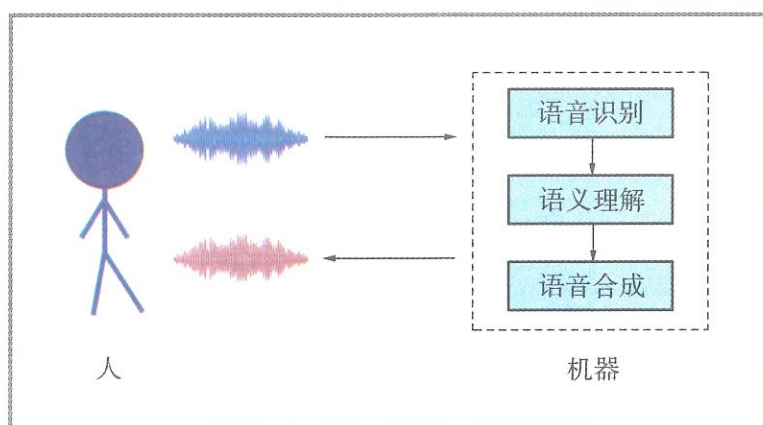


图 1-1-4 人机语音交互

语音识别技术所解决的问题是让计算机能够“听懂”人类的语音，将语音中包含的文字信息“提取”出来。语音识别技术在“能听会说”的智能计算机系统中扮演着重要角色，相当于给计算机系统安装上“耳朵”，使其具备“听懂”人说话的能力。

利用语音合成技术，机器会将文字转换成语音，相当于给计算机系统安装上“嘴巴”，使其具备“说话”的能力。

除了能“听懂”和会“说话”，计算机还要能“理解”人类在说什么，也就是计算机系统要能理解人类自然语言表述的意思。语音识别只是将声音转化成文字，语义理解则是提取文字中的相关信息和相应意图，从而“理解”人类在说什么，然后通过云端“大脑”决策，进行相应的问题回复或者动作反馈。

现今，机器人客服、语音输入法、语音助手、语音翻译、语音支付、智能家电等都使用了语音交互技术。例如，手机上的智能聊天 App，能够识别人的语音，并和人进行语音聊天。又如，语音控制的智能音箱，能根据人的语音指令来完成播放歌曲、预报天气等操作。



图 1-1-5 智能聊天 App

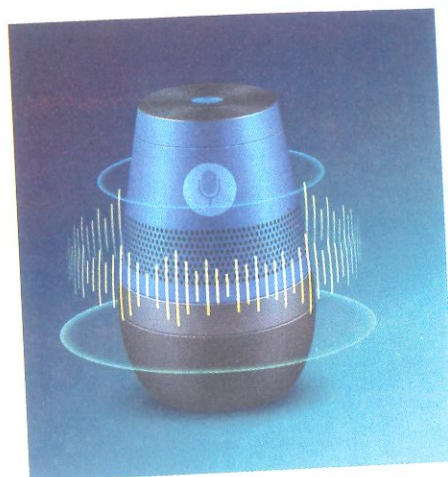


图 1-1-6 智能音箱

说说做做

交流更多生活中能与人直接进行语音对话的设备或软件，说说其功能和使用方法。

知识链接

语音识别技术的更多应用

2015 年，在某人工智能公司的产品发布会上，5 位持证上岗的专业速记员与机器记录员一起记录会议内容，会议穿插着主持人以及各位嘉宾的即席发言，最后测评显示，机器记录的平均准确率为 96%，专业速记员的平均准确率只有 70% 多。现在，语音识别的准确率已经提升至 98%，其中有一半因素是因为算法和深度学习的突破，还有一半因素是因为数据的积累，每天数十亿人次的语音数据会进入后台，机器通过深度学习训练得到的语音数据识别模型，使得语音识别表现更好。

随着识别准确率的提升，语音识别技术已经走进了家电、通信、汽车电子、消费电子产品等多个领域。例如，电视盒子可以通过语音控制来进行视频的检索、播放和换台等操作；又如，汽车的语音助手能够帮助驾驶员在开车的状态下打开一些常见应用。