

语音用户界面研究进展

韩 勇¹ 须 德¹ 戴国忠²

(北方交通大学计算机与信息技术学院 北京 100044)¹

(中国科学院软件研究所 人机交互及智能工程试验室 北京 100080)²

摘 要 语音是人们日常生活中高效、自然的交流方式之一。但是直到目前为止,语音交互方式在计算机技术上的应用还是比较少的。近年来,随着 Ubiquitous Computing 和便携式计算机的出现,再次对语音用户界面的应用提出了迫切的需求。而且语音识别、合成技术的发展也为语音交互界面的实现提供了技术基础。本文综合参考了国内外语音界面的一些应用系统实例以及语音这种独特的交流媒体的优点和局限性,总结了语音用户界面的适用环境和设计指导原则,并提出了对语音界面发展展望。

关键词 人机交互,语音界面,语音识别,语音合成,speech-as-data

Research on Development of Speech User Interface

HAN Yong¹ XU De¹ DAI Guo-Zhong²

(School of Computer & Information Technology, Northern Jiaotong University, Beijing 100044)¹

(Human-Computer Interaction & Intelligent Engineering Laboratory, Institute of Software, The Chinese Academy of Science, Beijing 100080)²

Abstract Speech is a most efficient and nature way of communication between people. Although speech interaction is seldom used in HCI at present, however, recent development of speech recognition, speech synthesis, ubiquitous computing and PDA provide feasibility and demand for speech interface. In this paper, we give some obstacles in design speech user interface and refer to some of the applications using speech interface. Finally, we summarize the virtue and limitation of speech as communication media, circumstance suitable for speech interface and guidelines for speech interface design.

Keywords Human computer interaction, Speech interface, Speech recognition, Speech synthesis, Speech-as-data

1 引言

语言和手势是人们每天生活中最自然、最有效的交互方式,这种交流方式在我们的生活和思想中根深蒂固。所以,根据这个特性建立的交互方式可以让用户使用时的感觉更自然。从计算机出现的那一天起,人们就幻想能够与计算机通过自然语言来交互。随着语音识别、自然语言处理和计算机硬件技术的不断发展进步,使得利用听觉通道的语音交互逐步成为人一机交互的重要形式。

语音技术在人一机系统中的出现及应用虽然历时不久,但其发展前景却得到了普遍乐观的估计,基于语音的用户界面的需求目前正在逐渐增加。我们都有过在电话上与电脑交互的经验,并且在许多情况下都已经习惯了这种方式,但是在以往的这些交互中,都是通过电话上的数字键来选择一些预先定义好的功能,现在利用语音识别等技术,用户可以进行复杂的查询、咨询和订票等任务。对于声音信号的存储与处理技术的发展也为基于语音的应用奠定了基础,有了很多应用于电话系统的语音邮件、语音中转系统。语音系统也可以在用户的手和眼都忙于处理其他事情的情况下使用,如车载系统以及医疗救治中。语音界面还可以应用于文档输入、地图查询、CAD 设计和显示屏很小的便携式计算机,供盲人使用的屏幕阅读器,以及用户屏幕显示的内容过于烦杂和用户在外出和移动中需要与计算机交互的场合。对于当前经常提到的无处不在的计算,语音也是一个重要的输入输出手段。Text to

Speech 功能使系统的输出信息不再仅仅局限于文本的形式,而是可以通过语音的方式通知用户,将用户从计算机屏幕前解放出来。

2 国内外研究情况

近几年来,语音识别、合成、编码和存储技术以及自然语言理解和生成等研究取得了较大的进展,已经具备了研究基于语音的人机交互的条件。语音有其独特的特点,所以将语音技术应用于交互也有其自身的特征。从目前国内外的研究看来,基于语音的交互技术可以分为两种,一种是基于语音识别的技术,即需要将语音输入信号转换为相应的文本信息,另一种是“speech-as-data”^[33],即不将语音数据识别,而是将这些语音记录直接作为数据处理。

2.1 基于识别的应用

虽然语音识别仍然是语音交互系统的瓶颈,但是通过上下文和精心设计的对话策略可以弥补识别引擎的不足,实现特定领域的语音交互。

• 远程应用 很多家庭用户还没有使用 Internet,但是几乎所有家庭都有电话,将语音交互技术应用于电话系统中可以实现自动应答、远程的信息访问等应用。例如 MIT 的 Jupiter^[9~11]和 CMU 的 Communicator^[12]系统,综合使用了语音识别、语音合成、自然语言理解与生成等技术,Jupiter 是通过电话查询天气,Communicator 实现了通过电话预定机票等功能。这两个系统都是通过电话与用户交谈,将用户的语音信

号经过识别后交给自然语言处理模块,了解用户的意图以确定是继续交流收集查询条件还是返回查询结果。另外还有 MailCall 系统,让用户可以用电话来收发电子邮件。

- **手眼忙于其他事务** 语音交互不需要一直占据用户的注意力,可以在用户做其他事务的时候同时使用,可以在用户手和眼忙于处理其他事务的情况下使用,如 MIT 的 Direction Assistance,该系统提供了波士顿地区的交通路线图,用户可以在驾车时使用语音查询^[25]。UMD 的 Grasso 研发的医疗信息的记录系统^[26~29],使医疗人员在检查、诊断的同时用语音作医疗记录等信息的采集。

- **残疾人使用** 语音输入作为命令控制的 Web 浏览器^[13~15]和邮件收发程序^[6,16],Cornell 大学 Raman 设计了科技文档阅读器和根据上下文的屏幕阅读工具以及使用语音作为命令控制、反馈的应用程序 Emacspeak 等^[19,20],其他人的工作还有 Berkeley 的 IC2D 系统利用声音反馈辅助用户选取特定位置的点以及将光标定位等任务^[21]。其他一些应用系统还包括三维应用程序(如 CAD)的增强工具、空间定位^[22~24]。

- **其他应用** 语音还可以应用于互动式的儿童娱乐,如 MIT 的 StoryRooms^[30]。其他语音系统还有电话、传真的语音拨号、转接、留言等应用。语音识别和合成还可以应用于语言教学中。

2.2 speech-as-data

- **信息捕捉与浏览** 语音的另一种应用技术并不将语音数据识别,而是将语音数据记录下来,并提供对这些数据的访问方法。“speech-as-data”有其自身的优缺点:虽然数据易于获取,但是不便于获取和浏览保存的数据中任意位置的内容。speech-as-data 的一个应用是 SpeechSkimmer^[33~35]。SpeechSkimmer 是利用声音的结构特性来访问语音数据,实现语音数据的快速浏览。快速阅读或者浏览通常是在视觉的任务中使用,当我们在阅读或者查找时常常不自觉地用到这些能力,但是,人们却没有自然的方式来快速扫描一段语音信息。这是由语音的短暂性的时域特性所决定的。SpeechSkimmer 探索了语音的这些限制,研究了用户快速扫描一段录音记录的方法。SpeechSkimmer 研究了语音数据特点,如是由许多声音的小片段组成的,说话中对于强调和重点内容,语调上会有变化,而且话题内容转换时语音间的停顿时间较长,根据这些结构特点将语音数据进行加工处理,可以在浏览时根据用户的需要进行不同的快速浏览。如去掉声音间的停顿、时间压缩、基于语调的浏览、基于话题的浏览,节省了用户访问这些数据的时间。

在会议中许多重要的信息都是通过语音来表达的,因此,能够捕捉并访问这些口头的信息就有很高的价值。通常人们使用笔和纸摘录一些要点。但是仅仅这些手写记录是不够的,录音记录是对手写记录的最好的补充^[1]。一个应用是将与语音同步获取的笔迹的空间信息作为声音信息的索引标记来访问所需的片段。Audio NoteBook^[35,36]是一个声音采集系统,可以根据写在纸上的信息以及纸张的页数将所采集到的语音数据同步起来,可以用笔选中纸上某一段笔迹记录来访问作笔迹记录时所采集的声音信息。与 Audio NoteBook 类似的系统有 HP Laboratories 的系统 FiloChat^[37],不同之处只是在访问语音数据时,FiloChat 是在相应的笔记上做手势动作。另外还有 MIT, Xerox 合作研究的系统 Dynamite^[38,39],作为电子笔记本系统,Dynamite 与别的系统不同之处是可以根据用户的需要将选择记录声音信息以便节省设备的存储空间。

Berkeley 也有类似的研究项目 NotePals^[40,41],但是该系统只是记录了笔迹的信息而没有记录声音的信息,但是可以把这些记录上载到一个服务器中去,使同一组的人共享、参考其他人的笔记。在 IBM 的 Classroom2000^[41~43]项目中参考了以上的各个原型系统,实现了基于 Web 的课堂语音、图像的访问以及课堂笔记、批注等功能的便携式笔记系统。

- **辅助增强功能** 人在每天的生活中都要通过语音来交流,虚拟现实作为自然人机交互的体现,必定不能缺少语音这种重要的交流方式。在虚拟现实环境中引入语音和声音信息可以达到增强效果和真实感的目的^[44~46]。此外,在系统中引入 earicon 和声音信息作为反馈提示可以增加输出通道,增强交互效果。

3 界面特点及设计要素

人机交互(Human-Computer Interaction)是研究人、计算机以及它们相互影响的技术。人机界面是计算机与人(使用者)之间通信和对话的接口,是计算机系统的重要组成部分,是计算机科学、心理学、认知科学和人素学的交叉研究领域^[1]。

计算机软、硬件技术的发展推动了人机接口技术和人机界面的发展,其发展经历了手工操作、命令语言和图形用户界面三个阶段^[1]。虽然计算机的运算能力有了很大的提高,但是我们目前仍然是以键盘和鼠标为输入工具,这使得输入方式成为了目前计算机系统的瓶颈,所以研究新的交互设备和交互手段的任务就显得日益紧迫了。这也推动了语音用户界面的出现。语音界面是指使用了语音或者是合成的语音的界面。

3.1 语音交互界面的优点

- 语音是一种高效的交流媒体^[1~4],比其他交流手段更加自然,包含更丰富的信息。

- 语音交互不需要一直占据用户的全部注意力,可以在用户做其他事务的时候同时使用。

- 语音交互方式特别适合在信息随机呈现并要求用户立即采取行动的任务中,或受工作环境限制(比如空间狭小,照明不佳),不能有效利用视觉通道传递信息的场合中采用。

- 语音对于说话的人效率较高,说话的表达方式比写字或者打字速度快。

- 语音交互适用于小型的移动设备,利用语音作为输入输出手段的设备,I/O 可以缩减为扬声器和话筒。

- 在三维建模等环境中,使用语音只需要一、两个单词就可以描绘出大家熟知的一些复杂形状(马、飞机、汽车、椅子、桌子)。

- 使用语音交互可以方便用户快速访问结构化组织的信息,如可以直接选中 WIMP 界面中的某个菜单项,而不必在菜单的树状结构中层层选择。

3.2 语音交互界面的缺点

- 语音是短暂性的^[2,3,7]。一旦发出来以后,语音信号就不可再得到了。因此,用户需要记住这些信息,消耗了用户大量的短时记忆资源,增加了使用者的记忆负担。

- 与图形化的用户界面相比,语音界面是串行的输出方式,速度较慢^[3]。

- 语音不同于 WIMP 界面,应用程序可以执行的操作都可以显示出来供用户选择,因此面对一个语音界面时,如果没有适当的提示,对该系统陌生的用户可能感觉无从下手,不知道该说些什么。

- 不同于指点式设备的交互,语音的交互包含一些模糊的信息,增加了输入处理的难度。
- 语音识别的准确性受环境的影响较大,在嘈杂的环境下,语音识别的准确性将大大降低。

3.3 语音界面分析

3.3.1 语音界面与传统界面的对比分析 人机界面从命令行的文本界面发展到了现在的基于 WIMP 的 GUI 界面。下表将这两种界面与语音界面作了对比。

表1 用户界面的优缺点比较

对话方式	优点	缺点
文本界面	灵活; 对计算机专业用户有吸引力; 支持用户主观能动作用; 便于建立用户定义的宏指令。	出错处理能力差; 需要较多的培训和记忆。
图形界面	以视觉方式呈现任务概念; 容易学习; 容易保持; 可以避免出错; 鼓励探索; 能提供高度的主观满足感。	编程难度可能较高; 可能需要图形显示和指示设备。
语音界面	使用作为语音命令控制任务,减少学习使用设备的负担	需要明晰的对话; 可能需要击键次数较多; 可能不显示上下文。

应该注意的是语音方式并不适用于所有的任务。如听觉通道适用于传递声音编码的信息,但对于空间定位等工作就不适合用语音来控制。其次,从通讯效率来看,人类听觉的通道容量远不及视觉容量,人所获得的大部分信息都是从视觉通道而来的。这在一定程度上限制了语音技术在人-机系统中的应用。研究语音界面,就是开发听觉通道,解决用户在使用传统交互时认知、视觉及手动负荷过重的问题,使各个通道的工作负荷能够合理分配,所以语音技术没有必要也不可能完全替代目前的视觉-手动交互方式。因此,围绕语音技术开展人机工程研究的出发点不在于用语音方式来实现目前人机交互中的所有功能,而是如何根据语音通讯的特点和任务特征,将其融合到现有的人-机交互系统中。比较成功的界面设计应该使用户在使用中感觉更加自然、方便。因此需要考虑如下两点^[3,5,6]:

- 界面符合人们希望完成任务所采用的方式。
- 交互方式的使用应该适合当时的交互环境条件。

3.3.2 适宜和不适宜使用语音输入的情景

表2 适宜和不适宜使用语音输入的情况

如下情况使用语音输入	如下情况避免语音输入
没有键盘等输入设备(如电话、一些移动和便携式设备)	任务需要用户在执行中与其他用户交谈
用户的手和眼忙于处理其他的任务(如驾车)	用户在嘈杂的环境中工作
命令位于较深层次的菜单结构中	任务用键盘和鼠标可以更加轻松地完成
用户不会或者不习惯键盘输入方式	
有残疾的用户使用	

3.3.3 适宜和不适宜使用语音输出的情景

表3 适宜和不适宜使用语音输出的情景

如下情况使用语音输出	如下情况避免语音输出
任务要占据用户的注意力,需要将用户从计算机屏幕前解放出来有残疾用户的使用	需要输出大量的信息
界面要体现人的一些特征	任务需要用户进行数据比较一些隐私或者绝密的信息

3.4 语言行为中人的因素

人们交流的目的就是为了使交流的双方对于所传递信息的内容达到一个共识。交流的模型如图1所示。

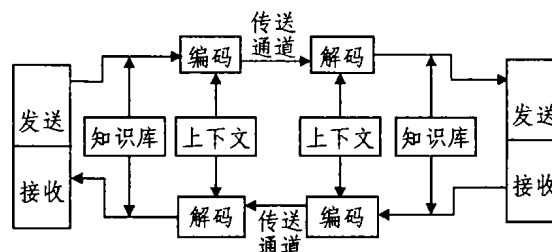


图1 语言交流模型

计算机界面的最终目标就是方便用户的交互操作,因此设计交互计算机系统时,必须把满足用户的使用要求作为重要依据,这就是交互计算机系统中的一个重要元素——人的因素。人的因素指的是用户操作模型,它与用户的各种特征有关。首先,用户是人,所以会不可避免地犯错误,如按错了键或拼写错误,因此,进行系统设计时要认真处理出错情况,并对各种操作给予提示帮助。其次,用户有不同的类型和文化层次,并且用户的年龄、文化程度、工作经历及职务各不相同,因而操作使用的要求也就不同。采用人机对话的交互计算机系统一定要考虑所服务对象的使用要求,以便设计好的用户界面。在更深入分析人的因素时,要设计用户的学习模型,认知心理学等内容。用户的认知模型等方面的研究兴趣和成果,随着人机对话的重要性的不断增长而有明显进展。

3.5 语音界面特点及设计要素

在面向应用的语音用户界面中需要考虑如下一些因素:是否需要对用户所使用的语言进行一些限制以便于语音识别软件的识别?这些限制是否会影响用户的工作效率?用户是否习惯于计算机合成的语音?除了这些因素以外,影响语音技术的应用还包括以下几个方面^[3,5]:

- 高错误率。语音技术中仍然错误率较高,尤其是语音识别系统的误识率比较高,而且这种错误是由计算机系统产生的,即使用户的输入是正确的。这不同于使用键盘或鼠标时用户的输入错误。

- 不可预知的错误。识别的错误是不可预测的,预先不可能知道识别系统产生的错误会是这样的一个识别结果,而且用户通常不愿意使用这种带有不确定性因素的产品。还有用户对于产品的期望值也是一个影响因素之一,应该使用户明白语音界面并不代表与计算机的交流完全等同于与人的交流,计算机有其局限性。而且用户通常也会采纳对于使用的语音进行一些限制来减少识别错误。

- 人们的预期。人们对计算机和语音有很高的期望。对于添加了语音技术的计算机,人们常常会期望能够与计算机进行自然的语音交谈。他们希望能够和计算机互相理解所表达的内容。如果他们的这种期望没有被满足,通常就会感觉失落并拒绝与计算机进行交谈。然而,如果计算机能够对不能理解的语言和词汇有个清楚的了解,人们就会对自己在交流中使

用的语言和词汇采取一定的限制。

- 人的短时记忆特点。因为纯语音的用户界面缺少视觉的反馈和确认,所以对记忆的负担较大,例如电话应用中很长的一串选择菜单就很难记住。

- 口语与书面用语的不同。人们使用的口头语言不同于书面语言,因此希望计算机系统也能这样。例如,用户看到 GUI 界面中标签为 Delete 和 Cancel 的按钮时就会理解这些术语的含义,但是在听计算机说出这些术语时,他们希望听到不是这么正规的语言。

- 在交谈中,人们通常会下意识的模仿别人的说话模式,因此我们可以利用计算机合成的语音来影响用户,使他们按照你希望的对话模式来说话,使用户使用的语言和词汇限定在某个范围以内。

- 用户控制。界面设计首先要注意用户控制。界面应该使用户感知到控制权并且允许适当的控制。如果用户感觉没有控制权,就会有比较情绪化的反应。而且人的认知、视觉、听觉、触觉等对信息的处理和获取都有其局限性。一个友好的界面不能对某一个处理通道过多的负载。

在设计中考虑人的因素就是为了减少用户的错误,使用户在使用过程中更加方便自然。确定一个语音应用软件是否成功的关键因素是看语音在该应用中是否有明显的优点。成功的语音应用程序是在这些应用中确实有必要利用语音技术而不是仅仅因为加入语音使该应用有新奇的功能。例如,当用户的手和眼忙于处理其他事务的时候,或者一些适合通过语音完成的任务,如电话系统中的应用,正是语音信号独特的特点决定了它所适用的具体环境。

结论 随着微电子技术的发展,计算机的尺寸越来越小,如穿戴式计算机和 PDA。由于这些设备的屏幕大小和尺寸限制,不适合使用目前台式机常用的 WIMP 界面以及鼠标、键盘等交互设备,所以需要探索新的输入输出手段。语音明显是较好的选择之一,使用语音作为输入输出的通道可以使输入输出设备简化为扬声器和话筒。缩减了输入输出设备的体积大小;使用语音可以使人机交互更加自然、灵活。语言是人类日常生活中的交流方式,使用语音作为交互手段,用户无需任何训练就可以进行操作,使用系统提供的功能;语音还可以提高交互的效率,用户的操作不必逐级查找树状结构的菜单,而是直接说出要进行的操作;通过语音以及语言的理解,还可以从用户模糊的描述中捕捉用户意图以协助用户完成期望的任务或通过交谈收集信息逐步引导用户完成操作;使用语音还使用户可以利用电话与系统远程交互,也可以在移动或手眼忙于其他事务时的交互,适应目前无处不在计算的研究发展趋势。所以语音交互是将来自然、高效的人机界面的重要交互方式之一。

虽然语音识别技术经过了半个多世纪的发展,但是语音识别仍然是语音交互系统的瓶颈。在使用这种有缺陷的技术作为交流手段时,需要精心地设计交互界面来回避技术上的局限。一个研究方向为我们可以根据语言交流的心理特点来设计对话模型,以此约束用户使用的词汇或者根据上下文确定用户的操作意图,从而提供动态的待识别词库,降低识别的复杂度,克服语音识别系统识别准确性的局限。另一个研究方向是识别错误的检测和纠正,我们可以根据上下文并结合自然语言理解估计用户的交互意图,发现识别错误,并根据上下文修正,给出最可能的识别结果进行处理。对话交流的系统需要准确理解用户的输入意图并产生响应,所以需要对自然语言的理解和生成技术有较高的要求。

在安静的办公室环境中使用语音界面会干扰影响别人的

工作。然而在其他的环境下或在移动中使用就会使采集的信号质量受到背景噪音、说话人离话筒距离等因素的干扰,影响识别准确性,所以语音界面能否成功应用还需要依靠语音识别技术在自适应性和鲁棒性方面的进一步研究和发展。

语音界面应用的另一个难题是当用户面对系统可能不知道该说什么好。在字符命令环境下,可以使用的操作和命令是隐藏起来的,到了 GUI 界面下,这些隐藏的命令以菜单和 I-CON 的形式呈现在桌面上由用户选择。到了语音界面环境下,又再次将操作命令隐藏了起来。如果让用户记住所有的操作命令必然增加认知负担。而且在语音界面下给出的帮助信息也不如图形界面下直观、快捷。所以语音交互界面的应用一般是比较简单的任务或者是人们日常生活中事务的处理,用户可以将机器系统看作人一样使用生活中的语言向系统发出命令和请求完成操作。建立语音交互系统的一个发展趋势是针对特定领域范围的某个任务,从语音识别模型和理解模块上整合起来。缩小识别的词汇量,降低语音识别与理解的复杂度。但是建立这样一个系统需要程序设计人员对大量的用户采集数据^[11,49],统计总结在各个任务步骤中用户会说些什么,再根据这些资料统计建立语言对话资料。这种设计方法需要耗费大量的时间。但是可以达到比较满意的交互效果和准确性。针对这种情况,CMU 提出了通用语音用户界面的思想^[50,51],使用一些通用的控制命令和语言格式,实现 WIMP 环境下一致的使用接口(如 WIMP 界面中滚动条、菜单、按钮的含义和使用方法都是一样的),因此对某个程序的使用经验可以很快地转移到其他程序中去。这种方式简化并加快了软件开发过程,但是给用户增加了一些使用约束,使用起来不如前一种方案自然,并且由于应用系统多种多样,这些通用的控制命令和语言格式可能无法描述某些复杂的任务。如果将以上两种解决方案一起采用,对于系统中复杂任务步骤的处理部分采用自然语言理解模块,对其他的步骤使用通用用户界面则可以利用两种解决方案的长处。

语音用于交互也有其自身的弱点,例如使用语音输入难于完成空间位置精确定位等指点式设备擅长的任务。但是语音与指点式输入设备这两种输入方式是互相弥补的^[8]。人们在每天的生活中主要是通过说话和基于纸、笔的交流。所以研究语音与笔的多通道整合技术也是当今自然、高效人机交互技术的研究热点。

综上所述,语音界面的应用发展依赖于解决语音识别的准确性、抗干扰能力,自然语言理解、生成技术以及对话交流的心理模型方面的研究发展。目前国外的语音界面的对话原型系统和商业应用展示了将语音应用于人机交互自然、高效的特性,揭示了未来人机界面的热点发展方向。

参考文献

- 1 Whittaker S, Hyland P, Wiley M. FILOCHAT: HANDWRITTEN NOTES PROVIDE ACCESS TO RECORDED CONVERSATION. ACM, 1994. 271~277
- 2 Wilcox L D, et al. Dynamite: A Dynamically Organized Ink and Audio Notebook. CHI97, Atlanta GA USA, 186~193
- 3 董士海,王坚,戴国忠,等. 人机交互和多通道用户界面. 科学出版社, 1999
- 4 Cohen P R, Oviatt S L. The Role of Voice Input for Human-Machine Communication
- 5 Design speech user interface
- 6 Roy D, et al. Wearable Audio Computing: A Survey of Interaction Techniques
- 7 Rudnick A I. Speech Interface Guidelines

(下转第 39 页)

Discard, EPD) 机制。由于绝大部分数据业务采用一定的帧格式以数据帧为单位进行信息传送, 这些帧进入 ATM 网络前大都需经过 AAL5 适配或固定长度的信元, 因此丢弃一个信元和丢弃整个帧的效果是一样的。AAL 在信元头 PT 域指示出该信元是否为承载的用户信息的最后一个信元。PPD 监测这一结束标记, 在发生信元丢弃时, 如果被丢弃的不是 AAL5 信元流的最后一个信元, 则将剩余的信元加以丢弃。EPD 也识别 AAL5 信元的结束标记, 在缓存队列的容量超过阈值时, 将后继的数据帧全部丢弃, 通过 PPD 及 EPD, 网络性能可以得到很大改善。

为了支持 ABR 业务, ATM 网络需要能够实现反馈控制机制。ABR 业务发送端以当前信元速率 (Current Cell Rate, CCR) 发送信息。CCR 可以在 PCR 及 MCR 间变化, 具体大小由用户与网络协商确定, 协商的过程便是反馈控制过程。反馈控制方式有两大类: ①端到端的 ABR 控制; ②分段的 ABR 控制。

分段的反馈控制比端到端的反馈控制更为灵敏、及时, 可进一步提高网络性能, 但同时也增加了实现的复杂度。

(上接第4页)

- 8 Yankelovich N, Levow G-A, Marx M. Designing Speech Acts: Issue in Speech User Interfaces, CHI95
- 9 Shneiderman B. The Limits of Speech Recognition: Understanding Acoustic memory and appreciating prosody
- 10 Oviatt S, et al. Designing the User Interface for Multimodal Speech and Pen-based Gesture Applications: State-of-the Art Systems and Future Research Directions. CHI2000
- 11 Zue V, et al. Jupiter: A Telephone-Based Conversational Interface for Weather Information
- 12 Polifroni J, et al. Evaluation Methodology for a Telephone-Based Conversational System
- 13 Hurley E, Polifroni J, Glass J. Telephone Data Collection Using the World Wide Web
- 14 Rudnicky A I, et al. Task and Domain Specific Modeling in the Carnegie Mellon Communicator System
- 15 Christian K, et al. A Comparison of Voice Controlled and Mouse Controlled Web Browsing
- 16 Wynblatt M, Benson D, Hsu A. Browsing the World Wide Web in a Non-Visual Environment
- 17 Petrucci L S, et al. WebSound: a generic Web sonification tool, and its application to an auditory Web browser for blind and visually impaired users
- 18 Walker M A, et al. What Can I Say?: Evaluating a Spoken Language Interface to Email. CHI98, 1998. 18~23
- 19 Wilcox L, Smith I, Bush M. Wordspotting for Voice Editing and Audio Indexing. CHI'92, 655~656
- 20 Oard D W. Speech-Based Information Retrieval for Digital Libraries
- 21 Raman T V. Audio System for Technical Readings: [PHD. Dissertation]. Cornell University
- 22 Raman T V. Emacspeak-Direct Speech Access
- 23 Kamel H M, Landay J A. Integrated Communication 2 Draw (IC2D): A Drawing Program for the Visually Impaired
- 24 Mereu S W, Kazman R. Audio Enhanced 3D Interfaces for Visually Impaired Users. CHI96, Electronic Proceedings
- 25 Myron D, Krueger W, Gilden D D. KnowWhere™: An Audio/Spatial Interface for Blind People
- 26 Marx M, Schmandt C. MailCall: Message Presentation and Navigation in a Nonvisual Environment
- 27 Davis J R. A voice interface to a direction giving program
- 28 Grasso M A, [Ph. D.]. Structured Speech Input for Clinical Data Collection
- 29 Grasso M A, Finin T. Task Integration in Multimodal Speech Recognition Environments
- 30 Grasso M A. Speech Input in Multimodal Environments: Effects of Perceptual Structure on Speed, Accuracy, and Acceptance: [PHD Thesis]

结束语 虽然 ATM 技术尚处于充分成熟过程之中, 但它以其高速、大容量的交换处理能力以及对不同类业务的适应性, 在广域网及局域网环境中得到了广泛的应用。ATM 对各种业务的管理能力更充分反映了 ATM 网络及系统的智能化水平, 体现了 ATM 技术的先进性, 同时也为 ATM 网络及系统的运营、维护及管理带来极大的便利。

参考文献

- 1 ATM Forum. Addendum to Traffic Management v4. 1 for an Optional Minimum Desired Cell Rate Indication for UBR. str-tm-dcr-01. 01, straw ballot, May 2000
- 2 Cerdan F, Casals O. Mapping an Internet Assured Service on the GFR ATM Service. Networking 2000, May 2000, no. 1815 in Springer Verlag LNCS
- 3 Bonaventure O. A Simulation Study of TCP with the GFR Service Category. High Perf. Networks for Multimedia Apps, Kluwer, 1998
- 4 Huang J, Lee B, Khorsandi S. Simulation Study of GFR Implementations. ATM Forum 98-1035, Dec. 1997
- 31 Grasso M A, Ebert D, Finin T. Acceptance of a Speech Interface for Biomedical Data Collection
- 32 Alborzi H, et al. Designing StoryRooms: Interactive Storytelling Spaces for Children
- 33 Lajoie S, et al. Application of Language Technology to Lotus Notes Based Messaging for Command and Control
- 34 Karl L, Pettey M, Shneiderman B. Speech-Activated versus Mouse-Activated Commands for Word Processing Applications: An Empirical Evaluation
- 35 Arons B. SpeechSkommer: Interactively Skimming Recorded Speech. In: Proc. of UIST'93, pp. 187~196
- 36 Arons B. SpeechSkimmer: A System for Interactively Skimming Recorded Speech. ACM Transactions on Computer-Human Interaction, 1997, 4(1): 3~38
- 37 Stifelman L J, et al. VoiceNotes: A Speech Interface for a Hand-Held Voice Notetaker. In: Proc. of INTERCHI, ACM, New York, 1993. 179~186
- 38 Stifelman L, Arons B, Schmandt C. The Audio Notebook Paper and Pen Interaction with Structured Speech. CHI, 2001. 182~189
- 39 Chiu P, Wilcox L. A Dynamic Grouping Technique for Ink and Audio Notes. UIST'98. San Francisco, CA, 195~202
- 40 Landay J A. Using Note-Taking Appliances for Student to Student Collaboration. 29th ASEE/IEEE Frontiers in Education Conference
- 41 Landay J A, Davis R C. Making Sharing Pervasive: Ubiquitous Computing for Shared Note Taking. IBM System Journal, 1999, 38 (4)
- 42 Davis R C, et al. NotePals: Lightweight Note Sharing by the Group, for the Group. CHI, 99. 338~345
- 43 Abowd G D, et al. Teaching and Learning as Multimedia Authoring: The Classroom 2000 Project
- 44 Darvishi A. A Visual User Interface for Creation and Manipulation of Auditory Scenes
- 45 McGlashan S. Speech Interfaces to Virtual Reality
- 46 Lyons K, Gandy M, Starner T. Guided by Voices: An Audio Augmented Reality System
- 47 Resnick P, Virzi R A. Relief from the Audio Interface Blues: Expanding the Spectrum of Menu, List, and Form Styles. ACM Transactions on Computer-Human Interaction, 1995, 2(2): 145~476
- 48 Resnick P, Virzi R A. Skip and Scan: Cleaning Up Telephone Interfaces. ACM, 1992. 419~426
- 49 Klemmer S R, et al. SUEDE: A Wizard of Oz prototyping tool for Speech User Interfaces
- 50 Rosenfeld R, Olsen D, Rudnicky A. Universal Human-Machine Speech Interface: A White paper
- 51 Shriver S, et al. A Unified Design for Human-Machine Voice Interaction