



语声研究、现代通信 与信息社会

刘 村 友

(电子工业部三〇研究所)

1984年6月22日收到

在当今“信息社会”里,用来传递信息的现代通信成为国民经济不可缺少的部分。语声在通信中占有重要地位,信息社会对语声研究提出了更高的要求。语声还能在通信以外的广阔领域发挥作用。我国的语声研究任重而道远。

一、引 言

信息社会需要进行大量的、各种各样的信息交换。广义讲,人类进行的任何信息交换活动都是“通信”,包括手势、表情之类。这里主要谈论狭义的“电通信”。

人类接收信息的途径有听觉(耳,听语声),视觉(眼,看图文),触觉(身,感知),嗅觉(鼻)和味觉(舌)。为盲人而设计的“触觉文字”早已问世。但目前普遍应用的还是听、视,即主要是语声和文字、图象的通信。

现代通信与老式通信有很多不同,它已发展为技术先进、设备齐全、能适应各种情况的复杂体系。

下面,首先谈谈现代通信的特点,然后讨论语声通信和语声研究。

二、现代通信的特点

1. 广泛,远距,网络化

通信现已渗透到国民经济和人民生活的各个角落。电报、电话、传真、图象等等愈来愈成为人们的“必需品”。没有现代通信,现代社会也就无法运转,这话并不过分。随着通信对象的

应用声学

迅速增长,大范围、远距离、组网配套成为必然趋势。通信卫星联接了地球上最远的用户;新的传输技术已开辟了星际通信的道路。现在,用户可以坐在自己家里与远隔重洋的朋友通电话、交换图片,甚至“现场”互访。

2. 多样、兼容、综合化

从信源讲,有语声、文字、图象、数据(含计算机数据)等多种型式;从传输通路讲,有电缆、光缆、短波、微波、卫星等恒参与变参信道;从信号格式讲,有模拟、数字及各类调制形式。不但多样,还要兼容。比如,要声、象对照,图、文并茂,要通信与分析、计算相结合等等。相应地,要求发展能传递语声、图象、数据等多种信息的综合性通信制式。

3. 经济、可靠、优质化

因为有大量信息通过信道,在规定信道容量内传递更多的信息就显得格外重要,这就是经济性。由于信道中流通着各类信号,自然界又存在着多种电磁干扰,所以要求在恶劣条件下保持信息传递的真确度就十分必要,这就是可靠性。有些特别敏感的信号,如控制核弹发射的信号,几乎不允许有任何差错,可靠性要非常高。进而,不但要保证通信,还要话音柔美,图文清晰,即优质。以上是从信息论观点提出的经济与可靠性要求。从实际设备的角度,则

要求“性能价格比”高,即价廉而物美。

4. 安全、保密、数字化

大量信号抛出,受到人为干扰、窃截、破译的机会增多,必须对消息采取更强有力的保护。不但军、政、经济重要情报要保密,而且商业、银行业、社会服务业的信息保密要求也日益强烈。银行从计算机里调用帐户档案就是不应让局外人知道其内容的。数字化给信号保密带来很大好处。当然,数字化的目的不只是为了保密。现代保密技术已可对各类信号实行强韧保护,使它们在很长的掩蔽时间内安全通信。目前,通信网和通信设备的数字化也已成为时髦而且实用的了。

5. 快速、自动、智能化

现代通信要求实时性好。甲方说话(电话)、画图(图象)、下指令(遥控),要求乙方立刻听到、看到和做出反应。在现代化社会,消息迟滞就会误大事。自动的记录、交换、显示、告警等等,可提高通信效率、加快速度和增强可靠性。此外,还要求通信设备具有一定的理解、辨别和独立处理问题的能力,即有某种“智能”或“智慧”。例如,要求专用电话能辨识讲话人,非规定人员打电话将被它拒绝。

三、语 声 通 信

用语声通信可以做到:简易(无需特别训练),迅速(能马上理解),准确(可准确地表达意思并可及时商谈与核实),识辨(识别讲话人,察辨讲话者的感情),高效率(不占用手脚)。

现代语声通信已不是单纯打普通电话那样的方式了。它面临着许多新的问题,主要是以下四方面。

1. 语声通信须适应各种环境——安静的、嘈杂的;固定的,移动的;有线的,无线的;空中、地面、水下的;常规的,特殊的。例如,有的要在舰船、飞机这样的高噪声环境鉴别和处理语声信号;有的要解决深潜员在氮氧混合气中的通信问题。为适应网系化和综合化,语声信号应呈现出与其它各类信号类似的“共性”,以便汇

入信息流和顺利地实现信号转接。例如,它要和其它信号一样,能容易地实现有/无线转接,时间或频率的分-合路,以及克服信号贮存、排序、延迟等带来的困难。

2. 经济、可靠和优质,是语声通信一直注目的课题。做到经济传输的根本出路是剔除语声信号的信息多余度。理论计算指出,一个典型的、带宽为 3KHZ 的信道,传输容量大约为 30000bits/sec。而实测表明,人类处理语声信息的能力则低于 50bits/sec。就是说,现在常用的 3KHZ 话音信道是太宽了,或者说,压缩语声频带(降低传输码率)的潜力是很大的。

要节省频带的理由是:首先,有些信道十分昂贵(如卫星),以宽的频带去传送多余信息是很不合算的;其次,现用的大量网系和设备是为传送模拟信号而设计的,用它们传送码化信号则相当困难(码化后频带拓宽了),因此要首先压缩原信号的频带,然后再码化;再次,有的信道干扰较大,为通信可靠而进行抗干扰处理也将拓宽频带,因而要用事先压缩频带来配合;最后,多余度剔除有利于保密,因为它减少了信号传输的重复性和暴露时间,且低码率易做到高度密化。低码率还大大有利于信息贮存,这在大量处理、交换信息的今天,无疑是十分可取的。

大量剔除信息多余度带来的问题是语声质量降低。因为代表个人发声特征的部分(如音色,音调变化)可能被当做“多余”而予以去除,使恢复后的语声变得粗糙、平淡、干枯,甚至怪声怪调和不清,这也相当于降低了可靠性。此外,剔除多余度后,单位信号的载信(息)量提高了,如不辅以抗干扰处理,单位信号受干扰的后果就可能很严重。大压缩比的设备也必然更复杂,造价也随之增高。

可见,传输的经济性与传输的可靠性有矛盾,与音质和设备的“性能·价格比”也有矛盾。用户对优质的要求是“象在对面讲话一样”,同时还要求省信道、造价和使用价都低。语声通信实用研究花了很大精力来寻找这些矛盾间的合理折衷,并一直受到这些矛盾的困扰。

3. 多功能、多用途、多种型式的需求日渐增长。用户已不满足于简单地拨号码打电话。他们可能希望用电话自动订票;用电话从银行自动取款;用电话从公用数据库内调出所需资料;用电话指挥家里的机器人干活,或操纵机房里计算机的运行。或许,他们希望用口述数字叫通电话;自己不在家时能自动录下外来电话并做出简捷回答。有的用户希望自己的谈话内容高度保密;有的想把电话装在衣袋里,随身使用;有的则要求打电话时能看到收话人,能留下对方的笔迹和照片。凡此等等,大大超出了常规的“市内电话”和“载波长途”的范畴。困难在于,这些需求不断“更新”,而原有的老信道、老设备、老网络则不可能迅速“跟进”。语声通信是在许多不利的限制条件下完成自己的改革,其艰辛可想而知。

4. 向数字化发展是现代语声通信的一大特色。数字化不但利于优质、智能、交换、复用、再生、监控、加密和综合业务,而且可与现代计算技术和微电路技术相结合,实现十分复杂的处理。比如,最常用的脉码调制器(PCM)、增量调制器(DM)及它们的改进型,已可在一个、两个微片上实现多路编解码,音质也可相当好。把语声码、数据码、图象码等组合在一起的全数字式综合交换网已付诸使用。用十来个片子已可做成码率为2400bits/sec的高质量声码器。语声数字化的阻力是信道传输有困难,老设备要改造,或数字化设备太贵。

四、语 声 研 究

信息社会对语声的多种需求决定了语声研究工作的迫切性、艰巨性。语声工作者多年来努力探索满足这些需求的途径和解决上述矛盾与困难的方法,并取得了可喜的进展。

早期的语声通信是以“波形保真”为依据的。在当时,这是一种自然而朴素的想法:既然声音是由介质振动表现出来的,那么,能把这种振动(波形)真实地传送和恢复,通信质量就有了保证。这一朴素想法成功而广泛地在实际通

信中应用了,并一直延用到现在。

与“波形保真”概念相伴随,对语声波形的研究——诸如它的长期统计特性,短时变化特性,前后相关性等——也蓬勃开展起来。在PCM中,以 μ 律分层法代替均匀分层法就是利用语声信号的幅度分布特性——小信号出现机会多,应予精细分层——把码率压低1/3左右而仍然保持高音质。差分脉码调制(DPCM)和线性预测编码(LPC)等,则充分利用了信号的前后相关性质,尽量少地传送信号中有关联的部分,从而节省码率好几倍,音质仍较好。

随着信息处理、传送量和各种需求的迅速增加,由“波形保真”原则带来的码率节省和产生的信息格式已不能满足需要了。语声研究进入一个新的、更深的阶段——寻找构成语声的基本要素或成分,它们的变化规律,统计性质,对音质的影响以及提取和恢复它们的方法等等。这就是“参数保真”观点。这一观点不担心语声经处理和传输后的波形变化,而着眼于保证语声主要参数的真确描述和保证由这些参数合理地“再制成”话音。

一个典型的例子是声码器。声码器从语声中提取出三个主要参数:声道传递特性或语声能量频谱,激励声道的信源性质(浊音,或清音)以及浊音时的基本频率(基调)。只要这三个参数提取准确,传送速率适当(能反映参数的变化),由它们重制(合成)话音的方法合理,就能得到清晰度很高的话音,自然度也可令人接受(现时还做不到令人满意),而码率(或频带)则可压缩到一般话音的几十分之一。

对语声参数的研究已取得很大成绩。现在,不但对语声的物理参数有了更多认识,而且已深入到语声产生的机理、生理等领域,制成了水平相当高的人类发声模型;电脑支配的话音应答机以及讲演者、歌唱家也已进入了社会。

人类用语声通信的两个终端是口(发声机构)和耳(接收机构)。语声研究几乎从一开始就把这两者联系起来考虑了。道理很简单:语声的通信效果只能由耳的接收效果来评判。如果把人的口比做厨师(制造各种语声“菜肴”),

那么人的耳就是顾客。所谓哪个语声参数重要或不重要,哪个语声参数需要多传送或少传送,最终都要以人耳的“听解”需求为依据。

对于人耳的动作机能和物理性质可以说是基本上搞清楚了。但更深的问题——人的听觉神经传送、接收和理解语声的奥秘,还正在探索之中。对人耳功能研究成果之一是由电脑构成的人工耳(语言识别器,语声感知器),它已可能“听懂”几百条人类发出的词汇,听解正确率达90%以上。只能听懂几十条命令的微电脑人工耳已进入市场和人们的生活中,价钱也不算贵。当然,和真正的人耳相比,它们的听辨能力和适应能力是太差了。

对语声(以及人类发声机构)和人耳功能的深入研究将引导人们更真确地仿出人的口和耳的模型,使它们能准确地“说话”,无误地“听话”,并以极低的速率(例如100bits/sec)实现语声传递。

达到上述目标的途程仍是艰难而漫长的,有很多事情还很难弄清楚。比如,人类理解语言的过程并不是孤立而简单地对声信号做出反应,它起码是前后连贯、意义相关并按习惯进行及时判断和纠错的复杂过程。可以想见,把如此高度的“智慧”发掘出来,并让电脑(即使是最近代的电脑)学会和熟练,绝不是件容易的事。

语声研究的另一课题是音质评定。人们为寻找一套既科学又实用的评定方法已努力几十年了。

语声的基础研究和应用研究都有很多事要做,任重而道远。

五、语声与国民经济

语声在国民经济中的作用,有时不易被人们理解和重视。原因之一是人们对“语声”已司空见惯,觉得很平常,没什么可研究的;原因之二是人们还没有普遍看到语声“新成就”在他们

工作、生活中引起的变化和带来的方便。在目前我国科学技术还比较落后的情况下,这一点尤为突出。

事实上,语声在国民经济中的作用,不限于上面谈到的通信领域,在遥感、遥控、自动变换方面都可发挥它的特长。它可为办事的高效能、高速度和高精度提供条件,而这是信息社会(或现代化社会)所必需的。语声感知器配以自动控制设备,就能按照人的口述命令行动。宇航员可用话音去操作;病残者用话音“取茶拿饭”;管理员用话音“填写清单”;……。这不但解放了手脚,还可与手脚同时工作,充分利用了时间。用语声自动变换技术有可能实现口述电报、声控打字和自动语言翻译等等,大大减少人的繁琐、重复劳动,加快工作速度。把话音自动变成文字、图象,或反过来把文字变成话音,就使聋哑人能利用他们的健康功能“听音”(看语声图)和“说话”(键控话音)。

当然,语声在各个领域的应用并不是一朝一夕可实现的。一方面,对语声和听觉的认识还不够深透,技术上还满足不了各类应用提出的要求;另一方面,推广应用受到社会情况和经济发展等的制约。比如,在劳动力充裕、社会生活还没有那么现代化的情况下,用语声去完成各种指挥、控制、应答,就要考虑有无必要和经济上合不合算的问题。以我国的现状来说,发展、改善和完备现代语声通信是势在必行,而且迫在眉睫。而语声的感知、遥控遥测和自动变换技术,看来应首先进入那些“非用不可”的场合。比如,操作者手脚忙碌,必须用语声帮助的场合;操作者在运动中,不使用手脚工作的场合;伤残人失去手脚动作能力的场合等等。

从语声在国民经济中已起到的和将要起到的作用来看,从信息社会对语声的各种需求来看,我国的语声研究工作需要进一步加强。研究工作的前途是广阔的,任务相当繁重。