

21st CENTURY
规划教材

高等院校信息与电子技术类规划教材
University Textbooks of Information Technology

语音信号及单片机处理

吴黎明 主 编

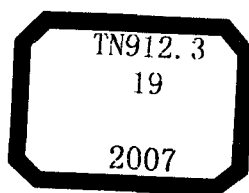
王桂棠 吴正光 副主编



21st CENTURY



科学出版社
www.sciencep.com



中国科学院教材建设专家委员会教材建设立项项目
高等院校信息与电子技术类规划教材



语音信号及单片机处理

吴黎明 主编
王桂棠 吴正光 副主编

科学出版社

北京

内 容 简 介

本书围绕语音信号的单片机处理技术,介绍了语音信号的时域和频域处理技术以及语音识别技术,并以凌阳 16 位单片机为例,系统介绍采用凌阳单片机系统组成、软硬件和应用系统的设计。特别是结合单片机对语音信号的处理和应用,使单片机控制器具备了更多更强的功能,实现应用系统形式多样并“能听会说”。通过本书的学习,能够掌握语音信号及其处理技术的应用,促进对单片机技术的深入理解,使单片机软硬件知识的学习能够有效结合工程实际需要、进而深化理解,使开发能力得到提高。

本书可以作为高校信息技术类专业,以及工程技术类专业学生学习的教材和参考书,还可作为工程技术人员学习单片机原理和语音信号处理应用技术的参考书。

图书在版编目 (CIP) 数据

语音信号及单片机处理/吴黎明主编. —北京:科学出版社, 2007

(高等院校信息与电子技术类规划教材)

ISBN 978-7-03-018628-7

I. 语… II. 吴… III. 单片微型计算机-语音数据处理-高等学校-教材 IV. TN912.3 TP368.1

中国版本图书馆 CIP 数据核字 (2007) 第 022960 号

责任编辑:李 伟/责任校对:刘彦妮

责任印制:吕春珉/封面设计:飞天创意

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

铭浩彩色印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2007 年 3 月第 一 版 开本:787×1092 1/16

2007 年 3 月第一次印刷 印张:23 1/2

印数:1—3 000 字数:537 000

定价:33.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

销售部电话 010-62136131 编辑部电话 010-62138978-8777(HI01)

高等院校信息与电子技术类规划教材

编 委 会

主 任 吴黎明（广东工业大学信息工程学院副院长、教授）

副主任 贺前华（华南理工大学电子与信息学院副院长、教授）

委 员 （按姓氏笔画排序）

马文华（广东外语外贸大学信息科学与技术学院副教授）

汤 庸（中山大学信息科学与技术学院副院长、教授）

杨振野（广东技术师范学院电子系教授）

吴正光（广州大学实验中心副主任、高级工程师）

周美娟（广东海洋大学信息学院院长、教授）

洪添胜（华南农业大学信息学院院长、教授）

徐 杜（广东工业大学信息工程学院院长、教授）

颜国正（上海交通大学电子信息与电气工程学院测控系主任、
教授）

前 言

语音信号的数字化处理和应用随着计算机技术的迅速发展而普及，并在结合单片机应用的过程中不断得到深入。单片机不但以其高可靠性、高性能价格比，在工业控制系统、数据采集系统、智能化仪器仪表、现代通信设备、嵌入式系统，以及测控领域等方面得到了广泛的应用，更由于 16 位单片机的出现，处理能力得到提高，具备 DSP 功能，可以采用单片机技术直接对语音信号进行处理和应用，使单片机作为微控制器具备了更多更强的功能，并使应用系统形式多样化，成为“能听会说”的机器，利用机器视觉也将成为可能。本书的编写以掌握语音信号及其处理技术的应用，促进对单片机技术的学习，使单片机软硬件知识的学习能够落实到有效结合工程实际需要，实现提高综合开发能力的目标。本书介绍的凌阳 16 位 61 系列单片机，可以作为学习单片机基础应用技术的入门技术。由于其功能强、性能好、上手快，也可以很快提高 MCS-51 等型号单片机的爱好者的应用能力。许多高校的同学在参加全国大学生电子设计竞赛时也可以此书为参考书。

本书内容兼顾语音信号处理和单片机原理应用，力图有效地帮助读者解决在学习单片机应用和数字语音处理技术过程中碰到的技术问题，同时也是理论教学和工程实践相结合的尝试。全书分为 13 章。

第 1 章介绍了语音信号分析与处理技术发展概况和单片机的发展趋势，讲述了嵌入式计算机语音处理的原理和微机原理的一些基础知识。

第 2、3 章介绍了语音信号的时域分析原理和记录文件、语音信号的频域分析原理和语音识别技术。

第 4、5 章介绍 16 位凌阳 61 系列单片机的组成结构、复位电路设计以及存储器配置等，介绍了 61 系列单片机的指令系统。

第 6、7、8 章分别介绍单片机系统硬件组成、单片机的中断系统和单片机应用系统接口扩展的原理及技术。

第 9 章介绍了单片机汇编语言伪指令、汇编语言程序设计。

第 10 章介绍了单片机 C 语言程序设计。

第 11 章介绍单片机系统应用技术设计，给出实际应用例子，提供了详细的电路和程序代码。

第 12 章介绍了 61 系列单片机应用的系统开发方法，简要介绍开发环境的功能，介绍了凌阳语音处理算法库及其应用程序接口 API，包括 A/D、编码处理、存储、解码处理以及 D/A 等功能。

第 13 章以语音处理为实例，进一步介绍单片机系统实例，重点介绍提供实例的技

术特点。

本书的附录列出凌阳单片机的指令集、伪指令集，并介绍了集成开发环境 IDE，方便读者查阅。

本书由广东工业大学、广州大学和广东技术师范学院的相关教师共同合作编写，这些老师都是从事多年信号与信息处理、单片机技术的教学、工程与实验技术的人员。吴黎明、吴正光编写了第 1、2、3、13 章，伍冯洁编写了第 4、12 章，张秀松编写了第 6、7、8 章以及第 10 章的部分内容，邓耀华、汤秀春编写了第 5、9 章和附录，王桂棠编写了第 11 章以及第 10 章的部分内容，全书由吴黎明统稿。此外，汤秀春高级实验师还负责本书的上机实验验证，研究生何仲凯、薛向东、刘润予、朱高峰为本书编写做了大量的文字工作和参加上机实验验证。本书的习题和电路图由李政广、李冬冬、王立萍、赖南辉编制。

谭胜波高级实验师多次参加本书编写提纲的讨论，提出不少的建议，并担任本书的主审。本书的编写得到凌阳大学计划的大力支持，并参考了专家、同行不少著作、教材和资料，在书后一一列举，谨在此一起致谢。书中存在错误和不足之处，恳请广大读者批评指正。

目 录

第 1 章 绪论	1
1.1 音频和语音处理	1
1.2 国外语音处理技术概况	1
1.3 我国语音处理技术的发展	4
1.4 语音处理技术的发展趋势	5
1.5 嵌入式语音处理技术	7
1.6 单片机的历史及发展.....	10
1.6.1 单片机的历史	10
1.6.2 单片机的发展趋势	10
1.7 凌阳单片机简介.....	12
1.7.1 凌阳 8 位单片机	13
1.7.2 凌阳 16 位单片机.....	13
习题	17
第 2 章 语音信号的时域分析	18
2.1 语音分析处理概述.....	18
2.2 语音信号的数字化和预处理.....	18
2.2.1 数字化	19
2.2.2 预加重处理	20
2.2.3 防混叠滤波	20
2.3 语音信号的加窗处理.....	20
2.4 短时平均能量与短时平均幅度.....	22
2.5 短时平均过零率.....	23
2.6 短时相关分析.....	24
2.7 计算机中音频的存储.....	26
习题	28
第 3 章 语音信号频域分析和语音识别	29
3.1 语音信号的频域分析概述.....	29
3.2 短时傅里叶变换.....	29
3.3 傅里叶变换的解释.....	30
3.3.1 频谱	30
3.3.2 滤波器	30

3.4 短时傅里叶反变换	31
3.4.1 滤波器组求和法	31
3.4.2 快速傅里叶变换求和法	32
3.5 语音识别的一般技术	34
3.5.1 语音识别基本原理	34
3.5.2 孤立词识别系统	36
3.5.3 语言模型	36
3.6 语音识别中的特征提取及其谱失真测度	37
3.6.1 带通滤波器组法的频谱参数及其失真测度	37
3.6.2 线性预测倒谱系数及其谱失真测度	38
3.6.3 Mel 频率倒谱系数及其失真测度	39
3.7 语音信号的矢量量化	40
3.7.1 矢量量化的基本原理	41
3.7.2 最佳矢量量化器和码书的设计	42
3.8 模板匹配法	44
习题	46
第4章 凌阳 16 位单片机系统结构	47
4.1 SPCE061A 单片机的引脚	47
4.2 $\mu'nSP^TM$ 的内核结构	51
4.2.1 用户寄存器	51
4.2.2 算术逻辑运算	51
4.3 寄存器	52
4.4 总线	54
4.5 时钟电路和工作模式	54
4.5.1 时钟电路	54
4.5.2 系统时钟	55
4.6 SPCE061A 最小系统与系统复位	57
4.6.1 SPCE061A 最小系统	57
4.6.2 系统复位	57
4.7 SPCE061A 存储器结构	60
4.7.1 片内 SRAM	60
4.7.2 片内 Flash 存储器	62
4.7.3 用并口扩展 SRAM	67
习题	73
第5章 指令系统	74
5.1 指令的格式与分类	74
5.2 数据传送指令	75
5.3 算术运算类指令	79

5.3.1 加法运算	79
5.3.2 减法运算	81
5.3.3 带进位的加减运算	82
5.3.4 取补运算	83
5.3.5 乘法运算	84
5.3.6 n 项内积运算	84
5.3.7 比较运算	85
5.4 逻辑操作类指令	86
5.4.1 逻辑与	86
5.4.2 逻辑或	88
5.4.3 逻辑异或	89
5.4.4 测试	89
5.4.5 移位操作	91
5.5 控制转移类指令	94
5.5.1 无符号数的跳转指令	95
5.5.2 有符号数的跳转指令	96
5.5.3 其他跳转指令	97
习题	98
第 6 章 SPCE061A 单片机的接口	99
6.1 SPCE061A 输入/输出接口	99
6.2 时基系统与定时器/计数器 CTC	108
6.2.1 时基系统	108
6.2.2 定时器/计数器	110
6.3 模数转换器 ADC	116
6.4 数模转换 DAC&PWM	125
6.5 串行接口	127
6.5.1 SIO	127
6.5.2 通用异步串行接口 UART	131
习题	135
第 7 章 单片机中断系统	136
7.1 中断系统概述	136
7.2 SPCE061A 中断系统	137
7.2.1 三种类型的中断	137
7.2.2 中断源	138
7.2.3 中断控制	141
7.2.4 中断控制指令	142
7.2.5 中断优先级	144
7.2.6 中断响应过程	145

7.2.7 中断响应时间	145
7.2.8 中断服务流程	146
7.3 中断系统的应用	148
7.3.1 定时器中断	148
7.3.2 触键唤醒中断	150
7.3.3 串行异步中断	152
习题	156
第8章 系统的扩展技术	157
8.1 简单并行接口	157
8.1.1 简单并行接口电路的结构	157
8.1.2 简单并行接口的特点	158
8.2 并行口的扩展	158
8.2.1 并行口扩展的电路连接	158
8.2.2 管理程序	160
8.2.3 用串行口扩展并行口	160
8.3 串行口的扩展	163
8.3.1 利用单片机并行口的位操作功能作为串行口	163
8.3.2 利用移位寄存器扩展串行口	164
8.4 A/D、D/A 的扩展	165
8.4.1 串行输出的 D/A 转换器	165
8.4.2 串行输出的 A/D 转换器	167
习题	168
第9章 单片机汇编语言程序设计	169
9.1 $\mu'nSP^TM$ IDE 的项目组织结构	169
9.2 伪指令	170
9.2.1 伪指令的语法格式及特点	171
9.2.2 伪指令符号约定	171
9.2.3 标准伪指令	171
9.2.4 宏定义与调用	182
9.2.5 段的定义与调用	185
9.2.6 结构的定义与调用	186
9.2.7 过程的定义与调用	189
9.2.8 伪指令的应用举例	190
9.3 汇编语言程序设计	200
9.3.1 一个简单的汇编代码	200
9.3.2 汇编的语法格式	202
9.3.3 汇编语言的程序结构	205
9.3.4 嵌套与递归	217

习题	220
第 10 章 单片机 C 语言程序设计	221
10.1 C 语言程序设计	221
10.1.1 μ^n SP TM 支持的 C 语言算术逻辑操作符	221
10.1.2 C 语言支持的数据类型	221
10.2 C 语言编程函数库及调用方法	222
10.2.1 C 语言编程函数库	222
10.2.2 库函数应用举例	226
10.2.3 程序调用协议	227
10.3 在 C 语言程序中调用汇编函数	228
10.4 在汇编程序中调用 C 语言函数	233
10.5 C 语言和汇编语言混合编程举例	234
10.6 C 语言的嵌入式汇编	237
习题	242
第 11 章 61 系列单片机的应用	243
11.1 μ^n SP TM 的应用领域	243
11.1.1 数字信号处理应用	243
11.1.2 用于开发研制便携移动式终端	244
11.1.3 用于开发嵌入式应用系统	244
11.2 三角波、正弦波、方波波形发生器	245
11.2.1 原理分析	246
11.2.2 硬件电路	246
11.2.3 程序说明	247
11.3 红外线遥控	248
11.3.1 原理分析	248
11.3.2 硬件电路	250
11.3.3 程序设计	250
11.4 FIR 滤波器	251
11.4.1 乘-累加 (MAC) 功能	251
11.4.2 低通 FIR 滤波器的实现	252
11.4.3 μ^n SP TM 实现 FIR 滤波需要注意的问题	256
11.4.4 滤波系数出现负数时的滤波运算	256
11.5 FFT 算法	257
11.5.1 FFT 算法原理	258
11.5.2 用 μ^n SP TM 实现 FFT 算法	259
习题	264
第 12 章 应用开发及支持工具	265
12.1 μ^n SP TM 系统开发设计方法	265

12.2 集成开发环境 IDE	268
12.3 编码算法库	269
12.3.1 概述	269
12.3.2 SACM_A2000	269
12.3.3 SACM_S480	275
12.3.4 SACM_S240	279
12.3.5 SACM_MS01	283
12.3.6 SACM_DVR	288
12.4 在线调试器	294
习题	295
第13章 语音综合应用系统	296
13.1 语音录放系统	296
13.1.1 系统结构	296
13.1.2 原理分析	296
13.1.3 硬件电路	297
13.1.4 程序设计	298
13.2 SPCE061A 语音遥控器	300
13.2.1 系统组成	301
13.2.2 硬件组成	301
13.2.3 软件设计	302
13.3 SPCE061A 在 IP 电话中的应用	303
13.3.1 IP Phone 手柄方案基本技术功能	304
13.3.2 基本参数	304
13.3.3 人机交互界面	304
13.3.4 系统硬件与功能概述	304
13.3.5 系统的软件设计	308
13.3.6 计算机端软件设计	314
13.4 多功能声控玩具车	315
13.4.1 系统的功能组成	315
13.4.2 系统的程序结构	315
13.4.3 乐音识别的软件设计	317
习题	322
附录 A $\mu'nSP^{\text{TM}}$ 的指令集	323
附录 B $\mu'nSP^{\text{TM}}$ 汇编器伪指令集	331
附录 C 集成开发环境 IDE	332
C.1 集成开发环境的界面	332
C.1.1 菜单栏	332
C.1.2 工具栏	338

C.1.3 窗口	339
C.2 项目	349
C.3 代码剖视器的使用及功能	357
主要参考文献	360

第 1 章 绪 论

1.1 音频和语音处理

我们所说的音频是指频率为 $20\text{Hz}\sim 20\text{kHz}$ 的声音信号,分为波形声音、语音和音乐三种,其中波形声音就是自然界中所有的声音,是声音数字化的基础。语音也可以表示为波形声音,但波形声音表示不出语言、语音学的内涵。语音是对讲话声音的一次抽象,是语言的载体,是人类社会特有的一种信息系统,是社会交际工具的符号。音乐与语音相比更规范一些,是符号化了的的声音,但音乐不能对所有的声音进行符号化。乐谱是符号化声音的符号组,表示比单个符号更复杂的声音信息内容。

语言是人类创造的,是人类区别于其他地球生命的本质特征之一。语音是语言最本质、最自然、最直接的表现形式或载体,其表现形式为声波——一种由空气分子振动而形成的机械波。人类用语言交流的过程可以看成是一个复杂的通信过程,为了获取便于分析和处理的语音信源,必须将在空气中传播的声波转变为包含语音信息并且记载着声波物理性质的模拟(或数字)电信号,即语音信号,因此语音信号就成为语音的表现形式或载体。

人们对语言的研究早已有之,其中语音学是研究人类语音的产生、传播及感知等过程机理的学科,包括发音语音学、声学语音学和听觉语音学三个分支。发音语音学研究发音器官在发音过程中的运动和语音的音位特性;声学语音学研究语音的物理属性(语音声波的振幅、频率和频谱特性等);听觉语音学研究听觉和语音感知。数字信号处理是一门通过计算机或其他专用设备,对离散信号用数字方式进行增强、压缩、滤波、变换及识别等处理的新兴学科。语音学和数字信号处理的交叉结合便形成了语音信号处理。语音信号处理(简称语音处理技术或语音技术)是建立在语音学和数字信号处理基础之上的,对语音信号模型进行分析、存储(编码)、传输、识别和合成等方面研究的一门综合性学科。它包括语音编码、语音识别、说话人识别和语音合成四大学科分支,并由此形成了语音分析技术、语音存储(编码)技术、语音识别技术和语音合成技术四大实用技术。

1.2 国外语音处理技术概况

自 1876 年 Bell 发明了采用声电转换技术实现远距离语音通信的电话开始,语音处理技术的发展大致经历了以下几个阶段。

1. 萌芽阶段

在这一阶段(20 世纪 30 年代~50 年代),人们对语音处理的研究主要是根据语音

学知识,提取若干特征参数,并利用这些参数制成模拟电路来模仿人的发音过程,实现简单的语音处理功能。

1930年,H. Dudley 首次成功研制出采用传输并合成的方法从语音中提取表征语音信息特征参量的声码器,其创造性的设计思路形成了语音产生模型的基本思想。1948年,美国 Haskins 实验室首先研制出由语谱图自动合成语音的语图回放机,这一发明直接孕育出共振峰合成法这一至今仍被认为是较好的语音合成的方法。1952年,Bell 实验室的 Davis 等人首次成功研制出可以识别 10 个英文数字的语音识别系统——Audry 系统,它根据语音第一、二共振峰提取若干特征参数以形成参考语音模式,然后通过计算参考语音模式的语音与未知语音之间的相关程度来达到识别目的。1958年,Duddley 和 Balashek 对其进行了改进,将语音分割为元音和辅音等语音单位,开创了音素识别的先河。1956年,Olson 和 Belar 等人采用了 8 个带通滤波器,提取频谱参量作为语音特征,成功研制出一台简单的声控打字机。

2. 发展阶段

20 世纪 60 年代~80 年代初,随着集成电路技术和计算机技术的发展,语音处理技术的理论和应用亦日趋完善和成熟。

1960年,Denes 和 Mathew 把数字计算机引入语音识别,从而改变了采用模拟电路进行语音处理的传统做法。计算机的应用推动了语音识别的发展,形成了线性预测分析技术 LP (Linear Prediction) 和动态规划 DP (Dynmic Programming) 两项重要成果。LP 较好地解决了语音信号产生模型的问题,对语音识别的发展产生了深远影响。1966年,美国麻省理工学院林肯实验室的 Gold 等人采用 16 个带通滤波器、基音检测器、浊音检测器和一台计算机构成了一个语音识别系统。在整个 20 世纪 60 年代,语音识别的研究主要是根据语音产生机理和人耳对不同频率语音的感知差别,采用硬件实现的滤波器组提取频谱特征,通过计算机进行匹配和判决。

近几十年来,语音处理技术得到了迅速的发展,在此期间产生了一些重大的理论突破。

- 20 世纪 70 年代,由 Itakura 提出的表示语音参数相似度测量的线性预测残差原理以及线性预测编码 (LPC) 较好地解决了语音特征的提取问题,从而使线性预测技术成功地应用于语音编码及语音识别。

- 把用于解决有序优化问题的动态规划技术应用到语音识别中。1972 年,由 Sakoe 提出的动态时间弯曲 DTW (Dynamic Time Warping) 算法有效地解决了语音两次发音之间的时间变形问题,对特定人孤立词的识别十分有效。从此,基于 LPC 分析及 DTW 算法的中、小字表孤立词特定人语音识别系统纷纷建立起来,语音识别开始走出实验室而进入实际应用阶段。

- 20 世纪 80 年代初,Gray 和 Markel 等人解决了矢量量化码书生成的方法。于是,一些学者就将这项原本用于信息压缩理论中的矢量量化 VQ (Vector Quantizaton) 技术成功用于语音编码以及语音识别之中。矢量量化的作用就是进行数据压缩,将连续的语音特征空间离散化,降低系统在时间及空间(存储)上的开销,从而减小语音处理

的复杂程度。矢量量化的另一个作用就是通过聚类分析获取一人或者多人的多次语音样本所共有的语音特征。

- 20 世纪 80 年代中期开始得到极大发展并成为语音处理研究热点的隐马尔可夫模型 HMM (Hidden Markov Model), 逐渐成为现代语音处理领域的重要理论基础之一, 并在语音处理的各个领域中得到十分广泛的应用。

3. 实用阶段

20 世纪 80 年代至今, 随着遵循摩尔定律的超大规模集成电路技术的迅速发展、个人计算机的迅速普及极大地促进了计算机技术和人工智能技术的迅猛发展, 使人类社会进入到数字信息时代。在此社会背景下, 人们对语音技术的实际需求愈发迫切, 促进了语音处理技术的发展和语音处理实用化产品的出现。

从 20 世纪 70 年代末开始, 由于大规模集成电路技术和语音理论与技术的成熟, 语音处理技术开始步入实用化阶段。1976 年, Votrax 公司推出的 Computalker 语音合成器投入市场。它采用 8080 微处理器, 并采用 S-100 总线与其他微计算机系统连接, 有 6KB 的存储器存储音素表和程序。当输入机器可读的标准语音表代码后, Computalker 产生合成语音。虽然合成语音的音质很差, 但是合成语音技术本身已经为人们所接受。同年, Votrax 公司推出另一款语音合成器——ML-I, 它是第一个由规则合成语音的产品。ML-I 采用 80 个音节、8 级音高和 4 级不同发音持续时间, 并提供了一份包含 625 个单词和短语的词典。1978 年, TI 公司首次推出采用超大规模集成电路技术的单片语音合成器 TMC0280。该产品成为语音 DSP 芯片的前身, 并使 TI 公司遥遥领先于同行。TI 公司用此芯片推出了一种产品——Speak'n Spell Toy, 使语音技术走出实验室进入市场。其硬件结构采用 4 位微处理器 TMS1000, 2 个 128Kb 的 ROM, 存储约 330 个单词和短语 (语音持续 3~4min), 数据传输率为 1200b/s; 采用线性预测合成法, 由格型滤波器实现, 10 级格型滤波器用 10 个反射系数表示; 语音合成的控制参数有 12 个: 10 个反射系数、1 个能量参数和 1 个音高参数。继 1978 年 TI 公司推出会讲话的 Speak'n Spell Toy 之后, 又出现了许多会讲话的产品, 如会讲话的怀表、会讲话的微波灶、会讲话的弹球机、会讲话的计算器等, 可以看成是嵌入式语音产品的雏形。

20 世纪 80 年代初, Votrax 公司推出了采用音素合成技术的大规模集成电路芯片 SC-01。随着计算机技术的飞速发展和小词汇量特定人孤立词语的语音识别技术日趋成熟, 特别是对 HMM 的深入研究和广泛应用, 出现了语音处理技术产品化的热潮。1985 年, 东京的 Matsushita 研究所研制了非特定人孤立系统, 包括 LPC 倒谱系数、辅音号分段、元音和半元音识别、辅音识别、音节序列和词的匹配, 该系统对 274 个词的识别率为 95.6%。在 Yamatokoriyama 的 Sharp 信息系统实验室, 根据日语的特点 (日语大约有 100 个音节, 语音以音节为单位), 用音节作为识别基元, 用音韵规则得到每个音节间的关系, 并采用 DTW 方法, 对 300 个孤立字的特定人的识别率是 94%。1985 年, Sharp 公司在超级市场放置了一个用声音操作的字处理系统, 引起人们的关注。ATR 将神经网络用于语音识别。1988 年, Waibel 用时延网络 TDNN 解决了难以区分的 B、D 和 G 的问题。网络能自学一些特征, 因此, 神经元的识别率是 98.5%, 而

HMM 方法的识别率为 93.7%。1989 年, NEC 公司的 Sokoe 将神经元用于孤立字非特定人的语音识别, 它利用动态规划技术 DTW 的优点, 通过自学改进性能, 对日语数字的识别率为 99.3%。

20 世纪 90 年代初, 国外许多研究机构都研制出词汇量达到几万的大词汇量识别系统, 例如能识别 70000 个词汇的 Dragon Dictate 词汇翻译系统, 识别率为 80% 以上。在非特定人连续词语识别方面有代表性的系统是卡内基梅隆大学研制成功 SPHINX 系统, 能识别包括 997 个词汇的连续语句, 识别率达到 95.8%。IBM 公司在语音识别领域的研究已有 30 年, 其语音技术一直处于世界领先地位, 并在这一领域拥有近百项专利。早在 1985 年, IBM 公司就成功研制出 5000 个词的英语听写机 Tangora-5; 20 世纪 80 年代末, 研制出能识别 20000 个词汇的 Tangora-20, 识别率达到 94.6%, 而且具有快速自适应说话者的特性; 1997 年, 又推出了汉语听写机产品 ViaVoice, 为语音识别技术在汉字输入方面的实际应用开辟了新的道路。此后, IBM 公司的 ViaVoice 抢占了我国 90% 的语音识别市场。

鉴于语音识别产品的鲁棒性较差, 对语音处理技术的研究将更加深入。美国 DARPA 战略计算计划局提出研究口语系统 (spoken language system), 该系统要求把语音识别与自然语言理解结合起来, 即让计算机像人一样具有语言理解能力, 而不需过多地在孤立词识别上下功夫, 从而形成了新一代语音识别系统。

1.3 我国语音处理技术的发展

我国对语音处理技术的研究起步要晚一些, 1980 年之前只有中国科学院声学所、清华大学计算机系等单位从事语音识别的研究工作。经过 20 多年的努力, 我国在语音处理领域取得了显著进步。

1. 汉语听写机方面

与英语相比, 由于汉语语音输入的迫切性, 以及汉语音节种类较少、结构很规则、协同发音和音变相对不严重, 便于以音节识别为基础实现无限词汇识别等特点, 因此很快获得了可与国际先进水平相比拟的成果, 一些汉语语音输入系统已经实用化。1988 年, 清华大学、中国科学院声学所首先研制成功无限词汇的汉语听写机。20 世纪 80 年代末, 四达技术开发中心率先推出汉语输入的实用产品, 并于 1991 年与哈尔滨工业大学合作推出具有自然语言理解处理能力的汉语听写机。这一时期的汉语听写机系统基本上都是基于特定人孤立音节识别技术的。孤立音节识别系统只能一个字一个字地读入, 断断续续的, 既不自然又很费力, 这样的听写机产品自然不能为广大用户所接受。在国家“863”计划支持下, 近几年来清华大学和中国科学院自动化所等单位研制的听写机原理样机, 不仅包含一个很大的多音节词表, 而且能用于非特定人连续语句输入, 用起来就方便多了。目前正在考虑改进性能、开发产品。近几年来, 汉语语音识别受到了前所未有的重视, 国外有多个公司投资巨款, 猛攻汉语听写机的研制, 台湾省也在听写机研究方面下了很大功夫。国家“863”计划最近几年也加大了这一研究的投资力度, 并