

语音识别与合成

杨家沅 主编

四川科学技术出版社



数据加载失败，请稍后重试！

DG65/17

语音识别与合成

杨家沅 主编

四川科学技术出版社

新登字(川)004号

书 名/ 语音识别与合成
编著者/ 杨家沅 主编

责任编辑· 喻瑞卿 任维丽
封面设计· 何 成
版面设计· 杨 杨
责任校对· 于 英

出 版 四川科学技术出版社
成都盐道街3号 邮编 610016
印 刷 四川省地矿局测绘队印刷厂
版 次 1994年9月成都第一版
1994年9月第一次印刷
规 格 787×1092毫米 1/16
印张 34 820千字
印 数 1—600册
定 价 50.00元
ISBN 7-5364-3002-7/H·33

内容简介

本书是第三届全国人机语音通讯学术会议的论文集,共收录论文 98 篇,内容包括综述、语音分析与特征提取、基音提取与声调识别、语音识别模型与方法、非特定人语音识别、连续语音识别与语言模型、噪声环境下的语音识别、说话人识别、神经网络应用、语音合成、语音编码、语音库与系统评价方法和语音技术应用与硬件实现等十三个部分。

本书反映了全国在语音合成、识别和理解方面学术研究和技术开发的最新成果,对我国人机语音通讯技术的发展有重要的参考价值。本书可供计算机、自动化、声学、通讯等专业的科研人员、工程技术人员和大专院校师生参考。

NCMMSC-94

第三届全国人机语音通讯学术会议

论文集

1994. 10. 24—1994. 10. 30

中国自动化学会	模式识别与机器智能专业委员会
中国计算机学会	人工智能与模式识别专业委员会
中国电子学会	信号处理学会语音图象通讯专业委员会
中国声学学会	语言听觉和音乐声学分科学会
中国中文信息学会	基础理论专业委员会
中国通信学会	通信理论专业委员会
国家高技术智能计算机系统专家组	

重 庆

一九九四年十月

1015272

第三届全国人机语音通讯学术会议

- 会议主席：**张家骥（中国声学学会 语音听觉和音乐声学分科学会）
- 副主席：**方棣棠（中国计算机学会 人工智能与模式识别专业委员会）
（中国中文信息学会 基础理论专业委员会）
- 黄泰翼（中国自动化学会 模式识别与机器智能专业委员会）
- 李昌立（中国电子学会 信号处理学会
语音图象通信专业委员会）
- 王作英（中国通信学会 通信理论专业委员会）
- 高文（国家高技术智能计算机系统专家组）
- 杨家沅（四川联合大学）
- 邱玉辉（西南师范大学）

程序委员会主任：杨家沅（四川联合大学）

委员：

- | | |
|--------------|--------------|
| 王仁华（中国科技大学） | 王作英（清华大学） |
| 方棣棠（清华大学） | 刘增寿（国防科技大学） |
| 吴文虎（清华大学） | 陈永彬（东南大学） |
| 陈道文（中科院自动化所） | 李昌立（中科院声学所） |
| 余崇智（南京大学） | 迟惠生（北京大学） |
| 严普强（清华大学） | 林茂灿（社科院语言所） |
| 张家骥（中科院声学所） | 柯有安（北京理工大学） |
| 俞铁城（中科院声学所） | 赵国田（哈尔滨工业大学） |
| 高文（哈尔滨工业大学） | 袁保宗（北方交通大学） |
| 徐秉铮（华南理工大学） | 徐近需（哈尔滨工业大学） |
| 柴佩琪（同济大学） | 黄泰翼（中科院自动化所） |

前 言

本书是第三届全国人机语音通讯学术会议的论文集。

全国人机语音通讯学术会议是由中国自动化学会模式识别与机器智能专业委员会、中国计算机学会人工智能与模式识别专业委员会、中国电子学会信号处理学会语音图象通信专业委员会、中国声学学会语言听觉和音乐声学分科学会、中国中文信息学会基础理论专业委员会、中国通信学会通信理论专业委员会等六个学术团体和国家高技术智能计算机系统专家组联合主办的全国性学术会议,是全国人机语音通讯界的重要学术会议,得到了全国人机语音通讯界同行的积极支持和参与。会议每两年举行一次。

根据第二届会议的决定,第三届全国人机语音通讯学术会议定于1994年10月24日—30日在四川举行,由四川大学筹办。本届会议的主要内容是交流全国在语音合成、识别和理解方面的学术研究和开发成果,研讨人机语音通讯技术的发展和实用化途径。由于人机语音通讯技术的发展受到国内外科技界的广泛重视和关注,会议的主办单位决定,将入选本届会议的98篇论文汇集出版,并取名为《语音识别与合成》。

本书收入的论文可分类为综述(3篇)、语音分析与特征提取(10篇)、基音提取与声调识别(11篇)、语音识别模型与方法(17篇)、非特定人语音识别(5篇)、连续语音识别与语言模型(7篇)、噪声环境下的语音识别(5篇)、说话人识别(3篇)、神经网络应用(4篇)、语音合成(11篇)、语音编码(3篇)、语音库与系统评价方法(4篇)、语音技术应用与硬件实现(15篇)等十三个专题。这些论文反映了我国近年来在人机语音通讯技术方面的研究开发工作和最新成果,对我国人机语音通讯技术的发展有重要的参考价值。我们相信,本书的出版,必将进一步推动我国人机语音通讯的研究和向实用化发展的进程。

杨家沅

一九九四年九月十日



北京声迅电子有限公司

北京声迅电子有限公司是一九九四年一月成立的股份公司。主要股东包括北京市科学技术研究院、北京市桑普技术公司、广东星河电子总公司等。

公司集上述股东各方面的优势，是专门从事语音技术研究及其相关系统、产品生产的专业化企业。在承担“国家 863 项目”的同时，公司致力于把自主研究的最新语音技术应用于通讯、计算机、多媒体等领域，使语言更好的成为人类驾驭信息时代的工具。

公司依靠新技术，立足实业，注重市场，现已推出：语音控制命令系统、DOS 和 WINDOWS 环境下语音控制命令开发系统、“数字语音电话 TS-946”及“数字语音拨号电话 TS-948”等产品。其中数字语音电话及数字语音拨号电话是新一代办公及家庭生活的全数字化电话产品，具有广阔的市场前景，目前已投入批量生产。

公司有自己独立的生产基地和完善的开发、生产配套设施，近几年内公司将把自主研究的具有较高水平的语音技术附加其中，使产品在国内形成规模生产能力，并打入国际市场。

公司的技术及管理人员由国内著名大学毕业的青年人组成。他们将新的技术思想、科学的管理手段和最真诚敬业精神奉献给公司的建设，在这种环境中孕育着公司美好的未来。

北京声迅电子有限公司

地址：北京市海淀区花园路 3 号

邮编：100083

电话：2014753

传真：2040741

目 录

综 述

- 汉语语音识别的当前任务与研究方向 方棟棠 (3)
- 汉语语言识别的声学模型和语言模型 张家骥 (13)
- 面向语音处理的神经网络发展综述 陈道文 黄泰翼 (25)

语音分析与特征提取

- 北京话声调分布区的感知实验研究 林茂烂 (35)
- 普通话双音节 V1—/z、zh、j/过渡的实验研究 孙国华 (39)
- 基于谱逼近准则的语音信号 ARMA 模型的参数估计
..... 毛德行 杨立志 赵 军 (44)
- 考虑误差积累效应的 LPC 参数的精确估计 杨立志 毛德行 文伏英 (48)
- 一种线谱对参数快速估计方法 蔡智威 关存太 陈永彬 (52)
- 连续小波变换在语音分形特征分析及语音分割中应用 秦 兵 沙宗先 (56)
- 汉语句子重音的韵律参数模型的研究 周俏峰 蔡莲红 陈 咏 (63)
- 离散小波变换 (DWT) 在语音处理中的应用 何其超 龙建中 周激流 (67)
- 一种开放型通用语音信号处理平台—EOGI 高 文 闫宝杰 张 晶 (73)
- 语音信号处理工作平台' 杨立志 苏文忠 (81)

基音提取与声调识别

- 基音长度的统计结果及分析 文伏英 毛德行 杨立志 (87)
- 一种基音提取方法 苏文忠 毛德行 (92)
- 多候选最优路径法基音提取 陈 翔 余崇智 (96)
- 一种基频提取的实用算法及其测试 陆 嚶 陈道文 杨子冬 (100)

一种新型汉语语音浊音基频提取器	张 凯 严普强 (104)
极值法 C/V 分割与基音序列的提取	李 柏 王作英 (109)
汉语声调判别的一种方法	史光敏 吴文虎 方棟棠 (112)
一种汉语四声检测算法的研究	刘大奔 徐秉铮 孙 蓐 (121)
基于小波变换的 HMM 声调判别方法	苏 俭 徐秉铮 李海洲 (125)
基于小波变换和规则的汉语声调识别	王 刚 黄泰翼 (129)
基于隐马尔可夫模型的汉语二字词音调模式识别 秦 勇 莫福源 李昌立 关定华 (135)

语音识别模型与方法

结合汉语语音学特点的大词汇词组语音识别	江 辉 王仁华 (141)
基于 HMM 的大词汇量孤立词汉语语音识别研究 徐 波 马斌 徐东昕 黄泰翼 陈道文 (146)
一具友善人机介面之高可靠度巨量词汇中文语音辩认系统 许志兴 王骏发 吴宗宪 (153)
无限词汇特定人汉语语音识别系统 苑宝生 高两青 洪小文 Gareth H. Loudon (157)
基于分段 K 平均训练的汉语全音节语音识别的研究 国立新 莫福源 李昌立 (161)
基于连续隐马尔可夫模型的汉语全音节识别系统 马 斌 徐 波 黄泰翼 (165)
大型线性预测 HMM 及其在汉语音素识别中的应用	谢锦辉 李 晖 (169)
基于稳态链语流模型的汉语语音识别方法	魏华武 吴文虎 (175)
基于汉语语音知识的识别技术和其特征提取	严普强 张 晶 张 凯 (179)
模糊统计在隐马尔可夫模型语音识别中的应用	李胜利 侯朝煊 (184)
使用最小差错准则训练隐含马尔可夫模型 (HMM) 刘 加 Steve Young (188)
一种基于模糊控制思想的声母分类识别方法 叶 军 刘大奔 李海洲 徐秉铮 (193)
基于语音波形过零点间时间间隔韵母识别 欧贵文 (198)
Windows 3.1 环境下基于 DTW 的孤立词识别快速搜索算法 周 浩 陈道文 (202)

一种基于 CDHMM 的语音识别系统的快速自适应算法	王文疆 黄泰翼 (208)
用统计的方法进行汉语语音理解和音字转换	
.....	吴 军 王作英 郭 进 孙甲松 (214)
语音识别中的音字转换.....	季 宏 刘志文 杜燕玲 黄曾阳 (221)

非特定人语音识别

不定人的语音识别系统.....	胡春静 吴善培 (233)
连续距离密度的分段概率模型 CDD—SPM	
.....	郑 方 杨红波 吴文虎 方棣棠 (238)
基于正态分布假设的非特定人语音识别 ...	余 华 方棣棠 李树青 卿 山 (242)
军事口令非特定人识别系统的实现	何为民 赵 锐 陈光发 (245)
非特定人电话数字语音识别及其应用	王跟东 吴德言 (250)

连续语音识别与语言模型

基于音节分割的连续语音多模板隐马尔可夫模型的研究	
.....	程兰颖 俞铁城 李忠香 (257)
语音辨认中阶层式以文法为主之马可夫语言模型	吴宗宪 王骏发 (265)
连续密度 HMM 用于非特定人的汉语连续语句识别的尝试	
.....	吴伟民 林道发 杨家沅 (269)
有限命令集连续语音识别	计天颖 王作英 陆大金 (273)
连续语音识别的研究和汉语数字连呼系统的实现	
.....	柴海新 吴文虎 方棣棠 (277)
基于本征值的汉语音节切分方法	柴佩琪 (288)
语音信号的自动分段方法研究	陈 韬 莫福源 李昌立 (292)

噪声环境下的语音识别

在中等噪声背景下的全音节汉语语音识别	
.....	朱思俞 李建人 余金森 王海燕 (299)
噪声背景下大词汇汉语语音识别	王海燕 朱思俞 李建人 余金森 (306)
噪声环境下语音识别的几个问题 (一)	葛余博 杨大利 曾德超 (312)
带噪语音信号的信号模型参数和噪声模型参数的辨别	
.....	沈亚强 程仲文 任德官 (316)

高噪环境下顽健语音特征抽取 徐近霁 刘明宝 杨子云 (322)

说话人识别

与文本无关的讲话人辨认研究 杨伟东 陈永彬 (331)

一种使用非限定词句的话者识别的系统 闫宝杰 高文 (335)

一种高性能的限定文本说话人确认系统 吴玺宏 迟惠生 (342)

神经网络应用

一种新的神经网络与隐马尔可夫模型混合的快速方法
..... 俞 栋 黄泰翼 陈道文 (349)

语音识别的神经网络自适应方法: 多层感知器与 LVQ 的比较
..... 徐东昕 马 茜 徐 波 陈道文 黄泰翼 (355)

音素分类用函数链网的训练 徐近霁 孟宪军 (360)

有限状态径向基函数网络结构在语音识别中的应用
..... 李苇营 汤笑梅 易克初 胡 征 (366)

语音合成

基于倒谱分析的汉语按规则合成系统 杨 鉴 (373)

调频信号的语音分析与合成 谭廷忠 余崇智 方 元 (377)

基于基音元的汉字单音节 LPC 合成方法 赵 军 毛德行 (380)

基于脉宽调制的语音合成技术研究 于鸿洋 杨红雨 朱维乐 (385)

音频合成与语言合成 蔡连红 周俏峰 (390)

利用波形基音同步叠加拼接技术合成汉语的研究
..... 初 敏 吕士楠 陆亚民 (394)

一种波形拼接的语音合成实验 敖其尔 巩 政 (398)

汉语语音合成二字词规则初步应用 唐涤飞 倪晋富 王仁华 (403)

汉语合成系统中音高音长规则研究 吕士楠 周同春 初 敏 陆亚民 (407)

基于词汇的汉语文语转换系统 倪 宏 李昌立 莫福源 (411)

基于单音节的汉语文语转换系统及其应用 李 彤 莫福源 李昌立 (415)

语音编码

一种 4Kbps CELP 编码器的实现 蔡智威 关存太 陈永彬 (421)

基音同步的语音波形矢量量化	高 晖	和应民	(425)
树判决在置乱语音信号时段拟合研究中的应用	张 军	张知易 王 璞	(429)

语音库与系统评价方法

按音节过渡组织词组音库语料	倪晋富	王仁华	(437)
全自动语音识别综合评测系统	王仁华 倪晋富	徐 超 刘必成	(442)
知音文书机语音数据库的建立和一种汉语语音特性的探讨			
.....	孙甲松	王作英 肖 熙	(447)
语音输入系统性能的评价标准及影响性能的主要因素	钟文青	徐秉铮	(452)

语音技术应用与硬件实现

一种高性能/价格比的特定人语音识别器	曾迎凡	(459)
一个实时汉语语音识别系统的硬件设计	刘大奔 叶 军 徐秉铮	(462)
基于快速模板匹配的 Windows 命令识别系统	张向东 申 凌 俞铁城	(466)
通用树状结构的实时语音查询系统	张向东 俞铁城 李忠香	(472)
有限命令集特定人汉语语音实时识别系统	曹 超 李昌立 莫福源	(478)
新一代语音识别、理解与合成系统——“知音文书机”		
.....	王作英 吴 军 肖 熙 王 侠	(482)
语音识别与语音合成在丝绸质量检验中的应用	吕林望 苟大举 林道发	(487)
一台数字程控调度机中的语音处理技术	王国胤 施鸿宝	(493)
笔记本式汉语语音合成器的研究	王炳锡 韩 华	(499)
在微机上运行的 T6668 语音合成系统	卢 化 岳东剑 柴佩琪	(502)
TSP53C30 语音合成系统	苟大举 杨启纲 贺德珏	(506)
一个计算机语音系统的硬件设计	盛珏新 郑纪蛟	(512)
汉语人机对话系统中基于浮点 DSP 芯片 TMS320C40 的高速并行处理系统		
.....	王宝利 李昌立 关定华	(518)
通用语音信号处理系统—BYDSP—C30	张延平 尹建琪 陈锡先	(522)
列车运行的语音记录装置	吴仲光 杨家沅	(527)

综 述



汉语语音识别的当前任务与研究方向

方棣棠

(清华大学计算机科学与技术系 北京,100084)

摘 要

本文在汉语语音识别的历史回顾基础上,阐述了语音识别的三个基本问题—特征抽取、模式划分和时间对准,以及语音识别中的语言模型。在分析汉语语音识别的研究现状基础上提出汉语语音识别的研究方向和当前研究任务的建议。

关键词 语音识别,语言模型,特征抽取,模式划分,时间对准。

Current Tasks and Research Orientation in Chinese Speech Recognition

Fang Ditang

Department of Computer Science & Technology, Tsinghua University

Abstract

In this paper, after reviewing the history of Chinese speech recognition research, we expound three main problems of speech recognition—feature extraction, pattern dividing and time alignment and the language model in speech recognition. On the basis of analyzing the present situation of Chinese speech recognition research, we propose the research orientation and current tasks in some suggestion as to the Chinese speech recognition.

Key words Speech Recognition, Language Model, Feature extraction, Pattern dividing, Time alignment.

一、引言

国内语音识别研究从1958年开始,由中国科学院声学所(当时属电子学研究所)用电子管电路识别十个元音。到1973年才由中国科学院声学研究所开始计算机语音识别。从七十年代末到八十年代,国内很多单位陆续参加到语音识别的研究行列,估计现在有近百个单位从事这方面的工作。

十多年来我国语音识别研究工作基本跟上国外的研究步伐,有些地方结合汉语的特点