

<<语音识别基本原理(英文)>>

图书基本信息

书名：<<语音识别基本原理(英文)>>

13位ISBN编号：9787302036401

10位ISBN编号：7302036403

出版时间：1999-08

出版时间：清华大学出版社

作者：罗宾纳

译者：阮平望

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<语音识别基本原理(英文)>>

### 内容概要

#### 内容简介

本书面向工程技术人员、科技工作者、语言学家、编程人员，主要讲解有关现代语音识别系统的基本知识、思路和方法。

本书共9章

分别为：1语音识别原理；2语音信号的产生、感知及声学语音学特征；3.用于语音识别的信号处理和分析方法；4模式对照技术；5语音识别系统的设计与实现结果；6隐马尔可夫模型的理论与实践；7.基于连接词模型的语音识别；8大词汇量连续语音识别；9适合不同任务的自动语音识别应用。

本书既可供研究工作者借鉴，也可供研究生在学习有关语音信号数字处理课程时参考。

<<语音识别基本原理(英文)>>

书籍目录

CONTENTS  
LIST OF FIGURES  
LIST OF TABLES  
PREFACE  
1 FUNDAMENTALS OF SPEECH RECOGNITION  
1.1 Introduction  
1.2 The Paradigm for Speech Recognition  
1.3 Outline  
1.4 A Brief History of Speech-Recognition Research  
2 THE SPEECH SIGNAL: PRODUCTION, PERCEPTION, AND ACOUSTIC-PHONETIC CHARACTERIZATION  
2.1 Introduction  
2.1.1 The Process of Speech Production and Perception in Human Beings  
2.2 The Speech-Production Process  
2.3 Representing Speech in the Time and Frequency Domains  
2.4 Speech Sounds and Features  
2.4.1 The Vowels  
2.4.2 Diphthongs  
2.4.3 Semivowels  
2.4.4 Nasal Consonants  
2.4.5 Unvoiced Fricatives  
2.4.6 Voiced Fricatives  
2.4.7 Voiced and Unvoiced Stops  
2.4.8 Review Exercises  
2.5 Approaches to Automatic Speech Recognition by Machine  
2.5.1 Acoustic-Phonetic Approach to Speech Recognition  
2.5.2 Statistical Pattern-Recognition Approach to Speech Recognition  
2.5.3 Artificial Intelligence (AI) Approaches to Speech Recognition  
2.5.4 Neural Networks and Their Application to Speech Recognition  
2.6 Summary  
3 SIGNAL PROCESSING AND ANALYSIS METHODS FOR SPEECH RECOGNITION  
3.1 Introduction  
3.1.1 Spectral Analysis Models  
3.2 The Bank-of-Filters Front-End Processor  
3.2.1 Types of Filter Bank Used for Speech Recognition  
3.2.2 Implementations of Filter Banks  
3.2.3 Summary of Considerations for Speech-Recognition Filter Banks  
3.2.4 Practical Examples of Speech-Recognition Filter Banks  
3.2.5 Generalizations of Filter-Bank Analyzer  
3.3 Linear Predictive Coding Model for Speech Recognition  
3.3.1 The LPC Model  
3.3.2 LPC Analysis Equations  
3.3.3 The Autocorrelation Method

<<语音识别基本原理(英文)>>

- 3.3.4 The Covariance Method
- 3.3.5 Review Exercise
- 3.3.6 Examples of LPC Analysis
- 3.3.7 LPC Processor for Speech Recognition
- 3.3.8 Review Exercises
- 3.3.9 Typical LPC Analysis Parameters
- 3.4 Vector Quantization
  - 3.4.1 Elements of a Vector Quantization Implementation
  - 3.4.2 The VQ Training Set
  - 3.4.3 The Similarity or Distance Measure
  - 3.4.4 Clustering the Training Vectors
  - 3.4.5 Vector Classification Procedure
  - 3.4.6 Comparison of Vector and Scalar Quantizers
  - 3.4.7 Extensions of Vector Quantization
  - 3.4.8 Summary of the VQ Method
- 3.5 Auditory-Based Spectral Analysis Models
  - 3.5.1 The EIH Model
- 3.6 Summary
- 4 PATTERN-COMPARISON TECHNIQUES
  - 4.1 Introduction
  - 4.2 Speech (Endpoint) Detection
  - 4.3 Distortion Measures--Mathematical Considerations
  - 4.4 Distortion Measures--Perceptual Considerations
  - 4.5 Spectral-Distortion Measures
    - 4.5.1 Log Spectral Distance
    - 4.5.2 Cepstral Distances
    - 4.5.3 Weighted Cepstral Distances and Liftering
    - 4.5.4 Likelihood Distortions
    - 4.5.5 Variations of Likelihood Distortions
    - 4.5.6 Spectral Distortion Using a Warped Frequency Scale
    - 4.5.7 Alternative Spectral Representations and Distortion Measures
    - 4.5.8 Summary of Distortion Measures--Computational Considerations
  - 4.6 Incorporation of Spectral Dynamic Features into the Distortion Measure
  - 4.7 Time Alignment and Normalization
    - 4.7.1 Dynamic Programming--Basic Considerations
    - 4.7.2 Time-Normalization Constraints
    - 4.7.3 Dynamic Time-Warping Solution
    - 4.7.4 Other Considerations in Dynamic Time Warping
    - 4.7.5 Multiple Time-Alignment Paths
  - 4.8 Summary
- 5 SPEECH RECOGNITION SYSTEM DESIGN AND IMPLEMENTATION ISSUES
  - 5.1 Introduction
  - 5.2 Application of Source-Coding Techniques to Recognition
    - 5.2.1 Vector Quantization and Pattern Comparison Without Time Alignment
    - 5.2.2 Centroid Computation for VQ Codebook Design
    - 5.2.3 Vector Quantizers with Memory

<<语音识别基本原理(英文)>>

- 5.2.4 Segmental Vector Quantization
- 5.2.5 Use of a Vector Quantizer as a Recognition Preprocessor
- 5.2.6 Vector Quantization for Efficient Pattern Matching
- 5.3 Template Training Methods
  - 5.3.1 Casual Training
  - 5.3.2 Robust Training
  - 5.3.3 Clustering
- 5.4 Performance Analysis and Recognition Enhancements
  - 5.4.1 Choice of Distortion Measures
  - 5.4.2 Choice of Clustering Methods and kNN Decision Rule
  - 5.4.3 Incorporation of Energy Information
  - 5.4.4 Effects of Signal Analysis Parameters
  - 5.4.5 Performance of Isolated Word-Recognition Systems
- 5.5 Template Adaptation to New Talkers
  - 5.5.1 Spectral Transformation
  - 5.5.2 Hierarchical Spectral Clustering
- 5.6 Discriminative Methods in Speech Recognition
  - 5.6.1 Determination of Word Equivalence Classes
  - 5.6.2 Discriminative Weighting Functions
  - 5.6.3 Discriminative Training for Minimum Recognition Error
- 5.7 Speech Recognition in Adverse Environments
  - 5.7.1 Adverse Conditions in Speech Recognition
  - 5.7.2 Dealing with Adverse Conditions
- 5.8 Summary
- 6 THEORY AND IMPLEMENTATION OF HIDDEN MARKOV MODELS
  - 6.1 Introduction
  - 6.2 Discrete-Time Markov Processes
  - 6.3 Extensions to Hidden Markov Models
    - 6.3.1 Coin-Toss Models
    - 6.3.2 The Um-and-Ball Model
    - 6.3.3 Elements of an HMM
    - 6.3.4 HMM Generator of Observations
  - 6.4 The Three Basic Problems for HMMs
    - 6.4.1 Solution to Problem 1-Probability Evaluation
    - 6.4.2 Solution to Problem 2--"Optimal" State Sequence
    - 6.4.3 Solution to Problem 3--Parameter Estimation
    - 6.4.4 Notes on the Reestimation Procedure
  - 6.5 Types of HMMs
  - 6.6 Continuous Observation Densities in HMMs
  - 6.7 Autoregressive HMMs
  - 6.8 Variants on HMM Structures-Null Transitions and Tied States
  - 6.9 Inclusion of Explicit State Duration Density in HMMs
  - 6.10 Optimization Criterion-ML, MMI, and MDI
  - 6.11 Comparisons of HMMs
  - 6.12 Implementation Issues for HMMs
    - 6.12.1 Scaling
    - 6.12.2 Multiple Observation Sequences

<<语音识别基本原理(英文)>>

- 6.12.3 Initial Estimates of HMM Parameters
- 6.12.4 Effects of Insufficient Training Data
- 6.12.5 Choice of Model
- 6.13 Improving the Effectiveness of Model Estimates
  - 6.13.1 Deleted Interpolation
  - 6.13.2 Bayesian Adaptation
  - 6.13.3 Corrective Training
- 6.14 Model Clustering and Splitting
- 6.15 HMM System for Isolated Word Recognition
  - 6.15.1 Choice of Model Parameters
  - 6.15.2 Segmental K-Means Segmentation into States
  - 6.15.3 Incorporation of State Duration into the HMM
  - 6.15.4 HMM Isolated-Digit Performance
- 6.16 Summary
- 7 SPEECH RECOGNITION BASED ON CONNECTED WORD MODELS
  - 7.1 Introduction
  - 7.2 General Notation for the Connected Word-Recognition Problem
  - 7.3 The Two-Level Dynamic Programming (Two-Level DP) Algorithm
    - 7.3.1 Computation of the Two-Level DP Algorithm
  - 7.4 The Level Building (LB) Algorithm
    - 7.4.1 Mathematics of the Level Building Algorithm
    - 7.4.2 Multiple Level Considerations
    - 7.4.3 Computation of the Level Building Algorithm
    - 7.4.4 Implementation Aspects of Level Building
    - 7.4.5 Integration of a Grammar Network
    - 7.4.6 Examples of LB Computation of Digit Strings
  - 7.5 The One-Pass (One-State) Algorithm
  - 7.6 Multiple Candidate Strings
  - 7.7 Summary of Connected Word Recognition Algorithms
  - 7.8 Grammar Networks for Connected Digit Recognition
  - 7.9 Segmental K-Means Training Procedure
  - 7.10 Connected Digit Recognition Implementation
    - 7.10.1 HMM-Based System for Connected Digit Recognition
    - 7.10.2 Performance Evaluation on Connected Digit Strings
  - 7.11 Summary
- 8 LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION
  - 8.1 Introduction
  - 8.2 Subword Speech Units
  - 8.3 Subword Unit Models Based on HMMs
  - 8.4 Training of Subword Units
  - 8.5 Language Models for Large Vocabulary Speech Recognition
  - 8.6 Statistical Language Modeling
  - 8.7 Perplexity of the Language Model
  - 8.8 Overall Recognition System Based on Subword Units

<<语音识别基本原理(英文)>>

- 8.8.1 Control of Word Insertion/Word Deletion Rate
- 8.8.2 Task Semantics
- 8.8.3 System Performance on the Resource Management Task
- 8.9 Context-Dependent Subword Units
  - 8.9.1 Creation of Context-Dependent Diphones and Triphones
  - 8.9.2 Using Interword Training to Create CD Units
  - 8.9.3 Smoothing and Interpolation of CD PLU Models
  - 8.9.4 Smoothing and Interpolation of Continuous Densities
  - 8.9.5 Implementation Issues Using CD Units
  - 8.9.6 Recognition Results Using CD Units
  - 8.9.7 Position Dependent Units
  - 8.9.8 Unit Splitting and Clustering
  - 8.9.9 Other Factors for Creating Additional Subword Units
  - 8.9.10 Acoustic Segment Units
- 8.10 Creation of Vocabulary-Independent Units
- 8.11 Semantic Postprocessor for Recognition
- 8.12 Summary
- 9 TASK ORIENTED APPLICATIONS OF AUTOMATIC SPEECH RECOGNITION
  - 9.1 Introduction
  - 9.2 Speech-Recognizer Performance Scores
  - 9.3 Characteristics of Speech-Recognition Applications
    - 9.3.1 Methods of Handling Recognition Errors
  - 9.4 Broad Classes of Speech-Recognition Applications
  - 9.5 Command-and-Control Applications
    - 9.5.1 Voice Repertory Dialer
    - 9.5.2 Automated Call-Type Recognition
    - 9.5.3 Call Distribution by Voice Commands
    - 9.5.4 Directory Listing Retrieval
    - 9.5.5 Credit Card Sales Validation
  - 9.6 Projections for Speech Recognition

<<语音识别基本原理(英文)>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>