The Algorithms Of Speech Recognition Programming And

Nonlinear Speech Modeling and Applications
Large Margin Algorithms for Discriminative Continuous Speech Recognition
A Comparative Study of Several Dynamic Time Warping Algorithms for Speech Recognition
Automatic Speech and Speaker Recognition
Experimental Evaluation of Algorithms for Connected Speech Recognition Using Hidden Markov Models
Efficient Search Algorithms for Continuous Speech Recognition
A Deep Learning Approach
Deep Learning for NLP and Speech Recognition
A Hybrid Approach
Algorithms and Applications
Spoken Language Processing
Speech Recognition and Processing
Large Margin and Kernel Methods
Algorithms for Low Vocabulary Speech Recognition for Handheld/portable Devices
Speech Recognition in Noise Using Weighted Matching Algorithms
Nonlinear Analyses and Algorithms for Speech Processing
Improving Training of Neural Networks for Speech Recognition
A Bridge to Practical Applications
Algorithms and Architectures
A Guide to Theory, Algorithm, and System Development
Comparison of Algorithms for Speech Recognition
Speech Recognition Algorithms based on Weighted Finite-State Transducers
Speech & Language Processing
An Efficient A* Stack Decoder Algorithm for Continuous Speech Recognition with a Stochastic Language Model
Efficient Algorithms for Speech Recognition
Speech Enhancement, Modeling and Recognition- Algorithms and Applications
Speech Recognition
Algorithms and Applied Principles
Automatic Speech Recognition
Recent Advances in Robust Speech Recognition Technology
Dynamic Speech Models
10th International Conference, ICCCI 2018, Bristol, UK, September 5-7, 2018, Proceedings, Part I
Connectionist Speech Recognition
Speech Processing in Embedded Systems
Performance Analysis of Object-oriented, Speech Recognition Algorithms in a Real-time, Multiprocessor Environment
Speech Recognition
Advanced Lectures and Revised Selected Papers
International Conference on Non-Linear Speech Processing, NOLISP 2005, Barcelona, Spain, April 19-22, 2005, Revised Selected Papers

Information-processing Algorithms for Speaker-independent Speech Recognition
Robust Automatic Speech Recognition Algorithms for Dealing with Noise and Accent

DOUGLAS DALTON

Nonlinear Speech Modeling and Applications BoD – Books on Demand
Robust Automatic Speech Recognition: A Bridge to Practical Applications establishes a solid foundation for automatic speech recognition that is robust against acoustic environmental distortion. It provides a thorough overview of classical and modern noise-and reverberation robust techniques that have been developed over the past thirty years, with an emphasis on practical methods that have been proven to be successful and which are likely to be further developed for future applications. The strengths and weaknesses of robustness-enhancing speech recognition techniques are carefully analyzed. The book covers noise-robust techniques designed for acoustic models which are based on both Gaussian mixture models and deep neural networks. In addition, a guide to selecting the best methods for practical applications is provided. The reader will: Gain a unified, deep and systematic understanding of the state-of-the-art technologies for robust speech recognition Learn the links and relationship between alternative technologies for robust speech recognition Be able to use the technology analysis and categorization detailed in the book to guide future technology development Be able to develop new noise-robust methods in the current era of deep learning for acoustic modeling in speech recognition The first book that provides a comprehensive review on noise and reverberation robust speech recognition methods in the era of deep neural networks Connects robust speech recognition techniques to machine learning paradigms with rigorous mathematical treatment Provides elegant and structural ways to categorize and analyze noise-robust speech recognition techniques Written by leading researchers who have been actively working on the subject matter in both industrial and academic organizations for many years
Large Margin Algorithms for Discriminative Continuous Speech Recognition Prentice Hall
Provides a theoretically sound, technically accurate, and complete description of the basic knowledge and ideas that constitute a modern system for speech recognition by machine. Covers production, perception, and acoustic-phonetic characterization of the speech signal; signal processing and analysis methods for speech recognition; pattern comparison techniques; speech recognition system design and implementation; theory and implementation of hidden Markov models; speech recognition based on connected word models; large vocabulary continuous speech recognition; and task- oriented application of automatic speech recognition. For practicing engineers, scientists, linguists, and programmers interested in speech recognition.
A Comparative Study of Several Dynamic Time Warping Algorithms for Speech Recognition John Wiley & Sons
Refereed postproceedings of the International Conference on Non-Linear Speech Processing, NOLISP 2005. The 30 revised full papers presented together with one keynote speech and 2 invited talks were carefully reviewed and selected from numerous submissions for inclusion in the book. The papers are organized in topical sections on speaker recognition, speech analysis, voice pathologies, speech recognition, speech enhancement, and applications.
Automatic Speech and Speaker Recognition Springer
This book provides a comprehensive overview of the recent advancement in the field of automatic speech recognition with a focus on deep learning models including deep neural networks and many of their variants. This is the first automatic speech recognition book dedicated to the deep learning approach. In addition to the rigorous mathematical treatment of the subject, the book also presents insights and theoretical foundation of a series of highly successful deep learning models.
Experimental Evaluation of Algorithms for Connected Speech Recognition Using Hidden Markov Models BoD – Books on Demand
This book presents the revised tutorial lectures given at the International Summer School on Nonlinear Speech Processing-Algorithms and Analysis held in Vietri sul Mare, Salerno, Italy in September 2004. The 14 revised tutorial lectures by leading international researchers are organized in topical sections on dealing with nonlinearities in speech signals, acoustic-to-articulatory modeling of speech phenomena, data driven and speech processing algorithms, and algorithms and models based on speech perception mechanisms. Besides the tutorial lectures, 15 revised reviewed papers are included presenting original research results on task oriented speech applications.
Efficient Search Algorithms for Continuous Speech Recognition Pearson Education India
Speech Processing has rapidly emerged as one of the most widespread and well-understood application areas in the broader discipline of Digital Signal Processing. Besides the telecommunications applications that have hitherto been the largest users of speech processing algorithms, several non-traditional embedded processor applications are enhancing their functionality and user interfaces by utilizing various aspects of speech processing. "Speech Processing in Embedded Systems" describes several areas of speech processing, and the various algorithms and industry standards that address each of these areas. The topics covered include different types of Speech Compression, Echo Cancellation, Noise Suppression, Speech Recognition and Speech Synthesis. In addition this book explores various issues and considerations related to efficient implementation of these algorithms on real-time embedded systems, including the role played by processor CPU and peripheral functionality.
A Deep Learning Approach Wiley-Blackwell
Current Automatic Speech Recognition devices attempt to solve the connected word recognition problem by assuming that an unknown phrase is the output of a sequence of statistical word-models. Typically, these models are constructed using examples of words spoken in isolation; however, the acoustic patterns corresponding to words as they occur in fluent speech are quite different from those representing the same words spoken in isolation, and so the use in speech recognizers of models based on isolated utterances severely limits the performance of such devices. A method of extracting training utterances from fluent speech and constructing Hidden Markov Models (HMMs) from these templates, known as Embedded Training, is investigated here, in

conjunction with a two-level algorithm for connected word recognition. The effects on recognition performance of various HMM training procedures are discussed, and experimental results are presented.

Deep Learning for NLP and Speech Recognition Morgan & Claypool Publishers

Speech dynamics refer to the temporal characteristics in all stages of the human speech communication process. This speech “chain” starts with the formation of a linguistic message in a speaker's brain and ends with the arrival of the message in a listener's brain. Given the intricacy of the dynamic speech process and its fundamental importance in human communication, this monograph is intended to provide a comprehensive material on mathematical models of speech dynamics and to address the following issues: How do we make sense of the complex speech process in terms of its functional role of speech communication? How do we quantify the special role of speech timing? How do the dynamics relate to the variability of speech that has often been said to seriously hamper automatic speech recognition? How do we put the dynamic process of speech into a quantitative form to enable detailed analyses? And finally, how can we incorporate the knowledge of speech dynamics into computerized speech analysis and recognition algorithms? The answers to all these questions require building and applying computational models for the dynamic speech process. What are the compelling reasons for carrying out dynamic speech modeling? We provide the answer in two related aspects. First, scientific inquiry into the human speech code has been relentlessly pursued for several decades. As an essential carrier of human intelligence and knowledge, speech is the most natural form of human communication. Embedded in the speech code are linguistic (as well as para-linguistic) messages, which are conveyed through four levels of the speech chain. Underlying the robust encoding and transmission of the linguistic messages are the speech dynamics at all the four levels. Mathematical modeling of speech dynamics provides an effective tool in the scientific methods of studying the speech chain. Such scientific studies help understand why humans speak as they do and how humans exploit redundancy and variability by way of multitiered dynamic processes to enhance the efficiency and effectiveness of human speech communication. Second, advancement of human language technology, especially that in automatic recognition of natural-style human speech is also expected to benefit from comprehensive computational modeling of speech dynamics. The limitations of current speech recognition technology are serious and are well known. A commonly acknowledged and frequently discussed weakness of the statistical model underlying current speech recognition technology is the lack of adequate dynamic modeling schemes to provide correlation structure across the temporal speech observation sequence. Unfortunately, due to a variety of reasons, the majority of current research activities in this area favor only incremental modifications and improvements to the existing HMM-based state-of-the-art. For example, while the dynamic and correlation modeling is known to be an important topic, most of the systems nevertheless employ only an ultra-weak form of speech dynamics; e.g., differential or delta parameters. Strong-form dynamic speech modeling, which is the focus of this monograph, may serve as an ultimate solution to this problem. After the introduction chapter, the main body of this monograph consists of four chapters. They cover various aspects of theory, algorithms, and applications of dynamic speech models, and provide a comprehensive survey of the research work in this area spanning over past 20~years. This monograph is intended as advanced materials of speech and signal processing for graudate-level teaching, for professionals and engineering practioners, as well as for seasoned researchers and engineers specialized in speech processing

A Hybrid Approach Springer Science & Business Media

Automatic Speech Recognition (ASR) is the enabling technology for hands-free dictation and voice-triggered computer menus. It is becoming increasingly prevalent in environments such as private telephone exchanges and real-time information services. Speech Recognition introduces the principles of ASR systems, including the theory and implementation issues behind multi-speaker continuous speech recognition. Focusing on the algorithms employed in commercial and laboratory systems, the treatment enables the reader to devise practical solutions for ASR system problems. It addresses in detail C++ programming techniques used to develop ASR applications, thus offering skills that will prove useful in any large C++ based software project. Possible extensions of the well-established ASR technology are highlighted, based on "Hidden Markov Models" applied to fields such as modelling and prediction of econometric series. Features include: * Accompanying website containing all C++ source code of a complete laboratory multi-speaker continuous-speech ASR system (e.g. Initialisation, Training, Recognition, Evaluation, etc.) www.wiley.com/go/becchetti_speech * Detailed theoretical, mathematical and technical explanations of ASR * A practical account of the functioning of ASR A crucial source of information for researchers, developers and project managers involved with ASR systems, Speech Recognition is also structured for use by students of digital signal processing, speech recognition and C++ programming techniques.

Algorithms and Applications Springer

This book reflects decades of important research on the mathematical foundations of speech recognition. It focuses on underlying statistical techniques such as hidden Markov models, decision trees, the expectation-maximization algorithm, information theoretic goodness criteria, maximum entropy probability estimation, parameter and data clustering, and smoothing of probability distributions. The author's goal is to present these principles clearly in the simplest setting, to show the advantages of self-organization from real data, and to enable the reader to apply the techniques.

Spoken Language Processing Bentham Science

This book introduces the theory, algorithms, and implementation techniques for efficient decoding in speech recognition mainly focusing on the Weighted Finite-State Transducer (WFST) approach. The decoding process for speech recognition is viewed as a search problem whose goal is to find a sequence of words that best matches an input speech signal. Since this process becomes computationally more expensive as the system vocabulary size increases, research has long been devoted to reducing the computational cost. Recently, the WFST approach has become an important state-of-the-art speech recognition technology, because it offers improved decoding speed with fewer recognition errors compared with conventional methods. However, it is not easy to understand all the algorithms used in this framework, and they are still in a black box for many people. In this book, we review the WFST approach and aim to provide comprehensive interpretations of WFST operations and decoding algorithms to help anyone who wants to understand, develop, and study

WFST-based speech recognizers. We also mention recent advances in this framework and its applications to spoken language processing. Table of Contents: Introduction / Brief Overview of Speech Recognition / Introduction to Weighted Finite-State Transducers / Speech Recognition by Weighted Finite-State Transducers / Dynamic Decoders with On-the-fly WFST Operations / Summary and Perspective

Speech Recognition and Processing Academic Press

This E-book is a collection of articles that describe advances in speech recognition technology. Robustness in speech recognition refers to the need to maintain high speech recognition accuracy even when the quality of the input speech is degraded, or when the acoustical, articulate, or phonetic characteristics of speech in the training and testing environments differ. Obstacles to robust recognition include acoustical degradations produced by additive noise, the effects of linear filtering, nonlinearities in transduction or transmission, as well as impulsive interfering sources, and diminished accuracy caused by changes in articulation produced by the presence of high-intensity noise sources. Although progress over the past decade has been impressive, there are significant obstacles to overcome before speech recognition systems can reach their full potential. Automatic speech recognition (ASR) systems must be robust to all levels, so that they can handle background or channel noise, the occurrence on unfamiliar words, new accents, new users, or unanticipated inputs. They must exhibit more 'intelligence' and integrate speech with other modalities, deriving the user's intent by combining speech with facial expressions, eye movements, gestures, and other input features, and communicating back to the user through multimedia responses. Therefore, as speech recognition technology is transferred from the laboratory to the marketplace, robustness in recognition becomes increasingly significant. This E-book should be useful to computer engineers interested in recent developments in speech recognition technology.

Large Margin and Kernel Methods Springer

Este documento tiene por objetivo recoger la información relativa al proyecto sobre la creación de un algoritmo para Start-and-End point detection de una señal pregrabada. La intención inicial del desarrollo de este algoritmo es que pueda ser utilizado en la entrada de una aplicación de reconocimiento de voz. En términos generales, el resultado de este trabajo es un algoritmo que puede detectar el comienzo y el fin de una señal previamente grabada basado en un algoritmo de detección de la actividad de la voz previamente desarrollado por la Czech Technical University, Faculty of Electrical Engineering. Hay dos temas principales de estudio en este proyecto: detección de la actividad de la voz (VAD algorithm) y determinar el punto de inicio y fin de la señal (Start-and-End point detection). El primer paso para la construcción del algoritmo final es ser capaz de identificar la actividad de la voz en una señal mediante el VAD algorithm para después ser capaz de detectar el inicio y final de la actividad de la voz y descartar los silencios de la señal mediante el Startand- End point detection algorithm. Con el fin de demostrar el modo de funcionamiento de dicho algoritmo se ha creado una aplicación en MATLAB que permite ver gráficamente una señal previamente grabada y posteriormente su punto inicial y final después de aplicar los algoritmos. Por último, para proporcionar resultados más gráficos y dar al proyecto un valor añadido y con vistas a convertirse en una futura aplicación posible se ha añadido el reconocimiento de dígitos basado en de un algoritmo DTW (Dinamic Time Warping). English: The objective of the project is the creation of an algorithm for Start-and-End point detection of a pre-recorded signal. The initial reason for developing this algorithm is so it can be used at the input of a voice recognition application. Overall, the result of this work is an algorithm that can detect the beginning and end of a previously recorded signal based on a detection algorithm of the voice activity previously developed by the Czech Technical University, Faculty of Electrical Engineering. Two main issues are studied in this project: Detecting the Voice Activity (VAD algorithm) and determining the start and end point of the signal (Start-and-End point detection). To demonstrate the mode of operation of the algorithm, I have created an application in MATLAB to show graphically the process for a previously recorded signal and then the start and end points after applying the algorithms. Finally, to provide better graphic performance and provide added value to the project, I have added a digit recognition algorithm based on a DTW (Dynamic Time Warping).

Algorithms for Low Vocabulary Speech Recognition for Handheld/portable Devices Academic Press

Intelligent Speech Signal Processing investigates the utilization of speech analytics across several systems and real-world activities, including sharing data analytics, creating collaboration networks between several participants, and implementing video-conferencing in different application areas. Chapters focus on the latest applications of speech data analysis and management tools across different recording systems. The book emphasizes the multidisciplinary nature of the field, presenting different applications and challenges with extensive studies on the design, development and management of intelligent systems, neural networks and related machine learning techniques for speech signal processing. Highlights different data analytics techniques in speech signal processing, including machine learning and data mining Illustrates different applications and challenges across the design, implementation and management of intelligent systems and neural networks techniques for speech signal processing Includes coverage of biomodal speech recognition, voice activity detection, spoken language and speech disorder identification, automatic speech to speech summarization, and convolutional neural networks

Speech Recognition in Noise Using Weighted Matching Algorithms Springer Science & Business Media

This book on Speech Processing consists of seven chapters written by eminent researchers from Italy, Canada, India, Tunisia, Finland and The Netherlands. The chapters covers important fields in speech processing such as speech enhancement, noise cancellation, multi resolution spectral analysis, voice conversion, speech recognition and emotion recognition from speech. The chapters contain both survey and original research materials in addition to applications. This book will be useful to graduate students, researchers and practicing engineers working in speech processing.

Nonlinear Analyses and Algorithms for Speech Processing Springer Science & Business Media

Connectionist Speech Recognition: A Hybrid Approach describes the theory and implementation of a method to incorporate neural network approaches into state of the art continuous speech recognition systems based on hidden Markov models (HMMs) to improve their performance. In this framework, neural networks (and in particular, multilayer perceptrons or MLPs) have been restricted to well-defined subtasks of the whole system, i.e. HMM emission probability estimation and feature extraction. The book describes a successful five-year international collaboration between the authors. The lessons learned form a case study that demonstrates how hybrid systems can be developed to combine neural networks with more traditional statistical approaches. The book illustrates both the advantages and limitations of neural networks in the framework of a statistical systems. Using standard databases and comparison with some conventional approaches, it is shown that MLP probability estimation can improve recognition performance. Other approaches are discussed, though there is no such unequivocal experimental result for these methods. Connectionist Speech Recognition is of use to anyone intending to use neural networks for speech recognition or within the framework provided by an existing successful statistical approach. This includes research and development groups working in the field of speech recognition, both with standard and neural network approaches, as well as other pattern recognition and/or neural network researchers. The book is also suitable as a text for advanced courses on neural networks or speech processing.

Improving Training of Neural Networks for Speech Recognition Morgan & Claypool Publishers
This textbook explains Deep Learning Architecture, with applications to various NLP Tasks, including Document Classification, Machine Translation, Language Modeling, and Speech Recognition. With the widespread adoption of deep learning, natural language processing (NLP),and speech applications in many areas (including Finance, Healthcare, and Government) there is a growing need for one comprehensive resource that maps deep learning techniques to NLP and speech and provides insights into using the tools and libraries for real-world applications. Deep Learning for NLP and Speech Recognition explains recent deep learning methods applicable to NLP and speech, provides state-of-the-art approaches, and offers real-world case studies with code to provide hands-on experience. Many books focus on deep learning theory or deep learning for NLP-specific tasks while others are cookbooks for tools and libraries, but the constant flux of new algorithms, tools, frameworks, and libraries in a rapidly evolving landscape means that there are few available texts that offer the material in this book. The book is organized into three parts, aligning to different groups of readers and their expertise. The three parts are: Machine Learning, NLP, and Speech Introduction The first part has three chapters that introduce readers to the fields of NLP, speech recognition, deep learning and machine learning with basic theory and hands-on case studies using Python-based tools and libraries. Deep Learning Basics The five chapters in the second part introduce deep learning and various topics that are crucial for speech and text processing, including word embeddings, convolutional neural networks, recurrent neural networks and speech recognition basics. Theory, practical tips, state-of-the-art methods, experimentations and analysis in using the methods discussed in theory on real-world tasks. Advanced Deep Learning Techniques for Text and Speech The third part has five chapters that discuss the latest and cutting-edge research in the areas of deep learning that intersect with NLP and speech. Topics including attention mechanisms, memory augmented networks, transfer learning, multi-task learning, domain adaptation, reinforcement learning, and end-to-end deep learning for speech recognition are covered using case studies.

A Bridge to Practical Applications Springer
This two-volume set (LNAI 11055 and LNAI 11056) constitutes the refereed proceedings of the 10th International Conference on Collective Intelligence, ICCCI 2018, held in Bristol, UK, in September 2018 The 98 full papers presented were carefully reviewed and selected from 240 submissions. The conference focuses on knowledge engineering and semantic web, social network analysis, recommendation methods and recommender systems, agents and multi-agent systems, text processing and information retrieval, data mining methods and applications, decision support and control systems, sensor networks and internet of things, as well as computer vision techniques.

Algorithms and Architectures Nonlinear Analyses and Algorithms for Speech ProcessingInternational Conference on Non-Linear Speech Processing, NOLISP 2005, Barcelona, Spain, April 19-22, 2005, Revised Selected Papers
This book discusses large margin and kernel methods for speech and speaker recognition Speech and Speaker Recognition: Large Margin and Kernel Methods is a collation of research in the recent advances in large margin and kernel methods, as applied to the field of speech and speaker recognition. It presents theoretical and practical foundations of these methods, from support vector machines to large margin methods for structured learning. It also provides examples of large margin based acoustic modelling for continuous speech recognizers, where the grounds for practical large margin sequence learning are set. Large margin methods for discriminative language modelling and text independent speaker verification are also addressed in this book. Key Features: Provides an up-to-date snapshot of the current state of research in this field Covers important aspects of extending the binary support vector machine to speech and speaker recognition applications Discusses large margin and kernel method algorithms for sequence prediction required for acoustic modeling Reviews past and present work on discriminative training of language models, and describes different large margin algorithms for the application of part-of-speech tagging Surveys recent work on the use of kernel approaches to text-independent speaker verification, and introduces the main concepts and algorithms Surveys recent work on kernel approaches to learning a similarity matrix from data This book will be of interest to researchers, practitioners, engineers, and scientists in speech processing and machine learning fields.

A Guide to Theory, Algorithm, and System Development Prentice Hall
Chapters in the first part of the book cover all the essential speech processing techniques for building robust, automatic speech recognition systems: the representation for speech signals and the methods for speech-features extraction, acoustic and language modeling, efficient algorithms for searching the hypothesis space, and multimodal approaches to speech recognition. The last part of the book is devoted to other speech processing applications that can use the information from automatic speech recognition for speaker identification and tracking, for prosody modeling in emotion-detection systems and in other speech processing applications that are able to operate in real-world environments, like mobile communication services and smart homes.