Speech Processing Solutions

Single Channel Phase-Aware Signal Processing in Speech Communication
Image, Video Processing and Analysis, Hardware, Audio, Acoustic and Speech Processing
Plunkett's Infotech Industry Almanac 2008
Selected Papers from the Thirteenth International Baltic Conference, DB&IS 2018
Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing
Advancements in Domain Adaptation for Speaker Recognition and Effective Speaker De-identification
Recent Advances in Nonlinear Speech Processing
Advances in Modern Blind Signal Separation Algorithms
Proceedings of the International Conference on Information Technology and Computer Application Engineering (ITCAE 2014), Hong Kong, China, 10-11 December 2014
Technological Innovation for Sustainability
Signals and Images
Whisper Speech Processing
Databases and Information Systems X
Academic Press Library in Signal Processing
Plunkett's Almanac of Middle Market Companies 2009
Proceedings of the 3rd International School on Neural Nets “Eduardo R. Caianiello”
Speech Technology
Fundamentals of Statistical Signal Processing
NASA Tech Briefs
Multilingual Speech Processing
The Only Comprehensive Guide to InfoTech Companies And Trends
Microphone Array Signal Processing
Applied Speech Processing
Theory and Applications
Artificial Neural Networks and Machine Learning - ICANN 2011
Noise Reduction in Speech Processing
1972 Conference on Speech Communication and Processing, April 24-26, 1972
Second IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2011, Costa de Caparica, Portugal, February 22-24, 2011, Proceedings
New Advances and Trends
Understanding Digital Signal Processing with MATLAB® and Solutions
Handbook On Computational Intelligence (In 2 Volumes)
Practical algorithm development
Privacy-Preserving Machine Learning for Speech Processing
The Application of Programmable DSPs in Mobile Communications
Assistive Technology for Visually Impaired and Blind People
Theory and Practice
Automatic Speech Recognition on Mobile Devices and over Communication Networks
Speech Recognition and Coding

PEARSON TRINITY

Single Channel Phase-Aware Signal Processing in Speech Communication Springer
The book discusses receiving signals that most electrical engineers detect and study. The vast majority of signals could never be detected due to random additive signals, known as noise, that distorts them or completely overshadows them. Such examples include an audio signal of the pilot communicating with the ground over the engine noise or a bioengineer listening for a fetus' heartbeat over the mother's. The text presents the methods for extracting the desired signals from the noise. Each new development includes examples and exercises that use MATLAB to provide the answer in graphic forms for the reader's comprehension and understanding.
Image, Video Processing and Analysis, Hardware, Audio, Acoustic and Speech Processing Springer Science & Business Media
Multilingual Speech ProcessingElsevier
Plunkett's Infotech Industry Almanac 2008 Springer Science & Business Media
With the introduction of WAP in Europe and I-mode in Japan, mobile terminals took their first steps out of the world of mobile telephony and into the world of mobile data. At the same time, the shift from 2nd generation to 3rd generation cellular technology has increased the potential data rate available to mobile users by tenfold as well as shifting data transport from circuit switched to packet data. These fundamental shifts in nature and the quantity of data available to mobile users has led to an explosion in the number of applications being developed for future digital terminal devices. Though these applications are diverse they share a common need for complex Digital Signal Processing (DSP) and in most cases benefit from the use of programmable DSPs (Digital Signal Processors). * Features contributions from experts who discuss the implementation and applications of programmable DSPs * Includes detailed introductions to speech coding, speech recognition, video and audio compression, biometric identification and their application for mobile communications devices * Discusses the alternative DSP technology which is attempting to unseat the programmable DSP from the heart of tomorrow's mobile terminals * Presents innovative new applications that are waiting to be discovered in the unique environment created when mobility meets signal processing The Application of Programmable DSPs in Mobile Communications provides an excellent overview for engineers moving into the area of mobile communications or entrepreneurs looking to understand state of the art in mobile terminals. It is also a must for students and professors looking for new application areas where DSP technology is being applied.
Selected Papers from the Thirteenth International Baltic Conference, DB&IS 2018 Springer Science & Business Media
This book constitutes of the major results of the EU COST (European Cooperation in the field of Scientific and Technical Research) Action 277: NSP, Nonlinear Speech Processing, running from April 2001 to June 2005. Coverage includes such areas as speech analysis for speech synthesis, speech recognition, speech-non speech discrimination and voice quality assessment, speech enhancement, and emotional state detection.
Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing CRC Press
Users of signal processing systems are never satis?ed with the system they currently use. They are constantly asking for higher quality, faster perf- mance, more comfort and lower prices. Researchers and developers should be appreciative for this attitude. It justi?es their constant e?ort for improved systems. Better knowledge about biological and physical interrelations c- ing along with more powerful technologies are their engines on the endless road to perfect systems. This book is an impressive image of this process. After "Acoustic Echo 1 and Noise Control" published in 2004 many new results lead to "Topics in 2 Acoustic Echo and Noise Control" edited in 2006 . Today – in 2008 – even morenew?ndingsandsystemscouldbecollectedinthisbook.Comparin gthe contributions in both edited volumes progress in knowledge and technology becomesclearlyvisible:Blindmethodsandmultiinputsystemsreplace "h- ble" low complexity systems. The functionality of new systems is less and less limited by the processing power available under economic constraints. The editors have to thank all the authors for their contributions. They cooperated readily in our e?ort to unify the layout of the chapters, the ter- nology, and the symbols used. It was a pleasure to work with all of them. Furthermore, it is the editors concern to thank Christoph Baumann and the Springer Publishing Company for the encouragement and help in publi- ing this book.
Advancements in Domain Adaptation for Speaker Recognition and Effective Speaker De-identification Academic Press
Recent advancements in machine learning and artificial intelligence have significantly impacted the way humans interact with machines. Voice assistant based solutions are examples of emerging technology advancements that impact human-machine interaction. Since, speech is the most natural form of human communication, voice assistant devices have received wide user acceptance, and have become a pleasant way to facilitate and address everyday living needs, including access to the current news, events, etc. These voice-based technologies have been made possible through advanced robust processing of speech signals. Depending on the application, various speech processing techniques are required to achieve an effective overall robust solution. Speech recognition is required when text content of spoken words is needed; for example adding text captions to broadcast news or YouTube videos. If a service should become available based on who is interacting with the device, speaker recognition becomes a required step; for example, if an individual gains access to a data account (e.g., music, voice-mail, health or financial records), effective speaker recognition is needed for that service. Overall, a range of solutions in speech processing can be required to address an overall request. Other areas of speech processing that benefit the human-machine interaction include language/dialect recognition, speech enhancement, machine translation, speech synthesis, voice conversion, general diarization, etc. The environment where a person interacts with a device and input tools employed (such as phone or microphone) can impact performance. It is common to have intrinsic/extrinsic mismatch between train data and application data; in other words, data used for training the speech processing tasks is often different than those at the test time. These variations need to be

considered while developing effective speech systems, especially when performance is impacted significantly due to mismatch conditions. In this dissertation, we study the problem of speaker recognition for domain mismatch. Recognizing the identity of a speaker is an important task in speaker-dependent applications, and providing robust performance regardless of how data is captured for model training and considering environmental/extrinsic changes within the application phase is very important. In this dissertation, we propose two categories of solutions to address the mismatch problem in speaker recognition: discriminant analysis based adaptation methods (generalized discriminant analysis-GDA, and support vector discriminant analysis-SVDA) and deep learning based adaptation technique (a-vector speaker embeddings). The proposed solutions are evaluated on NIST SRE-10, NIST SRE-16 and NIST SRE-18 tasks. The GDA and SVDA achieved 20% and 32% improvement in terms of EER for SRE-10 task. A-Vectors with incorporating SVDA achieved up to 18% improvement over the previous best performing solution on SRE-16 task. In addition, we propose a solution for speaker de-identification task. In more detail, the first category of solutions we propose is based on domain mismatch compensation with discriminant analysis methods. Traditional speaker recognition use linear discriminant analysis to reduce the dimensionality of speaker embeddings and provide a better discriminant feature representations for speaker classes. We propose non-linear discriminant analysis to compensate for variabilities included during recording through generalized discriminant analysis. In addition, domain adaptation is also incorporated through our proposed support vector discriminant analysis method; which also provides improved discrimination by considering the boundary structure of speaker classes. The second category of solutions are based on domain mismatch compensation with deep learning approaches. We propose a deep learning based technique to compensate for unwanted directions and information included in speaker embeddings, and provide domaininvariant speaker representations. Finally, we address speaker de- identification advancements to help protect confidential speaker or text-content within a given audio stream. Taken collectively, these three domains highlight technological advancement, which strengthen and make speaker recognition more useful in commercial, personal, and governmental/society applications, which incorporate human-speech engagement. The environment where a person interacts with a device and input tools employed (such as phone or microphone) can impact performance. It is common to have intrinsic/extrinsic mismatch between train data and application data; in other words, data used for training the speech processing tasks is often different than those at the test time. These variations need to be considered while developing effective speech systems, especially when performance is impacted significantly due to mismatch conditions. In this dissertation, we study the problem of speaker recognition for domain mismatch. Recognizing the identity of a speaker is an important task in speaker-dependent applications, and providing robust performance regardless of how data is captured for model training and considering environmental/extrinsic changes within the application phase is very important. In this dissertation, we propose two categories of solutions to address the mismatch problem in speaker recognition: discriminant analysis based adaptation methods (generalized discriminant analysis-GDA, and support vector discriminant analysis-SVDA) and deep learning based adaptation technique (a-vector speaker embeddings). The proposed solutions are evaluated on NIST SRE-10, NIST SRE-16 and NIST SRE-18 tasks. The GDA and SVDA achieved 20% and 32% improvement in terms of EER for SRE-10 task. A-Vectors with incorporating SVDA achieved up to 18% improvement over the previous best performing solution on SRE-16 task. In addition, we propose a solution for speaker de-identification task. In more detail, the first category of proposed solutions we propose are based on domain mismatch compensation with discriminant analysis methods. Traditional speaker recognition use linear discriminant analysis to reduce the dimensionality of speaker embeddings and provide a better discriminant feature representations for speaker classes. We propose non-linear discriminant analysis to compensate for variabilities included during recording through generalized discriminant analysis. In addition, domain adaptation is also incorporated through our proposed support vector discriminant analysis method; which also provides improved discrimination by considering boundary structure of speaker classes. The second category of solutions are based on domain mismatch compensation with deep learning approaches. We propose a deep learning based technique to compensate for unwanted directions and information included in speaker embeddings, and provide domain-invariant speaker representations. Finally, we address speaker de-identification advancements to help protect confidential speaker or text-content within a given audio stream. Taken collectively, these three domains highlight technological advancement, which strengthen and make speaker recognition more useful in commercial, personal, and governmental/society applications, which incorporate human-speech engagement.

Recent Advances in Nonlinear Speech Processing Springer Science & Business Media

Equal accessibility to public places and services is now required by law in many countries. For the vision-impaired, specialised technology often can provide a fuller enjoyment of the facilities of society, from large scale meetings and public entertainments to reading a book or making music. This volume explores the engineering and design principles and techniques used in assistive technology for blind and vision-impaired people. This book maintains the currency of knowledge for engineers and health workers who develop devices and services for people with sight loss, and is an excellent source of reference for students of assistive technology and rehabilitation.

Advances in Modern Blind Signal Separation Algorithms IOS Press

Signals and Images: Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing cohesively combines contributions from field experts to deliver a comprehensive account of the latest developments in signal processing. These experts detail the results of their research related to audio and speech enhancement, acoustic image estimation, video compression, biometric recognition, hyperspectral image analysis, tensor decomposition with applications in communications, adaptive sparse-interpolated filtering, signal processing for power line communications, bio-inspired signal processing, seismic data processing, arithmetic transforms for spectrum computation, particle filtering in cooperative networks, three-dimensional television, and more. This book not only shows how signal processing theory is applied in current and emerging technologies, but also demonstrates how to tackle key problems such as how to enhance speech in the time domain, improve audio quality, and meet the desired electrical consumption target for controlling carbon emissions. Signals and Images: Advances and Results in Speech, Estimation,

Compression, Recognition, Filtering, and Processing serves as a guide to the next generation of signal processing solutions for speech and video coding, hearing aid devices, big data processing, smartphones, smart digital communications, acoustic sensors, and beyond.
Proceedings of the International Conference on Information Technology and Computer Application Engineering (ITCAE 2014), Hong Kong, China, 10-11 December 2014 World Scientific
Audio Signal Processing for Next-Generation Multimedia Communication Systems presents cutting-edge digital signal processing theory and implementation techniques for problems including speech acquisition and enhancement using microphone arrays, new adaptive filtering algorithms, multichannel acoustic echo cancellation, sound source tracking and separation, audio coding, and realistic sound stage reproduction. This book's focus is almost exclusively on the processing, transmission, and presentation of audio and acoustic signals in multimedia communications for telecollaboration where immersive acoustics will play a great role in the near future.
Technological Innovation for Sustainability Springer Science & Business Media
Based on a NATO Advanced Study Institute held in 1993, this book addresses recent advances in automatic speech recognition and speech coding. The book contains contributions by many of the most outstanding researchers from the best laboratories worldwide in the field. The contributions have been grouped into five parts: on acoustic modeling; language modeling; speech processing, analysis and synthesis; speech coding; and vector quantization and neural nets. For each of these topics, some of the best-known researchers were invited to give a lecture. In addition to these lectures, the topics were complemented with discussions and presentations of the work of those attending. Altogether, the reader is given a wide perspective on recent advances in the field and will be able to see the trends for future work.
Signals and Images John Wiley & Sons
Robust Speech Recognition in Embedded Systems and PC Applications provides a link between the technology and the application worlds. As speech recognition technology is now good enough for a number of applications and the core technology is well established around hidden Markov models many of the differences between systems found in the field are related to implementation variants. We distinguish between embedded systems and PC-based applications. Embedded applications are usually cost sensitive and require very simple and optimized methods to be viable. Robust Speech Recognition in Embedded Systems and PC Applications reviews the problems of robust speech recognition, summarizes the current state of the art of robust speech recognition while providing some perspectives, and goes over the complementary technologies that are necessary to build an application, such as dialog and user interface technologies. Robust Speech Recognition in Embedded Systems and PC Applications is divided into five chapters. The first one reviews the main difficulties encountered in automatic speech recognition when the type of communication is unknown. The second chapter focuses on environment-independent/adaptive speech recognition approaches and on the mainstream methods applicable to noise robust speech recognition. The third chapter discusses several critical technologies that contribute to making an application usable. It also provides some design recommendations on how to design prompts, generate user feedback and develop speech user interfaces. The fourth chapter reviews several techniques that are particularly useful for embedded systems or to decrease computational complexity. It also presents some case studies for embedded applications and PC-based systems. Finally, the fifth chapter provides a future outlook for robust speech recognition, emphasizing the areas that the author sees as the most promising for the future. Robust Speech Recognition in Embedded Systems and PC Applications serves as a valuable reference and although not intended as a formal University textbook, contains some material that can be used for a course at the graduate or undergraduate level. It is a good complement for the book entitled Robustness in Automatic Speech Recognition: Fundamentals and Applications co-authored by the same author.
Whisper Speech Processing CRC Press
"For those involved in the design and implementation of signal processing algorithms, this book strikes a balance between highly theoretical expositions and the more practical treatments, covering only those approaches necessary for obtaining an optimal estimator and analyzing its performance. Author Steven M. Kay discusses classical estimation followed by Bayesian estimation, and illustrates the theory with numerous pedagogical and real-world examples."--Cover, volume 1.
Databases and Information Systems X Academic Press
An overview on the challenging new topic of phase-aware signal processing Speech communication technology is a key factor in human-machine interaction, digital hearing aids, mobile telephony, and automatic speech/speaker recognition. With the proliferation of these applications, there is a growing requirement for advanced methodologies that can push the limits of the conventional solutions relying on processing the signal magnitude spectrum. Single-Channel Phase-Aware Signal Processing in Speech Communication provides a comprehensive guide to phase signal processing and reviews the history of phase importance in the literature, basic problems in phase processing, fundamentals of phase estimation together with several applications to demonstrate the usefulness of phase processing. Key features: Analysis of recent advances demonstrating the positive impact of phase-based processing in pushing the limits of conventional methods. Offers unique coverage of the historical context, fundamentals of phase processing and provides several examples in speech communication. Provides a detailed review of many references and discusses the existing signal processing techniques required to deal with phase information in different applications involved with speech. The book supplies various examples and MATLAB® implementations delivered within the PhaseLab toolbox. Single-Channel Phase-Aware Signal Processing in Speech Communication is a valuable single-source for students, non-expert DSP engineers, academics and graduate students.
Academic Press Library in Signal Processing Springer Science & Business Media
This proceedings volume brings together peer-reviewed papers presented at the International Conference on Information Technology and Computer Application Engineering, held 10-11 December 2014, in Hong Kong, China. Specific topics under consideration include Computational Intelligence, Computer Science and its Applications, Intelligent Information Processing and Knowledge Engineering, Intelligent Networks and Instruments, Multimedia Signal Processing and Analysis, Intelligent Computer-Aided Design Systems and other related topics. This book provides readers a state-of-the-art survey of recent innovations and research worldwide in Information

Technology and Computer Application Engineering, in so-doing furthering the development and growth of these research fields, strengthening international academic cooperation and communication, and promoting the fruitful exchange of research ideas. This volume will be of interest to professionals and academics alike, serving as a broad overview of the latest advances in the dynamic field of Information Technology and Computer Application Engineering.

Plunkett's Almanac of Middle Market Companies 2009 Multilingual Speech Processing

Innovations and Advances in Computer Sciences and Engineering includes a set of rigorously reviewed world-class manuscripts addressing and detailing state-of-the-art research projects in the areas of Computer Science, Software Engineering, Computer Engineering, and Systems Engineering and Sciences. Innovations and Advances in Computer Sciences and Engineering includes selected papers form the conference proceedings of the International Conference on Systems, Computing Sciences and Software Engineering (SCSS 2008) which was part of the International Joint Conferences on Computer, Information and Systems Sciences and Engineering (CISSE 2008).

Proceedings of the 3rd International School on Neural Nets "Eduardo R. Caianiello" Springer Science & Business Media

Noise is everywhere and in most applications that are related to audio and speech, such as human-machine interfaces, hands-free communications, voice over IP (VoIP), hearing aids, teleconferencing/telepresence/telecollaboration systems, and so many others, the signal of interest (usually speech) that is picked up by a microphone is generally contaminated by noise. As a result, the microphone signal has to be cleaned up with digital signal processing tools before it is stored, analyzed, transmitted, or played out. This cleaning process is often called noise reduction and this topic has attracted a considerable amount of research and engineering attention for several decades. One of the objectives of this book is to present in a common framework an overview of the state of the art of noise reduction algorithms in the single-channel (one microphone) case. The focus is on the most useful approaches, i.e., filtering techniques (in different domains) and spectral enhancement methods. The other objective of Noise Reduction in Speech Processing is to derive all these well-known techniques in a rigorous way and prove many fundamental and intuitive results often taken for granted. This book is especially written for graduate students and research engineers who work on noise reduction for speech and audio applications and want to understand the subtle mechanisms behind each approach. Many new and interesting concepts are presented in this text that we hope the readers will find useful and inspiring.

Speech Technology Springer

This book presents recent advances in nonlinear speech processing beyond nonlinear techniques. It shows that it exploits heuristic and psychological models of human interaction in order to succeed in the implementations of socially believable VUIs and applications for human health and psychological support. The book takes into account the multifunctional role of speech and what is "outside of the box" (see Björn Schuller's foreword). To this aim, the book is organized in 6 sections, each collecting a small number of short chapters reporting advances "inside" and "outside" themes related to nonlinear speech research. The themes emphasize theoretical and practical issues for modelling socially believable speech interfaces, ranging from efforts to capture the nature of sound changes in linguistic contexts and the timing nature of speech; labors to identify and detect speech features that help in the diagnosis of psychological and neuronal disease, attempts to improve the effectiveness and performance of Voice User Interfaces, new front-end algorithms for the coding/decoding of effective and computationally efficient acoustic and linguistic speech representations, as well as investigations capturing the social nature of speech in signaling personality traits, emotions and improving human machine interactions.

Fundamentals of Statistical Signal Processing Plunkett Research, Ltd.

With the Internet, the proliferation of Big Data, and autonomous systems, mankind has entered into an era of 'digital obesity'. In this century, computational intelligence, such as thinking machines, have been brought forth to process complex human problems in a wide scope of areas — from social sciences, economics and biology, medicine and social networks, to cyber security.The Handbook of Computational Intelligence (in two volumes) prompts readers to look at these problems from a non-traditional angle. It takes a step by step approach, supported by case studies, to explore the issues that have arisen in the process. The Handbook covers many classic paradigms, as well as recent achievements and future promising developments to solve some of these very complex problems. Volume one explores the subjects of fuzzy logic and systems, artificial neural networks, and learning systems. Volume two delves into evolutionary computation, hybrid systems, as well as the applications of computational intelligence in decision making, the process industry, robotics, and autonomous systems.This work is a 'one-stop-shop' for beginners, as well as an inspirational source for more advanced researchers. It is a useful resource for lecturers and learners alike.

NASA Tech Briefs Elsevier

Applied Speech Processing: Algorithms and Case Studies is concerned with supporting and enhancing the utilization of speech analytics in several systems and real-world activities, including sharing data analytics related information, creating collaboration networks between several participants, and the use of video-conferencing in different application areas. The book provides a well-standing forum to discuss the characteristics of the intelligent speech signal processing systems in different domains. The book is proposed for professionals, scientists, and engineers who are involved in new techniques of intelligent speech signal processing methods and systems. It provides an outstanding foundation for undergraduate and post-graduate students as well. Includes basics of speech data analysis and management tools with several applications, highlighting recording systems Covers different techniques of big data and Internet-of-Things in speech signal processing, including machine learning and data mining Offers a multidisciplinary view of current and future challenges in this field, with extensive case studies on the design, implementation, development and management of intelligent systems, neural networks, and related machine learning techniques for speech signal processing

Multilingual Speech Processing Springer Science & Business Media

In the past few years we have written and edited several books in the area of acousticandspeechsignalprocessing. Thereasonbehindthisendeavoristhat there were almost no books available in the literature when we ?rst started while there was (and still is) a real need to publish manuscripts summarizing the most useful ideas, concepts, results, and state-of-the-art algorithms in this important area of research. According to all the feedback we have received so far, we can say that we were right in doing this. Recently, several other researchers have followed us in this journey and have published interesting books with their own visions and perspectives. The idea of writing a book on Microphone Array Signal Processing comes from discussions we have had with many colleagues and friends. As a c- sequence of these discussions, we came up with the conclusion that, again, there is an urgent need for a monograph that carefully explains the theory and implementation of microphone arrays. While there are many manuscripts on antenna arrays from a narrowband perspective (narrowband signals and narrowband processing), the literature is quite scarce when it comes to s- sor arrays explained from a truly broadband perspective. Many algorithms for speech applications were simply borrowed from narrowband antenna - rays. However, a direct application of narrowband ideas to broadband speech processing may not be necessarily appropriate and can lead to many m- understandings.