Applied Speech And Audio Processing With Matlab Examples

When Speech and Audio Signal Processing published in 1999,it stood out from its competition in its breadth of coverage andits accessible, intutiont-based style. This book was aimed atindividual students and engineers excited about the broad span ofaudio processing and curious to understand the availabletechniques. Since then, with the advent of the iPod in 2001,the field of digital audio and music has exploded, leading to amuch greater interest in the technical aspects of audioprocessing. This Second Edition will update and revise the originalbook to augment it with new material describing both the enablingtechnologies of digital music distribution (most significantly theMP3) and a range of exciting new research areas in automatic musiccontent processing (such as automatic transcription, musicsimilarity, etc.) that have emerged in the past five years, drivenby the digital music revolution. New chapter topics include: Psychoacoustic Audio Coding, describing MP3 and relatedaudio coding schemes based on psychoacoustic masking ofquantization noise Music Transcription, including automatically derivingnotes, beats, and chords from music signals. Music Information Retrieval, primarily focusing onaudio-based genre classification, artist/style identification, andsimilarity estimation. Audio Source Separation, including multi-microphonebeamforming, blind source separation, and the perception-inspiredtechniques usually referred to as Computational Auditory SceneAnalysis (CASA).

With this comprehensive and accessible introduction to the field, you will gain all the skills and knowledge needed to work with current and future audio, speech, and hearing processing technologies. Topics covered include mobile telephony, human-computer interfacing through speech, medical applications of speech and hearing technology, electronic music, audio compression and reproduction, big data audio systems and the analysis of sounds in the environment. All of this is supported by numerous practical illustrations, exercises, and hands-on MATLAB® examples on topics as diverse as psychoacoustics (including some auditory illusions), voice changers, speech compression, signal analysis and visualisation, stereo processing, low-frequency ultrasonic scanning, and machine learning techniques for big data. With its pragmatic and application driven focus, and concise explanations, this is an essential resource for anyone who wants to rapidly gain a practical understanding of speech and audio processing and technology.

Commercial applications of speech processing and recognition are fast becoming a growth industry that will shape the next decade. Now students and practicing engineers of signal processing can find in a single volume the fundamentals essential to understanding this rapidly developing field. IEEE Press is pleased to publish a classic reissue of Discrete-Time Processing of Speech Signals. Specially featured in this reissue is the addition of valuable World Wide Web links to the latest speech data references. This landmark book offers a balanced discussion of both the mathematical theory of digital speech signal processing and critical contemporary applications. The authors provide a comprehensive view of all major modern speech processing areas: speech production physiology and modeling, signal analysis techniques, coding, enhancement, quality assessment, and recognition. You will learn the principles needed to understand advanced technologies in speech processing -- from speech coding for communications systems to biomedical applications of speech analysis and recognition. Ideal for self-study or as a course text, this far-reaching reference book offers an extensive historical context for concepts under discussion, end-of-chapter problems, and practical algorithms. Discrete-Time Processing of Speech Signals is the definitive resource for students, engineers, and scientists in the speech processing field. An Instructor's Manual presenting detailed solutions to all the problems in the book is available upon request from the Wiley Makerting Department.

This book offers an overview of audio processing, including the latest advances in the methodologies used in audio processing and speech recognition. First, it discusses the importance of audio indexing and classical information retrieval problem and presents two major indexing techniques, namely Large Vocabulary Continuous Speech Recognition (LVCSR) and Phonetic Search. It then offers brief insights into the human speech production system and its modeling, which are required to produce artificial speech. It also discusses various components of an automatic speech recognition (ASR) system. Describing the chronological developments in ASR systems, and briefly examining the statistical models used in ASR as well as the related mathematical deductions, the book summarizes a number of state-of-the-art classification techniques and their application in audio/speech classification. By providing insights into various aspects of audio/speech processing and speech recognition, this book appeals a wide audience, from researchers and postgraduate students to those new to the field.

This book describes the basic principles underlying the generation, coding, transmission and enhancement of speech and audio signals, including advanced statistical and machine learning techniques for speech and speaker recognition with an overview of the key innovations in these areas. Key research undertaken in speech coding, speech enhancement, speech recognition, emotion recognition and speaker diarization are also presented, along with recent advances and new paradigms in these areas.

An in-depth treatment of algorithms and standards for perceptual coding of high-fidelity audio, this self-contained reference surveys and addresses all aspects of the field. Coverage includes signal processing and perceptual (psychoacoustic) fundamentals, details on relevant research and signal models, details on standardization and applications, and details on performance measures and perceptual measurement systems. It includes a comprehensive bibliography with over 600 references, computer exercises, and MATLAB-based projects for use in EE multimedia, computer science, and DSP courses. An ftp site containing supplementary material such as wave files, MATLAB programs and workspaces for the students to solve some of the numerical problems and computer exercises in the book can be found at ftp://ftp.wiley.com/public/sci_tech_med/audio_signal

Intended to anyone interested in numerical computing and data science: students, researchers, teachers, engineers, analysts, hobbyists... Basic knowledge of Python/NumPy is recommended. Some skills in mathematics will help you understand the theory behind the computational methods.

Over the last 20 years, approaches to designing speech and language processing algorithms have moved from methods based on linguistics and speech science to data-driven pattern recognition techniques. These techniques have been the focus of intense, fast-moving research and have contributed to significant advances in this field. Pattern Reco

This hands-on, one-stop resource describes the key techniques of speech and audio processing illustrated with extensive MATLAB examples.

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

Introduction to Digital Speech Processing highlights the central role of DSP techniques in modern speech communication research and applications. It presents a comprehensive overview of digital speech processing that ranges from the basic nature of the speech signal, through a variety of methods of representing speech in digital form, to applications in voice communication and automatic synthesis and recognition of speech. Introduction to Digital Speech Processing provides the reader with a practical introduction to the wide range of important concepts that comprise the field of digital speech processing. It serves as an invaluable reference for students embarking on speech research as well as the experienced researcher already working in the field, who can utilize the book as a reference guide.

This concise overview of digital signal generation will introduce you to powerful, flexible and practical digital waveform generation techniques. These techniques, based on phase-accumulation and phase-amplitude mapping, will enable you to generate sinusoidal and arbitrary real-time digital waveforms to fit your desired waveshape, frequency, phase offset and amplitude, and to design bespoke digital waveform generation systems from scratch. Including a review of key definitions, a brief explanatory introduction to classical analogue waveform generation and its basic conceptual and mathematical foundations, coverage of recursion, DDS, IDFT and dynamic waveshape and spectrum control, a chapter dedicated to detailed examples of hardware design, and accompanied by downloadable Mathcad models created to help you explore 'what if?' design scenarios, this is essential reading for practitioners in the digital signal processing community, and for students who want to understand and apply digital waveform synthesis techniques.

Humans are remarkable in processing speech, audio, image and some biomedical signals. Artificial neural networks are proved to be successful in performing several cognitive, industrial and scientific tasks. This peer reviewed book presents some recent advances and surveys on the applications of artificial neural networks in the areas of speech, audio, image and biomedical signal processing. It chapters are prepared by some reputed researchers and practitioners around the globe.

This book aims to convey to engineering students and researchers alike the relevant knowledge about the nature of acoustics, sound and hearing that will enable them to develop new technologies in this area through acquiring a thorough understanding of how sound and hearing works. There is currently no technical book available covering the communication path from sound sources through medium to the formation of auditory events in the brain – this book will fill this gap in the current book literature. It discusses the multidisciplinary area of acoustics, hearing, psychoacoustics, signal processing, speech and sound quality and is suitable for use as a main course textbook for senior undergraduate and graduate courses related to audio communication systems. It covers the basics of signal processing, traditional acoustics as well as the human hearing system and how to build audio techniques based on human hearing resolution. It discusses the technologies and applications for sound synthesis and reproduction, and for speech and audio quality evaluation.

This lecture is a review of what is known about modeling human speech recognition (HSR). A model is proposed, and data are tested against the model.There seem to be a large number of theories, or points of view, on how human speech recognition functions, yet few of these theories are comprehensive. What is needed is a set of models that are supported by experimental observation, that characterize how human speech recognition really works. Finally there is the practical problem of building a machine recognizer. One way to do this is to build a machine recognizer based on the reversed engineering of human recognition. This has not been the traditional approach to automatic speech recognition (ASR).What is needed is some insight into why this large difference between human performance and present day machine performance exists. Author Jont Allen addresses this and other questions.

This book is primarily intended for the undergraduate students of electronics and communication engineering and audiology. The objective of the book is to give a hands-on experience in speech and audio signal processing, starting from the recording process to the much involved signal processing aspects. The book gives a minimal treatment for the theoretical aspects. More importance is given to the experimental method for understanding the subject by doing simple experiments using Octave/Matlab, universally accepted platforms for signal processing.KEY FEATURES • Brief theoretical description fosters ability to understand the process of human speech production and perception. • Illustrative examples give hands-on experience in application development. • Exercises and problems develop skills on problem solving and assessment of level of understanding.

Learn the technology behind hearing aids, Siri, and Echo Audio source separation and speech enhancement aim to extract one or more source signals of interest from an audio recording involving several sound sources. These technologies are among the most studied in audio signal processing today and bear a critical role in the success of hearing aids, hands-free phones, voice command and other noise-robust audio analysis systems, and music post-production software. Research on this topic has followed three convergent paths, starting with sensor array processing, computational auditory scene analysis, and machine learning based approaches such as independent component analysis, respectively. This book is the first one to provide a comprehensive overview by presenting the common foundations and the differences between these techniques in a unified setting. Key features: Consolidated perspective on audio source separation and speech enhancement. Both historical perspective and latest advances in the field, e.g. deep neural networks. Diverse disciplines: array processing, machine learning, and statistical signal processing. Covers the most important techniques for both single-channel and multichannel processing. This book provides both introductory and advanced material suitable for people with basic knowledge of signal processing and machine learning. Thanks to its comprehensiveness, it will help students select a promising research track, researchers leverage the acquired cross-domain knowledge to design improved techniques, and engineers and developers choose the right technology for their target application scenario. It will also be useful for practitioners from other fields (e.g., acoustics, multimedia, phonetics, and musicology) willing to exploit audio source separation or speech enhancement as pre-processing tools for their own needs.

Applied Speech and Audio ProcessingWith Matlab ExamplesCambridge University Press

Starting with essential maths, fundamentals of signals and systems, and classical concepts of DSP, this book presents, from an application-oriented perspective, modern concepts and methods of DSP including machine learning for audio acoustics and engineering. Content highlights include but are not limited to room acoustic parameter measurements, filter design, codecs, machine learning for audio pattern recognition and machine audition, spatial audio, array technologies and hearing aids. Some research outcomes are fed into book as worked examples. As a research informed text, the book attempts to present DSP and machine learning from a new and more relevant angle to acousticians and audio engineers. Some MATLAB® codes or frameworks of algorithms are given as downloads available on the CRC Press website. Suggested exploration and mini project ideas are given for "proof of concept" type of exercises and directions for further study and investigation. The book is intended for researchers, professionals, and senior year students in the field of audio acoustics.

Speech processing addresses various scientific and technologicalareas. It includes speech analysis and variable rate coding, inorder to store or transmit speech. It also covers speech synthesis,especially from text, speech recognition, including speaker andlanguage identification, and spoken language understanding. This book covers the following topics: how to realize speechproduction and perception systems, how to synthesize and understandspeech using state-of-the-art methods in signal processing, patternrecognition, stochastic modelling computational linguistics andhuman factor studies.

Audio signal processing is a highly active research field where digital signal processing theory meets human sound perception and real-time programming requirements. It has a wide range of applications in computers, gaming, and music technology, to name a few of the largest areas. Successful applications include, for example, perceptual audio coding, digital music synthesizers, and music recognition software. The fact that music is now often listened to using headphones from a mobile device leads to new problems related to background noise control and signal enhancement. Developments in processor technology, such as parallel computing, are changing the way signal-processing algorithms are designed for audio. Topics covered, but were not limited to, the following areas: - Audio signal analysis - Music information retrieval - Enhancement and restoration of audio - Audio equalization and filtering - Audio effects processing - Sound synthesis and modeling - Audio coding - Sound capture and noise control - Sound source separation - Room acoustics and spatial audio - Signal processing for headphones and loudspeakers - High-performance computing in audio

Speech and audio processing has undergone a revolution in preceding decades that has accelerated in the last few years generating game-changing technologies such as truly successful speech recognition systems; a goal that had remained out of reach until very recently. This book gives the reader a comprehensive overview of such contemporary speech and audio processing techniques with an emphasis on practical implementations and illustrations using MATLAB code. Core concepts are firstly covered giving an introduction to the physics of audio and vibration together with their representations using complex numbers, Z transforms and frequency analysis transforms such as the FFT. Later chapters give a description of the human auditory system and the fundamentals of psychoacoustics. Insights, results, and analyses given in these chapters are subsequently used as the basis of understanding of the middle section of the book covering: wideband audio compression (MP3 audio etc.), speech recognition and speech coding. The final chapter covers musical synthesis and applications describing methods such as (and giving MATLAB examples of) AM, FM and ring modulation techniques. This chapter gives a final example of the use of time-frequency modification to implement a so-called phase vocoder for time stretching (in MATLAB). Features A comprehensive overview of contemporary speech and audio processing techniques from perceptual and physical acoustic models to a thorough background in relevant digital signal processing techniques together with an exploration of speech and audio applications. A carefully paced progression of complexity of the described methods; building, in many cases, from first principles. Speech and wideband audio coding together with a description of associated standardised codecs (e.g. MP3, AAC and GSM). Speech recognition: Feature extraction (e.g. MFCC features), Hidden Markov Models (HMMs) and deep learning techniques such as Long Short-Time Memory (LSTM) methods. Book and computer-based problems at the end of each chapter. Contains numerous real-world examples backed up by many MATLAB functions and code.

Karlheinz Brandenburg and Mark Kahrs With the advent of multimedia, digital signal processing (DSP) of sound has emerged from the shadow of bandwidth limited speech processing. Today, the main appli cations of audio DSP are high quality audio coding and the digital generation and manipulation of music signals. They share common research topics including percep tual measurement techniques and analysis/synthesis methods. Smaller but nonetheless very important topics are hearing aids using signal processing technology and hardware architectures for digital signal processing of audio. In all these areas the last decade has seen a significant amount of application oriented research. The topics covered here coincide with the topics covered in the biannual work shop on “Applications of Signal Processing to Audio and Acoustics”. This event is sponsored by the IEEE Signal Processing Society (Technical Committee on Audio and Electroacoustics) and takes place at Mohonk Mountain House in New Paltz, New York. A short overview of each chapter will illustrate the wide variety of technical material presented in the chapters of this book. John Beerends: Perceptual Measurement Techniques. The advent of perceptual measurement techniques is a byproduct of the advent of digital coding for both speech and high quality audio signals. Traditional measurement schemes are bad estimates for the subjective quality after digital coding/decoding. Listening tests are subject to sta tistical uncertainties and the basic question of repeatability in a different environment.

This second edition focuses on audio, image and video data, the three main types of input that machines deal with when interacting with the real world. A set of appendices provides the reader with self-contained introductions to the mathematical background necessary to read the book. Divided into three main parts, From Perception to Computation introduces methodologies aimed at representing the data in forms suitable for computer processing, especially when it comes to audio and images. Whilst the second part, Machine Learning includes an extensive overview of statistical techniques aimed at addressing three main problems, namely classification (automatically assigning a data sample to one of the classes belonging to a predefined set), clustering (automatically grouping data samples according to the similarity of their properties) and sequence analysis (automatically mapping a sequence of observations into a sequence of human-understandable symbols). The third part Applications shows how the abstract problems defined in the second part underlie technologies capable to perform complex tasks such as the recognition of hand gestures or the transcription of handwritten data. Machine Learning for Audio, Image and Video Analysis is suitable for students to acquire a solid background in machine learning as well as for practitioners to deepen their knowledge of the state-of-the-art. All application chapters are based on publicly available data and free software packages, thus allowing readers to replicate the experiments.

An accessible introduction to speech and audio processing with numerous practical illustrations, exercises, and hands-on MATLAB examples.

Multimedia Signal Processing is a comprehensive and accessible text to the theory and applications of digital signal processing (DSP). The applications of DSP are pervasive and include multimedia systems, cellular communication, adaptive network management, radar, pattern recognition, medical signal processing, financial data forecasting, artificial intelligence, decision making, control systems and search engines. This book is organised in to three major parts making it a coherent and structured presentation of the theory and applications of digital signal processing. A range of important topics are covered in basic signal processing, model-based statistical signal processing and their applications. Part 1: Basic Digital Signal Processing gives an introduction to the topic, discussing sampling and quantization, Fourier analysis and synthesis, Z-transform, and digital filters. Part 2: Model-based Signal Processing covers probability and information models, Bayesian inference, Wiener filter, adaptive filters, linear prediction hidden Markov models and independent component analysis. Part 3: Applications of Signal Processing in Speech, Music and Telecommunications explains the topics of speech and music processing, echo cancellation, deconvolution and channel equalization, and mobile communication signal processing. Covers music signal processing, explains the anatomy and psychoacoustics of hearing and the design of MP3 music coder Examines speech processing technology including speech models, speech coding for mobile phones and speech recognition Covers single-input and multiple-inputs denoising methods, bandwidth extension and the recovery of lost speech packets in applications such as voice over IP (VoIP) Illustrated throughout, including numerous solved problems, Matlab experiments and demonstrations Companion website features Matlab and C++ programs with electronic copies of all figures. This book is ideal for researchers, postgraduates and senior undergraduates in the fields of digital signal processing, telecommunications and statistical data analysis. It will also be a valuable text to professional engineers in telecommunications and audio and signal processing industries.

A strong reference on the problem of signal and speech enhancement, describing the newest developments in this exciting field. The general emphasis is on noise reduction, because of the large number of applications that can benefit from this technology.

A fully updated second edition of the excellent Digital Audio Signal Processing Well established in the consumer electronics industry, Digital Audio Signal Processing (DASP) techniques are used in audio CD, computer music and multi-media components. In addition, the applications afforded by this versatile technology now range from real-time signal processing to room simulation. Digital Audio Signal Processing, Second Edition covers the latest signal processing algorithms for audio processing. Every chapter has been completely revised with an easy to understand introduction into the basics and exercises have been included for self testing. Additional Matlab files and Java Applets have been provided on an accompanying website, which support the book by easy to access application examples. Key features include: A thoroughly updated and revised second edition of the popular Digital Audio Signal Processing, a comprehensive coverage of the topic as whole Provides basic principles and fundamentals for Quantization, Filters, Dynamic Range Control, Room Simulation, Sampling Rate Conversion, and Audio Coding Includes detailed accounts of studio technology, digital transmission systems, storage media and audio components for home entertainment Contains precise algorithm description and applications Provides a full account of the techniques of DASP showing their theoretical foundations and practical solutions Includes updated computer-based exercises, an accompanying website, and features Web-based Interactive JAVA-Applets for audio processing This essential guide to digital audio signal processing will serve as an invaluable reference to audio engineering professionals, R&D engineers, researchers in consumer electronics industries and academia, and Hardware and Software developers in IT companies. Advanced students studying multi-media courses will also find this guide of interest.

This book collects a wealth of information about spatial audio coding into one comprehensible volume. It is a thorough reference to the 3GPP and MPEG Parametric Stereo standards and the MPEG Surround multi-channel audio coding standard. It describes key developments in coding techniques, which is an important factor in the optimization of advanced entertainment, communications and signal processing applications. Until recently, technologies for coding audio signals, such as redundancy reduction and sophisticated source and receiver models did not incorporate spatial characteristics of source and receiving ends. Spatial audio coding achieves much higher compression ratios than conventional coders. It does this by representing multi-channel audio signals as a downmix signal plus side information that describes the perceptually-relevant spatial information. Written by experts in spatial audio coding, Spatial Audio Processing: reviews psychoacoustics (the relationship between physical measures of sound and the corresponding percepts) and spatial audio sound formats and reproduction systems; brings together the processing, acquisition, mixing, playback, and perception of spatial audio, with the latest coding techniques; analyses algorithms for the efficient manipulation of multiple, discrete and combined spatial audio channels, including both MP3 and MPEG Surround; shows how the same insights on source and receiver models can also be applied for manipulation of audio signals, such as the synthesis of virtual auditory scenes employing head-related transfer function (HRTF) processing and stereo to N-channel audio upmix. Audio processing research engineers and audio coding research and implementation engineers will find this an insightful guide. Academic audio and psychoacoustic researchers, including post-graduate and third/fourth year students taking courses in signal processing, audio and speech processing, and telecommunications, will also benefit from the information inside.

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

With the proliferation of mobile devices and hearing devices, including hearing aids and cochlear implants, there is a growing and pressing need to design algorithms that can improve speech intelligibility without sacrificing quality. Responding to this need, Speech Enhancement: Theory and Practice, Second Edition introduces readers to the basic pr

This textbook presents an introduction to signal processing for audio applications. The author’s approach posits that math is at the heart of audio processing and that it should not be simplified. He thus retains math as the core of signal processing and includes concepts of difference equations, convolution, and the Fourier Transform. Each of these is presented in a context where they make sense to the student and can readily be applied to build artifacts. Each chapter in the book builds on the previous ones, building a linear, coherent story. The book starts with a definition of sound and goes on to discuss digital audio signals, filters, The Fourier Transform, audio effects, spatial effects, audio equalizers, dynamic range control, and pitch estimation. The exercises in each chapter cover the application of the concepts to audio signals. The exercises are made specifically for Pure Data (Pd) although traditional software, such as MATLAB, can be used. The book is intended for students in media technology bachelor programs. The book is based on material the author developed teaching on the topic over a number of years.

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.Comparingthe 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.

Written by the world's top experts in the field, this multidisciplinary book explores all phases of speech technology. Topics covered include: Conversion of computerized (keyboarded) text into synthesized speech, aimed at developing "talking computers" Development of automatic speech recognition, allowing electronic devices to process verbal commands Speech training and the use of synthesized speech for the hearing- and speech-impaired In-depth discussions of specific speech technologies are included, as well as a treatment of the issues and challenges of human-computer interfaces. Oriented toward state-of-the-art applications, the book emphasizes the practical utilization of emerging technologies and includes numerous case studies.