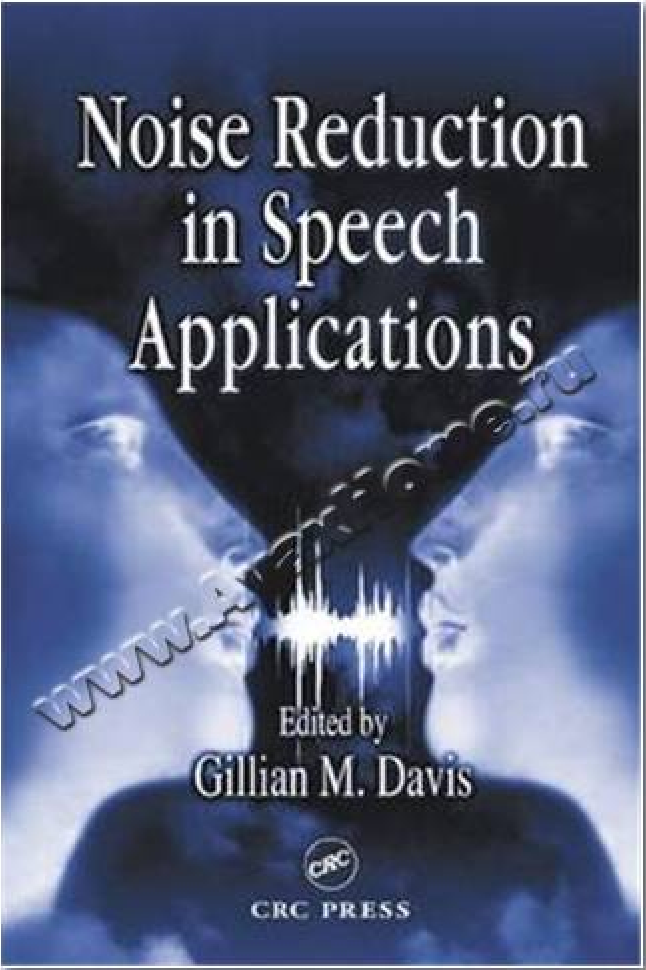
Noise Reduction in Speech Applications
Edited by
Gillian M. Davis
CRC
CRC PRESS

Noise Reduction in Speech Applications

THE ELECTRICAL ENGINEERING AND APPLIED SIGNAL PROCESSING SERIES

Edited by Alexander Poularikas

The Advanced Signal Processing Handbook: Theory and Implementation for Radar, Sonar, and Medical Imaging Real-Time Systems
Stergios Stergiopoulos

The Transform and Data Compression Handbook
K.R. Rao and P.C. Yip

Handbook of Multisensor Data Fusion
David Hall and James Llinas

Handbook of Neural Network Signal Processing
Yu Hen Hu and Jenq-Neng Hwang

Handbook of Antennas in Wireless Communications
Lal Chand Godara

Noise Reduction in Speech Applications
Gillian M. Davis

Forthcoming Titles

Propagation Data Handbook for Wireless Communications
Robert Crane

The Digital Color Imaging Handbook
Guarav Sharma

Applications in Time Frequency Signal Processing
Antonia Papandreou-Suppappola

Signal Processing Noise
Vyacheslav P. Tuzlukov

Digital Signal Processing with Examples in MATLAB®
Samuel Stearns

Smart Antennas
Lal Chand Godara

Pattern Recognition in Speech and Language Processing
Wu Chou and Bing Huang Juang

Nonlinear Signal and Image Processing: Theory, Methods, and Applications
Kenneth Barner and Gonzalo R. Arce

Noise Reduction in Speech Applications

Edited by

Gillian M. Davis

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Noise reduction in speech applications / edited by Gillian M. Davis.
p. cm. — (The electrical engineering and applied signal processing series)
Includes bibliographical references and index.
ISBN 0-8493-0949-2 (alk. paper)
1. Speech processing systems. 2. Telephone systems. 3. Electronic noise—Prevention. 4. Noise control. 5. Signal processing—Digital techniques. I. Davis, Gillian M. II. Series.

TK7882.S65 N65 2002
621.382′8—dc21 2002017483

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

International Standard Book Number 0-8493-0949-2
Library of Congress Card Number 2002017483
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

A wide range of potential sources of noise and distortion can degrade the quality of the speech signal in a communication system. *Noise Reduction in Speech Applications* explores the effects of these interfering sounds on speech applications and introduces a range of techniques for reducing their influence and enhancing the acceptability, intelligibility, and speaker recognizability of the communications signal. A systems approach to noise reduction is taken that emphasizes the advantage of minimizing noise pickup and creation, in the first instance, in addition to choosing the most appropriate noise reduction technique at each stage to obtain the best overall result. This handbook aims to make the available technologies better known and to set expectations of what can actually be achieved in practice at a realistic level. Sufficient detail is given for readers to decide which, if any, of the noise reduction techniques discussed is an appropriate solution for their own systems and also to help them make the best use of these technologies.

The timing of this book is particularly appropriate. Although much of the technology required for noise reduction has existed for some time, it is only with the recent development of powerful but inexpensive digital signal processing (DSP) hardware that implementation of the technology in everyday systems has started to become practical.

Noise Reduction in Speech Applications begins with a tutorial chapter covering background material on digital signal processing and adaptive filtering. Emphasis is placed on techniques relevant to noise reduction in speech applications that are referenced by the authors of later chapters. This tutorial chapter is written at a level suitable for students studying DSP techniques as part of their electrical engineering or acoustics courses and for master's degree students taking specialist digital signal processing courses.

The remainder of the book is divided into three sections:

Systems Aspects addresses the need to consider the complete system and apply the most appropriate noise reduction technique at each stage to achieve the best overall result.

Digital Algorithms and Implementation looks at three types of digital noise reduction algorithms in detail: single-channel speech enhancers, microphone arrays, and echo cancellers. Example code and audio wavefiles illustrating the noise problems and solutions are provided to accompany these chapters. These files are available at http://www.crcpress.com/e_products/download.asp?cat_no=0949. The example code will be of particular interest to students of the subject, whereas the audio wavefiles will be of interest to a wider audience, including readers with limited technical knowledge but an interest in the noise reduction achievable with such algorithms.

Special Applications investigates the use of noise reduction techniques in eight application areas, including speech recognition, Internet telephony, and digital hearing aids. This final section is aimed at potential commercial customers of this technology and focuses on the sorts of results that can be achieved in practice. Audio wavefiles are provided to accompany these chapters at http://www.crcpress.com/e_products/download.asp?cat_no=0949.

Each chapter of this book concludes with a list of references that provide guidance for readers wishing to examine the subject of the chapter in more detail. In addition, because many of the chapters use acronyms that may be unfamiliar to the reader, a general list of acronyms is provided at the front of the book.

It has been a great pleasure working with the chapter authors on the production of this book. The willingness of these specialists to sacrifice valuable research time to prepare this review material is greatly appreciated. Without their commitment, this book would not have been possible. I have enjoyed watching this book develop, and I hope that those who read it will find it to be a valuable source of information.

Many thanks are due to NCT Group, Inc. U.S.A. for supporting my involvement in the preparation of this book. A substantial debt of gratitude is due also to James Elburn of NCT (Europe) Ltd., U.K. who provided invaluable IT support throughout this project, and to Stephen Leese, who advised on suitable subjects and contributors.

Gillian M. Davis

The Editor

Gillian M. Davis, D.Phil., is Managing Director of Noise Cancellation Technologies (Europe) Ltd. and a Vice President of the parent company, NCT Group, Inc. Previously she held research positions at Sharp Laboratories of Europe Ltd., NTT, Japan, and Rutherford Appleton Laboratory. Dr. Davis received her D.Phil. from the Clarendon Laboratory, Oxford University and her M.B.A. from the Open University, Milton Keynes.

Contributors

Victor Bray Auditory Research Department, Sonic Innovations, Inc., Salt Lake City, Utah, U.S.A.

Douglas M. Chabries College of Engineering and Technology, Brigham Young University, Provo, Utah, U.S.A.

Ingvar Claesson Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, Ronneby, Sweden

John W. Cook BTexact Technologies, Martlesham Heath, Ipswich, U.K.

Mattias Dahl Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, Ronneby, Sweden

Gillian M. Davis NCT (Europe) Ltd., Cambridge, U.K.

Graham P. Eatwell Adaptive Audio, Inc., Annapolis, Maryland, U.S.A.

Craig Fancourt Sarnoff Corporation, Princeton, New Jersey, U.S.A.

Lars Håkansson Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, Ronneby, Sweden

Dennis Hardman Agilent Technologies, Inc., Colorado Springs, Colorado, U.S.A.

Malcolm J. Hawksford Department of Electronic Systems Engineering, University of Essex, Colchester, Essex, U.K.

Sven Johansson Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, Ronneby, Sweden

Elizabeth G. Keate Texas Instruments, Santa Barbara, California, U.S.A.

George Keratiotis BTexact Technologies, Martlesham Heath, Ipswich, U.K.

Stephen J. Leese NCT (Europe) Ltd., Cambridge, U.K.

Larry Lind Department of Electronic Systems Engineering, University of Essex, Colchester, Essex, U.K.

Robert S. Oshana Software Development Systems, Texas Instruments, Dallas, Texas, U.S.A.

Ira L. Panzer Dynastat, Inc., Austin, Texas, U.S.A.

Lucas Parra Sarnoff Corporation, Princeton, New Jersey, U.S.A.

Minesh Patel BTexact Technologies, Martlesham Heath, Ipswich, U.K.

Bhiksha Raj Mitsubishi Electric Research Laboratories, Cambridge, Massachusetts, U.S.A.

Alan D. Sharpley Dynastat, Inc., Austin, Texas, U.S.A.

Rita Singh School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.

Per Sjösten National Institute for Working Life, Göteborg, Sweden

Richard M. Stern Department of Electrical and Computer Engineering and School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.

Robert W. Stewart Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, Scotland, U.K.

William D. Voiers Dynastat, Inc., Austin, Texas, U.S.A.

Darren B. Ward Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London, U.K.

Stephan Weiss Department of Electronics and Computer Science, University of Southampton, Southampton, U.K.

Pete Whelan BTexact Technologies, Martlesham Heath, Ipswich, U.K.

Acronyms

ACRM	Absolute Category Rating Method
ADC	analog-to-digital converter
AFC	alternative forced choice
AGC	automatic gain control
ALU	arithmetic logic unit
ANC	active noise control
ANCU	adaptive noise cancellation unit
ANN	artificial neural network
API	application program interface
APLB	adaptive phase-locked buffer algorithm
AR	autoregressive
ASIC	application-specific integrated circuit
ATM	asynchronous transfer mode; automated teller machine
BB	background buzz (specific to the DAM)
BC	background chirping (specific to the DAM)
BJT	bipolar junction transistor
BNH	background noise high frequency (specific to the DAM)
BNM	background noise mid frequency (specific to the DAM)
BNL	background noise low frequency (specific to the DAM)
BS	background static (specific to the DAM)
BSS	blind source separation
CAE	composite acceptability estimate (specific to the DAM)
CBA	composite background acceptability (specific to the DAM)
CCS	crosstalk cancellation system
CDCN	codeword-dependent cepstral normalization
CE	common mode error
CER	command error rate
CIA	composite isometric acceptability (specific to the DAM)
CM	common mode
CMN	cepstral mean normalization
CNG	comfort noise generator
COTS	commercial off-the-shelf
CPA	composite perceptual acceptability (specific to the DAM)
CPU	central processing unit
CSA	composite signal acceptability (specific to the DAM)
CSU	critical speech unit

DAC	digital-to-analog converter
DACS	Digital Access Carrier Systems
DALT	Diagnostic Alliteration Test
DAM	Diagnostic Acceptability Measure
DAT	Diagnostic Alliteration Test
DCRM	Degradation Category Rating Method
DCT	discrete cosine transform
DE	differential error
DFT	discrete Fourier transform
DM	differential mode
DMA	direct memory access
DMCT	Diagnostic Medial Consonant Test
DMOS	Degradation Mean Opinion Score
DPLL	digital phase-locked loop
DRT	Diagnostic Rhyme Test
DSL	digital subscriber line
DSP	digital signal processing/processor
DSRT	Diagnostic Speaker Recognizability Test
DST	discrete sine transform
DTFT	discrete-time Fourier transform
DWT	discrete wavelet transform
EM	expectation maximization
EMAP	extended MAP
emf	electromagnetic field
EPQ	elementary perceived qualities
ERL	echo return loss
ERLE	echo return loss enhancement
ETSI	European Telecommunications Standards Institute
FEC	front-end clipping
FFT	fast Fourier transform
FIR	finite impulse response
FXLMS	filtered-x least mean squares
FXO	foreign exchange office
FXS	foreign exchange station
GCC	generalized cross correlation
GJB	Griffiths-Jim beamformer
GPP	general-purpose processor
GSC	generalized sidelobe canceller
GSD	generalized sidelobe decorrelator
GSS	geometric source separation
GUI	graphical user interface

HINT	Hearing in Noise Test
HMM	hidden Markov model
HOS	higher-order statistics
HOT	hold-over time
HPI	host port interface
HRTF	head-related transfer function
IBA	isometric background acceptability (specific to the DAM)
IIR	infinite impulse response
IMC	internal model control
INT	induction neutralizing transformers
ISA	instruction set architecture; isometric signal acceptability (specific to the DAM)
ISDN	integrated services digital network
ISR	interrupt service routine
ITU	International Telecommunication Union
JFET	junction field-effect transistor
KLT	Karhunen-Loeve transform
LAN	local area network
LCMV	linearly constrained minimum variance
LMS	least mean square(s)
LP	linear prediction
LPF	low-pass filter
LSS	linear spectral subtraction
LTI	linear and time invariant
MAC	multiply and accumulate
MAP	maximum *a posteriori*
MFCC	Mel frequency cepstral coefficients
MLLR	maximum likelihood linear regression
MMSE	minimum mean square(d) error
MNRU	modulated noise reference unit
MOS	mean opinion score
MPLS	multiprotocol label switching
MRT	Modified Rhyme Test
MSE	mean squared error
MSUB	magnitude spectrum of noise
NG	noise generator
NLP	nonlinear processing
NOC	network operations center

NOP	null operations
NSS	nonlinear spectral subtraction
OS	operating system
OSI	open system interconnection
PAC	physical acoustical correlates
PAMS	perceptual analysis measurement system
PB	phonetically balanced
PBA	perceptual background acceptability (specific to the DAM)
PCM	pulse code modulated
PCU	pipeline control unit
pdf	probability density function
PDF	probability distribution function
PESQ	perceptual evaluation of speech quality
PMC	parallel model combination
POF	probabilistic optimal filtering
PSA	perceptual signal acceptability (specific to the DAM)
PSD	power spectral density
PSTN	public-switched telephone network
PSUB	power spectrum of noise
PVT	perceived voice trait
QoS	quality of service
RASTA	relative spectral processing
RFI	radio frequency interference
RISC	reduced instruction set computer
RLS	recursive least squares (algorithm)
RMA	rate-monotonic analysis
RMS	root-mean square
RSVP	resource reservation protocol
RTCP	real-time transport control protocol
RTOS	real-time operating system
RTP	real-time transport protocol
SB	signal babble (specific to the DAM)
SD	signal distortion (specific to the DAM)
SF	signal flutter (specific to the DAM)
SH	signal high pass (specific to the DAM)
SI	signal interrupted (specific to the DAM)
SIP	Session Initiation Protocol
SIR	signal-to-interference ratio
SL	signal low pass (specific to the DAM)

SN	signal nasal (specific to the DAM)
SNR	signal-to-noise ratio; speech-to-noise ratio
SPINE	speech in noisy environments (database)
SQNR	signal-to-quantization-noise ratio
SRAM	static RAM
SS	spectral subtraction
ST	signal thin (specific to the DAM)
TCB	task control block
TCL	terminal coupling loss
TCP	transport control protocol
TF	transfer function
THD	total harmonic distortion
TIA	Telecommunications Industry Association
TOS	type of service
TRI	transformed rating intelligibility (specific to the DAM)
TRP	transformed rating pleasantness (specific to the DAM)
UDP	user datagram protocol
VAD	voice activity detector
VCA	voltage-controlled amplifier
VLIW	very long instruction word
VoFR	voice over frame relay
VoIP	Voice over Internet Protocol
VOX	voice-operated switch
VTS	vector Taylor series
WAN	wide area network
WDRC	wide dynamic range compression
WS	waveform synthesis

Contents

Section IV Special Applications

Section I:

Tutorial

1

Noise and Digital Signal Processing

Stephan Weiss, Robert W. Stewart, and Gillian M. Davis

CONTENTS

Introduction

Electronic systems in the context of audio communication perform transmission, recording, playback, analysis, or synthesis of speech signals. When designing a system for any of these purposes, noise influences must be carefully considered. Different types of noise and distortion can be characterized, and a number of signal processing concepts exist that can assist in mitigating their effect, thus enhancing the quality or intelligibility of the speech signal. Digital signal processing (DSP) offers a number of powerful tools that, depending on the circumstances, may or may not be applicable to specific types of noise corruptions. Therefore, the purpose of this introductory chapter is to create some awareness of the "noise chain" in a speech communication system and the relevant fundamental digital signal processing concepts, which can be exploited to design a system that is robust toward noise.

"Analog/Digital Interfacing and Noise Chain" gives an overview of the different stages in a general digital speech communication system and their exposure to, and inherent production of, noise. To characterize noise, which is generally assumed to be random, some background of stochastic signals and their quantities are reviewed in "Stochastic Signals and Their Characteristics." If the noise does not share the same frequency range as the speech signal, digital filtering can be a promising technique for noise reduction; this is discussed in "Digital Filtering." If noise and speech overlap in frequency but the speech signal exhibits specific features, processing in a transform domain may present a viable solution. The fundamentals of transforms and transform-domain processing are detailed in "Discrete Signal Transforms." Finally, if a probe of the corrupting noise is available or the noise is periodic, powerful adaptive DSP algorithms can be applied; these are reviewed in "Adaptive Digital Filtering."

The scope of this chapter is to provide a simple and generally intuitive overview of the DSP fundamentals upon which noise reduction methods can be based. Examples are provided that give additional insight into the working and application of the techniques discussed. For further information, the reader should consult the reference list at the end of this chapter and the specialized DSP applications discussed elsewhere in Sections III and IV of this book.

Analog/Digital Interfacing and Noise Chain

Any system involving the transmission, acquisition, or generation of speech is subject to a wide range of influences that may deteriorate the quality of the speech signal. This can include external interferences such as background