SpringerBriefs in Speech Technology

Series Editor:

Amy Neustein

For further volumes:

<http://www.springer.com/series/10043>

Editor's Note

The authors of this series have been hand-selected. They comprise some of the most outstanding scientists –drawn from academia and private industry –whose research is marked by its novelty, applicability, and practicality in providing broad based speech solutions. The SpringerBriefs in Speech Technology series provides the latest findings in speech technology gleaned from comprehensive literature reviews and *empirical investigations* that are performed in both laboratory and *real life* settings. Some of the topics covered in this series include the presentation of real life commercial deployment of spoken dialog systems, contemporary methods of speech parameterization, developments in information security for automated speech, forensic speaker recognition, use of sophisticated speech analytics in call centers, and an exploration of new methods of soft computing for improving human-computer interaction. Those in academia, the private sector, the self service industry, law enforcement, and government intelligence, are among the principal audience for this series, which is designed to serve as an important and essential reference guide for speech developers, system designers, speech engineers, linguists and others. In particular, a major audience of readers will consist of researchers and technical experts in the automated call center industry where speech processing is a key component to the functioning of customer care contact centers.

Amy Neustein, Ph.D., serves as Editor-in-Chief of the International Journal of Speech Technology (Springer). She edited the recently published book "Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics" (Springer 2010), and serves as guest columnist on speech processing for Womensenews. Dr. Neustein is Founder and CEO of Linguistic Technology Systems, a NJ-based think tank for intelligent design of advanced natural language based emotion-detection software to improve human response in monitoring recorded conversations of terror suspects and helpline calls. Dr. Neustein's work appears in the peer review literature and in industry and mass media publications. Her academic books, which cover a range of political, social and legal topics, have been cited in the Chronicles of Higher Education, and have won her a pro Humanitate Literary Award. She serves on the visiting faculty of the National Judicial College and as a plenary speaker at conferences in artificial intelligence and computing. Dr. Neustein is a member of MIR (machine intelligence research) Labs, which does advanced work in computer technology to assist underdeveloped countries in improving their ability to cope with famine, disease/illness, and political and social affliction. She is a founding member of the New York City Speech Processing Consortium, a newly formed group of NY-based companies, publishing houses, and researchers dedicated to advancing speech technology research and development.

Fathi E. Abd El-Samie

Information Security for Automatic Speaker Identification

Fathi E. Abd El-Samie
Department of Electronics and Electrical Communications
Faculty of Electronic Engineering
Menoufia University
Menouf, 32952, Egypt
Fathi_sayed@yahoo.com

ISSN 2191-737X e-ISSN 2191-7388
ISBN 978-1-4419-9697-8 e-ISBN 978-1-4419-9698-5
DOI 10.1007/978-1-4419-9698-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011929038

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To my grand mother, my mother, my wife,
and my daughter.*

Preface

I know what you are asking yourself – “there are a lot of books available in speech processing, what is novel in this book?” Well, I can summarize the answer for this question in the following points:

1. You always see different algorithms for speech enhancement, deconvolution, signal separation, watermarking, and encryption, separately, without specific applications for these algorithms.
2. You also see literature books and research papers on speaker identification concentrating on how to extract features, and the comparison between feature extraction methods.
3. How to make use of speech enhancement, deconvolution, and signal separation to enhance the performance of speaker identification systems is a missing issue. This book presets this issue and gives comparison studies between different algorithms that can be used for this purpose.
4. Speech watermarking and encryption are studied for the first time in this book in a framework that enhances the security of speaker identification systems.
5. Performance enhancement and security enhancement of speaker identification systems are contradicting objectives. How they affect each other is also studied in this book.

Finally, I hope that this book will be a starting step towards an extensive study to build speaker identification systems with multilevels of security.

Fathi E. Abd El-Samie

Acknowledgements

I would like to thank my students Amira Shafik, Hossam Hammam, Emad Mosa, and Marwa Abdelfattah for their cooperation in software arrangement.

Contents

Information Security for Automatic Speaker Identification	1
1 Introduction	2
2 Speaker Identification.....	3
2.1 Feature Extraction.....	5
2.2 Feature Matching	9
3 Feature Extraction from Discrete Transforms.....	10
3.1 The Discrete Wavelet Transform	11
3.2 The DCT	14
3.3 The DST.....	14
3.4 Speaker Identification with Discrete Transforms.....	14
4 Speech Enhancement	16
4.1 Speech Quality Metrics	16
4.2 Spectral Subtraction.....	18
4.3 Wiener Filter	21
4.4 The Adaptive Wiener Filter.....	23
4.5 Wavelet Denoising.....	25
5 Blind Signal Separation.....	28
6 Deconvolution of Speech Signals	36
6.1 LMMSE Deconvolution	36
6.2 Inverse Filter Deconvolution	46
6.3 Regularized Deconvolution.....	47
6.4 Comparison Study	49
7 Speech Watermarking	52
7.1 Singular Value Decomposition	56
7.2 The SVD Speech Watermarking Algorithm.....	57
7.3 Chaotic Encryption.....	58
7.4 The Segment-by-Segment SVD Speech Watermarking Algorithm	59
7.5 Evaluation of SVD Speech Watermarking	61

8 Speech Encryption..... 78

8.1 Framing and Reshaping into 2D Blocks 91

8.2 Block Randomization 91

8.3 Steps of the Rounds 92

8.4 Masking..... 92

8.5 Discrete Transforms 93

8.6 Performance Evaluation of Speech Encryption 94

9 Simultaneous Watermarking and Encryption 112

10 Conclusion..... 112

11 Directions for Future Research..... 117

References 118

About the Author

Fathi E. Abd El-Samie received the B.Sc. (Honors), M.Sc., and Ph.D. from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1998, 2001, and 2005, respectively. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 2005. He is a coauthor of about 130 papers in international conference proceedings and journals. He has received the most cited paper award from Digital Signal Processing journal for 2008. His current research areas of interest include image processing, speech processing, and digital communications.