

Preface

Speech and hearing sciences are fundamental to numerous technological advances of the digital world in the past decade, from music compression in MP3 to digital hearing aids, from network based voice enabled services to speech interaction with mobile phones. Mathematics and computation are intimately related to these leaps and bounds. On the other hand, speech and hearing are strongly interdisciplinary areas where dissimilar scientific and engineering publications and approaches often coexist and make it difficult for newcomers to enter.

The aim of our book is to give an accessible introduction of mathematical models and signal processing methods in speech and hearing sciences for senior undergraduate and beginning graduate students with basic knowledge of linear algebra, differential equations, numerical analysis, and probability. The models and methods are selected based on their physical and biological origin, mathematical simplicity, and their utility for signal processing. Connections are drawn as much as possible between model solutions and speech/hearing phenomena. Concepts such as critical bands, sound masking, and hearing loss are introduced in terms of both model solutions and experimental data. It is our hope that the self-contained presentation of hidden Markov models and the associated Matlab codes for isolated words recognition in chapter four will help make speech recognition accessible to beginners. We include representative Matlab programs and a moderate number of exercises in each chapter to help the readers gain hands-on experience and consolidate their understanding. Speech data for the Matlab programs are either clean signals or recorded mixtures downloadable from the first author's website. Matlab signal processing and statistics toolboxes are needed for some of the programs. The mathematical tools consist of elementary analysis of differential equations, asymptotic and numerical methods, transform techniques, filtering and clustering methods, statistical and optimization methods. Some of these tools show up multiple times in the book especially in the context of solving concrete model and real world problems.

The first chapter of the book presents background materials on function spaces, Fourier and z-transforms, filtering-clustering-spectral analysis of data, optimization and statistical methods. Chapter two is on modeling speech production with mechanical and digital source-filter models. Chapter three discusses partial differential

equation (PDE) models of the peripheral auditory system, their analysis and computation, their applications in sound transform and processing, and hearing aids. Chapter four introduces the hidden Markov concept, the framework of speech recognition, and the related learning and searching algorithms. Chapter five studies blind source separation and speech enhancement (noise reduction) methods based on statistical criteria, sparsity and feature clustering in time-frequency domain. The order of chapter two to chapter five follows logically the human speech chain: speech production, audition, recognition and signal processing.

The book is based on the authors' decade long collaborations with graduate students, postdoctoral fellows and colleagues in mathematics, speech and hearing sciences, and signal processing. We are grateful to Professor Stanley Osher for his constant support and his pioneering work on image processing that inspired us. We thank the following colleagues (in alphabetical order) for their interest, encouragement and assistance that helped us embark on our journey and pursue our goals: Professors Luis Caffarelli, Russel Caflisch, Emmanuel Candès, Tony Chan, Ingrid Daubechies, Susan Friedlander, Irene Gamba, James Hyman, Joe Keller, Peter Lax, Jerry Marsden, Tinsley Oden, George Papanicolaou, Charles Peskin, George Pollak, Donald Saari, Charles Steele, Ronald Stern, Howard Tucker, Frederick Wan, Shing-Tung Yau, and Hongkai Zhao. We thank Professors Li Deng, Deliang Wang, Yang Wang, and Fan-Gang Zeng for many fruitful discussions on speech and hearing research and applications. Progress would not have been possible without the opportunity of working with creative and energetic students, postdoctoral fellows and visiting scholars (in chronological order): M. Drew LaMar, Yongsam Kim, Jie Liu, Hsin-I Yang, Meng Yu, J. Ernie Esser, Yuanchang Sun, Wenye Ma, Ryan Ritch, Penghang Yin, Daniel Quang, Yifei Lou, He Qi and Xiaohua Shi. Part of the book has been used for training and supervised research experience of the undergraduate students of the NSF supported PRISM (Proactive Recruitment in Introductory Science and Mathematics) program at UC Irvine (iCAMP) with the help of Dr. Ernie Esser (2009 –2013). We benefited from the IMA Speech Processing Workshop at the University of Minnesota in 2000, and from organizing and interacting with the participants of the IPAM workshop on "Mathematics of the Ear and Sound Signal Processing" at UCLA in 2005. Part of the materials is drawn from lectures at the Beijing Summer School in 2010 organized by Professor Zhimin Chen at Academia Sinica. We thank Professor Thomas Hou for kindly hosting one of us at Caltech while our work was ongoing, and for suggesting this book project. We thank Dandan Yu for the cover figure design. Finally, we acknowledge the financial support from the National Science Foundation (NSF), the Guggenheim Foundation, the Army Research Office, the National Institute of Health, the University of Texas at Austin, and the University of California at Irvine.

Irvine, California
September 2013

Jack Xin
Yingyong Qi

Mathematical Modeling and Signal Processing in
Speech and Hearing Sciences

Xin, J.; Qi, Y.

2014, XI, 208 p. 10 illus., Hardcover

ISBN: 978-3-319-03085-2