Computational Models of Auditory Function

Edited by

Steven Greenberg

International Computer Science Institute, Berkeley, California, USA

and

Malcolm Slaney

IBM Almaden Research Center, San Jose, California, USA

IOS Press

Ohmsha

Amsterdam • Berlin • Oxford • Tokyo • Washington, DC

Published in cooperation with NATO Scientific Affairs Division

Proceedings of the NATO Advanced Study Institute on
Computational Hearing
Il Ciocco, Italy
1–12 July 1998

© 2001, IOS Press

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior written permission from the publisher.

ISBN 90 5199 457 5 (IOS Press)
ISBN 4 274 90465 2 C3045 (Ohmsha)
Library of Congress Control Number: 2001094183

Publisher
IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 620 3419
e-mail: order@iospress.nl

Distributor in the UK and Ireland
IOS Press/Lavis Marketing
73 Lime Walk
Headington
Oxford OX3 7AD
England
fax: +44 1865 75 0079

Distributor in the USA and Canada
IOS Press, Inc.
5795-G Burke Centre Parkway
Burke, VA 22015
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

Distributor in Germany, Austria and Switzerland
IOS Press/LSL.de
Gerichtsweg 28
D-04103 Leipzig
Germany
fax: +49 341 995 4255

Distributor in Japan
Ohmsha, Ltd.
3-1 Kanda Nishiki-cho
Chiyoda-ku, Tokyo 101
Japan
fax: +81 3 3233 2426

LEGAL NOTICE
The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

Contents

Preface ix
Steven Greenberg and Malcolm Slaney

Physiology of the Auditory Periphery

Introduction 3
Steven Greenberg

Coding Efficiency of Inner Hair Cells at the Threshold of Hearing 5
Ilse C. Gebeshuber and Frank Rattay

A Method for Evaluation of Multiparameter Nonlinear Models Illustrated on a Computational Model of Auditory-Nerve Rate–Level Curves 17
Konstantina M. Stankovic

Characteristics of Cochlear Nucleus Onset Units Studied with a Model 29
Sridhar Kalluri and Bertrand Delgutte

Modeling the Cochlea

Introduction 47
Malcolm Slaney

An Analysis/Synthesis Auditory Filterbank Based on an IIR Gammachirp Filter 49
Toshio Irino and Masashi Unoki

Advances in Computational Modeling of Cochlear Implant Physiology and Perception 65
I. C. Bruce, M. W. White, L. S. Irlicht, S. J. O'Leary and G. M. Clark

Sound Localization and Binaural Mechanisms

Introduction 79
Jens Blauert

Preliminary Model of Auditory Distance Perception for Nearby Sources 83
Douglas S. Brungart

A Computational Model of Auditory Sound Localization 97
Kazuhito Ito and Masato Akagi

A Computational Model of Sound Localization Based on Neurophysiological Data 113
Klaus Hartung and Susanne J. Sterbing

The Bat as a Computational System

Introduction 129
Steven Greenberg

Time and Frequency Information Used in the Computation of Elevation by an Echolocating Bat 131
Janine M. Wotton, Michael J. Ferragamo, Rick L. Jenison and James A. Simmons

Computational Assessment of an Acoustic Flow Hypothesis in CF-Bats 141
Rolf Müller and Hans-Ulrich Schnitzler

Auditory Processing of Pitch

Introduction 153
Malcolm Slaney

A Computational Auditory Model of the Lateralization of the Huggins Pitch 157
Michael A. Akeroyd and A. Quentin Summerfield

Processing of Pitch Information in Complex Stimuli by a Model of the Octopus Cells in the Cochlear Nucleus 171
Yidao Cai, JoAnn McGee and Edward J. Walsh

Temporal Processing and Periodicity Analysis

Introduction 185
Roy D. Patterson

Aspects of Envelope Coding in the Auditory System 189
Peter Heil

Functional Significance of Latencies in the Auditory System 205
Stefan Bleeck and Gerald Langner

A Computational Model of Co-Modulation Masking Release 221
Masashi Unoki and Masato Akagi

Neural Timing Nets for Auditory Computation 233
Peter Cariani

Auditory Dynamics

Introduction 251
Shihab Shamma

Synchronous Oscillations in the Thalamocortical System and the Effects of Naturalistic Ripple Stimuli 253
L. M. Miller, M. Escabí and C. E. Schreiner

Auditory Scene Analysis

Introduction 267
Malcolm Slaney

A Procedure for Identification and Segregation of Multiple Auditory Objects 269
Uwe Baumann

Cortical Synaptic Depression and Auditory Perception 281
Sue L. Denham

Applying a Model of Concurrent Vowel Segregation to Real Speech 297
Georg F. Meyer, Dekun Yang and William A. Ainsworth

Auditory Processing of Speech

Introduction 313
Steven Greenberg and Malcolm Slaney

Modeling the Perception of Pitch-Rate Amplitude Modulation in Noise 315
Brian P. Strope and Abeer A. Alwan

Front-End Design by Using Auditory Modeling in Speech Recognition 329
Jilei Tian, Kari Laurila, Ramalingam Hariharan and Imre Kiss

STRAIGHT: An Extremely High-Quality VOCODER for Auditory and Speech Perception Research 343
Hideki Kawahara

Subject Index 355

Author Index 361

Computational Models of Auditory Function
S. Greenberg and M. Slaney (eds.)
IOS Press, 2001

PREFACE

"The purpose of computing is insight, not numbers."
Richard Hamming

Homo sapiens is considered to be, above all else, a visually oriented species, with the other senses viewed more as supporting players than as star performers. This "bias" in perceived sensory function has had the historical consequence of casting the auditory modality into the "back seat" of both experimental and computational neuroscience for many decades — or so it would appear.

In actuality, auditory neuroscience has been at the vanguard of both disciplines for many years. Important scientific landmarks include:

(1) the first biological application of Fourier's theorem [5]
(2) the first systematic application of electrical recording technology in sensory neurophysiology [8]
(3) the original use of micro-electrodes for recording the electrical activity of single neurons [2]
(4) the initial application of computer technology for presenting experimental signals [3]
(5) the original utilization of computers to collect and analyze neurophysiological data [4]
(6) the first application of entirely digital technology for signal presentation, data collection and analysis [6][7]
(7) the first apparent application of non-linear modeling to behavioral function [1]

As we start the twenty-first century (and the third millennium) the dawn of a new scientific era is approaching, one that melds traditional experimental and descriptive methodology with the emerging power of computational and quantitative approaches. The current volume serves to define the shape, texture and scope of this important, new field of scientific inquiry, as well as to delineate its likely technological contribution to such fields as telephony, automatic speech recognition, hearing prostheses, speech synthesis, high-quality voice/audio reproduction and transmission.

The volume is divided into nine sections, each focusing on a specific topic germane to computational hearing.

The first section discusses computational approaches to the physiology of the auditory periphery, ranging from the cochlea (the chapter by Gebeshuber and Rattay) to the auditory nerve (Stankovic) and up through the ventral cochlear nucleus (Kalluri and Delgutte).

The second section applies computational approaches to two areas germane to processing in the cochlea. Irino and Unoki describe a model for spectral analysis based on gammachirp filtering, while Bruce and colleagues describe models for processing of sound by cochlear implants used in the profoundly hearing-impairing.

The third section of the book focuses on the localization of sound from a variety of different perspectives. Brungart describes a model for the perception of auditory distance, while Ito and Akagi apply sophisticated computational techniques to the problem of sound localization in general. Hartung and Sterbing, in their chapter, use physiological data to predict behavioral performance.

The following section discusses one particular model system — the echolocating bat — as a means of melding computational approaches to behavior and physiology. Wotton and colleagues consider the cues used for computation of elevation, while Müller and Schnitzler discuss the concept of "acoustic flow" in bats.

Section five focuses on pitch perception from the behavioral (Akeroyd and Summerfield) and physiological (Cai and colleagues) perspectives.

Temporal processing and periodicity analysis has been a controversial area of research for over a century. The sixth section focuses on several issues germane to this topic. The first two chapters (by Heil, and by Bleeck and Langner) focus on the importance of the waveform envelope (particularly at the beginning of a signal) for evoking neural excitation. Unoki and Akagi, in their chapter, model the perceptual phenomenon of "co-modulation masking release," a topic of intense behavioral research over the past two decades. Finally, Cariani discusses the importance of neural networks specialized for extracting timing cues in the perception of pitch and timbre.

The seventh section contains a paper by Miller and colleagues that examines the relationship between the thalamic and cortical regions of the auditory pathway, using dynamic signals to deduce the interconnections between these parts of the brain.

Auditory scene analysis, the ability to pick out specific "objects" from a background based on acoustic cues, has been a topic of keen investigation over the past decade. Baumann describes a model for identification and segregation of musical tones. Denham proposes a model of cortical activity (and inhibition) as the basis for some of the segregation ability observed in human listeners. Meyer and colleagues examine the ability of listeners to segregate two streams of speech as an example of auditory scene analysis.

Much of the interest in auditory computational models pertains to their utility for speech processing. The final section of the book examines three different approaches to speech processing using auditory models. Strope and Alwan are concerned with potential robustness of the speech signal in noisy environments based on pitch-relevant, amplitude-modulation cues. Tian and colleagues apply a model of the auditory periphery for robust speech recognition by computer, while Kawahara uses an auditory-inspired model to create realistic talking voices.

This volume is based on a NATO Advanced Study Institute, held at Il Ciocco, in the mountains of Tuscany, between July 1–12, 1998. Over a hundred scientists, representing 17 countries in Europe, North America and Asia, participated in the meeting (for further details, see http://www.icsi.berkeley.edu/real/comhear). The ASI's intent was to provide a rigorous, scientific overview of auditory function in concert with a critical examination of specific strategic issues that potentially hold the key to understanding how the brain portrays the world in terms of sound. As far we know, this was the first scientific meeting to specifically focus on melding computational approaches with the traditional venues of auditory neuroscience and psychoacoustics.

We would like to express our appreciation and gratitude to the ASI Faculty (Jont Allen, Jens Blauert, Ellen Covey, Dan Ellis, Ted Evans, Phil Green, Hynek Hermansky, Gerald Langner, Roy Patterson, Christoph Schreiner, Shihab Shamma, Jim Simmons, Quentin Summerfield, Marianne Vater, Jeff Winer and Eric Young) for their excellent lectures and to all of the other participants for helping to make the meeting an outstanding success.

We would also like to express our appreciation to NATO, which provided the lion's share of funding required to support the meeting through its Office of Scientific and Environmental Affairs. A particular debt of gratitude is owed to Dr. L. Veiga da Cunha, then-director of the Advanced Study Institute Program at NATO, for his help and guidance. We are also grateful to the Turkish and Portuguese divisions of NATO for providing additional financial support to defray the travel expenses of participants from their respective countries.

The ASI also received generous financial support from the (U.S.) Office of Naval Research and the (U.S.) Air Force Office of Scientific Research, for which we are deeply appreciative. We thank, in particular, Dr. Harold Hawkins of ONR and Dr. John Tangney of AFOSR for their efforts in supporting the ASI. We are also grateful for the support of the (U.S.) National Science Foundation which subsidized the travel of several of the younger American scientists. We thank Ms. Rosa Knox and Dr. Robert Metcalfe of that foundation's Graduate Education and Research Development office for their support and assistance. We also thank Barbara and Tilo Kester of International Transfer of Science and Technology (Brussels) who work closely with NATO to provide the infrastructure support for Advanced Study Institutes. The Kesters suggested Il Ciocco in Tuscany as the site most ideal for the purposes of the ASI and helped to coordinate our initial contacts with the hotel and conference center. We are especially grateful to Bruno Gianassi and his staff at Il Ciocco, who continually went beyond the call of duty to insure that everything ran smoothly during the course of the meeting.

Finally, we would like to express our deepest appreciation to the authors for taking the time to prepare their chapters for this volume, as well to thank them for their patience and understanding during the lengthy preparation of the book.

References

[1] Chistovich, L. A., Kozhevnikov, V. A. and Alyakrinskii, V. V. *Speech, Articulation and Perception*. Nauka: Acad. Sci. U. S. S. R., 1965.

[2] Galambos, R. and Davis, H. "The response of single auditory-nerve fibers in acoustic stimulation." *J. Neurophysiol.,* 6: 39–58, 1943.

[3] Goldstein, M. H. Jr., Kiang, N. Y. S. and Brown, R. M. "Responses of the auditory cortex to repetitive acoustic stimuli." *J. Acoust. Soc. Am.,* 31: 356–364, 1959.

[4] Kiang, N. Y.-S. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MA: M.I.T. Press, 1965.

[5] Ohm, G. S. "Über die definition des Tones, nebst daran geknupfter Theorie der Sirene und ahnlicher Tonbildener Vorrichtungen." *Ann. D. Phys.,* 59: 497–565, 1843.

[6] Rhode, W. S. "Neurophysiological techniques and the minicomputer." In *Use of Minicomputers in Research on Sensory and Information Processing*, M. S. Mayzner and T. R. Dolan (eds.), Hilldale, NJ: Lawrence Erlbaum, pp. 229–260, 1978.

[7] Rhode, W. S. and Olsen, R. *Digital Stimulus System for Auditory Neurophysiology*. Technical Report, Department of Neurophysiology, University of Wisconsin, Madison, WI, 1976.

[8] Wever, E. G. and Bray, C. W. "Action currents in the auditory nerve in response to acoustical stimulation." *Proc. Nat. Acad. Sci.,* 16: 344–350, 1930.

Steven Greenberg and Malcolm Slaney
May, 2001