Audio Bandwidth Extension

Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design. E. Larsen and R. M. Aarts
 ISBN 0-470-85864-8

Audio Bandwidth Extension

Application of Psychoacoustics, Signal Processing and Loudspeaker Design

Erik Larsen
MIT, Speech and Hearing Bioscience and Technology, USA

Ronald M. Aarts
Philips Research Laboratories, The Netherlands

John Wiley & Sons, Ltd

Copyright © 2004 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-85864-8

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by TJ International, Padstow, Cornwall
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Preface

Bandwidth extension (BWE) refers to methods that increase the frequency spectrum, or bandwidth, of electronic signals. Such frequency extension is desirable if at some point the frequency content of the signal has been reduced, as can happen, for example, during recording, transmission (including storage), or reproduction, mostly because of economical constraints. In this text, we limit the discussion to audio applications, which include music and speech. Most of the BWE methods heavily use signal processing – in fact, it is almost a premise that BWE is a signal-processing tool for achieving what is otherwise physically not possible. As Chapter 4 shows, a combination of mechanical engineering and signal processing can lead to interesting results as well.

BWE is a field that has seen increasing attention in recent years. Although some work was done in the early years of the twentieth century, a much more systematic and large-scale approach did not occur until recently. BWE for speech is the most mature area in this field, as the primary application (telephony) has existed for a long time. It is the objective of this book to gather most of the recent work into a single volume and present a coherent framework to the reader. It is the first time an entire book has been devoted to BWE theory, applications, and algorithms. It is intended as a broad introduction to BWE topics, but also discusses in detail various applications, thereby including material from scientific and patent literature, and also presents some previously unpublished work. The latter can be found in Sec. 2.4 (BWE using frequency tracking), most of Chapter 3 (low-frequency physical BWE systems), and Sec. 5.6 (high-frequency BWE using instantaneous compression).

Bandwidth reduction implies a decrease in perceptual quality, and therefore BWE algorithms are employed as tools to enhance the perceived quality of reproduced sound. In most cases, BWE methods are therefore post-processing algorithms, occurring just before sound reproduction, and the processing aims to compensate for the limited bandwidth that is available in a prior part of the chain. Sometimes, however, bandwidth reduction is actually desirable, for example, to enhance the coding efficiency of perceptual audio codecs. With a little additional complexity and data storage, required bit rates can be drastically reduced while maintaining subjective audio quality.

The main application areas discussed are bass enhancement for sound reproduction (Chapters 2, 3, and 4), high-frequency enhancement for general audio (Chapter 5) and speech (Chapter 6) applications. We include a short discussion on how BWE can be used as a very effective noise-abatement technology (Chapter 7), and present an overview of BWE patents (Chapter 8). Chapter 6, on BWE for speech, is contributed entirely by Peter Jax, who did most of the presented work as part of his doctoral dissertation [128].

In Appendix A, we present a brief overview of univariate and multidimensional scaling, which can be very useful techniques for analyzing the outcome of subjective experiments; an application example can be found in Sec. 2.5.

Although the BWE algorithms for the various application areas have many similarities on a conceptual level, there are interesting differences. For example, BWE to enhance bass reproduction on small loudspeakers, making use of the psychoacoustic effect of the 'missing fundamental', is largely focused on devising and analysing non-linear functions that are used to generate harmonics of very low-frequency signals. The goal of this technology is to allow a small loudspeaker to be used, giving a good percept of the entire audible bass range (down to 20 Hz), while the loudspeaker may only be able to physically radiate sound down to 100 Hz. A completely different approach is taken in Chapter 4, where inherently inefficient small loudspeakers are mechanically transformed into units that are highly efficient at only one frequency. Bass sounds from an entire bass frequency band are then mapped to the particular frequency at which the loudspeaker can radiate a considerable amount of energy. Thus, the problems of small loudspeakers at low frequencies can be ameliorated with different BWE methods. As another example, high-frequency enhancement of band-limited signals can be handled in two fundamentally different ways. In the first case, nothing is known about the missing high-frequency band, and one must resort to general procedures for recreating a 'reasonable' signal. The emphasis is again on devising proper non-linear functions, the output signals of which are then filtered to have a spectral envelope that is 'reasonable'. On the other hand, if some information is available about the missing high frequencies, the emphasis shifts to modeling the high-frequency spectral envelope as accurately as possible, as it is known that this is the dominating factor in achieving a high-quality signal. Both these approaches are discussed in Chapter 5. Chapter 6 deals with high-frequency enhancement as well, but for a very particular application, namely, telephonic speech. This application also demands an algorithm that works without information about the missing high frequencies, but because the signal is restricted to speech (and the band limitation is very well defined), a very specialized algorithm can be developed, which works well on speech (but not on other signals). It appears that the available low-frequency band contains information on the spectral envelope of the high-frequency band. Note that a considerable portion of what is presented in Chapter 2 (low-frequency psychoacoustic BWE systems) also applies to material of Chapter 3 (low-frequency physical BWE systems), and to a lesser degree, also to material of Chapter 5 (high-frequency BWE for audio). Therefore, Chapter 2 is considerably larger in size than the latter two chapters and much cross-referencing will be used to avoid repetition.

In most of the work that we present, we have tried to justify the approaches by considering psychoacoustical models of auditory perception. Because this may not be familiar to all readers, some psychoacoustics is reviewed in Chapter 1. This chapter is titled 'From physics to psychophysics' as we have presented a little background in all of the required disciplines: signal processing, statistics of audio signals, loudspeakers, and psychoacoustics. Of course, these background materials cover only the basic concepts that would be helpful for understanding the BWE topics in this book and are definitely not sufficient to cover all that may be of interest. Therefore, references are given, as has been done throughout the book, so that much of the relevant literature (both for background as well as for more specific BWE material) can easily be found.

As always, the work that was done and presented in this book, as well as the writing itself, is the product of the influence and support of many people. A special acknowledgement goes to Dr Janssen for valuable mathematical contributions that had considerable impact on Secs. 1.3.3.2, 2.4.2, 2.6, and 5.6.2. Of the others who contributed, we can only mention a few names here: Arie Kaizer for introducing us to electro-acoustics; Erik Druyvesteyn for always being enthusiastic; Paul Boers, Erik van der Tol, Okke Ouweltjes, and Daniël Schobben for valuable support; Cathy Polisset, Stefan Willems, and Gerrit DePoortere for development work; John Vanderkooy for the cooperation on high-Bl loudspeakers and many interesting discussions; Nathan Thomas and Michael Danessis from Salford University for their help; Ronald van Rijn for proofreading the manuscript; and Jeannie and Doortje for being patient, supportive, and for just being there.

E. Larsen and R. M. Aarts
Champaign (IL, USA) & Geldrop (The Netherlands)