John A. Sørensen

High-Fidelity Multichannel Audio Coding, by Dai Tracy Yang, Chris Kyriakakis, and C.-C. Jay Kuo. EURASIP Book Series on Signal Processing and Communications, vol. 1. Hindawi Publishing, 2004, 252 pp., ISBN 977-5945-08-9 (hardbound).

The aim of this book is to present advanced audio-coding principles at the senior undergraduate or graduate level and extend these principles into innovative audio-coding methods. The book is organized into three parts, the first of which covers the necessary fundamentals of digital audio codecs, digital signal processing, multichannel audio, signal quantization and coding methods, psychoacoustics, and objective and subjective quality measures for digital audio codecs. The second part presents MPEG-4 audio-coding tools and advanced audio-coding tools. Finally, the third part presents a selection of new coding methods, developed by the authors, that expand the present audio-coding system's quality and functionality.

The book is organized into 12 chapters, accompanied by a bibliography with 126 references and a four-page index register.

The text introduces digital audio concepts in Chapter 1, "Introduction to Digital Audio." Here the structure of a digital audio codec is presented, the goal of codec design is outlined, and potential sources of coding errors are identified. The fundamentals of signal sampling are described, together with a review of the history of multichannel audio, including perceptual cues and surround sound systems and standards. Chapter 2, "Quantization," then describes the scalar uniform and nonuniform quantization process, the vector quantization and the family of nearest neighbor quantization algorithms, and the problem of optimal bit allocation. This is followed by Chapter 3, "Entropy Coding," which presents a selection of Huffman coding algorithms, including an adaptive coding algorithm. This chapter also presents arithmetic coding, including adaptive arithmetic coding and QM coding. The latter is an adaptive coding algorithm, a successor of arithmetic coding.

In Chapter 4, "Introduction to Psychoacoustics," the fundamentals of sound measurements, sound pressure level, loudness, and hearing threshold in quiet are presented. This chapter presents the human hearing attributes described by the properties of masking in frequency and time. The first portion of the book concludes with Chapter 5, "Quality Measurement of Perceptual Audio Codecs," which introduces experimental design methods for subjective assessment of systems with small impairments. It describes how to construct the experimental design for quality assessment using a panel of listeners. It also delineates which reproducing devices, listening conditions, and data analysis methods should be used.

The second part of the text is introduced by Chapter 6, "MPEG-4 Audio Coding Tools." This chapter starts with an overview of MPEG-4 as a general standard for advanced sound compression, synthesis, manipulation, or playback. This standard also enables new opportunities in object-based audio coding, interactive presentation, and dynamic soundtracks. Then follows a presentation of the MPEG-4 audio-coding capabilities, consisting of speech, audio, synthesis, composition, and scalability tools. MPEG-4 audio coding is then presented in more detail, including the issues of error robustness, low-delay audio coding, fine granularity scalability, parametric audio coding, CELP silence compression, environmental spatialization, back channel, and audio transport stream. The general introduction to MPEG-4 is then followed by Chapter 7, "MPEG Advanced Audio Coding." MPEG Advanced Audio Coding (ACC) is currently the most powerful member of the MPEG family for high-quality, multichannel, digital audio compression. MPEG AAC supports a wide range of applications, from low-bit-rate Internet audio to multichannel broadcasting services. Initially, MPEG-2 AAC is presented and verified to achieve near-transparent subjective audio quality at a bit rate of 256 to 320 kb/s for five channels. The AAC encoder and decoder are described, including the forward modified discrete cosine transform (MDCT) used in the encoder and the inverse MDCT used in the decoder. To enable the encoder control of the fine temporal structure of the quantization noise, a temporal noise shaping (TNS) is used. Different methods for stereo coding are then described, including joint stereo coding, M/S stereo coding, and intensity stereo coding. This is followed by a description of the quantization and coding in the MPEG-2 AAC standard. This includes nonuniform quantization, coding of quantized spectral values, noise shaping, and iterative processes for optimal quantization. Having presented MPEG-2 AAC, the text provides a description of the MPEG-4 additions to AAC for improving the coding efficiency. Additions include perceptual noise substitution (PNS), long-term prediction, and transform domain weighted interleave vector quantization (TwinVQ). Chapter 7 offers a detailed description of MPEG-4 scalable audio coding, including

(continued on page 153)

1053-5888/05/$20.00©2005IEEE

 continued from page 150

large-step scalable audio-coding and bit-sliced arithmetic coding; the latter is an alternative method used when fine-grain scalability is required.

The third part of the book begins with Chapter 8, "Introduction to New Audio-Coding Tools." Here the authors describe their work on further improvements of multichannel audio coding methods. The improvements are obtained by combining several approaches, including the exploitation of interchannel redundancy removal approaches; by further work on audio concealment and channel transmission strategies for heterogeneous networks; and by exploiting quantization efficiency for the adaptive Karhunen-Loève transform (KLT). A novel, progressive, syntax-rich multichannel audio compression algorithm is proposed. The algorithm consists of a subband selection strategy, layered-coefficient coding, multiple context lossless coding, and three user-defined profiles. Finally, the authors propose a new error-resilient scalable audio-coding algorithm for WCDMA channels.

Chapter 9, "Interchannel Redundancy Removal and Channel-Scalable Decoding," discusses evidence for interchannel decorrelation, energy compaction effect, frequency-domain versus time-domain KLT transform, and temporal adaptive KLT. Eigen channel coding and transmission, as well as the modified AAC with KLT (MAACKLT) encoder, are then described. These are presented together with experimental results and comparison with MPEG AAC using lossy channels where packets are dropped in a coded bit stream. It is demonstrated that MAACKLT exhibits improved objective and subjective performance at the typical low bit rate of 64 kb/s per channel, at a similar complexity for encoder and decoder, compared to AAC.

Chapter 10, "Adaptive Karhunen-Loève Transform and Its Quantization Efficiency," compares scalar and vector quantization techniques for the KLT coefficients. Chapter 11, "Progressive Syntax-Rich Multichannel Audio Codec," presents a scalable compression algorithm, PSMAC, that is able to transmit and decode the bit stream with a bit rate that can be adapted to a dynamically varying environment. The encoder uses successive approximation quantization and a context-based QM coder. Implementation issues, such as frame, subband, and channel skipping, and determination of the mask-to-noise ratio (MNR) are discussed. The chapter concludes by describing the subjective listening test. Finally, in Chapter 12, "Error-Resilient Scalable Audio Codec Design," a coder-targeted WCDMA channels presented. The chapter starts by describing the characteristics of a WCDMA channel. Then, the layered coding structure is described together with error-resilient design, adaptive segmentation, frequency interleaving, bit-stream architecture, and error-control strategy. The algorithm is tested experimentally in a single channel.

This book is highly valuable to audio-coding researchers as well as to those seeking background material within this area. SP