AES

OURNAL OF THE AUDIO ENGINEERING SOCIETY
UDIO/ACOUSTICS/APPLICATIONS
olume 42 Number 10 1994 October

In this issue...

Perceptive Coding Standard

Diffuse Reflection Models

Noise Reduction

Virtual Voice Images

Draft Standard:
Listening Tests

Features...

Review of Society's Sustaining Members

AUDIO
A
ES
®

REGIONAL OFFICES

Europe
AES Europe Region
Zevenbunderslaan 142/9
B-1190 Brussels, Belgium
Tel: 32-2-345-7971
Fax: 32-2-345-3419

United Kingdom
British Section
Audio Engineering Society Ltd.
P. O. Box 645
Slough, SL1 8BJ
United Kingdom
Tel: Burnham 44 (0) 628 663725
Fax: 44 (0) 628 667002

Japan
Japan Section
c/o Japan Audio Society, Mori Building
14-34, 1-chome, Jingumae, Shibuya-ku,
Tokyo 158, Japan
Tel: 81-3-3403-6649
Fax: 81-3-3404-6549

AES REGIONS AND SECTIONS

Eastern Region
Sections: Atlanta, Boston, District of Columbia, New York, Philadelphia, Toronto.
Student Sections: Duquesne University, Fredonia, Full Sail Center for Recording Arts, McGill University, Pennsylvania State University, University of Massachusetts-Lowell, University of Miami.

Central Region
Sections: Chicago, Detroit, Indianapolis, Kansas City, Nashville, St. Louis, Upper Midwest, West Michigan.
Student Sections: Belmont University, Hutchinson Technical College, Lawrence Technological University, Michigan Technological University, Middle Tennessee State University, Ohio University.

Western Region
Sections: Albuquerque/Santa Fe, Los Angeles, Pacific Northwest, Portland, San Diego, San Francisco, Southern Arizona, Utah, Vancouver.
Student Sections: American River, California State University–Chico, Cogswell Polytechnical College, Denver, Long Beach City College, Loyola Marymount University, San Francisco State University, University of Southern California.

Europe Region
Sections: Austrian, Belgian, Belgrade, British, Bulgarian, Central German, Croatian, Czech, Danish, Finnish, French, Greek, Hungarian, Iberian, Israel, Italian, Netherlands, North German, Norwegian, Polish, Romanian, Russian, South German, Slovakian Republic, Slovenian, Swedish, Swiss.

International Region
Sections: Adelaide, Japan, Melbourne, Mexican, Philippine, Sydney, Venezuelan.

PURPOSE
The Audio Engineering Society is organized for the purpose of: uniting persons performing professional services in the audio engineering field and its allied arts; collecting, collating, and disseminating scientific knowledge in the field of audio engineering and its allied arts; advancing such science in both theoretical and practical applications; and preparing, publishing, and distributing literature and periodicals relative to the foregoing purposes and policies.

MEMBERSHIP
Individuals who are interested in audio engineering may become members of the Society. Applications are considered by the Admissions Committee.

Grades and annual dues are as follows: Full member, $65.00; Associate member, $65.00; Student member, $35.00. A membership application form may be obtained from headquarters.

Sustaining memberships are available to persons, corporations, or organizations who wish to support the Society.

AES Journal of the Audio Engineering Society

(ISSN 0004-7554), Volume 42, Number 10, 1994 October
Published monthly, except January/February and July/August when published bi-monthly, by the Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA, Telephone 212-661-2355. Fax 212-682-0477. Second-class postage paid at New York, New York, and at an additional mailing office. Postmaster: Send address corrections to Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520.

SUBSCRIPTIONS
The *Journal* is available by subscription. Annual rates are $125.00 surface mail, $170.00 air mail. For further information, contact AES Headquarters.

BACK ISSUES
Selected back issues are available: From Vol. 1 (1953) through Vol. 12 (1964), $10.00 per issue (members), $15.00 (nonmembers); Vol. 13 (1965) to present, $6.00 per issue (members), $11.00 (nonmembers). For information, contact AES Headquarters office.

MICROFILM
Copies of Vol. 19, No. 1 (1971 January) to the present edition are available on microfilm from University Microfilms International, 300 North Zeeb Rd., Ann Arbor, MI 48106, USA.

ADVERTISING
Contact the AES Editorial office.

ISO-MPEG-1 Audio: A Generic Standard for Coding of High-Quality Digital Audio*

KARLHEINZ BRANDENBURG, *AES Fellow*

FhG-IIS, Erlangen, Germany

AND

GERHARD STOLL

Institut für Rundfunktechnik, Munich, Germany

Coauthors:
Yves-François Dehéry, C.C.E.T.T, France
James D. Johnston, AES Member, AT&T, Bell Laboratories, USA
Leon v.d. Kerkhof, Philips, Netherlands
Ernst F. Schröder, AES Member, Thomson Consumer Electronics, Germany

The standardization body ISO/IEC/JTC1/SC29/WG11 (Moving Pictures Expert Group, MPEG) was drafting a standard for compressing the high bit rate of moving pictures and associated audio down to 1.5 Mbit/s. The audio part of the proposed standard is described. Three layers of the audio coding scheme with increasing complexity and performance were defined. These layers were developed in collaboration mainly with AT&T, CCETT, FhG/University of Erlangen, Philips, IRT, and Thomson Consumer Electronics. The generic coding system is suitable for different applications, such as storage on inexpensive storage media or transmission over channels with limited capacity (such as digital audio broadcasting or ISDN audio transmission).

0 INTRODUCTION

The necessity to specify a generic video and audio coding scheme for many applications dealing with digitally coded video and audio and requiring low data rates has led the ISO/IEC standardization body to establish the ISO/IEC JTC1/SC29/WG11, called MPEG (Moving Pictures Experts Group). This group had the task to compare and assess several digital audio low-bit-rate coding techniques in order to develop an international standard for the coded representation of moving pictures, associated audio, and their combination when used for storage and retrieval on digital storage media (DSM). The DSM targeted by MPEG include CD-ROM, DAT, magneto-optical disks, and computer disks, and it is expected that MPEG-based bit-rate reduction techniques will be used in a variety of communication channels such as ISDN and local area networks and in broadcasting applications. The international standard ISO/IEC 11172 "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s" was finalized in November 1992 and consists of three parts: system, video, and audio [1]. The system part (11172-1) deals with synchronization and multiplexing of audio–visual information, whereas the video (11172-2) and audio (11172-3) parts address the video and the audio bit-rate reduction techniques, respectively. This standard is also known as the MPEG-1 standard.

to MPEG-1. The main aspects are high quality of five (+1) audio channels, low bit rate and backward compatibility—the key to insuring that existing 2-channel decoders will still be able to decode compatible stereo information from five (+1) multichannel signals.

This standard, which is expected in November 1994, is based on standards and recommendations from international organizations such as ITU-R, SMPTE, and EBU. International standardization bodies will insure the highest audio signal quality by extensive testing. For audio reproduction the loudspeaker positions left, center, right, left and right surround are used, according to the 3/2-standard.

1 STANDARDIZATION AND QUALITY ASSESSMENTS WITHIN MPEG-1 AUDIO

Since 1988 ISO/MPEG has been undertaking the standardization of compression techniques for video and associated audio. The main topic for standardization in MPEG was video coding together with audio coding for DSM. On the other hand the audio coding standard developed by this group was the first international standard in the field of digital audio compression and is expected to be followed in different applications. Beside several subgroups such as video, system, test, implementation, requirement, and DSM, the audio subgroup of MPEG had the responsibility for developing a standard for coding of PCM audio signals with sampling rates of 32, 44.1, and 48 kHz at bit rates in a range of 32–192 kbit/s per mono and 64–384 kbit/s per stereo audio channel. The operating modes are

- Single channel
- Dual channel, like bilingual
- Stereo
- Joint stereo (combined coding of left and right channels of a stereophonic audio program)

Table 1 gives a short general view of the milestones of the MPEG-AUDIO group. This group asked for proposals for the audio coding standard in mid-1989, and 14 proposals were submitted for this purpose. The original proposals were grouped into four clusters according to algorithmic similarities. The clustered candidate algorithms were called ASPEC, ATAC, MUSICAM, and SB/ADPCM.

A number of subjective tests were performed [2]–[4] since mid-1990 to assess the audio quality of the ISO/MPEG/Audio coding standard. During this time period several improvements have been made to meet the present audio quality. The important milestones in the development of the standard have been the official tests organized by the Swedish Broadcasting Corporation in Stockholm under the auspices of ISO and EBU. In July 1990 large listening tests and objective evaluations, such as basic audio quality at different bit rates, sensitivity to transmission bit errors, encoder and decoder complexity,

Both the ASPEC and MUSICAM proposals have shown a very high subjective quality at bit rates of about 100 kbit/s per channel.

Due to the result that the proposals of the ASPEC and MUSICAM groups have been subjectively nearly equally rated, and were judged relatively close in their overall performance, the official decision was as follows [5]:

> . . . the MPEG standardization committee decided to approve a collaborative development of the draft audio coding standard between the ASPEC and MUSICAM groups, because the ASPEC codec was slightly superior with respect to the audio quality, especially for lower bit rates (64 kbit/s/channel), and the MUSICAM codec was slightly superior with respect to implementation complexity and decoding delay. The decision was that MUSICAM should be the basis for the low-complexity first layer, and algorithmic refinements including contributions of ASPEC should be used in the subsequent layers.

Table 1. Milestones of ISO/MPEG-Audio group during the development of audio part of the International Standard 11172.

Date	Activities
1988 December	First audio meeting in Hanover
1989 January to 1990 March	Preparation of tests
1989 May	Determining requirements and weighting procedure
1989 June	Proposal of 14 algorithms to be tested
1989 October	Clustering of proponents into four groups
1989 December	Detailed description of four clustered proposals: ASPEC, ATAC, MUSICAM, and SB-ADPCM
1990 May	Exchange of tapes with coded audio sequences between four clusters
1990 June	Subjective and objective tests at SR, Stockholm
1990 August	Presentation of results and decision to follow a layer concept
1990 December	First draft of part 3, "Audio Coding" of International Standard ISO 11172 was prepared.
1991 May	Verification of three layers by subjective testing, again at SR in Stockholm
1991 June	Layers I and II are frozen; Layer III and 'joint stereo coding' are still under discussion
1991 November	Second verification of Layer III and first checking of Joint Stereo Coding by subjective testing at NDR in Hanover
1991 December	Draft of International Standard (DIS) ready for balloting at national standardization bodies
1992 November	International Standard ISO/IEC 11172-3 accepted by

A three-layer coding algorithm has been defined. These three layers were tested again in April 1991 by the Swedish Broadcasting Corporation [3], and a last verification test for the very low bit rate of 64 kbit/s/channel and "joint stereo coding" was carried out by the University of Hanover under the auspices of NDR in November 1991 [4]. In November 1991 the final proposal, consisting of three modes of operation called "Layers," was adapted by ISO/MPEG [6].

2 BASIC STRUCTURE OF A GENERIC AUDIO CODING SCHEME USING PERCEPTUAL CRITERIA

The basic structure of a perceptual audio coding scheme is shown in Fig. 1.

1) A time–frequency mapping (filter bank) is used to decompose the input signal into subsampled spectral components. Depending on the filter bank used, these are called subband values or frequency lines.

2) The output of this filter bank, or the output of a parallel transform, is used to calculate an estimate of the actual (time-dependent) masking threshold using rules known from psychoacoustics.

3) The subband samples or frequency lines are quantized and coded with the aim of keeping the noise, which is introduced by quantizing, below the masking threshold. Depending on the algorithm, this step is done in very different ways. The complexity varies from block companding to analysis-by-synthesis systems using additional noiseless compression.

4) A frame packing is used to assemble the bit stream, which typically consists of the quantized and coded mapped samples and some side information, such as bit allocation information.

Depending on the focus on either low frequency resolution together with high time resolution or high frequency resolution which leads to only limited time resolution, the systems are usually called subband coders or transform coders.

2.1 Filter Banks

The following list provides a short overview over the most common filter banks used for coding of high-quality audio signals:

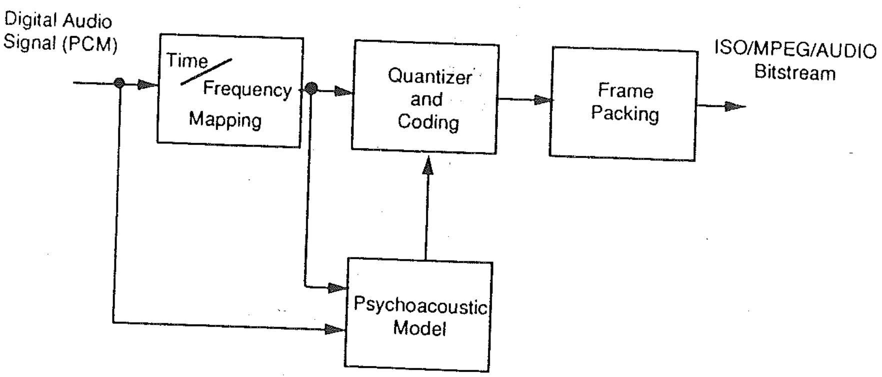
1) *QMF-Tree Filter Banks*: Different frequency resolution at different frequencies is possible. Typical QMF-tree filter banks use from 4 to 24 bands. The computational complexity is high.

2) *Polyphase Filter Banks*: These are equally spaced filter banks which combine the filter design flexibility of generalized QMF banks with low computational complexity [7]. It is possible to design the prototype filter in a way that achieves both good frequency resolution (stop-band attenuation better than 96 dB) and good control of possible time-domain artifacts. A polyphase filter bank using 32 bands is used for Layers I and II of the ISO/MPEG audio coder.

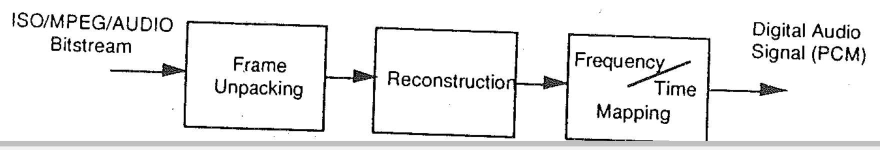
3) *DFT, DCT with Sine-Taper Window*: These were the first transforms used in transform coding of audio signals. They implement equally spaced filter banks with 128–512 bands at a low computational complexity. They do not provide critical sampling, that is, the number of time–frequency components is greater than the number of time samples represented by one block length. Another disadvantage of these transforms are possible blocking artifacts.

4) *Modified Discrete Cosine Transform (MDCT, using time-domain aliasing cancellation as proposed in [8])*: This transform combines critical sampling with a good frequency resolution provided by a sine window (compared to a sine-taper window) and the computational efficiency of a fast FFT-like algorithm. Typically 128–512 equally spaced bands are used.

5) *Hybrid Structures* (such as polyphase and MDCT): Using hybrid structures as first proposed in [9] it is possible to combine different frequency resolutions at different frequencies with moderate implementation complexity. A hybrid system consisting of a polyphase filter bank and an MDCT is used in Layer III.



(a)



Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.