

**Implementation of
EN ISO/IEC 13818-3:1996**

Information technology — Generic coding of moving pictures and associated audio information —

Part 3: Audio

The European Standard EN ISO/IEC 13818-3:1996 has the status of a British Standard

Committees responsible for this British Standard

The preparation of this British Standard was entrusted to Technical Committee IST/37, Coding of pictures, audio, multimedia and hypermedia information, upon which the following bodies were represented:

- British Broadcasting Corporation
- British Telecommunications plc
- Elysium Ltd.
- GEC Plessey Semiconductors
- IBM United Kingdom Ltd.
- Institution of Electrical Engineers
- International Computers Limited
- National Health Services
- National Transcommunications Ltd.
- PICON
- Philips Research Laboratories
- Pioneer Digital Design Centre Ltd.
- SGS Thomson Microelectronics Ltd.
- UKERMA

This British Standard, having been prepared under the direction of the Information Systems Technology Assembly, was published under the authority of the Standards Board and comes into effect on 15 October 1995

© BSI 05-2000

The following BSI references relate to the work on this standard:
Committee reference IST/37
Draft for comment 94/645298 DC

ISBN 0 580 24811 9

Amendments issued since publication

Amd. No.	Date	Comments
9285	December 1996	Indicated by a sideline in the margin

Contents

	Page
Committees responsible	Inside front cover
National foreword	ii
Foreword	2
Foreword	iii
Text of ISO/IEC 13818-3	1

National foreword

This British Standard has been prepared by Technical Committee IST/37 and is the English language version of EN ISO/IEC 13818-3 — *Information technology — Generic of moving coding pictures associated audio information — Part 3: Audio*, published by the European Committee for Standardization (CEN). It is identical with ISO/IEC 13818-3:1995 published by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC).

This British Standard is published under the direction of the Information Systems Technology Assembly whose Technical Committee IST/37 has the responsibility to:

- aid enquirers to understand the text;
- present to the responsible international committee any enquiries on interpretation, or proposals for change, and keep UK interests informed;
- monitor related international and European developments and promulgate them in the UK.

NOTE International and European Standards, as well as overseas standards, are available from Customer Services, BSI, 389 Chiswick High Road, London W4 4AL.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, the EN ISO/IEC title page, page 2, the ISO/IEC title page, pages ii to iv, pages 1 to 110 and a back cover.

This standard has been updated (see copyright date) and may have had amendments incorporated. This will be indicated in the amendment table on the inside front cover.

English version

Information technology — Generic of moving coding pictures associated audio information — Part 3: Audio

(ISO/IEC 13818-3:1995)

Technologies de l'information — Codage générique des images animées et des informations sonores associées — Partie 3: Son (ISO/IEC 13818-3:1995)

Informationstechnik — Codierung von bewegten Bildern und damit verbundenen Toninformationen — Teil 3: Audio (ISO/IEC 13818-3:1995)

This European Standard was approved by CEN on 1996-04-18. CEN members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the Central Secretariat or to any CEN member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN member into its own language and notified to the Central Secretariat has the same status as the official versions.

CEN members are the national standards bodies of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

CEN

European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Central Secretariat: rue de Stassart 36, B-1050 Brussels

Foreword

The text of the International Standard from ISO/IEC/JTC 1, Information technology, of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) has been taken over as a European Standard by the Technical Board of CEN.

This European Standard shall be given the status of a national standard, either by publication of an identical text or by endorsement, at the latest by November 1996, and conflicting national standards shall be withdrawn at the latest by November 1996.

According to CEN/CENELEC Internal Regulations, the national standards organizations of the following countries are bound to announce this European Standard: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom.

**Information technology — Generic coding
of moving pictures and associated audio
information —**

Part 3:
Audio

*Technologies de l'information — Codage générique des images animées
et des informations sonores associées —*

Partie 3: Son



Reference number
ISO/IEC 13818-3:1995(E)

Contents

	Page
Foreword	iii
Introduction	1
Section 1. General	
1.1 Scope	7
1.2 Normative References	7
Section 2. Technical elements	
2.1 Definitions	8
2.2 Symbols and abbreviations	21
2.3 Method of describing bit stream syntax	24
2.4 Requirements for Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies	25
2.5 Requirements for low bitrate coding of multichannel audio	31
Annex A (normative) Diagrams	70
Annex B (normative) Tables	72
Annex C (informative) The encoding process	76
Annex D (informative) Psychoacoustic models	86
Annex E (informative) List of patent holders	109
Figure A.1 — Structure of the ISO 13818-3 Layer II multichannel extension, backwards compatible with ISO 11172-3 Layer II	70
Figure A.2 — Structure of the ISO 13818-3 Layer II multichannel extension, using the ISO/IEC 11172-3 compatible bit stream as well as the extension bit stream	71
Table B.1 — Possible quantisation per subband, Layer II	72
Table B.2 — Layer III scalefactor bands	73
Table B.3 — Low-pass filter description	76
Table D.1a — Frequencies, critical band rates and absolute threshold	92
Table D.1b — Frequencies, critical band rates and absolute threshold	93
Table D.1c — Frequencies, critical band rates and absolute threshold	94
Table D.1d — Frequencies, critical band rates and absolute threshold	95
Table D.1e — Frequencies, critical band rates and absolute threshold	96
Table D.1f — Frequencies, critical band rates and absolute threshold	97
Table D.2a — Critical band boundaries	98
Table D.2b — Critical band boundaries	98
Table D.2c — Critical band boundaries	99
Table D.2d — Critical band boundaries	99
Table D.2e — Critical band boundaries	100
Table D.2f — Critical band boundaries	100
Table D.3.a — Sampling_frequency = 24 kHz long blocks	101
Table D.3.b — Sampling_frequency = 22,05 kHz long blocks	102
Table D.3.c — Sampling_frequency = 16 kHz long blocks	103
Table D.3.d — Sampling_frequency = 24 kHz short blocks	104
Table D.3.e — Sampling_frequency = 22,05 kHz short blocks	105
Table D.3.f — Sampling_frequency = 16 kHz short blocks	106
Table D.4 — Tables for converting threshold calculation partitions to scalefactor bands	107

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

International Standard ISO/IEC 13818-3 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 13818 consists of the following parts, under the general title *Information technology — Generic coding of moving pictures and associated audio information*:

- *Part 1: Systems*;
- *Part 2: Video*;
- *Part 3: Audio*.

Annex A and Annex B form an integral part of this part of ISO/IEC 13818. Annex C to Annex E are for information only.

Introduction

ISO/IEC 13818 was prepared by SC29/WG 11, also known as MPEG (Moving Pictures Expert Group). MPEG was formed in 1988 to establish a standard for the coded representation of moving pictures and associated audio stored on digital storage media.

ISO/IEC 13818 is published in three parts. Part 1 — systems — specifies the system coding layer of the standard. It defines a multiplexed structure for combining audio and video data and means of representing the timing information needed to replay synchronised sequences in real-time. Part 2 — video — specifies the coded representation of video data and the decoding process required to reconstruct pictures.

Part 3 — audio — specifies the coded representation of audio data and the decoding process required to decode audio signals.

0.1 Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies

In order to achieve better audio quality at very low bit rates (< 64 kbit/s per audio channel), in particular if compared with CCITT Standard G-722 performance, three additional sampling frequencies are provided for ISO/IEC 11172-3 layers I, II and III. The additional sampling frequencies are 16 kHz, 22,05 kHz and 24 kHz. This allows corresponding audio bandwidths of approximately 7,5 kHz, 10,3 kHz and 11,25 kHz. The syntax, semantics, and coding techniques of ISO/IEC 11172-3 are maintained except for a new definition of the sampling frequency field, the bitrate index field, and the bit allocation tables. These new definitions are valid if the ID bit in the ISO/IEC 11172-3 header equals zero. To obtain the best audio performance, the parameters of the psychoacoustic model used in the encoder have to be changed accordingly.

With these sampling frequencies, the duration of the audio frame corresponds to:

Layer	Sampling Frequency in kHz		
	16	22,05	24
I	24 ms	17,41.. ms	16 ms
II	72 ms	52,24.. ms	48 ms
III	36 ms	26,12.. ms	24 ms

0.2 Low bitrate coding of multichannel audio

0.2.1 Universal multichannel audio system

A standard on low bit rate coding for mono or stereo audio signals was established by MPEG-1 Audio in ISO/IEC 11172-3. This standard is applicable for carrying of high quality digital audio signals associated with or without picture information on storage media or transmission channels with limited capacity.

The ISO/IEC 11172-3 audio coding standard can be used together with both MPEG-1 and MPEG-2 Video as long as only two-channel stereo is required. MPEG-2 Audio (ISO/IEC 13818-3) provides the extension to 3/2 multichannel audio and an optional low frequency enhancement channel (LFE).

Multichannel audio systems provide enhanced stereophonic stereo performance compared to conventional two channel audio systems. It is recognised that improved presentation performance is desirable not only for applications with accompanying picture but also for audio-only applications. A universal and compatible multichannel audio system applicable to satellite or terrestrial television broadcasting, digital audio broadcasting (terrestrial and satellite), as well as other non-broadcasting media, e.g.,

CATV	Cable TV Distribution
CDAD	Cable Digital Audio Distribution
ENG	Electronic News Gathering (including Satellite News Gathering)
IPC	Interpersonal Communications (video conference, videophone, etc.)
ISM	Interactive Storage Media (optical disks, etc.)
NDB	Network Database Services (via ATM, etc.)
DSM	Digital Storage Media (digital VTR, etc.)
EC	Electronic Cinema

HTT	Home Television Theatre
ISDN	Integrated Services Digital Network

seems to be very attractive to the manufacturer, producer, and consumer.

This part of ISO/IEC 13818 describes an audio subband coding system called ISO/MPEG-Audio Multichannel, which can be used to transfer high quality digital multichannel and/or multilingual audio information on storage media or transmission channels with limited capacity. One of the basic features is the backwards compatibility to ISO/IEC 11172-3 coded mono, stereo or dual channel audio programmes. It is designed for use in different applications as considered by the ISO/MPEG audio group and the specialist groups TG 10/1, 10/2 and 10/3 of the ITU-R (previously CCIR).

0.2.2 Representation of multichannel audio

0.2.2.1 The 3/2-stereo plus LFE format

Regarding stereophonic presentation, specialist groups of ITU-R, SMPTE, and EBU recommend the use of an additional centre loudspeaker channel C and two surround loudspeaker channels LS and RS, augmenting the front left and right loudspeaker channels L and R. This reference audio format is referred to as “3/2-stereo” (3 front/2 surround loudspeaker channels) and requires the transmission of five appropriately formatted audio signals.

For audio accompanying picture applications (e.g. HDTV), the three front loudspeaker channels ensure sufficient directional stability and clarity of the picture related frontal images, according to the common practice in the cinema. The dominant benefit is the “stable centre”, which is guaranteed at any location of the listener and important for most of the dialogue.

Additionally, for audio-only applications, the 3/2-stereo format has been found to be an improvement over two-channel stereophony. The addition of one pair of surround loudspeaker channels allows improved realism of auditory ambience.

A low frequency enhancement channel (in this part of ISO/IEC 13818 called LFE channel) can, optionally, be added to any of these configurations. The purpose of this channel is to enable listeners to extend the low frequency content of the reproduced programme in terms of both frequency and level. In this way it is the same as the LFE channel proposed by the film industry for their digital sound systems.

The LFE channel should not be used for the entire low frequency content of the multichannel sound presentation. The LFE channel is optional at the receiver, and thus should only carry low frequency sound effects, which may have a high level. The LFE channel is not included in any dematrixing operation in the decoder. The sampling frequency of the LFE channel corresponds to the sampling frequency of the main channels, divided by a factor of 96. This provides 12 LFE samples within one audio frame. The LFE channel is capable of handling signals in the range from 15 Hz to 120 Hz.

0.2.2.2 Compatibility

Downwards compatibility.

A hierarchy of audio formats providing a lower number of loudspeaker channels and reduced presentation performance (down to 2/0-stereo or even mono) and a corresponding set of downwards mixing equations are recommended in ITU-R Recommendation 775: “Multichannel stereophonic audio system with and without accompanying picture”, November 1992. Alternative lower level audio formats which may be used in circumstances where economic or channel capacity constraints apply, are 3/1, 3/0, 2/2, 2/1, 2/0, and 1/0. Corresponding loudspeaker arrangements are 3/2, 3/1, 3/0, 2/2, 2/1, 2/0, and 1/0.

Backwards compatibility.

For several applications, the intention is to extend the existing 2/0-stereo sound system by transmitting additional audio channels (centre, surround) without making use of simulcast operation. This provision of backwards compatibility with existing receivers implies the use of compatibility matrices: the decoder of the previous generation must reproduce the two conventional basic stereo signals L_o/R_o , and the multichannel decoder produces the complete 3/2-stereo presentation $L'/C'/R'/LS'/RS'$ from the basic stereo signal and the extension signals.

It is recognised that backward compatibility may not be required for all applications of MPEG-2 Audio. Therefore, nonbackward compatible (NBC) audio coding systems free of the constraints of backward compatibility are being evaluated for optional use with the standard.

0.2.2.3 Multilingual capability

Particularly for HDTV applications, multichannel stereo performance and bilingual programmes or multilingual commentaries are required. This standard provides for alternative audio channel configurations in the five-channel sound system, for example a bilingual 2/0 stereo programme or one 2/0, 3/0 stereo sound plus accompanying services (e.g. “clean dialogue” for the hard-of-hearing, commentary for the visually impaired, multilingual commentary etc.). An important configuration is the reproduction of commentary dialogue (e.g. via centre loudspeaker) together with the common music/effect stereo downmix (examples are documentation film, sport reports).

0.2.3 Basic Parameters of the Multichannel Audio Coding System

The transmission of the five audio signals of a 3/2 sound system requires five transmission channels (although, in the context of bitrate reduced signals, these are not necessarily independent). In order that two of the transmitted signals can provide a stereo service on their own, the source sound signals are generally combined in a linear matrix prior to encoding. These combined signals (and their transmission channels) are identified by the notation T0, T1, T2, T3 and T4.

0.2.3.1 Compatibility with ISO/IEC 11172-3

Backwards and forwards compatibility with an ISO/IEC 11172-3 decoder is provided.

For a multichannel audio bit stream, backwards compatibility means, that an ISO/IEC 11172-3 audio decoder properly decodes the basic stereo information. The basic stereo information consists of a left and right channel that constitute an appropriate downmix of the audio information in all channels, or, optionally, the basic stereo information may consist only of the left and right channel of the multichannel audio configuration. Appropriate downmix equations are given by equation pairs (1) and (2), (3) and (4), and (5) and (6).

$$Lo = L + x * C + y * LS \quad (1)$$

$$Ro = R + x * C + z * RS \quad (2)$$

or

$$Lo = L \quad (3)$$

$$Ro = R \quad (4)$$

or

$$Lo = L + x * C - y * jS \quad (5)$$

$$Ro = R + x * C + y * jS \quad (6)$$

where jS is derived from LS and RS by calculation of the mono component, bandwidth limitation to the range 100–7 000 Hz, half Dolby^{®1)} B-type encoding, and 90 degrees phase shifting (Prologic^{®1)} surround matrixing).

Compatibility with existing surround sound decoders by use of equations (5) and (6) has not been verified at the time of printing of this part of ISO/IEC 13818.

Forwards compatibility means that an MPEG 2 multichannel audio decoder is able to decode properly an ISO/IEC 11172-3 audio bit stream.

The following combinations are possible:

Basic Lo, Ro Stereo	Multichannel Extension
Layer II	Layer II mc
Layer III	Layer III mc
Layer I	Layer II mc

This part of ISO/IEC 13818 describes the combinations of the basic Lo, Ro stereo of Layer I, II and III and the multichannel extension of Layer II mc and Layer III mc.

¹⁾ Dolby and Prologic are registered trademarks of Dolby Laboratories Licensing Corp.

The ISO/MPEG-Audio Multichannel system provides full compatibility with the ISO Standard 11172-3. This compatibility is realised by coding the basic stereo information in conformance with ISO/IEC 11172-3 and exploiting the ancillary data field of the ISO/IEC 11172-3 audio frame and an optional extension bit stream for the multichannel extension.

The complete ISO/IEC 11172-3 frame incorporates four different types of information:

- Header information within the first 32 bits of the ISO/IEC 11172-3 audio frame.
- Cyclic Redundancy Check (CRC), consisting of 16 bits, just after the header information (optional).
- Audio data, for Layer II consisting of bit allocation (BAL), scalefactor select information (SCFSI), scalefactors (SCF), and the subband samples.
- Ancillary data. Due to the large number of different applications which will use the ISO/IEC 11172-3 Standard, the length and usage of this field are not specified.

The variable length of the ancillary data field enables packing the complete extension information of the channels T2/T3/T4 into the first part of the ancillary data field. If the MC encoder does not use all of the ancillary data field for the multichannel extension information, the remaining part of the field can be used for other ancillary data.

The bit rate required for the multichannel extension information may vary on a frame by frame basis, depending on the sound signals. The overall bit rate may be increased above that provided for in ISO/IEC 11172-3 by the use of an optional extension bit stream. The maximum bit rate, including the extension bit stream, is given by the following table:

Sampling Frequency	Layer	Maximum Total Bit Rate
32 kHz	I	903 kbit/s
32 kHz	II	839 kbit/s
32 kHz	III	775 kbit/s
44.1 kHz	I	1 075 kbit/s
44.1 kHz	II	1 011 kbit/s
44.1 kHz	III	947 kbit/s
48 kHz	I	1 130 kbit/s
48 kHz	II	1 066 kbit/s
48 kHz	III	1 002 kbit/s

0.2.3.2 Audio Input/Output Format

Sampling frequencies: 48, 44.1 or 32 kHz

Quantisation: up to 24 bits/sample PCM resolution

The following combinations of audio channels can be applied as inputs to the audio encoder:

- a) Five channels, using the 3/2 configuration
L, C, R plus two surround channels LS, RS
- b) Five channels, using the 3/0 + 2/0 configuration
L, C, R of first programme plus L2, R2 of second programme
- c) Four channels, using the 3/1 configuration
L, C, R plus single surround channel S
- d) Four channels, using the 2/2 configuration
L, R plus two surround channels LS, RS
- e) Four channels, using the 2/0 + 2/0 configuration
L, R of first programme plus L2, R2 of second programme

- f) Three channels using the 3/0 configuration
L, C, R without surround
- g) Three channels using the 2/1 configuration
L, R with single surround channel S
- h) Two channels, using the 2/0 configuration
Stereo or dual channel mode (as in ISO/IEC 11172-3)
- i) One channel, using the 1/0 configuration
Single channel mode (as in ISO/IEC 11172-3)

The different combinations of audio input signals are encoded and transmitted within the up to five available transmission channels T0, T1, T2, T3, and T4, of which channels T0 and T1 are the two basic channels of ISO/IEC 11172-3 and convey the backwards compatible signals Lo and Ro. Transmission channels T2, T3, and T4 together form the multichannel extension information, which is compatibly transmitted within the ISO/IEC 11172-3 ancillary data field and an optional extension bit stream.

After multichannel decoding, the up to five audio channels are recovered and can then be presented in any convenient format at the choice of the listeners:

- a) Five channels, using the 3/2 configuration
Front: Left (L) and right (R) channels plus centre channel (C)
Surround: Left surround (LS) and right surround (RS)
- b) Four channels, using the 3/1 configuration
Front: Left (L) and right (R) channels plus centre channel (C)
Surround: Mono surround (S)
- c) Four channels, using the 2/2 configuration
Front: Left (L) and right (R) channel
Surround: Left surround (LS) and right surround (RS)
- d) Three channels, using the 2/1 configuration
Front: Left (L) and right (R) channels
Surround: Mono surround (S)
- e) Three channels using the 3/0 configuration
Front: Left (L) and right (R) channel plus centre channel (C)
Surround: No surround
- f) Two channels, using the 2/0 configuration
Front: Left (L) and right channel (R)
Surround: No surround
- g) One channel output, using the 1/0 configuration
Front: Mono channel (Mo)
Surround: No surround

A low frequency enhancement channel can, optionally, be added to any of these configurations.

Outputs may be required to provide discrete signals, or may be combined in accordance with downward mixing, or upwards conversion equations, as defined in ITU-R Recommendation 775.

0.2.3.3 Composite Coding Modes

Dynamic Transmission Channel Switching

In order to provide a better orthogonality between the two compatible signals T0 and T1, and the three additionally transmitted signals T2, T3 and T4, it is necessary to have flexibility in the choice of the channels T2, T3 and T4. ISO/IEC 13818-3 allows, independently for a number of frequency regions, the selection of a number of combinations of three out of the five signals L, C, R, LS, RS to be transmitted in T2, T3, T4.

Dynamic Crosstalk

According to a binaural hearing model, it is possible to determine the portion of the stereophonic signal which is irrelevant with respect to the spatial perception of the stereophonic presentation. The stereo-irrelevant signal components are not masked, but they do not contribute to the localisation of sound sources. They are ignored in the binaural processor of the human auditory system. Therefore, stereo-irrelevant components of any stereo signal (L, C, R, LS or RS) may be reproduced via any loudspeaker, or via several loudspeakers of the arrangement, without affecting the stereophonic impression. This can be done independently for a number of frequency regions.

Adaptive Multichannel Prediction

In order to make use of the statistical inter-channel dependencies, adaptive multichannel prediction is used for redundancy reduction. Instead of transmitting the actual signals in the transmission channels T2, T3, T4, the corresponding prediction error signals are transmitted. A predictor of up to 2nd order with delay compensation is used.

Phantom Coding of Centre

Due to the fact that the human auditory system uses only intensity cues of the audio signal for localisation at higher frequencies, it is possible to transmit the high frequency part of the centre channel in the front left and right channels, constituting a phantom source at the location of the centre loudspeaker.

0.2.3.4 Encoder and Decoder Parameters

Encoding and decoding:

similar to ISO/IEC 11172-3.

Coding modes:

3/2, 3/0 + 2/0, 3/1, 2/0 + 2/0, 3/0, 2/2, 2/1, 2/0, 1/0

second stereo programme,

up to 7 additional multilingual or commentary channels,

associated services

Subband filter transforms:

Number of subbands: 32

Sampling frequency: $F_s/32$

Bandwidth of subbands: $F_s/64$

Additional decomposition by MDCT (Layer III only):

Frequency Resolution: 6 or 18 components per subband

LFEC filter transform:

Number of LFECs: 1

Sampling frequency: $F_s/96$

Bandwidth of LFEC: 125 Hz

Dynamic range: more than 20 bits

Section 1. General

1.1 Scope

This part of ISO/IEC 13818 specifies the extension of ISO/IEC 11172-3 to lower sampling frequencies, the coded representation of multichannel and multilingual high quality audio for broadcasting, transmission and storage media, and the method for decoding of multichannel and multilingual high quality audio signals. The input of the encoder and the output of the decoder are compatible with existing PCM standards.

1.2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 13818. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO/IEC 13818 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 11172-3:1993, *Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s — Part 3: Audio*.

CCIR Recommendation 601-1:1990, *Encoding parameters of digital television for studios*.

CCIR Recommendation 648:1986, *Recording of audio signals*.

CCIR Recommendation 775:1992, *Multichannel stereophonic sound system with and without accompanying picture*.

CCIR Report 955-2:1990, *Sound broadcasting by satellite for portable and mobile receivers*, including Annex IV *Summary description of Advanced Digital System II*.

IEC 908:1987, *Compact disc digital audio system*.

IEEE Draft Standard P1180/D2:1990, *Specification for the implementation of 8x 8 inverse discrete cosine transform*.

Section 2. Technical elements

2.1 Definitions

For the purposes of ISO/IEC 13818, the following definitions apply. If specific to a part, this is noted in square brackets.

- 2.1.1**
ac coefficient [video]
any DCT coefficient for which the frequency in one or both dimensions is non-zero
- 2.1.2**
access unit [system]
in the case of compressed audio an access unit is an audio access unit. In the case of compressed video an access unit is the coded representation of a picture
- 2.1.3**
adaptive segmentation [audio]
a subdivision of the digital representation of an audio signal in variable segments of time
- 2.1.4**
adaptive bit allocation [audio]
the assignment of bits to subbands in a time and frequency varying fashion according to a psychoacoustic model
- 2.1.5**
adaptive multichannel prediction [audio]
a method of multichannel data reduction exploiting statistical inter-channel dependencies
- 2.1.6**
adaptive noise allocation [audio]
the assignment of coding noise to frequency bands in a time and frequency varying fashion according to a psychoacoustic model
- 2.1.7**
alias [audio]
mirrored signal component resulting from sub-Nyquist sampling
- 2.1.8**
analysis filterbank [audio]
filterbank in the encoder that transforms a broadband PCM audio signal into a set of subsampled subband samples
- 2.1.9**
ancillary data [audio]
part of the bit stream that might be used for transmission of ancillary data
- 2.1.10**
audio access unit [audio]
for Layers I and II, an audio access unit is defined as the smallest part of the encoded bit stream which can be decoded by itself, where decoded means “fully reconstructed sound”. For Layer III, an audio access unit is part of the bit stream that is decodable with the use of previously acquired main information
- 2.1.11**
audio buffer [audio]
a buffer in the system target decoder for storage of compressed audio data
- 2.1.12**
audio sequence [audio]
a non-interrupted series of audio frames in which the following parameters are not changed:
- ID
 - Layer

- Sampling Frequency
- For Layer I and II: Bitrate index

2.1.13**backward motion vector [video]**

a motion vector that is used for motion compensation from a reference picture at a later time in display order

2.1.14**bark [audio]**

unit of critical band rate. The Bark scale is a non-linear mapping of the frequency scale over the audio range closely corresponding with the frequency selectivity of the human ear across the band

2.1.15**bidirectionally predictive-coded picture; B-picture [video]**

a picture that is coded using motion compensated prediction from a past and/or future reference picture

2.1.16**bitrate**

the rate at which the compressed bit stream is delivered from the storage medium to the input of a decoder

2.1.17**block companding [audio]**

normalising of the digital representation of an audio signal within a certain time period

2.1.18**block [video]**

an 8-row by 8-column orthogonal block of pels

2.1.19**bound [audio]**

the lowest subband in which intensity stereo coding is used

2.1.20**byte**

sequence of 8-bits

2.1.21**byte aligned**

a bit in a coded bit stream is byte-aligned if its position is a multiple of 8-bits from the first bit in the stream

2.1.22**centre channel [audio]**

an audio presentation channel used to stabilise the central component of the frontal stereo image

2.1.23**channel**

a digital medium that stores or transports a CD 13818 bit stream

2.1.24**channel [audio]**

a sequence of data representing an audio signal being transported

2.1.25**chrominance (component) [video]**

a matrix, block or single pel representing one of the two colour difference signals related to the primary colours in the manner defined in CCIR Rec 601. The symbols used for the colour difference signals are Cr and Cb

2.1.26

coded audio bit stream [audio]

a coded representation of an audio signal as specified in this part of the CD

2.1.27

coded video bit stream [video]

a coded representation of a series of one or more pictures as specified in this CD

2.1.28

coded order [video]

the order in which the pictures are stored and decoded. This order is not necessarily the same as the display order

2.1.29

coded representation

a data element as represented in its encoded form

2.1.30

coding parameters [video]

the set of user-definable parameters that characterise a coded video bit stream. Bit streams are characterised by coding parameters. Decoders are characterised by the bit streams that they are capable of decoding

2.1.31

component [video]

a matrix, block or single pel from one of the three matrices (luminance and two chrominance) that make up a picture

2.1.32

compression

reduction in the number of bits used to represent an item of data

2.1.33

constant bitrate coded video [video]

a compressed video bit stream with a constant average bitrate

2.1.34

constant bitrate

operation where the bitrate is constant from start to finish of the compressed bit stream

2.1.35

constrained parameters [video]

the values of the set of coding parameters defined in **2.4.3.2** of ISO/IEC 11172-2

2.1.36

constrained system parameter stream (CSPS) [system]

an ISO/IEC 11172 multiplexed stream for which the constraints defined in **2.4.6** of ISO/IEC 11172-1 apply

2.1.37

CRC

cyclic redundancy check

2.1.38

critical band rate [audio]

psychoacoustic function of frequency. At a given audible frequency, it is proportional to the number of critical bands below that frequency. The units of the critical band rate scale are Barks

2.1.39

critical band [audio]

psychoacoustic measure in the spectral domain which corresponds to the frequency selectivity of the human ear. This selectivity is expressed in Bark

2.1.40**data element**

an item of data as represented before encoding and after decoding

2.1.41**dc-coefficient [video]**

the DCT coefficient for which the frequency is zero in both dimensions

2.1.42**dc-coded picture; D-picture [video]**

a picture that is coded using only information from itself. Of the DCT coefficients in the coded representation, only the dc-coefficients are present

2.1.43**DCT coefficient**

the amplitude of a specific cosine basis function

2.1.44**decoded stream**

the decoded reconstruction of a compressed bit stream

2.1.45**decoder input buffer [video]**

the first-in first-out (FIFO) buffer specified in the video buffering verifier

2.1.46**decoder input rate [video]**

the data rate specified in the video buffering verifier and encoded in the coded video bit stream

2.1.47**decoder**

an embodiment of a decoding process

2.1.48**decoding (process)**

the process defined in ISO/IEC 11172 that reads an input coded bit stream and produces decoded pictures or audio samples

2.1.49**decoding time-stamp; DTS [system]**

a field that may be present in a packet header that indicates the time that an access unit is decoded in the system target decoder

2.1.50**de-emphasis [audio]**

filtering applied to an audio signal after storage or transmission to undo a linear distortion due to emphasis

2.1.51**dequantisation [video]**

the process of rescaling the quantised DCT coefficients after their representation in the bit stream has been decoded and before they are presented to the inverse DCT

2.1.52**digital storage media; DSM**

a digital storage or transmission device or system

2.1.53**discrete cosine transform; DCT [video]**

either the forward discrete cosine transform or the inverse discrete cosine transform. The DCT is an invertible, discrete orthogonal transformation. The inverse DCT is defined in Annex A of ISO/IEC 11172-2

2.1.54

display order [video]

the order in which the decoded pictures should be displayed. Normally this is the same order in which they were presented at the input of the encoder

2.1.55

downmix [audio]

a matrixing of n channels to obtain less than n channels

2.1.56

dual channel mode [audio]

a mode, where two audio channels with independent programme contents (e.g. bilingual) are encoded within one bit stream. The coding process is the same as for the stereo mode

2.1.57

dynamic crosstalk [audio]

a method of multichannel data reduction in which stereo-irrelevant signal components are copied to another channel

2.1.58

dynamic transmission channel switching [audio]

a method of multichannel data reduction by allocating the most orthogonal signal components to the transmission channels

2.1.59

editing

the process by which one or more compressed bit streams are manipulated to produce a new compressed bit stream. Conforming edited bit streams must meet the requirements defined in this CD

2.1.60

elementary stream [system]

a generic term for one of the coded video, coded audio or other coded bit streams

2.1.61

emphasis [audio]

filtering applied to an audio signal before storage or transmission to improve the signal-to-noise ratio at high frequencies

2.1.62

encoder

an embodiment of an encoding process

2.1.63

encoding (process)

a process, not specified in this CD, that reads a stream of input pictures or audio samples and produces a valid coded bit stream as defined in this CD

2.1.64

entropy coding

variable length lossless coding of the digital representation of a signal to reduce redundancy

2.1.65

extension bit stream [audio]

information contained in an additional bit stream related to the basic audio bit stream at the system level, to support bit rates beyond those defined in ISO/IEC 11172-3. The additional bitstream contains the remainder of the multichannel and multilingual data

2.1.66**fast forward playback [video]**

the process of displaying a sequence, or parts of a sequence, of pictures in display-order faster than real-time

2.1.67**FFT**

fast fourier transformation. A fast algorithm for performing a discrete Fourier transform (an orthogonal transform)

2.1.68**filterbank [audio]**

a set of band-pass filters covering the entire audio frequency range

2.1.69**fixed segmentation [audio]**

a subdivision of the digital representation of an audio signal into fixed segments of time

2.1.70**forbidden**

the term “forbidden” when used in the clauses defining the coded bit stream indicates that the value shall never be used. This is usually to avoid emulation of start codes

2.1.71**forced updating [video]**

the process by which macroblocks are intra-coded from time-to-time to ensure that mismatch errors between the inverse DCT processes in encoders and decoders cannot build up excessively

2.1.72**forward motion vector [video]**

a motion vector that is used for motion compensation from a reference picture at an earlier time in display order

2.1.73**frame [audio]**

a part of the audio signal that corresponds to audio PCM samples from an Audio Access Unit

2.1.74**free format [audio]**

any bitrate other than the defined bitrates that is less than the maximum valid bitrate for each layer

2.1.75**future reference picture [video]**

the future reference picture is the reference picture that occurs at a later time than the current picture in display order

2.1.76**granules [Layer II] [audio]**

the set of 3 consecutive subband samples from all 32 subbands that are considered together before quantisation. They correspond to 96 PCM samples

2.1.77**granules [Layer III] [audio]**

576 frequency lines that carry their own side information

2.1.78**group of pictures [video]**

a series of one or more coded pictures intended to assist random access. The group of pictures is one of the layers in the coding syntax defined in ISO/IEC 11172-2

2.1.79

hann window [audio]

a time function applied sample-by-sample to a block of audio samples before Fourier transformation

2.1.80

huffman coding

a specific method for entropy coding

2.1.81

hybrid filterbank [audio]

a serial combination of subband filterbank and MDCT

2.1.82

IMDCT [audio]

inverse modified discrete cosine transform

2.1.83

intensity stereo [audio]

a method of exploiting stereo irrelevance or redundancy in stereophonic audio programmes based on retaining at high frequencies only the energy envelope of the right and left channels

2.1.84

interlace [video]

the property of conventional television pictures where alternating lines of the picture represent different instances in time

2.1.85

intra coding [video]

coding of a macroblock or picture that uses information only from that macroblock or picture

2.1.86

intra-coded picture; I-picture [video]

a picture coded using information only from itself

2.1.87

ISO/IEC 11172 (multiplexed) stream [system]

a bit stream composed of zero or more elementary streams combined in the manner defined in ISO/IEC 11172-1

2.1.88

joint stereo coding [audio]

any method that exploits stereophonic irrelevance or stereophonic redundancy

2.1.89

joint stereo mode [audio]

a mode of the audio coding algorithm using joint stereo coding

2.1.90

layer [audio]

one of the levels in the coding hierarchy of the audio system defined in this part of the CD

2.1.91

layer [video and systems]

one of the levels in the data hierarchy of the video and system specifications defined in ISO/IEC 11172-1 and ISO/IEC 11172-2

2.1.92

low frequency enhancement channel [audio]

a limited bandwidth channel for low frequency audio effects in a multichannel system

2.1.93**luminance (component) [video]**

a matrix, block or single pel representing a monochrome representation of the signal and related to the primary colours in the manner defined in CCIR Rec 601. The symbol used for luminance is Y

2.1.94**macroblock [video]**

the four 8 by 8 blocks of luminance data and the two corresponding 8 by 8 blocks of chrominance data coming from a 16 by 16 section of the luminance component of the picture. Macroblock is sometimes used to refer to the pel data and sometimes to the coded representation of the pel values and other data elements defined in the macroblock layer of the syntax defined in ISO/IEC 11172-2. The usage is clear from the context

2.1.95**mapping [audio]**

conversion of an audio signal from time to frequency domain by subband filtering and/or by MDCT

2.1.96**masking [audio]**

a property of the human auditory system by which an audio signal cannot be perceived in the presence of another audio signal

2.1.97**masking threshold [audio]**

a function in frequency and time below which an audio signal cannot be perceived by the human auditory system

2.1.98**MDCT [audio]**

modified discrete cosine transform which corresponds to the Time Domain Aliasing Cancellation Filter Bank

2.1.99**motion compensation [video]**

the use of motion vectors to improve the efficiency of the prediction of pel values. The prediction uses motion vectors to provide offsets into the past and/or future reference pictures containing previously decoded pel values that are used to form the prediction error signal

2.1.100**motion estimation [video]**

the process of estimating motion vectors during the encoding process

2.1.101**motion vector [video]**

a two-dimensional vector used for motion compensation that provides an offset from the coordinate position in the current picture to the coordinates in a reference picture

2.1.102**MS stereo [audio]**

a method of exploiting stereo irrelevance or redundancy in stereophonic audio programmes based on coding the sum and difference signal instead of the left and right channels

2.1.103**multichannel [audio]**

a combination of audio channels used to create a spatial sound field

2.1.104**multilingual [audio]**

a presentation of dialogue in more than one language

2.1.105

non-intra coding [video]

coding of a macroblock or picture that uses information both from itself and from macroblocks and pictures occurring at other times

2.1.106

non-tonal component [audio]

a noise-like component of an audio signal

2.1.107

nyquist sampling

sampling at or above twice the maximum bandwidth of a signal

2.1.108

pack [system]

a pack consists of a pack header followed by one or more packets. It is a layer in the system coding syntax described in ISO/IEC 11172-1

2.1.109

packet data [system]

contiguous bytes of data from an elementary stream present in a packet

2.1.110

packet header [system]

the data structure used to convey information about the elementary stream data contained in the packet data

2.1.111

packet [system]

a packet consists of a header followed by a number of contiguous bytes from an elementary data stream. It is a layer in the system coding syntax described in ISO/IEC 11172-1

2.1.112

padding [audio]

a method to adjust the average length in time of an audio frame to the duration of the corresponding PCM samples, by conditionally adding a slot to the audio frame

2.1.113

past reference picture [video]

the past reference picture is the reference picture that occurs at an earlier time than the current picture in display order

2.1.114

pel aspect ratio [video]

the ratio of the nominal vertical height of pel on the display to its nominal horizontal width

2.1.115

pel [video]

picture element

2.1.116

picture period [video]

the reciprocal of the picture rate

2.1.117

picture rate [video]

the nominal rate at which pictures should be output from the decoding process

2.1.118**picture [video]**

source, coded or reconstructed image data. A source or reconstructed picture consists of three rectangular matrices of 8-bit numbers representing the luminance and two chrominance signals. The Picture layer is one of the layers in the coding syntax defined in this CD. Note that the term “picture” is always used in this CD in preference to the terms field or frame

2.1.119**polyphase filterbank [audio]**

a set of equal bandwidth filters with special phase interrelationships, allowing for an efficient implementation of the filterbank

2.1.120**prediction [audio]**

the use of a predictor to provide an estimate of the subband sample in one channel from the subband samples in other channels

2.1.121**prediction [video]**

the use of a predictor to provide an estimate of the pel value or data element currently being decoded

2.1.122**predictive-coded picture; P-picture [video]**

a picture that is coded using motion compensated prediction from the past reference picture

2.1.123**prediction error [video]**

the difference between the actual value of a pel or data element and its predictor

2.1.124**predictor [video]**

a linear combination of previously decoded pel values or data elements

2.1.125**presentation channel [audio]**

audio channels at the output of the decoder corresponding to the loudspeaker positions left, centre, right, left surround and right surround

2.1.126**presentation time-stamp; PTS [system]**

a field that may be present in a packet header that indicates the time that a presentation unit is presented in the system target decoder

2.1.127**presentation unit; PU [system]**

a decoded audio access unit or a decoded picture

2.1.128**psychoacoustic model [audio]**

a mathematical model of the masking behaviour of the human auditory system

2.1.129**quantisation matrix [video]**

a set of sixty-four 8-bit values used by the dequantiser

2.1.130**quantised DCT coefficients [video]**

DCT coefficients before dequantisation. A variable length coded representation of quantised DCT coefficients is stored as part of the compressed video bit stream

2.1.131

quantiser scalefactor [video]

a data element represented in the bit stream and used by the decoding process to scale the dequantisation

2.1.132

random access

the process of beginning to read and decode the coded bit stream at an arbitrary point

2.1.133

reference picture [video]

reference pictures are the nearest adjacent I- or P-pictures to the current picture in display order

2.1.134

reorder buffer [video]

a buffer in the system target decoder for storage of a reconstructed I-picture or a reconstructed P-picture

2.1.135

requantisation [audio]

decoding of coded subband samples in order to recover the original quantised values

2.1.136

reserved

the term “reserved” when used in the clauses defining the coded bit stream indicates that the value may be used in the future for ISO/IEC defined extensions

2.1.137

reverse playback [video]

the process of displaying the picture sequence in the reverse of display order

2.1.138

scalefactor band [audio]

a set of frequency lines in Layer III which are scaled by one scalefactor

2.1.139

scalefactor index [audio]

a numerical code for a scalefactor

2.1.140

scalefactor [audio]

factor by which a set of values is scaled before quantisation

2.1.141

sequence header [video]

a block of data in the coded bit stream containing the coded representation of a number of data elements

2.1.142

side information

information in the bit stream necessary for controlling the decoder

2.1.143

skipped macroblock [video]

a macroblock for which no data are stored

2.1.144

slice [video]

a series of macroblocks. It is one of the layers of the coding syntax defined in ISO/IEC 11172-2

2.1.145**slot [audio]**

a slot is an elementary part in the bit stream. In Layer I a slot equals four bytes, in Layers II and III one byte

2.1.146**source stream**

a single non-multiplexed stream of samples before compression coding

2.1.147**spreading function [audio]**

a function that describes the frequency spread of masking effects

2.1.148**start codes [system and video]**

32-bit codes embedded in that coded bit stream that are unique. They are used for several purposes including identifying some of the layers in the coding syntax

2.1.149**STD input buffer [system]**

a first-in first-out buffer at the input of the system target decoder for storage of compressed data from elementary streams before decoding

2.1.150**stereo-irrelevant [audio]**

a portion of a stereophonic audio signal which does not contribute to spatial perception

2.1.151**stereo mode [audio]**

mode, where two audio channels which form a stereo pair (left and right) are encoded within one bit stream. The coding process is the same as for the dual channel mode

2.1.152**stuffing (bits); stuffing (bytes)**

code-words that may be inserted into the compressed bit stream that are discarded in the decoding process. Their purpose is to increase the bitrate of the stream

2.1.153**subband [audio]**

subdivision of the audio frequency band

2.1.154**subband filterbank [audio]**

a set of band filters covering the entire audio frequency range. In this part of the CD, the subband filterbank is a polyphase filterbank

2.1.155**subband samples [audio]**

the subband filterbank within the audio encoder creates a filtered and subsampled representation of the input audio stream. The filtered samples are called subband samples. From 384 time-consecutive input audio samples, 12 time-consecutive subband samples are generated within each of the 32 subbands

2.1.156**surround channel [audio]**

an audio presentation channel added to the front channels (L and R or L, R, and C) to enhance the spatial perception

2.1.157

syncword [audio]

a 12-bit code embedded in the audio bit stream that identifies the start of a frame

2.1.158

synthesis filterbank [audio]

filterbank in the decoder that reconstructs a PCM audio signal from subband samples

2.1.159

system header [system]

the system header is a data structure defined in ISO/IEC 11172-1 that carries information summarising the system characteristics of the ISO/IEC 11172 multiplexed stream

2.1.160

system target decoder; STD [system]

a hypothetical reference model of a decoding process used to describe the semantics of an ISO/IEC 11172 multiplexed bit stream

2.1.161

time-stamp [system]

a term that indicates the time of an event

2.1.162

triplet [audio]

a set of 3 consecutive subband samples from one subband. A triplet from each of the 32 subbands forms a granule

2.1.163

tonal component [audio]

a sinusoid-like component of an audio signal

2.1.164

variable bitrate

operation where the bitrate varies with time during the decoding of a compressed bit stream

2.1.165

variable length coding; VLC

a reversible procedure for coding that assigns shorter code-words to frequent events and longer code-words to less frequent events

2.1.166

video buffering verifier; VBV [video]

a hypothetical decoder that is conceptually connected to the output of the encoder. Its purpose is to provide a constraint on the variability of the data rate that an encoder or editing process may produce

2.1.167

video sequence [video]

a series of one or more groups of pictures. It is one of the layers of the coding syntax defined in ISO/IEC 11172-2

2.1.168

zigzag scanning order [video]

a specific sequential ordering of the DCT coefficients from (approximately) the lowest spatial frequency to the highest

2.2 Symbols and abbreviations

The mathematical operators used to describe this part of ISO/IEC 13818 are similar to those used in the C programming language. However, integer division with truncation and rounding are specifically defined. The bitwise operators are defined assuming twos-complement representation of integers. Numbering and counting loops generally begin from zero.

2.2.1 Arithmetic operators

+	Addition.
−	Subtraction (as a binary operator) or negation (as a unary operator).
++	Increment.
− −	Decrement.
*	Multiplication.
^	Power.
/	Integer division with truncation of the result toward zero. For example, $7/4$ and $-7/4$ are truncated to 1 and $-7/4$ and $7/-4$ are truncated to -1 .
//	Integer division with rounding to the nearest integer. Half-integer values are rounded away from zero unless otherwise specified. For example $3//2$ is rounded to 2, and $-3//2$ is rounded to -2 .
DIV	Integer division with truncation of the result towards $-\infty$.
	Absolute value. $ x = x$ when $x > 0$ $ x = 0$ when $x == 0$ $ x = -x$ when $x < 0$
%	Modulus operator. Defined only for positive numbers.
Sign()	Sign. $\text{Sign}(x) = 1$ when $x > 0$ $\text{Sign}(x) = 0$ when $x == 0$ $\text{Sign}(x) = -1$ when $x < 0$
NINT ()	Nearest integer operator. Returns the nearest integer value to the real-valued argument. Half-integer values are rounded away from zero.
sin	Sine.
cos	Cosine.
exp	Exponential.
√	Square root.
$\log_{10}$	Logarithm to base ten.
$\log_e$	Logarithm to base e.
$\log_2$	Logarithm to base 2.

2.2.2 Logical operators

	Logical OR.
&&	Logical AND.
!	Logical NOT

2.2.3 Relational operators

>	Greater than.
>=	Greater than or equal to.
<	Less than.
<=	Less than or equal to.
==	Equal to.
!=	Not equal to.
max [...]	the maximum value in the argument list.
min [...]	the minimum value in the argument list.

2.2.4 Bitwise operators

A twos complement number representation is assumed where the bitwise operators are used.

&	AND
	OR
>>	Shift right with sign extension.
<<	Shift left with zero fill.

2.2.5 Assignment

=	Assignment operator.
---	----------------------

2.2.6 Mnemonics

The following mnemonics are defined to describe the different data types used in the coded bit stream.

bslbf	Bit string, left bit first, where “left” is the order in which bit strings are written in ISO/IEC 13818. Bit strings are written as a string of 1s and 0s within single quote marks, e.g. ‘1000 0001’. Blanks within a bit string are for ease of reading and have no significance.
centre_chan	Index of centre channel.
centre_limited	Variable which indicates whether a subband of the centre is not transmitted. It is used in the case of Phantom coding of centre channel.
ch	Channel. If ch has the value 0, the left channel of a stereo signal or the first of two independent signals is indicated.
dyn_cross	dyn_cross means that dynamic crosstalk is used for a certain transmission channel and a certain subband.
gr	Granule of 3 * 32 subband samples in audio Layer II, 18 * 32 subband samples in audio Layer III.
Lo, Ro	Compatible stereo audio signals.
L, C, R, LS, RS	Left, centre, right, left surround and right surround audio signals.
L ^W , C ^W , R ^W , LS ^W , RS ^W	Weighted left, centre, right, left surround and right surround audio signals. The weighting is necessary for two reasons: <ul style="list-style-type: none"> 1) All signals have to be attenuated prior to encoding to avoid overload when calculating the compatible stereo signal. 2) The matrix equations contain attenuation factors and other processing like phase shifting.

	The weighted and processed signals are actually coded and transmitted, and denormalised in the decoder.
left_sur_chan	Index of left surround channel.
main_data	The main_data portion of the bit stream contains the scalefactors, Huffman encoded data, and ancillary information.
mono_sur_chan	Index of the mono surround channel. This index is identical to the index of the left surround channel.
msblimit	Maximum used subband.
nch	Number of channels; equal to 1 for single_channel mode, 2 in other modes.
nmch	Number of channels in the multichannel extension part.
npred	Number of allowed predictors according to the tables in 2.5.2.10.
npredcoeff	Number of prediction coefficients used.
part2_length	The number of main_data bits used for scalefactors.
pci	index of predictor[0, 1, 2].
px	index of predictor[0, 1, ..., npred-1].
right_sur_chan	Index of right surround channel.
rpchof	Remainder polynomial coefficients, highest order first.
sb	Subband.
sbgr	Groups of individual subbands according to subbandgroup table in 2.5.2.10.
sblimit	The number of the lowest subband for which no bits are allocated.
scfsi	Scalefactor selection information.
switch_point_l	Number of scalefactor band (long block scalefactor band) from which point on window switching is used.
switch_point_s	Number of scalefactor band (short block scalefactor band) from which point on window switching is used.
T0, T1, T2, T3, T4	Audio transmission channels. The assignment of audio signals to transmission channels is determined by the dematrixing procedure and the transmission channel allocation information.
tc	Transmitted channel.
uimsbf	Unsigned integer, most significant bit first.
vlclbf	Variable length code, left bit first, where "left" refers to the order in which the VLC codes are written.
window	Number of the actual time slot in case of block_type==2, 0 <= window <= 2. (Layer III)

The byte order of multi-byte words is most significant byte first.

2.2.7 Constants

π	3,14159265358...
e	2,71828182845...

2.3 Method of describing bit stream syntax

The bit stream retrieved by the decoder is described in 2.4.1 and 2.5.1. Each data item in the bit stream is in bold type. It is described by its name, its length in bits, and a mnemonic for its type and order of transmission.

The action caused by a decoded data element in a bit stream depends on the value of that data element and on data elements previously decoded. The decoding of the data elements and the definition of the state variables used in their decoding are described in 2.4.2, 2.4.3, 2.5.2 and 2.5.3. The following constructs are used to express the conditions when data elements are present, and are in normal type:

Note this syntax uses the 'C'-code convention that a variable or expression evaluating to a non-zero value is equivalent to a condition that is true.

while (condition) {	If the condition is true, then the group of data elements occurs next in the data stream. This repeats until the condition is not true.
data_element	
...	
}	
do {	The data element always occurs at least once. The data element is repeated until the condition is not true.
data_element	
...	
} while (condition)	
if (condition) {	If the condition is true, then the first group of data elements occurs next in the data stream
data_element	
...	
}	
else {	If the condition is not true, then the second group of data elements occurs next in the data stream.
data_element	
...	
}	
for (expr1; expr2; expr3)	Expr1 is an expression specifying the initialisation of the loop. Normally it specifies the initial state of the counter. Expr2 is a condition specifying a test made before each iteration of the loop. The loop terminates when the condition is not true. Expr3 is an expression that is performed at the end of each iteration of the loop, normally it increments a counter.
{	
data_element	
...	
}	

Note that the most common usage of this construct is as follows:

for (i = 0; i < n; i++) {	The group of data elements occurs n times. Conditional constructs within the group of data elements may depend on the value of the loop control variable i, which is set to zero for the first occurrence, incremented to one for the second occurrence, and so forth.
data_element	
...	
}	

As noted, the group of data elements may contain nested conditional constructs. For compactness, the {} may be omitted when only one data element follows.

data_element []	data_element [] is an array of data. The number of data elements is indicated by the context.
data_element [n]	data_element [n] is the n+1 th element of an array of data.
data_element [m][n]	data_element [m][n] is the m+1,n+1 th element of a two-dimensional array of data.
data_element [l][m][n]	data_element [l][m][n] is the l+1,m+1,n+1 th element of a three-dimensional array of data.
data_element [m..n]	data_element [m..n] is the inclusive range of bits between bit m and bit n in the data_element.

While the syntax is expressed in procedural terms, it should not be assumed that 2.4.3 implements a satisfactory decoding procedure. In particular, it defines a correct and error-free input bit stream. Actual decoders must include a means to look for start codes in order to begin decoding correctly.

Definition of bytealigned function

The function bytealigned () returns 1 if the current position is on a byte boundary, that is the next bit in the bit stream is the first bit in a byte. Otherwise it returns 0.

Definition of nextbits function

The function nextbits () permits comparison of a bit string with the next bits to be decoded in the bit stream.

Definition of next_start_code function

The next_start_code function removes any zero bit and zero byte stuffing and locates the next start code.

Syntax	No. of bits	Mnemonic
next_start_code() { while (!bytealigned()) zero_bit while (nextbits() != '0000 0000 0000 0000 0000 0001') zero_byte }	'1' 8	'0' '00000000'

This function checks whether the current position is bytealigned. If it is not, zero stuffing bits are present. After that any number of zero bytes may be present before the start-code. Therefore start-codes are always bytealigned and may be preceded by any number of zero stuffing bits.

2.4 Requirements for Extension of ISO/IEC 11172-3 Audio Coding to Lower Sampling Frequencies

2.4.1 Specification of the Coded Audio Bit stream Syntax

2.4.1.1 Layer I, II

See ISO/IEC 11172-3, 2.4.1.

2.4.1.2 Layer III

Syntax	No. of bits	Mnemonic
audio_data() { main_data_begin if (mode==single_channel) private_bits else private_bits for (ch=0; ch<nch; ch++) { part2_3_length[ch] big_values[ch] global_gain[ch] scalefac_compress[ch] window_switching_flag[ch] if (window_switching_flag[ch]) { block_type[ch] mixed_block_flag[ch] for (region=0; region<2; region++) table_select[ch][region] for (window=0; window<3; window++) subblock_gain[ch][window] } } }	8 1 2 12 9 8 9 1 2 1 5 3	uimbsf bslbf bslbf uimbsf uimbsf uimbsf bslbf bslbf bslbf uimbsf bslbf uimbsf

The main data bit stream is defined below. The `main_data` field in the `audio_data()` syntax contains bytes from the main data bit stream. However, because of the variable nature of Huffman coding used in Layer III, the main data for a frame does not generally follow the header and side information for that frame. The `main_data` for a frame starts at a location in the bit stream preceding the header of the frame at a negative offset given by the value of `main_data_begin`. (See definition of `main_data_begin` in ISO/IEC 11172-3).

Huffman code bits — see ISO/IEC 11172-3, 2.4.1.7.

2.4.2 Semantics for the Audio Bit stream Syntax

2.4.2.1 *Audio Sequence General*

See ISO/IEC 11172-3, 2.4.2.1.

2.4.2.2 Audio Frame

See ISO/IEC 11172-3, 2.4.2.2.

2.4.2.3 Header

The first 32 bits (four bytes) are header information which is common to all layers.

syncword — See ISO/IEC 11172-3, 2.4.2.3.

ID — One bit to indicate the ID of the algorithm. Equals '1' for ISO/IEC 11172-3, '0' for extension to lower sampling frequencies.

Layer — See ISO/IEC 11172-3, 2.4.2.3.

protection_bit — See ISO/IEC 11172-3, 2.4.2.3.

bitrate_index — Four bits to indicate the bitrate. The all zero value indicates the 'free format' condition, in which a fixed bitrate which does not need to be in the list can be used. Fixed means that a frame contains either N or N+1 slots, depending on the value of the padding bit. The bitrate_index is an index to a table, which is different for each layer.

The bitrate_index indicates the total bitrate irrespective of the mode (stereo, joint_stereo, dual_channel, single_channel), according to the following table for ID=0.

bitrate_index	bitrate specified (kbit/s) for Fs = 16, 22,05, 24 kHz	
	Layer I	Layer II, Layer III
'0000'	free	free
'0001'	32	8
'0010'	48	16
'0011'	56	24
'0100'	64	32
'0101'	80	40
'0110'	96	48
'0111'	112	56
'1000'	128	64
'1001'	144	80
'1010'	160	96
'1011'	176	112
'1100'	192	128
'1101'	224	144
'1110'	256	160
'1111'	forbidden	forbidden

The decoder is not required to support bitrates higher than 256 kbit/s, 160 kbit/s, 160 kbit/s in respect to Layer I, II and III when in free format mode.

sampling_frequency — Indicates the sampling frequency for ID='0', according to the following table.

sampling_frequency	frequency specified (kHz)
'00'	22,05
'01'	24
'10'	16
'11'	reserved