

---

---

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

**Part 5:  
Software simulation**

**iTeh STANDARD PREVIEW**  
(standards.iteh.ai)

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

*Partie 5: Simulation de logiciel*

ISO/IEC TR 13818-5:2005

<https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652bf4a6b8/iso-iec-tr-13818-5-2005>

**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC TR 13818-5:2005](https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652bff4a6b8/iso-iec-tr-13818-5-2005)

<https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652bff4a6b8/iso-iec-tr-13818-5-2005>

© ISO/IEC 2005

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

**Contents**

Page

Foreword .....	iv
Introduction.....	vi
1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions .....	2
4 Symbols and abbreviations.....	17
5 Systems simulation.....	17
6 Video simulation.....	18
7 Audio simulation .....	18
7.1 Layer 1, Layer 2 and Layer 3 .....	18
7.2 AAC.....	18
8 MPEG-2 IPMP Reference Software .....	19
8.1 Architecture .....	19
8.2 Core Components .....	20
8.3 Usage of the Reference Software .....	25
Annex A (normative) Electronic annex containing software.....	29
Annex B (informative) List of patent holders .....	30
Bibliography.....	32

## 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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. 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.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

- type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;
- type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;
- type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/IEC 13818-5, which is a Technical Report of type 3, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 13818-5:1997), which has been technically revised. It also incorporates the Amendments ISO/IEC TR 13818-5:1997/Amd.1:1999 and ISO/IEC TR 13818-5:1997/Amd.2:2005, and the Technical Corrigenda ISO/IEC TR 13818-5:1997/Amd.1:1999/Cor.1:2003 and ISO/IEC TR 13818-5:1997/Amd.1:1999/Cor.2:2004.

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*
- *Part 4: Conformance testing*
- *Part 5: Software simulation* [Technical Report]

- *Part 6: Extensions for DSM-CC*
- *Part 7: Advanced Audio Coding (AAC)*
- *Part 9: Extension for real time interface for systems decoders*
- *Part 10: Conformance extensions for Digital Storage Media Command and Control (DSM-CC)*
- *Part 11: IPMP on MPEG-2 systems*

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC TR 13818-5:2005](https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652bff4a6b8/iso-iec-tr-13818-5-2005)

<https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652bff4a6b8/iso-iec-tr-13818-5-2005>

## Introduction

This Part of ISO/IEC 13818 was developed in response to the growing need for a generic coding method of moving pictures and of associated sound for various applications such as digital storage media, television broadcasting and communication. The use of this specification means that motion video can be manipulated as a form of computer data and can be stored on various storage media, transmitted and received over existing and future networks and distributed on existing and future broadcasting channels.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of patents.

The ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the ISO and IEC that he is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with the ISO and IEC. Information may be obtained from the companies listed in Annex B.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified in Annex B. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC TR 13818-5:2005](https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652b4a6b8/iso-iec-tr-13818-5-2005)

<https://standards.iteh.ai/catalog/standards/sist/ce0411d7-64c6-4e6b-b276-6652b4a6b8/iso-iec-tr-13818-5-2005>

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

## Part 5: Software simulation

### 1 Scope

This Technical Report provides a C language software simulation of an encoder and decoder for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), Part 7 (AAC) and Part 11 (IPMP) of ISO/IEC 13818.

### 2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 639 (all parts), *Code for the representation of names of languages*

ISO 8859-1, *Information processing - 8-bit single-byte coded graphic character sets - Part 1: Latin alphabet No. 1*

ISO/IEC 10918-1:1994, *Information technology - Digital compression and coding of continuous-tone still images: Requirements and guidelines* (See also ITU-T Rec. T.81.)

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

ISO/IEC 11172-2:1993, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 2: Video*

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*

ISO/IEC 11172-4:1995, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 4: Compliance testing*

ISO/IEC 11172-5:1998, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 5: Software simulation*

ISO/IEC 11172-6, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 5: Specification for implementation of Inverse Discrete Cosine Transform*

ITU-T Rec. H.222.0 (2000) | ISO/IEC 13818-1:2000, *Information technology - Generic coding of moving pictures and associated audio information : Systems*

ITU-T Rec. H.262 (2000) | ISO/IEC 13818-2:2000, *Information technology - Generic coding of moving pictures and associated audio information : Video* (See also ITU-T Rec. H.262.)

ISO/IEC 13818-3:1998, *Information technology - Generic coding of moving pictures and associated audio information - Part 3: Audio*

ISO/IEC 13818-4:2004, *Information technology - Generic coding of moving pictures and associated audio information - Part 4: Conformance testing*

ISO/IEC 13818-7:2004, *Information technology - Generic coding of moving pictures and associated audio information - Part 7: Advanced Audio Coding (AAC)*

### 3 Terms and definitions

For the purposes of this document, the following definitions apply.

**3.1 16x8 prediction [video]:** A prediction mode similar to field-based prediction but where the predicted block size is 16x8 luminance samples.

**3.2 AC coefficient [video]:** Any DCT coefficient for which the frequency in one or both dimensions is non-zero.

**3.3 access unit [systems]:** A coded representation of a presentation unit. In the case of audio, an access unit is the coded representation of an audio frame.

In the case of video, an access unit includes all the coded data for a picture, and any stuffing that follows it, up to but not including the start of the next access unit. If a picture is not preceded by a `group_start_code` or a `sequence_header_code`, the access unit begins with the picture start code. If a picture is preceded by a `group_start_code` and/or a `sequence_header_code`, the access unit begins with the first byte of the first of these start codes. If it is the last picture preceding a `sequence_end_code` in the bitstream all bytes between the last byte of the coded picture and the `sequence_end_code` (including the `sequence_end_code`) belong to the access unit.

**3.4 adaptive bit allocation [audio]:** The assignment of bits to subbands in a time and frequency varying fashion according to a psychoacoustic model.

**3.5 adaptive multichannel prediction [audio]:** A method of multichannel data reduction exploiting statistical inter-channel dependencies.

**3.6 adaptive noise allocation [audio]:** The assignment of coding noise to frequency bands in a time and frequency varying fashion according to a psychoacoustic model.

**3.7 adaptive segmentation [audio]:** A subdivision of the digital representation of an audio signal in variable segments of time.

**3.8 alias [audio]:** Mirrored signal component resulting from sub-Nyquist sampling.

**3.9 analysis filterbank [audio]:** Filterbank in the encoder that transforms a broadband PCM audio signal into a set of subsampled subband samples.

**3.10 ancillary data [audio]:** part of the bitstream that might be used for transmission of ancillary data.

**3.11 audio access unit [audio]:** For Layers I and II, an audio access unit is defined as the smallest part of the encoded bitstream which can be decoded by itself, where decoded means "fully reconstructed sound". For Layer III, an audio access unit is part of the bitstream that is decodable with the use of previously acquired main information.

**3.12 audio buffer [audio]:** A buffer in the system target decoder for storage of compressed audio data.

**3.13 audio sequence [audio]:** A non-interrupted series of audio frames (base frames plus optional extension frames) in which the following parameters are not changed:



- ID
- Layer
- Sampling Frequency

For Layer I and II, a decoder is not required to support a continuously variable bitrate (change in the bitrate index) of the base stream. Such a relaxation of requirements does not apply to the extension stream.

**3.14 B-field picture [video]:** A field structure B-Picture.

**3.15 B-frame picture [video]:** A frame structure B-Picture.

**3.16 B-picture; bidirectionally predictive-coded picture [video]:** A picture that is coded using motion compensated prediction from past and/or future reference fields or frames.

**3.17 backward compatibility:** A newer coding standard is backward compatible with an older coding standard if decoders designed to operate with the older coding standard are able to continue to operate by decoding all or part of a bitstream produced according to the newer coding standard.

**3.18 backward motion vector [video]:** A motion vector that is used for motion compensation from a reference frame or reference field at a later time in display order.

**3.19 backward prediction [video]:** Prediction from the future reference frame (field).

**3.20 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.

**3.21 base layer [video]:** First, independently decodable layer of a scalable hierarchy.

**3.22 big picture [video]:** A coded picture that would cause VBV buffer underflow as defined in C.7 Annex C of ISO/IEC 13818-2. Big pictures can only occur in sequences where low\_delay is equal to 1. "Skipped picture" is a term that is sometimes used to describe the same concept.

**3.23 bitrate [audio]:** The rate at which the compressed bitstream is delivered to the input of a decoder.

**3.24 bitstream; stream:** An ordered series of bits that forms the coded representation of the data.

**3.25 bitstream verifier [video]:** A process by which it is possible to test and verify that all the requirements specified in ISO/IEC 13818-2 are met by the bitstream.

**3.26 block [video]:** An 8-row by 8-column matrix of samples, or 64 DCT coefficients (source, quantised or dequantised).

**3.27 block companding [audio]:** Normalising of the digital representation of an audio signal within a certain time period.

**3.28 bottom field [video]:** One of two fields that comprise a frame. Each line of a bottom field is spatially located immediately below the corresponding line of the top field.

**3.29 bound [audio]:** The lowest subband in which intensity stereo coding is used.

**3.30 byte aligned:** A bit in a coded bitstream is byte-aligned if its position is a multiple of 8-bits from the first bit in the stream.

**3.31 byte:** Sequence of 8-bits.

**3.32 centre channel [audio]:** An audio presentation channel used to stabilise the central component of the frontal stereo image.

**3.33 channel [audio]:** A sequence of data representing an audio signal being transported.

**3.34 chroma simulcast [video]:** A type of scalability (which is a subset of SNR scalability) where the enhancement layer (s) contain only coded refinement data for the DC coefficients, and all the data for the AC coefficients, of the chrominance components.

**3.35 chrominance format [video]:** Defines the number of chrominance blocks in a macroblock.

**3.36 chrominance component [video]:** A matrix, block or single sample representing one of the two colour difference signals related to the primary colours in the manner defined in the bitstream. The symbols used for the chrominance signals are Cr and Cb.

**3.37 coded audio bitstream [audio]:** A coded representation of an audio signal as specified in part 3 of ISO/IEC 13818.

**3.38 coded B-frame [video]:** A B-frame picture or a pair of B-field pictures.

**3.39 coded frame [video]:** A coded frame is a coded I-frame, a coded P-frame or a coded B-frame.

**3.40 coded I-frame [video]:** An I-frame picture or a pair of field pictures, where the first field picture is an I-picture and the second field picture is an I-picture or a P-picture.

**3.41 coded order [video]:** The order in which the pictures are transmitted and decoded. This order is not necessarily the same as the display order.

**3.42 coded P-frame [video]:** A P-frame picture or a pair of P-field pictures.

**3.43 coded picture [video]:** A coded picture is made of a picture header, the optional extensions immediately following it, and the following picture data. A coded picture may be a coded frame or a coded field.

**3.44 coded representation:** A data element as represented in its encoded form.

**3.45 coded video bitstream [video]:** A coded representation of a series of one or more pictures as defined in ISO/IEC 13818-2.

**3.46 coding parameters [video]:** The set of user-definable parameters that characterise a coded bitstream. Bitstreams are characterised by coding parameters. Decoders are characterised by the bitstreams that they are capable of decoding.

**3.47 component [video]:** A matrix, block or single sample from one of the three matrices (luminance and two chrominance) that make up a picture.

**3.48 compression:** Reduction in the number of bits used to represent an item of data.

**3.49 constant bitrate:** Operation where the bitrate is constant from start to finish of the coded bitstream.

- 3.50 constrained parameters [video]:** The values of the set of coding parameters defined in 2.4.3.2 of ISO/IEC 11172-2.
- 3.51 constrained system parameter stream; CSPS [systems]:** A Program Stream for which the constraints defined in 2.7.9 of ISO/IEC 13818-1 apply.
- 3.52 CRC:** The Cyclic Redundancy Check to verify the correctness of data.
- 3.53 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.
- 3.54 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.
- 3.55 data element:** An item of data as represented before encoding and after decoding.
- 3.56 data partitioning [video]:** A method for dividing a bitstream into two separate bitstreams for error resilience purposes. The two bitstreams have to be recombined before decoding.
- 3.57 DC coefficient [video]:** The DCT coefficient for which the frequency is zero in both dimensions.
- 3.58 DCT coefficient [video]:** The amplitude of a specific cosine basis function.
- 3.59 de-emphasis [audio]:** Filtering applied to an audio signal after storage or transmission to undo a linear distortion due to emphasis.
- 3.60 decoded stream:** The decoded reconstruction of a compressed bitstream.
- 3.61 decoder input buffer [video]:** The first-in first-out (FIFO) buffer specified in the video buffering verifier.
- 3.62 decoder:** An embodiment of a decoding process.
- 3.63 decoder sub-loop [video]:** Stages within encoder which produce numerically identical results to the decode process described in ISO/IEC 13818-2 clause 7. Encoders capable of producing more than just I-pictures embed a decoder sub-loop to create temporal predictions and to model the behaviour of downstream decoders.
- 3.64 decoding (process):** The process defined in ISO/IEC 13818 parts 1, 2 and 3 that reads an input coded bitstream and outputs decoded pictures or audio samples.
- 3.65 decoding time-stamp; DTS [systems]:** A field that may be present in a PES packet header that indicates the time that an access unit is decoded in the system target decoder.
- 3.66 dequantisation:** The process of rescaling the quantised DCT coefficients after their representation in the bitstream has been decoded and before they are presented to the inverse DCT.
- 3.67 digital storage media; DSM:** A digital storage or transmission device or system.
- 3.68 discrete cosine transform; DCT:** Either the forward discrete cosine transform or the inverse discrete cosine transform. The DCT is an invertible, discrete orthogonal transformation.

- 3.69 display aspect ratio [video]:** The ratio height/width (in SI units) of the intended display.
- 3.70 display order [video]:** The order in which the decoded pictures are displayed. Normally this is the same order in which they were presented at the input of the encoder.
- 3.71 display process [video]:** The (non-normative) process by which reconstructed frames are displayed.
- 3.72 downmix [audio]:** A matrixing of  $n$  channels to obtain less than  $n$  channels.
- 3.73 drift [video]:** Accumulation of mismatch between the reconstructed output produced by the hypothetical decoder sub-loop embedded within an encoder (see definition of "decoder sub-loop") and the reconstructed outputs produced by a (downstream) decoder.
- 3.74 DSM-CC:** digital storage media command and control.
- 3.75 dual channel mode [audio]:** A mode, where two audio channels with independent programme contents (e.g. bilingual) are encoded within one bitstream. The coding process is the same as for the stereo mode.
- 3.76 dual-prime prediction [video]:** A prediction mode in which two forward field-based predictions are averaged. The predicted block size is  $16 \times 16$  luminance samples. Dual-prime prediction is only used in interlaced P-pictures.
- 3.77 dynamic crosstalk [audio]:** A method of multichannel data reduction in which stereo-irrelevant signal components are copied to another channel.
- 3.78 dynamic transmission channel switching [audio]:** A method of multichannel data reduction by allocating the most orthogonal signal components to the transmission channels.
- 3.79 editing:** The process by which one or more coded bitstreams are manipulated to produce a new coded bitstream. Conforming edited bitstreams must meet the requirements defined in parts 1, 2, and 3 of ISO/IEC 13818.
- 3.80 Elementary Stream Clock Reference; ESCR [systems]:** A time stamp in the PES Stream from which decoders of PES streams may derive timing.
- 3.81 elementary stream; ES [systems]:** A generic term for one of the coded video, coded audio or other coded bitstreams in PES packets. One elementary stream is carried in a sequence of PES packets with one and only one stream\_id.
- 3.82 emphasis [audio]:** Filtering applied to an audio signal before storage or transmission to improve the signal-to-noise ratio at high frequencies.
- 3.83 encoder:** An embodiment of an encoding process.
- 3.84 encoding (process):** A process, not specified in ISO/IEC 13818, that reads a stream of input pictures or audio samples and produces a valid coded bitstream as defined in parts 1, 2, and 3 of ISO/IEC 13818.
- 3.85 enhancement layer [video]:** A relative reference to a layer (above the base layer) in a scalable hierarchy. For all forms of scalability, its decoding process can be described by reference to the lower layer decoding process and the appropriate additional decoding process for the enhancement layer itself.

- 3.86 entitlement control message; ECM [systems]:** Entitlement Control Messages are private conditional access information which specify control words and possibly other, typically stream-specific, scrambling and/or control parameters.
- 3.87 entitlement management message; EMM [systems]:** Entitlement Management Messages are private conditional access information which specify the authorisation levels or the services of specific decoders. They may be addressed to single decoders or groups of decoders.
- 3.88 entropy coding:** Variable length lossless coding of the digital representation of a signal to reduce redundancy.
- 3.89 event [systems]:** An event is defined as a collection of elementary streams with a common time base, an associated start time, and an associated end time.
- 3.90 evil bitstreams:** Bitstreams orthogonal to reality.
- 3.91 extension bitstream [audio]:** Information contained in an optional additional bit stream related to the audio base bit stream at the system level, to support bit rates beyond those defined in ISO/IEC 11172-3. The optional extension bit stream contains the remainder of the multichannel and multilingual data.
- 3.92 fast reverse playback [video]:** The process of displaying the picture sequence in the reverse of display order faster than real-time.
- 3.93 fast forward playback [video]:** The process of displaying a sequence, or parts of a sequence, of pictures in display-order faster than real-time.
- 3.94 FFT:** Fast Fourier Transformation. A fast algorithm for performing a discrete Fourier transform (an orthogonal transform).
- 3.95 field [video]:** For an interlaced video signal, a "field" is the assembly of alternate lines of a frame. Therefore an interlaced frame is composed of two fields, a top field and a bottom field.
- 3.96 field period [video]:** The reciprocal of twice the frame rate.
- 3.97 field picture; field structure picture [video]:** A field structure picture is a coded picture with picture\_structure is equal to "Top field" or "Bottom field".
- 3.98 field-based prediction [video]:** A prediction mode using only one field of the reference frame. The predicted block size is 16x16 luminance samples. Field-based prediction is not used in progressive frames.
- 3.99 filterbank [audio]:** A set of band-pass filters covering the entire audio frequency range.
- 3.100 fixed segmentation [audio]:** A subdivision of the digital representation of an audio signal into fixed segments of time.
- 3.101 flag:** A variable which can take one of only the two values defined in this specification.
- 3.102 FLC:** Fixed Length Code.
- 3.103 forbidden:** The term "forbidden", when used in the clauses defining the coded bitstream, indicates that the value shall never be used. This is usually to avoid emulation of start codes.