SLOVENSKI STANDARD
SIST EN ISO/IEC 11172-1:1997

01-december-1997

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 11172-1:1993)

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 11172-1:1993)

Informationstechnik - Codierung von bewegten Bildern und damit verbundenen Tonsignalen für digitale Speichermedien bis zu 1,5 Mbit/s - Teil 1: Systeme (ISO/IEC 11172-1:1993)

Technologies de l'information - Codage de l'image animée et du son associé pour les supports de stockage numérique jusqu'a environ 1,5 Mbit/s - Partie 1: Systemes (ISO/IEC 11172-1:1993)

Ta slovenski standard je istoveten z: EN ISO/IEC 11172-1:1995

ICS:

35.040	Nabori znakov in kodiranje informacij	Character sets and information coding

SIST EN ISO/IEC 11172-1:1997   **en**

iTeh STANDARD PREVIEW
(standards.iteh.ai)

EUROPEAN STANDARD

NORME EUROPÉENNE

EUROPÄISCHE NORM

EN ISO/IEC 11172-1

February 1995

ICS 35.040

Descriptors: data processing, moving pictures, video data, audio data, video recording, data storage devices, digital storage, coded representation, coding (data conversion), system architecture

English version

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 11172-1:1993)

Technologies de l'information - Codage de l'image animée et du son associé pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s - Partie 1: Systèmes (ISO/IEC 11172-1:1993)

Informationstechnik – Codierung von bewegten Bildern und damit verbundenen Tonsignalen für digitale Speichermedien bis zu 1,5 Mbit/s – Teil 1: Systeme (ISO/IEC 11172-1:1993)

This European Standard was approved by CEN on 1994-11-25. 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.

The European Standards exist 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

Ref. No. EN ISO/IEC 11172-1:1995 E

Foreword

This European Standard was taken over by CEN from the work of ISO/IEC/JTC 1 "Information technology".

The Technical Board had decided to submit the final draft for Formal Vote. The result was positive.

For the time being, this European Standard exists in English and French only.

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 August 1995, and conflicting national standards shall be withdrawn at the latest by August 1995.

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

Endorsement notice

The text of the International Standard ISO/IEC 11172-1:1993 was approved by CEN as a European Standard without any modification.

INTERNATIONAL STANDARD

ISO/IEC
11172-1

First edition
1993-08-01

Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s —

Part 1:
Systems

Technologies de l'information — Codage de l'image animée et du son associé pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s —

Partie 1: Systèmes

Reference number
ISO/IEC 11172-1:1993(E)

ISO/IEC 11172-1: 1993 (E)

Contents

Page

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 11172-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Sub-Committee SC 29, *Coded representation of audio, picture, multimedia and hypermedia information*.

ISO/IEC 11172 consists of the following parts, under the general title *Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s*:

— *Part 1: Systems*

— *Part 2: Video*

— *Part 3: Audio*

— *Part 4: Compliance testing*

Annexes A and B of this part of ISO/IEC 11172 are for information only.

Introduction

> Note -- Readers interested in an overview of the MPEG Systems layer should read this Introduction and then proceed to annex A, before returning to the clauses 1 and 2. Since the system target decoder concept is referred to throughout both the normative and informative clauses of this part of ISO/IEC 11172, it may also be useful to refer to clause 2.4, and particularly 2.4.2, where the system target decoder is described.

The systems specification addresses the problem of combining one or more data streams from the video and audio parts of this International Standard with timing information to form a single stream. Once combined into a single stream, the data are in a form well suited to digital storage or transmission. The syntactical and semantic rules imposed by this systems specification enable synchronized playback without overflow or underflow of decoder buffers under a wide range of stream retrieval or receipt conditions. The scope of syntactical and semantic rules set forth in the systems specification differ: the syntactical rules apply to systems layer coding only, and do not extend to the compression layer coding of the video and audio specifications; by contrast, the semantic rules apply to the combined stream in its entirety.

The systems specification does not specify the architecture or implementation of encoder or decoders. However, bitstream properties do impose functional and performance requirements on encoders and decoders. For instance, encoders must meet minimum clock tolerance requirements. Notwithstanding this and other requirements, a considerable degree of freedom exists in the design and implementation of encoders and decoders.

A prototypical audio/video decoder system is depicted in figure 1 to illustrate the function of an ISO/IEC 11172 decoder. The architecture is not unique -- System Decoder functions including decoder timing control might equally well be distributed among elementary stream decoders and the Medium Specific Decoder -- but this figure is useful for discussion. The prototypical decoder design does not imply any normative requirement for the design of an ISO/IEC 11172 decoder. Indeed non-audio/video data is also allowed, but not shown.
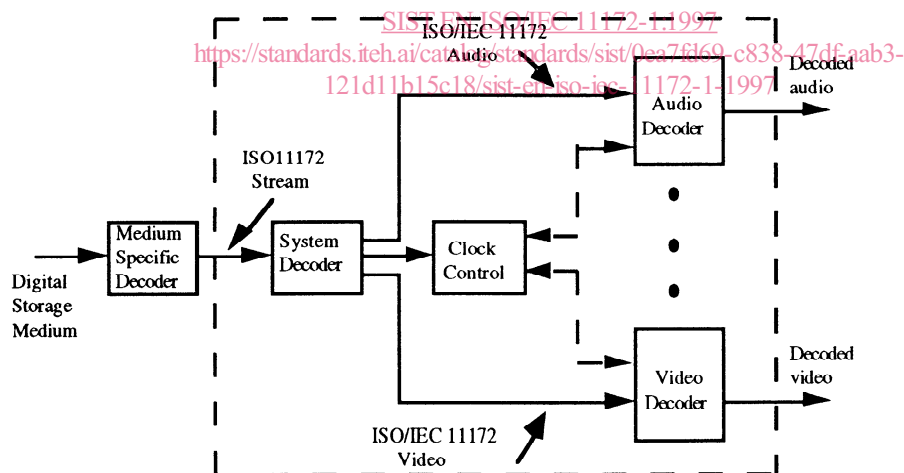

Figure 1 -- Prototypical ISO/IEC 11172 decoder

The prototypical ISO/IEC 11172 decoder shown in figure 1 is composed of System, Video, and Audio decoders conforming to Parts 1, 2, and 3, respectively, of ISO/IEC 11172. In this decoder the multiplexed coded representation of one or more audio and/or video streams is assumed to be stored on a digital storage medium (DSM), or network, in some medium-specific format. The medium specific format is not governed by this International Standard, nor is the medium-specific decoding part of the prototypical ISO/IEC 11172 decoder.

The prototypical decoder accepts as input an ISO/IEC 11172 multiplexed stream and relies on a System Decoder to extract timing information from the stream. The System Decoder demultiplexes the stream, and the elementary streams so produced serve as inputs to Video and Audio decoders, whose outputs are decoded video and audio signals. Included in the design, but not shown in the figure, is the flow of timing information among the System Decoder, the Video and Audio Decoders, and the Medium Specific Decoder.

ISO/IEC 11172-1: 1993 (E)

The Video and Audio Decoders are synchronized with each other and with the DSM using this timing information.

ISO/IEC 11172 multiplexed streams are constructed in two layers: a system layer and a compression layer. The input stream to the System Decoder has a system layer wrapped about a compression layer. Input streams to the Video and Audio decoders have only the compression layer.

Operations performed by the System Decoder either apply to the entire ISO/IEC 11172 multiplexed stream ("multiplex-wide operations"), or to individual elementary streams ("stream-specific operations"). The ISO/IEC 11172 system layer is divided into two sub-layers, one for multiplex-wide operations (the pack layer), and one for stream-specific operations (the packet layer).

0.1   Multiplex-wide operations (pack layer)

Multiplex-wide operations include the coordination of data retrieval off the DSM, the adjustment of clocks, and the management of buffers. The tasks are intimately related. If the rate of data delivery off the DSM is controllable, then DSM delivery may be adjusted so that decoder buffers neither overflow nor underflow; but if the DSM rate is not controllable, then elementary stream decoders must slave their timing to the DSM to avoid overflow or underflow.

ISO/IEC 11172 multiplexed streams are composed of packs whose headers facilitate the above tasks. Pack headers specify intended times at which each byte is to enter the system decoder from the DSM, and this target arrival schedule serves as a reference for clock correction and buffer management. The schedule need not be followed exactly by decoders, but they must compensate for deviations about it.

An additional multiplex-wide operation is a decoder's ability to establish what resources are required to decode an ISO/IEC 11172 multiplexed stream. The first pack of each ISO/IEC 11172 multiplexed stream conveys parameters to assist decoders in this task. Included, for example, are the stream's maximum data rate and the highest number of simultaneous video channels.

0.2   Individual stream operations (packet layer)

The principal stream-specific operations are 1) demultiplexing, and 2) synchronizing playback of multiple elementary streams. These topics are discussed next.

0.2.1 Demultiplexing

On encoding, ISO/IEC 11172 multiplexed streams are formed by multiplexing elementary streams. Elementary streams may include private, reserved, and padding streams in addition to ISO/IEC 11172 audio and video streams. The streams are temporally subdivided into packets, and the packets are serialized. A packet contains coded bytes from one and only one elementary stream.

Both fixed and variable packet lengths are allowed subject to constraints in 2.4.3.3 and in 2.4.5 and 2.4.6.

On decoding, demultiplexing is required to reconstitute elementary streams from the ISO/IEC 11172 multiplexed stream. This is made possible by stream_id codes in packet headers.

0.2.2 Synchronization

Synchronization among multiple streams is effected with presentation time stamps in the ISO/IEC 11172 multiplexed stream. The time stamps are in units of 90kHz. Playback of N streams is synchronized by adjusting the playback of all streams to a master time base rather than by adjusting the playback of one stream to match that of another. The master time base may be one of the N decoders' clocks, the DSM or channel clock, or it may be some external clock.

Because presentation time-stamps apply to the decoding of individual elementary streams, they reside in the packet layer. End-to-end synchronization occurs when encoders record time-stamps at capture time, when the time stamps propagate with associated coded data to decoders, and when decoders use those time-stamps to schedule presentations.

Synchronization is also possible with DSM timing time stamps in the multiplexed data stream.

0.2.3 Relation to compression layer

The packet layer is independent of the compression layer in some senses, but not in all. It is independent in the sense that packets need not start at compression layer start codes, as defined in parts 2 and 3. For example, a video packet may start at any byte in the video stream. However, time stamps encoded in packet headers apply to presentation times of compression layer constructs (namely, presentation units).

0.3   System reference decoder

Part 1 of ISO/IEC 11172 employs a "system target decoder," (STD) to provide a formalism for timing and buffering relationships. Because the STD is parameterized in terms of fields defined in ISO/IEC 11172 (for example, buffer sizes) each ISO/IEC 11172 multiplexed stream leads to its own parameterization of the STD. It is up to encoders to ensure that bitstreams they produce will play in normal speed, forward play on corresponding STDs. Physical decoders may assume that a stream plays properly on its STD; the physical decoder must compensate for ways in which its design differs from that of the STD.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

INTERNATIONAL STANDARD © ISO/IEC   ISO/IEC 11172-1: 1993 (E)

Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s —

Part 1:
Systems

Section 1: General

1.1   Scope

This part of ISO/IEC 11172 specifies the system layer of the coding. It was developed principally to support the combination of the video and audio coding methods defined in ISO/IEC 11172-2 and ISO/IEC 11172-3. The system layer supports five basic functions:

   a) the synchronization of multiple compressed streams on playback,
   b) the interleaving of multiple compressed streams into a single stream,
   c) the initialization of buffering for playback start up,
   d) continuous buffer management, and
   e) time identification.

An ISO/IEC 11172 multiplexed bit stream is constructed in two layers: the outermost layer is the system layer, and the innermost is the compression layer. The system layer provides the functions necessary for using one or more compressed data streams in a system. The video and audio parts of this specification define the compression coding layer for audio and video data. Coding of other types of data is not defined by the specification, but is supported by the system layer provided that the other types of data adhere to the constraints defined in clause 2.4.

1.2   Normative references

The following International Standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 11172. 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 11172 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 11172-2:1993 *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.*

CCIR Recommendation 601-2 *Encoding parameters of digital television for studios.*

CCIR Report 624-4 *Characteristics of systems for monochrome and colour television.*

CCIR Recommendation 648 *Recording of audio signals.*

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

1

ISO/IEC 11172-1: 1993 (E)   © ISO/IEC

CCITT Recommendation J.17 *Pre-emphasis used on Sound-Programme Circuits.*

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

IEC publication 908:1987 *CD Digital Audio System.*

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Section 2: Technical elements

2.1 Definitions

For the purposes of ISO/IEC 11172, 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 noise allocation [audio]: The assignment of coding noise to frequency bands in a time and frequency varying fashion according to a psychoacoustic model.

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

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

2.1.8 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.

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

2.1.10 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

2.1.11 backward motion vector [video]: A motion vector that is used for motion compensation from a reference picture at a later time in display order.

2.1.12 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.

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

2.1.14 bitrate: The rate at which the compressed bitstream is delivered from the storage medium to the input of a decoder.

2.1.15 block companding [audio]: Normalizing of the digital representation of an audio signal within a certain time period.

2.1.16 block [video]: An 8-row by 8-column orthogonal block of pels.

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

2.1.18 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.

2.1.19 byte: Sequence of 8-bits.

2.1.20 channel: A digital medium that stores or transports an ISO/IEC 11172 stream.

2.1.21 channel [audio]: The left and right channels of a stereo signal

2.1.22 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.

2.1.23 coded audio bitstream [audio]: A coded representation of an audio signal as specified in ISO/IEC 11172-3.

2.1.24 coded video bitstream [video]: A coded representation of a series of one or more pictures as specified in ISO/IEC 11172-2.

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

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

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

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

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

2.1.30 constant bitrate coded video [video]: A compressed video bitstream with a constant average bitrate.

2.1.31 constant bitrate: Operation where the bitrate is constant from start to finish of the compressed bitstream.

2.1.32 constrained parameters [video]: The values of the set of coding parameters defined in 2.4.3.2 of ISO/IEC 11172-2.

2.1.33 constrained system parameter stream (CSPS) [system]: An ISO/IEC 11172 multiplexed stream for which the constraints defined in 2.4.6 of this part of ISO/IEC 11172 apply.

2.1.34 CRC: Cyclic redundancy code.

2.1.35 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.

2.1.36 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.

2.1.37 data element: An item of data as represented before encoding and after decoding.

2.1.38 dc-coefficient [video]: The DCT coefficient for which the frequency is zero in both dimensions.

ISO/IEC 11172-1: 1993 (E)

2.1.39 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.

2.1.40 DCT coefficient: The amplitude of a specific cosine basis function.

2.1.41 decoded stream: The decoded reconstruction of a compressed bitstream.

2.1.42 decoder input buffer [video]: The first-in first-out (FIFO) buffer specified in the video buffering verifier.

2.1.43 decoder input rate [video]: The data rate specified in the video buffering verifier and encoded in the coded video bitstream.

2.1.44 decoder: An embodiment of a decoding process.

2.1.45 decoding (process): The process defined in ISO/IEC 11172 that reads an input coded bitstream and produces decoded pictures or audio samples.

2.1.46 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.

2.1.47 de-emphasis [audio]: Filtering applied to an audio signal after storage or transmission to undo a linear distortion due to emphasis.

2.1.48 dequantization [video]: The process of rescaling the quantized DCT coefficients after their representation in the bitstream has been decoded and before they are presented to the inverse DCT.

2.1.49 digital storage media; DSM: A digital storage or transmission device or system.

2.1.50 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.

2.1.51 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.

2.1.52 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.

2.1.53 editing: The process by which one or more compressed bitstreams are manipulated to produce a new compressed bitstream. Conforming edited bitstreams must meet the requirements defined in ISO/IEC 11172.

2.1.54 elementary stream [system]: A generic term for one of the coded video, coded audio or other coded bitstreams.

2.1.55 emphasis [audio]: Filtering applied to an audio signal before storage or transmission to improve the signal-to-noise ratio at high frequencies.

2.1.56 encoder: An embodiment of an encoding process.

2.1.57 encoding (process): A process, not specified in ISO/IEC 11172, that reads a stream of input pictures or audio samples and produces a valid coded bitstream as defined in ISO/IEC 11172.

2.1.58 entropy coding: Variable length lossless coding of the digital representation of a signal to reduce redundancy.

2.1.59 fast forward playback [video]: The process of displaying a sequence, or parts of a sequence, of pictures in display-order faster than real-time.