INTERNATIONAL STANDARD

ISO/IEC 14496-1

Second edition
2001-10-01

AMENDMENT 7
2004-08-15

Information technology — Coding of audio-visual objects —

Part 1: **Systems**

AMENDMENT 7: Use of AVC (Advanced Video Coding) in MPEG-4 systems

Technologies de l'information — Codage des objets audiovisuels —

Partie 1: Systèmes

AMENDEMENT 7: Utilisation du codage vidéo avancé dans les systèmes MPEG-4

Reference number
ISO/IEC 14496-1:2001/Amd.7:2004(E)

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
Web www.iso.org

Published in Switzerland

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.

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

Amendment 7 to ISO/IEC 14496-1:2001 was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Introduction

The Advanced Video Coding (AVC), jointly developed by the ITU-T SG 16 VCEG (Video Coding Expert Group) and ISO/IEC JTC 1/SC 29/WG 11, offers not only increased coding efficiency and enhanced robustness, but also many features for the systems that use it. To enable the best visibility of, and access to, those features, and to enhance the opportunities for the interchange and interoperability of media, this standard specifies the usage of AVC within the ISO/IEC 14496-1 Framework (ISO/IEC 14496-1:2001 and its various amendments). This specification enables AVC video streams to:

- be used in conjunction with other media streams, such as audio; and
- be used in an MPEG-4 systems environment, if desired.

Information technology — Coding of audio-visual objects —

Part 1:
Systems

AMENDMENT 7: Use of AVC (Advanced Video Coding) in MPEG-4 systems

Add the following to Clause 2 Normative references:

ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

ISO/IEC 14496-15, *Information technology — Coding of audio-visual objects — Part 15: Advanced Video Coding (AVC) file format*

Insert the following definitions alphabetically:

AVC Parameter Set
either a sequence parameter set or a picture parameter set.

NOTE This term is used to refer to both types of parameter sets.

AVC Access Unit
an access unit made up of NAL Units as defined in ITU-T Recommendation H.264 | ISO/IEC 14496-10 with the structure defined in 5.2.3 of ISO/IEC 14496-15

AVC Parameter Set Access Unit
an Access Unit made up only of sequence parameter set NAL units or picture parameter set NAL units having same timestamps to be applied

AVC Parameter Set Elementary Stream
an elementary stream made up only of AVC parameter set access units

AVC Video Elementary Stream
an elementary stream containing access units made up of NAL units for coded picture data

Add the following to Clause 5 Abbreviations and Symbols alphabetically:

AVC	Advanced Video Coding, ITU-T Recommendation H.264 | ISO/IEC 14496-10
HRD	Hypothetical Reference Decoder
IDR	Instantaneous Decoding Refresh
NAL	Network Abstraction Layer
SEI	Supplementary Enhancement Information

Replace Table 6 with following table:

Table 6 - visualProfileLevelIndication Values

Value	Profile	Level
0x00-0x7E	defined in ISO/IEC 14496-2 Annex G	-
0x7F	ISO/IEC 14496-10 Advanced Video Coding	-
0x80-0xFD	defined in ISO/IEC 14496-2 Annex G	-
0xFE	no visual profile specified	-
0xFF	no visual capability required	
NOTE — Usage of the value 0x7F indicates the use of any profile and level of ISO/IEC 14496-10 AVC. For the real profile and level numbers for ISO/IEC 14496-10 refer to the DecoderSpecificInfo. NOTE — Usage of the value 0xFE indicates that the content described by this InitialObjectDescriptor does not comply to any visual profile specified in ISO/IEC 14496-2 or -10. Usage of the value 0xFF indicates that none of the visual profile capabilities are required for this content.		

Replace Table 8 with the following table:

Table 8 — objectTypeIndication Values

Value	`ObjectTypeIndication` Description
0x00	Forbidden
0x01	Systems ISO/IEC 14496-1 [a]
0x02	Systems ISO/IEC 14496-1 [b]
0x03-0x1F	reserved for ISO use
0x20	Visual ISO/IEC 14496-2
0x21	Visual ITU-T Recommendation H.264 | ISO/IEC 14496-10 [e]
0x22	Parameter Sets for ITU-T Recommendation H.264 | ISO/IEC 14496-10 [e]
0x23-0x3F	reserved for ISO use
0x40	Audio ISO/IEC 14496-3 [d]
0x41-0x5F	reserved for ISO use
0x60	Visual ISO/IEC 13818-2 Simple Profile
0x61	Visual ISO/IEC 13818-2 Main Profile
0x62	Visual ISO/IEC 13818-2 SNR Profile
0x63	Visual ISO/IEC 13818-2 Spatial Profile
0x64	Visual ISO/IEC 13818-2 High Profile
0x65	Visual ISO/IEC 13818-2 422 Profile
0x66	Audio ISO/IEC 13818-7 Main Profile
0x67	Audio ISO/IEC 13818-7 LowComplexity Profile
0x68	Audio ISO/IEC 13818-7 Scaleable Sampling Rate Profile
0x69	Audio ISO/IEC 13818-3
0x6A	Visual ISO/IEC 11172-2
0x6B	Audio ISO/IEC 11172-3
0x6C	Visual ISO/IEC 10918-1
0x6D - 0xBF	reserved for ISO use
0xC0 - 0xFE	user private
0xFF	No object type specified

[a] This object type shall be used for all streamTypes defined in ISO/IEC 14496-1 except IPMP streams.

[b] Includes associated Amendment(s) and Corrigendum(a).

[c] Includes associated Amendment(s) and Corrigendum(a). The actual object types are defined in ISO/IEC 14496-2 and are conveyed in the DecoderSpecificInfo as specified in ISO/IEC 14496-2, Annex K.

[d] Includes associated Amendment(s) and Corrigendum(a). The actual object types are defined in ISO/IEC 14496-3 and are conveyed in the DecoderSpecificInfo as specified in ISO/IEC 14496-3, 6.2.1.

[e] Includes associated Amendment(s) and Corrigendum(a). The actual object types are defined in ITU-T Recommendation H.264 | ISO/IEC 14496-10 and are conveyed in the DecoderSpecificInfo as specified in this amendment, subclause 14.2.

Add the following after Clause 13:

14 Usage of ITU-T Recommendation H.264 | ISO/IEC 14496-10 AVC

14.1 SL packet encapsulation of AVC Access Unit

The definition of AVC Access Unit is specified in 7.4.1.2 of ITU-T Recommendation H.264 | ISO/IEC 14496-10. Following restrictions and recommendation are applied when it is encapsulated as an SL packet.

- Start Codes shall not be present in the stream. The field indicating the size of each following NAL unit shall be added before NAL unit. The size of this field is defined in DecoderSpecificInfo.
- SL packet whose randomAccessPointFlag in the header is set to '1' and subsequent SL packets shall carry access units that parameter sets required to decode are provided prior to their use.
- The Picture Timing SEI message that defines the timing information may be present in the video elementary stream, as this message contains other information than timing, and may be required for conformance testing of decoder. However, when it is encapsulated as SL packets, those time information carried by the Picture Timing SEI message shall not be used to decide decoding time or composition time of access unit. Timing information for decoding and composition shall be provided by SL packet header.
- It is recommended encapsulating one NAL unit in one SL packet when it is delivered over lossy environment.

14.2 Handling of Parameter Sets

14.2.1 Usage of DecoderSpecificInfo

Parameter Sets of AVC contents may be updated dynamically. However, DecoderSpecificInfo carrying Parameter Sets shall not be changed through the session. Parameter Sets carried in the DecoderSpecificInfo shall be updated by one of two way as follows:

- Sequence or Picture Parameter Set NAL units may be inserted in the video stream;
- A parameter set elementary stream, containing only parameter set access units, may be used to carry parameter sets separately from AVC video elementary stream. When the parameter set elementary stream is used, access units in AVC video elementary stream shall not carry parameter sets. The parameter sets shall be updated when the decoding time defined in the header of SL packet carrying those parameter sets is reached.

14.2.1.1 Decoder Specific Information

This subclause defines the `DecoderSpecificInfo` descriptor for an AVC elementary stream.

14.2.1.1.1 Syntax

```
aligned(8) class AVCDecoderSpecificInfo extends DecoderSpecificInfo : bit(8)
tag=DecSpecificInfoTag {
     AVCVideoConfigurationRecord config;
}
```

14.2.1.1.2 Semantics

The decoder specific information for an AVC stream contains an AVC video stream decoder configuration record, which is defined in ISO/IEC 14496-15 subclause 5.2.4.

`config` contains the decoder configuration record for the AVC elementary stream decoder configuration.

14.2.1.2 Object type indication

The `DecoderConfigDescriptor` shall set the value of `streamtype` equal to 0x04 for both AVC video elementary stream and the AVC parameter set elementary streams.

The `DecoderConfigDescriptor` for an AVC video elementary stream, possibly including in-line sequence or picture parameter sets, shall set `objectTypeIndication` to be 0x21 (ITU-T Recommendation H.264 | ISO/IEC 14496-10). For the AVC parameter set elementary stream, the `objectTypeIndication` value shall equal 0x22 (Parameter Sets from ITU-T Recommendation H.264 | ISO/IEC 14496-10).

14.3 Stream dependency

If the parameter set elementary stream is present, the elementary stream descriptors for the two streams shall satisfy the following conditions:

(1) The video elementary stream is dependent on the parameter set elementary stream and the `ES_Descriptor` for the video elementary stream shall have a `streamDependenceFlag` equal to true and indicate the ESID of the parameter set elementary stream in the `dependsOnESID` field. The `streamDependenceFlag` in the `ES_Descriptor` for the parameter set elementary stream shall be false.

(2) The elementary stream clocks for the parameter set elementary stream and the video elementary stream shall be the same and synchronized. The `OCRstreamflag` and `OCR_ES_Id` fields in the `ESDescriptor` for the video elementary stream and parameter set elementary streams shall be used to indicate that both streams share the same OCR.

14.4 Usage of ISO/IEC 14496-14 AVC File Format in MPEG-4 Systems

This subclause specifies how the AVC file format shall be used when the file is marked as being compatible with the MPEG-4 file format specified in ISO/IEC 14496-14. This subclause applies when the file is branded with the MPEG-4 file format brand of 'mp41' or 'mp42', and the AVC video data must be used in an MPEG-4 systems context.

14.4.1 Forming the Elementary Stream Descriptor

As is normal for MPEG-4 streams, the TrackID is related to the ElementaryStreamID, and the SLConfigDescriptor is generated following the rules for any MPEG-4 stream The format of the ES Descriptor is specified in clause 13 of the MPEG-4 systems specification [ISO/IEC 14496:2001].

If the Elementary Stream Descriptor should contain any other descriptors than SLConfigDescriptor or DecoderConfigDescriptor, they are stored in the Sample Description as defined in ISO/IEC 14496-15 subclause 5.3.4.1.

14.4.2 Forming the DecoderConfigDescriptor

The buffersizeDB, maxBitrate, and avgBitrate fields can be filled by inspection of the VUI Sequence Parameters, if present in the sequence parameter set for the AVC stream.

The DecoderSpecificInfo is formed using the contents of the AVCConfigurationBox.

14.4.3 Elementary Stream Descriptors

The use of multiple non-dependent elementary stream descriptors may also be indicated by the presence of more than one independent AVC video elementary stream descriptors in an object descriptor.

14.4.4 Switching Picture Tracks

Switching picture tracks are not MPEG-4 elementary streams and shall not be included within an MPEG-4 object descriptor.