INTERNATIONAL STANDARD

ISO/IEC 14496-4

Second edition
2004-12-15

AMENDMENT 41
2014-08-01

Information technology — Coding of audio-visual objects —

Part 4:
Conformance testing

AMENDMENT 41: Conformance testing of MVC plus depth extension of AVC

Technologies de l'information — Codage des objets audiovisuels —

Partie 4: Essai de conformité

AMENDEMENT 41: Essai de conformité de MVC et extension de profondeur de l'AVC

Reference number
ISO/IEC 14496-4:2004/Amd.41:2014(E)

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This Amendment establishes conformance test requirements for conformance to ITU-T Rec. H.264 | ISO/IEC 14496-10.

In this Amendment, additional text to ITU-T Rec. H.264 | ISO/IEC 14496-10 is specified for testing the conformance of ITU-T Rec. H.264 | ISO/IEC 14496-10 video decoders including in particular the Multiview Depth High Profiles.

The following subclauses specify the normative tests for verifying conformance of ITU-T Rec. H.264 | ISO/IEC 14496-10 video bitstreams and decoders. These normative tests make use of test data (bitstream test suites) provided as an electronic annex to this document, and of the reference software decoder specified in ITU-T Rec. H.264.2 | ISO/IEC 14496-5 with source code available in electronic format.

Information technology — Coding of audio-visual objects —

Part 4:
Conformance testing

AMENDMENT 41: Conformance testing of MVC plus depth extension of AVC

Replace 10.6.4 with the following:

10.6.4 Procedure to test bitstreams

A bitstream that claims conformance with ITU-T H.264 | ISO/IEC 14496-10 shall pass the following normative test:

The bitstream shall be decoded by processing it with the reference software decoder. When processed by the reference software decoder, the bitstream shall not cause any error or non-conformance messages to be reported by the reference software decoder. This test should not be applied to bitstreams that are known to contain errors introduced by transmission, as such errors are highly likely to result in bitstreams that lack conformance to ITU-T H.264 | ISO/IEC 14496-10.

Successfully passing the reference software decoder test provides only a strong presumption that the bitstream under test is conforming to the video layer, i.e., that it does indeed meet all the requirements for the video layer (except Annexes C, D and E and clauses G.12, H.12 and I.12) specified in ITU-T H.264 | ISO/IEC 14496-10 that are tested by the reference software decoder.

Additional tests may be necessary to more thoroughly check that the bitstream properly meets all the requirements specified in ITU-T H.264 | ISO/IEC 14496-10 including the hypothetical reference decoder (HRD) conformance (based on Annexes C, D and E and clauses G.12, H.12 and I.12). These complementary tests may be performed using other video bitstream verifiers that perform more complete tests than those implemented by the reference software decoder.

ITU-T H.264 | ISO/IEC 14496-10 contains several informative recommendations that are not an integral part of that Recommendation | International Standard. When testing a bitstream for conformance, it may also be useful to test whether or not the bitstream follows those recommendations.

To check correctness of a bitstream, it is necessary to parse the entire bitstream and to extract all the syntax elements and other values derived from those syntactic elements and used by the decoding process specified in ITU-T H.264 | ISO/IEC 14496-10.

A verifier may not necessarily perform all stages of the decoding process specified in ITU-T H.264 | ISO/IEC 14496-10 in order to verify bitstream correctness. Many tests can be performed on syntax elements in a state prior to their use in some processing stages.

Replace 10.6.5.1 with the following:

10.6.5.1 Conformance bitstreams

A bitstream has values of profile_idc, level_idc, and constraint_setX_flag (where X is a number in the range of 0 to 6, inclusive) corresponding to a set of specified constraints on a bitstream for which a decoder conforming to a specified profile and level is required in Annex A, clause G.10, H.10, or I.10 of ITU-T H.264 | ISO/IEC 14496-10 to properly perform the decoding process.

Replace 10.6.5.3 with the following:

10.6.5.3 Requirements on output of the decoding process and timing

Two classes of decoder conformance are specified:

— output order conformance; and

— output timing conformance.

The output of the decoding process is specified in clauses 8, G.8, G.12, H.8, H.12, I.8, I.12 and Annex C of ITU-T H.264 | ISO/IEC 14496-10.

For output order conformance, it is a requirement that all of the decoded pictures specified for output in Annex C, clause G.12, H.12, or I.12 of ITU-T H.264 | ISO/IEC 14496-10 shall be output by a conforming decoder in the specified order and that the values of the decoded samples in all of the pictures that are output shall be (exactly equal to) the values specified in clause 8, clause G.8, H.8, or I.8 of ITU-T H.264 | ISO/IEC 14496-10.

For output timing conformance, it is a requirement that a conforming decoder shall also output the decoded samples at the rates and times specified in Annex C, clause G.12, H.12, or I.12 of ITU-T H.264 | ISO/IEC 14496-10.

The display process, which ordinarily follows the output of the decoding process, is outside the scope of this Recommendation | International Standard.

Replace 10.6.5.6 with the following:

10.6.5.6 Dynamic tests for output timing conformance

Dynamic tests are applied to check that all the decoded samples are output and that the timing of the output of the decoder's decoded samples conforms to the specification of clauses 8, G.8, G.12, H.8, H.12, I.8, I.12; and Annex C of ITU-T H.264 | ISO/IEC 14496-10, and to verify that the HRD models (as specified by the CPB and DPB specification in Annex C, clause G.12, H.12, or I.12 of ITU-T H.264 | ISO/IEC 14496-10) are not violated when the bits are delivered at the proper rate.

The dynamic test is often easier to perform on a complete decoder system, which may include a systems decoder, a video decoder and a display process. It may be possible to record the output of the display process and to check that display order and timing of fields or frames are correct at the output of the display process. However, since the display process is not within the normative scope of ITU-T H.264 | ISO/IEC 14496-10, there may be cases where the output of the display process differs in timing or value even though the video decoder is conforming. In this case, the output of the video decoder itself (before the display process) would need to be captured in order to perform the dynamic tests on the video decoder. In particular the field or frame order and timing shall be correct.

If buffering period SEI and picture timing SEI are included in the test bitstream, HRD conformance shall be verified using the values of initial_cpb_removal_delay, initial_cpb_removal_delay_offset, cpb_removal_delay and dpb_removal_delay that are included in the bitstream.

If buffering period SEI and picture timing SEI are not included in the bitstream, the following inferences shall be made to generate the missing parameters:

— fixed_frame_rate_flag shall be inferred to be 1.

— low_delay_hrd_flag shall be inferred to be 0.

— cbr_flag shall be inferred to be 0.

— The frame rate of the bitstream shall be inferred to be the frame rate value specified in the corresponding table of clause 6.7, where the bitstream is listed. If this is missing, then a frame rate of either 25 or 30000 ÷ 1001 can be inferred.

— time_scale shall be set to 90,000 and the value of num_units_in_tick shall be computed based on field rate (twice the frame rate).

— The bit rate of the bitstream shall be inferred to be the maximum value for the level specified in Table A-1 in ITU-T H.264 | ISO/IEC 14496-10.

— CPB and DPB sizes shall be inferred to be the maximum value for the level specified in Table A-1 in ITU-T H.264 | ISO/IEC 14496-10.

With the above inferences, the HRD shall be operated as follows.

— The CPB is filled starting at time t = 0, until it is full, before removal of the first access unit. This means that the initial_cpb_removal_delay shall be inferred to be equal to the total CPB buffer size divided by the bit rate divided by 90000 (rounded downwards) and initial_cpb_removal_delay_offset shall be inferred to be equal to zero.

— The first access unit is removed at time t = initial_cpb_removal_delay ÷ 90000 and subsequent access units are removed at intervals based on the frame distance, i.e., 2 * (90000 ÷ num_units_in_tick) or the field distance, i.e., (90000 / num_units_in_tick), depending on whether the access unit is coded as a frame picture or field picture.

— Using these inferences, the CPB will not overflow or underflow and the DPB will not overflow.

In 10.6.5.7, add the following at the end of the sub clause 10.6.5.7:

A decoder that conforms to the Multiview Depth High profile at a specific level shall be capable of decoding the specified bitstreams in Table AMD41.1. A decoder that conforms to the Multiview Depth High profile shall also be capable of decoding all bitstreams that are required to be decoded by a Main, High or Stereo High profile decoder of the same level. In addition to the specified bitstreams in Table AMD41.1, a decoder that conforms to the Multiview Depth High profile shall be capable of decoding the bitstreams in Tables AMD9-1, AMD9-2 and AMD38.1 that correspond to these requirements.

Add the following after the sub clause 10.6.6.34.2:

10.6.6.35 Test Bitstreams – Multiview Depth High Profile

10.6.6.35.1 Test bitstream #MVDDR-1. #MVDDR-2

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. The width and the height of depth view components are half of the texture view components. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode lower resolution depth view component.

10.6.6.35.2 Test bitstream #MVDDR-3. #MVDDR-4

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. The width and the height of depth view components are equal to the texture view components. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode the same resolution depth view component.

10.6.6.35.3 Test bitstream #MVDDR-5

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode various combinations of the resolution of texture and depth view component

10.6.6.35.4 Test bitstream #MVDVC-1, #MVDVC-2

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode various view configurations.

10.6.6.35.5 Test bitstream #MVDVC-3, #MVDVC-4

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 2. NumDepthViews is equal to 3. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode various view configurations.

10.6.6.35.6 Test bitstream #MVDIV-1, #MVDIV-2

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 2. NumDepthViews is equal to 3. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode texture and depth view components with inter view prediction.

10.6.6.35.7 Test bitstream #MVDIV-3, #MVDIV-4

Specification: All slices are coded as I , P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 2. NumDepthViews is equal to 3. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode texture and depth view components with inter frame prediction.

10.6.6.35.8 Test bitstream #MVDIV-5

Specification: All slices are coded as I , P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 2. NumDepthViews is equal to 3. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode texture and depth view components with various combinations of inter view and inter frame prediction.

10.6.6.35.9 Test bitstream #MVDIL-1

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. field_pic_flag is equal to 1 in texture view components. field_pic_flag is equal to 0 in depth view components. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode interlace texture view components and progressive depth components.

10.6.6.35.10 Test bitstream #MVDIL-2

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 1. NumDepthViews is equal to 2. field_pic_flag is equal to 1 in both texture and depth view components. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10.

Functional stage: Decoding of texture view component and depth view component.

Purpose: Check that the decoder can properly decode interlace texture and depth view components.

10.6.6.35.11 Test bitstream #MVDRS-1

Specification: All slices are coded as I, P or B slices. Only the first picture is coded as an IDR access unit. Each view component contains only one slice. num_views_minus1 is equal to 2. NumDepthViews is equal to 2. The width and the height of depth view components are half of the texture view components. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10. Bytestream consists of supplemental enhancement information (SEI) metadata describing texture and depth acquisition information and depth representation,

Functional stage: Decoding of texture and depth components, number of texture views is not equal to a number of depth views.

Purpose: Check that the decoder can properly decode MVD with non-equal numbers of views for texture and depth components and decode relevant SEI messages.

Add the following after Table AMD38.1:

Table AMD41.1 — Bitstreams for the Multiview Depth High Profile

Categories	Bitstream	Donated by	File Name	Multi-view Depth High	Level	Frame Rate (Frame/Sec)
Depth Resolution	MVDDR-1	Nokia	MVDDR-1	X	3 and higher	30
	MVDDR-2	Nokia	MVDDR-2	X	4 and higher	25
	MVDDR-3	Sony	MVDDR-3	X	3 and higher	30
	MVDDR-4	Sony	MVDDR-4	X	4 and higher	25
Views Configuration	MVDVC-1	Sony	MVDVC-1	X	3 and higher	30
	MVDVC-2	Sony	MVDVC-2	X	4 and higher	25
	MVDVC-3	Nokia	MVDVC-3	X	3 and higher	30
	MVDVC-4	Nokia	MVDVC-4	X	4 and higher	25
Inter View Prediction / temporal inter prediction	MVDIV-1	Nokia	MVDIV-1	X	4 and higher	25
	MVDIV-2	Sony	MVDIV-2	X	3 and higher	30
	MVDIV-3	MERL, Qualcomm	MVDIV-3	X	5.1 and higher	25
	MVDIV-4	MERL, Qualcomm	MVDIV-4	X	5.1 and higher	30
Interlaced Coding Tools	MVDIL-1	ITRI, Sony	MVDIL-1	X	3.1 and higher	30
	MVDIL-2	ITRI	MVDIL-2	X	4.1 and higher	15
MVD representation and SEI	MVDRS-1	Nokia	MVDRS-1	X	3 and higher	30

ICS 35.040

Price based on 6 pages