INTERNATIONAL STANDARD

ISO/IEC 14496-4

Second edition
2004-12-15

AMENDMENT 43
2015-12-01

Information technology — Coding of audio-visual objects —

Part 4: Conformance testing

AMENDMENT 43: 3D-AVC conformance testing

Technologies de l'information — Codage des objets audiovisuels —

Partie 4: Essai de conformité

AMENDEMENT 43: Essai de conformité 3D-AVC

Reference number
ISO/IEC 14496-4:2004/Amd.43:2015(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 14496-4 is specified for testing the conformance of ITU-T Rec. H.264.1 | ISO/IEC 14496-10 video decoders including in particular the MFC 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 43: 3D-AVC conformance testing

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 G.12, H.12, I.12, and J.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 G.12, H.12, I.12, and J.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, G.10, H.10, I.10, or J.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 Clause 8, G.8, G.12, H.8, H.12, I.8, I.12, J.8, J.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, G.12, H.12, I.12, or J.12 of ITU-T H.264 | ISO/IEC 14496-10 shall be the 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, G.8, H.8, I.8, or J.12 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, G.12, H.12, I.12, or J.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 Clause 8, G.8, G.12, H.8, H.12, I.8, I.12, J.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, G.12, H.12, I.12, or J.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 and picture timing SEI messages 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 and picture timing SEI messages are not included in the bitstream, the following inferences shall be made to generate the missing parameters.

— The fixed_frame_rate_flag shall be inferred to be 1.

— The low_delay_hrd_flag shall be inferred to be 0.

— The 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 6.7, where the bitstream is listed. If this is missing, then a frame rate of either 25 or 30 000 ÷ 1 001 can be inferred.

— The 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 the 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 90 000 (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 ÷ 90 000 and subsequent access units are removed at intervals based on the frame distance, i.e. 2 × (90 000 ÷ num_units_in_tick) or the field distance, i.e. (90 000 ÷ 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 10.6.5.7:

A decoder that conforms to the 3D-AVC profile at a specific level shall be capable of decoding the specified bitstreams in Table 6. A decoder that conforms to the 3D-AVC profile shall also be capable of decoding all bitstreams that are required to be decoded by a Main or High profile decoder of the same level. In addition to the specified bitstreams in Table 6, a decoder that conforms to the 3D-AVC profile shall be capable of decoding the bitstreams in Tables 1, 2, and 4 that correspond to these requirements.

Add the following after the 10.6.6.36.6:

10.6.6.37 Test bitstreams – 3D-AVC profile

10.6.6.37.1 Test bitstream #MVDDR3D-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. The width and the height of depth view components are quarter 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. The tools specified for the 3D-AVC profile are enabled, including depth-based MVP, VSP, adaptive luminance compensation, RLE Skip, and slice header prediction.

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

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

10.6.6.37.2 Test bitstream #MVDDR3D-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 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. The tools specified for the 3D-AVC profile are enabled, including depth-based MVP, VSP, adaptive luminance compensation, RLE Skip, and slice header prediction.

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.

10.6.6.37.3 Test bitstream #MVDCT-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 3. All NAL units are encapsulated into the byte stream format specified in Annex B in ITU-T H.264 | ISO/IEC 14496-10. View synthesis prediction (VSP) for dependent texture views is enabled.

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

Purpose: Check that the decoder can properly decode bitstreams if VSP is utilized.

10.6.6.37.4 Test bitstream #MVDCT-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. Depth-based motion vector prediction (DMVP) is enabled.

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

Purpose: Check that the decoder can properly decode bitstreams if DMVP is utilized.

10.6.6.37.5 Test bitstream #MVDCT-3

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. Adaptive luminance compensation (ALC) is enabled for coded texture views.

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

Purpose: Check that the decoder can properly decode bitstreams if ALC is utilized.

10.6.6.37.6 Test bitstream #MVDCT-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. Non-linear depth representation (NLDR) is enabled for coded depth views.

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

Purpose: Check that the decoder can properly decode bitstreams if NLDR is utilized.

10.6.6.37.7 Test bitstream #MVDCT-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. Slice header prediction (SHP) is enabled for texture component of dependent views and depth component of the base texture view.

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

Purpose: Check that the decoder can properly decode bitstreams if SHP is utilized.

10.6.6.37.8 Test bitstream #MVDCT-6

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. Multiview video plus depth (MVD) data is coded in order when texture component precedes depth component of the same view. NBDV disparity derivation method is utilized for the coding of texture components of dependent views.

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

Purpose: Check that the decoder can properly decode bitstreams if texture precedes the depth component and NBDV is used for disparity derivation of dependent texture views.

10.6.6.37.9 Test bitstream #MVDCT-7

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. Multiview video plus depth (MVD) data is coded in order when texture component precedes depth component of the same view. Depth oriented NBDV (DoNBDV) disparity derivation method is utilized for coding of texture component of dependent views.

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

Purpose: Check that the decoder can properly decode bitstreams if texture precedes the depth component and DoNBDV is used for disparity derivation of dependent texture views.

10.6.6.37.10 Test bitstream #MVDCT-8

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. Multiview video plus depth (MVD) data is with IPP inter-view prediction structure, with central view being predicted from left view (put in LIST0), right view being predicted from left and central views (both put in LIST0), which utilizes prediction of dependent views from multiple inter-view references.

Functional stage: Decoding of texture view component and depth view component with IPP inter-view prediction.

Purpose: Check that the decoder can properly decode bitstreams with dependent views which are coded with prediction from multiple inter-view references.

10.6.6.37.11 Test bitstream #MVDCT-9

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. Multiview video plus depth (MVD) data is with utilization of run-based arithmetic coding of the skip flag (RSAC) for dependent views.

Functional stage: Decoding of texture view component and depth view component with RSAC enabled

Purpose: Check that the decoder can properly decode bitstreams with dependent views which are coded with RSAC enabled.

10.6.6.37.12 Test bitstream #MVDCT-10

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. Multiview video plus depth (MVD) data is coded with IBP inter-view prediction structure, with right view being predicted from left view, central view being bi-predicted from left view (put in LIST0), and right view (put in LIST1).

Functional stage: Decoding of texture and depth view components with IBP interview prediction.

Purpose: Check that the decoder can properly decode bitstreams with IBP inter-view prediction structure.

Add the following Table AMD43.1: