INTERNATIONAL STANDARD ISO/IEC 14496-15:2014

TECHNICAL CORRIGENDUM 1

Published 2015-04-01

Information technology — Coding of audio-visual objects —

Part 15:
Carriage of network abstraction layer (NAL) unit structured video in ISO base media file format

TECHNICAL CORRIGENDUM 1

Technologies de l'information — Codage des objets audiovisuels —

Partie 15: Transport de vidéo structuré en unités NAL au format ISO de base pour les fichiers médias

RECTIFICATIF TECHNIQUE 1

Technical Corrigendum 1 to ISO/IEC 14496-15:2014 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information.*

In section 5.3.2, replace:
- **Sub-sequence and layering SEI messages.** Sub-sequence or layering SEI messages shall not occur in the AVC elementary stream. Specifically, the sub-sequence information, sub-sequence layer characteristics, and sub-sequence characteristics SEI messages shall not occur in the stored AVC video elementary stream.

with:
- **Sub-sequence and layering SEI messages.** Sub-sequence or layering SEI messages should not occur in the AVC elementary stream. Specifically, the sub-sequence information, sub-sequence layer characteristics, and sub-sequence characteristics SEI messages should not occur in the stored AVC video elementary stream.

ICS 35.040   **Ref. No. ISO/IEC 14496-15:2014/Cor.1:2015(E)**

In section 5.3.3.1.2 replace:
```
aligned(8) class AVCDecoderConfigurationRecord {
      unsigned int(8) configurationVersion = 1;
      unsigned int(8) AVCProfileIndication;
[…]
      if( profile_idc  ==  100  ||  profile_idc  ==  110  ||
          profile_idc  ==  122  ||  profile_idc  ==  144 )
```
with:
```
aligned(8) class AVCDecoderConfigurationRecord {
      unsigned int(8) configurationVersion = 1;
      unsigned int(8) AVCProfileIndication;
[…]
      if( AVCProfileIndication == 100 || AVCProfileIndication == 110 ||
          AVCProfileIndication == 122 || AVCProfileIndication == 144 )
```

In section 5.4.7 replace
The reserved field is defined for AVC as follows:
with
The `codec_specific_parameters` field of the Subsample Information box is defined for AVC as follows:

In section 6.5.9 replace
The reserved field is defined for SVC as follows:
with
The `codec_specific_parameters` field of the Subsample Information box is defined for SVC as follows:

Modify section 8 as follows:

8   HEVC elementary streams and sample definitions

8.1   Introduction

The High Efficiency Video Coding (HEVC) standard, jointly developed by the ITU-T and ISO/IEC JTC 1/SC 29/WG 11 (MPEG), 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 part of ISO/IEC 14496-15 defines a storage format for video streams compressed using HEVC.
This clause of ISO/IEC 14496-15 specifies the storage format for HEVC (ISO/IEC 23008-2) video streams.

The storage of HEVC content uses the existing capabilities of the ISO base media file format but also defines extensions to support the following features of the HEVC codec.

⎯ Parameter sets:
   The video, sequence and picture parameter set mechanism decouples the transmission of infrequently changing information from the transmission of coded block data. Each slice containing the coded block data references the picture parameter set containing its decoding parameters. In turn, the picture parameter set references a sequence parameter set that contains sequence level decoding parameter information, and the sequence parameter set references a video parameter set that contains global decoding parameter information (across layers or view in potential scalable and 3DV extensions).

This specification includes the following tools for supporting of HEVC contents:

⎯ Temporal scalability sample grouping:
   a structuring and grouping mechanism to indicate the association of access units with different hierarchy levels of temporal scalability.

⎯ Temporal sub-layer access sample grouping:
   a structuring and grouping mechanism to indicate the identification of access units as temporal sub-layer access (TSA) samples.

— Step-wise temporal sub-layer access sample grouping:
a structuring and grouping mechanism to indicate the identification of access units as step-wise temporal sub-layer access (STSA) samples.

8.2   Elementary Stream Structure

A video stream is represented by one video track in a file.
Two types of elementary streams are defined for storing HEVC content:

* **Video Elementary Stream** that does not contain any parameter sets; all parameter sets are stored in a sample entry or sample entries;

* **Video and Parameter set elementary stream** that may contain parameter sets, and may also have parameter sets stored in their sample entry or sample entries.

8.3   Sample and configuration definition

8.3.1   Introduction

HEVC sample: An HEVC sample is an access unit as defined in ISO/IEC 23008-2.

8.3.2   Canonical order and restrictions

The canonical stream format is an HEVC elementary stream that satisfies the following conditions in addition to the general conditions in 4.3.2

* **Access unit delimiter NAL units**: The constraints obeyed by access unit delimiter NAL units are defined in ISO/IEC 23008-2.

* **Parameter sets:** A parameter set to be used in a picture must be sent prior to the sample containing that picture or in the sample for that picture. For a video stream that a particular sample entry applies to, the video parameter set, sequence parameter sets, and picture parameter sets, shall be stored only in the sample entry when the sample entry name is 'hvc1', and may be stored in the sample entry and the samples when the sample entry name is 'hev1'.

  NOTE   Storing parameter sets in the sample entries of a video stream provides a simple and static way to supply parameter sets. Storing parameters in samples on the other hand is more complex but allows for more dynamism in the case of parameter set updates (a particular parameter set's content is changed but using the same ID) and in the case of adding additional parameter sets. A decoder initializes with the parameter sets in the sample entry, and then updates using the parameter sets as they occur in the stream, starting from any sample marked as a sync sample. Such updating may replace parameter sets with a new definition using the same identifier. Each time the sample entry changes, the decoder re-initializes with the parameter sets included in the sample entry.

* **SEI messages:** SEI messages of declarative nature may be stored in the sample entry; there is no prescription about removing such SEI messages from the samples.

* **Filler data.** Video data is naturally represented as variable bit rate in the file format and should be filled for transmission if needed. Filler Data NAL units and Filler Data SEI messages shall not be present in the file format stored stream when the sample entry does not also permit parameter sets.

  NOTE   The removal or addition of Filler Data NAL units, start codes, SEI messages or Filler Data SEI messages may change the bitstream characteristics with respect to conformance with the HRD when operating the HRD in CBR mode as specified in ISO/IEC 23008-2, Annex C.

8.3.3 Decoder configuration information

8.3.3.1 HEVC decoder configuration record

8.3.3.1.1 Definition

This subclause specifies the decoder configuration information for ISO/IEC 23008-2 video content.
This record contains the size of the length field used in each sample to indicate the length of its contained NAL units as well as the parameter sets, if stored in the sample entry. This record is externally framed (its size must be supplied by the structure which contains it).
This record contains a version field. This version of the specification defines version 1 of this record. Incompatible changes to the record will be indicated by a change of version number. Readers must not attempt to decode this record or the streams to which it applies if the version number is unrecognised.
Compatible extensions to this record will extend it and will not change the configuration version code. Readers should be prepared to ignore unrecognised data beyond the definition of the data they understand.
The values for general_profile_space, general_tier_flag, general_profile_idc, general_profile_compatibility_flags, general_constraint_indicator_flags, general_level_idc, min_spatial_segmentation_idc, chroma_format_idc, bit_depth_luma_minus8 and bit_depth_chroma_minus8 must be valid for all parameter sets that are activated when the stream described by this record is decoded (referred to as "all the parameter sets" in the following sentences in this paragraph). Specifically, the following restrictions apply:

- The value of general_profile_space in all the parameter sets must be identical.

- The tier indication general_tier_flag must indicate a tier equal to or greater than the highest tier indicated in all the parameter sets.

- The profile indication general_profile_idc must indicate a profile to which the stream associated with this configuration record conforms.

  NOTE   If the sequence parameter sets are marked with different profiles, then the stream may need examination to determine which profile, if any, the entire stream conforms to. If the entire stream is not examined, or the examination reveals that there is no profile to which the entire stream conforms, then the entire stream must be split into two or more sub-streams with separate configuration records in which these rules can be met.

- Each bit in general_profile_compatibility_flags may only be set if all the parameter sets set that bit.

- Each bit in general_constraint_indicator_flags may only be set if all the parameter sets set that bit.

- The level indication general_level_idc must indicate a level of capability equal to or greater than the highest level indicated for the highest tier in all the parameter sets.

- The min_spatial_segmentation_idc indication must indicate a level of spatial segmentation equal to or less than the lowest level of spatial segmentation indicated in all the parameter sets.

- The value of chroma_format_idc in all the parameter sets must be identical.

- The value of bit_depth_luma_minus8 in all the parameter sets must be identical.

- The value of bit_depth_chroma_minus8 in all the parameter sets must be identical.

Explicit indication can be provided in the HEVC Decoder Configuration Record about the chroma format and bit depth as well as other important format information used by the HEVC video elementary stream. Each type of such information must be identical in all parameter sets, if present, in a single HEVC configuration record. If two sequences differ in any type of such information, two different HEVC sample entries must be used. If the

two sequences differ in color space indications in their VUI information, then two different HEVC sample entries are also required.

There is a set of arrays to carry initialization NAL units. The NAL unit types are restricted to indicate SPS, PPS, VPS, and SEI NAL units only. NAL unit types that are reserved in ISO/IEC 23008-2 and in this specification may acquire a definition in future, and readers should ignore arrays with reserved or unpermitted values of NAL unit type.

NOTE – this 'tolerant' behaviour is designed so that errors are not raised, allowing the possibility of backwards-compatible extensions to these arrays in future specifications.

It is recommended that the arrays be in the order VPS, SPS, PPS, SEI.

When `general_non_packed_constraint_flag` (bit 3 of the 6-byte general_constraint_indicator_flags) is equal to 0 and some of the samples referring to this sample entry represent frame-packed content and any of the default display windows specified by the active SPSs for the samples referring to this sample entry covers more than one constituent frame of the frame-packed content, the techniques described in 8.15 of ISO/IEC 14496-12 ('Post-decoder requirements on media') using the scheme type "stvi" shall be used. In this case, the `stereo_scheme` in the Stereo Video Box should be set to 1, to indicate that the frame packing scheme used in HEVC is the same as in AVC.

8.3.3.1.2   Syntax

```
aligned(8) class HEVCDecoderConfigurationRecord {
   unsigned int(8) configurationVersion = 1;
   unsigned int(2) general_profile_space;
   unsigned int(1) general_tier_flag;
   unsigned int(5) general_profile_idc;
   unsigned int(32) general_profile_compatibility_flags;
   unsigned int(48) general_constraint_indicator_flags;
   unsigned int(8) general_level_idc;
   bit(4) reserved = '1111'b;
   unsigned int(12) min_spatial_segmentation_idc;
   bit(6) reserved = '111111'b;
   unsigned int(2) parallelismType;
   bit(6) reserved = '111111'b;
   unsigned int(2) chroma_format_idc;
   bit(5) reserved = '11111'b;
   unsigned int(3) bit_depth_luma_minus8;
   bit(5) reserved = '11111'b;
   unsigned int(3) bit_depth_chroma_minus8;
   bit(16) avgFrameRate;
   bit(2) constantFrameRate;
   bit(3) numTemporalLayers;
   bit(1) temporalIdNested;
   unsigned int(2) lengthSizeMinusOne;
   unsigned int(8) numOfArrays;
   for (j=0; j < numOfArrays; j++) {
      bit(1) array_completeness;
      unsigned int(1) reserved = 0;
      unsigned int(6) NAL_unit_type;
      unsigned int(16) numNalus;
      for (i=0; i< numNalus; i++) {
         unsigned int(16) nalUnitLength;
         bit(8*nalUnitLength) nalUnit;
      }
   }
}
```

8.3.3.1.3   Semantics

general_profile_space, general_tier_flag, general_profile_idc,
   general_profile_compatibility_flags, general_constraint_indicator_flags,
   general_level_idc, min_spatial_segmentation_idc, chroma_format_idc,
   bit_depth_luma_minus8 and bit_depth_chroma_minus8 contain the matching values for the
   fields general_profile_space, general_tier_flag, general_profile_idc,
   general_profile_compatibility_flag[i] for i from 0 to 31, inclusive, the 6 bytes starting with the byte
   containing the general_progressive_source_flag, general_level_idc, min_spatial_segmentation_idc,