INTERNATIONAL STANDARD

ISO/IEC 14496-15

Sixth edition
2022-10

AMENDMENT 1
2023-10

Information technology — Coding of audio-visual objects —

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

AMENDMENT 1: Support for LCEVC

Technologies de l'information — Codage des objets audiovisuels —

Partie 15: Transport de vidéo structurée en unités NAL sur la couche réseau au format ISO de base pour les fichiers médias

AMENDEMENT 1: Support pour LCEVC

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.

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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 14496 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Information technology — Coding of audio-visual objects —

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

AMENDMENT 1: Support for LCEVC

Normative references

Add the following reference:

ISO/IEC 23094-2:2021, *Information technology – General Video Coding – Part 2: Low Complexity Enhancement Video Coding*

3.1 Terms and definitions

Add the following terms and definitions:

3.1.67
parameter sets for LCEVC
<LCEVC> sequence_configuration, global_configuration, or additional_info

Note 1 to entry: As defined in ISO/IEC 23094-2:2021, 7.3.4, 7.3.5, 7.3.10.

3.1.68
picture dimensions for LCEVC
width and height of the decoded picture as specified by the referenced global_configuration

Note 1 to entry: As defined in ISO/IEC 23094-2:2021, 7.3.5.

3.2 Abbreviated terms

Add the following abbreviated terms:

LCEVC   Low Complexity Enhancement Video Coding [ISO/IEC 23094-2]

GC   Global Configuration for LCEVC [ISO/IEC 23094-2]

SC   Sequence Configuration for LCEVC [ISO/IEC 23094-2]

AI   Additional Information for LCEVC [ISO/IEC 23094-2]

4.2.3.3

Replace

"The syntax of a NAL unit is defined in the appropriate specification (e.g. ISO/IEC 14496-10) and includes both the one byte NAL header and the variable length encapsulated byte stream payload."

with

"NALUnit contains a single NAL unit. The syntax of a NAL unit is defined in the appropriate specification (e.g. ISO/IEC 14496-10) and includes both the NAL unit header and the variable length NAL unit payload."

5.4.2.1.1

Replace:

"The sample entry name `'avc1'` or `'avc3'` may only be used when the stream to which this sample entry applies is a compliant and AVC stream as viewed by an AVC decoder operating under the configuration (including profile and level) given in the AVCConfigurationBox. The file format specific structures that resemble NAL units (see Annex A) may be present but shall not be used to access the AVC base data; that is, the AVC data shall not be contained in Aggregators (though they may be included within the bytes referenced by the `additional_bytes` field) nor referenced by Extractors."

with:

"The sample entry name `'avc1'` or `'avc3'` may only be used when the stream to which this sample entry applies is a compliant and AVC stream as viewed by an AVC decoder operating under the configuration (including profile and level) given in the `AVCConfigurationBox`. Extractor and aggregator NAL-unit-like structures (see Annex A) shall not be present."

6.5.3.1.1

Replace "Extractors or aggregators may be used for SVC VCL NAL units in `'avc1'`, `'avc2'`, `'avc3'`, `'avc4'`, `'svc1'` or `'svc2'` tracks" with "Extractors or aggregators may be used for SVC VCL NAL units `'avc2'`, `'avc4'`, `'svc1'` or `'svc2'` tracks."

11.2.4.1.2

Replace "`ptl_multi_layer_enabled_flag`" with "`ptl_multilayer_enabled_flag`".

11.2.4.1.3

Replace "`ptl_multi_layer_enabled_flag`" with "`ptl_multilayer_enabled_flag`".

12.5.4.2

Replace subclause 12.5.4.2 with the following:

12.5.4.2 Sample entry for EVC slice base track

12.5.4.2.1 Definition

Sample Entry Type: 'evm1'

Container: Sample Table Box ('stbl')

Mandatory: An 'evm1' sample entry is mandatory for EVC slice base track

Quantity: One or more sample entries may be present

An EVC visual sample entry shall contain an EVC Configuration Box as defined in the subclause 12.4.1.1 when a track is an EVC slice base track. This includes an EVCDecoderConfigurationRecord defined in the subclause 12.3.3.

An optional BitRateBox may be present in the EVC visual sample entry to signal the bit rate information of the EVC slice data in this track.

Multiple sample entries may be used, as permitted by the ISO Base Media File Format specification, to indicate sections of video that use different configurations or parameter sets.

12.5.4.2.2 Syntax

```
class EVCSampleEntry() extends VisualSampleEntry('evm1'){
   EVCConfigurationBox  config;
   MPEG4ExtensionDescriptorsBox();  // optional
}
```

12.5.4.2.3 Semantics

Compressorname in the base class VisualSampleEntry indicates the name of the compressor used with the value "\012EVC Coding" being recommended (\012 is 10, the length of the string in bytes).

EVCDecoderConfigurationRecord is defined in subclause 12.3.3.

Clause 13

Add the following clause after Clause 12, before Annex A:

13 LCEVC elementary streams and sample definitions

13.1 Overview

The Low Complexity Enhancement Video Coding (LCEVC) standard, specified in ISO/IEC 23094-2:2021, is a low complexity solution to apply enhancement to existing video coding bitstreams generated using other video coding systems (e.g. AVC, HEVC, EVC, VVC).

Since the LCEVC elementary streams carry enhancement to a "base" codec such as the ones listed above, the LCEVC elementary stream in its own track makes reference to a "base" codec elementary stream in a separate track, so that the LCEVC stream can be decoded in conjunction with the "base" stream, while the "base" stream can be decoded independently of the LCEVC stream.

This clause defines the carriage of LCEVC elementary streams in the ISO base media file format as defined in this document.

The *Elementary Stream Structure* is provided in subclause 13.2.

The *Sample and Configuration Definitions* are provided in subclause 13.3.

Deviations from ISO base media file format are provided in subclause 13.4.

Internet media applications require defined values for the codecs parameter specified in IETF RFC 6381 for ISOBMFF Media tracks. The **'codecs' parameter string** for the LCEVC codec is defined in subclause E.10.

13.2 Elementary stream structure

LCEVC elementary streams are structured as NAL units and their storage in the ISO Base Media File Format follows principles similar to other NAL structured video formats.

The storage of LCEVC elementary streams is subdivided into two parts: static information that is globally used in the elementary stream and dynamic information that may vary per sample.

The sequence_configuration (SC), global_configuration (GC), and additional_info (AI), are considered to be part of the information that rarely changes and is considered to be static.

In case these parameter sets change in the elementary stream, a sample grouping is defined that indicates at which sample the parameter sets change.

13.3 Sample and configuration definitions

13.3.1 Overview

An LCEVC Sample follows the same structure of the "General Definitions" as defined in subclause 4.2.3 of this document.

The only additional constraint is that the `DecoderConfigurationRecord` shall follow the syntax and semantics specific for LCEVC, i.e. `LCEVCDecoderConfigurationRecord`.

An LCEVC Sample contains an access unit as defined in ISO/IEC 23094-2:2021, 3.1.

13.3.2 Canonical order

The canonical stream format is an LCEVC elementary stream that satisfies the general conditions in subclauses 6.1 and 7.3 of ISO/IEC 23094-2:2021.

The following additional constraints apply:

— **SC, GC, AI**: sequence_configuration, global_configuration, and additional_info to be used in a picture must be sent prior to the sample containing that picture or in the sample for that picture. At least the one sequence_configuration and one global_configuration must be stored in the sample entry of the track that contains the LCEVC elementary stream. The sequence_configuration and global_configuration, when carried in a sample for a specific picture, shall be present at least in each sync sample.

— **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.

13.3.3 Decoder Configuration Information

13.3.3.1 Definition

This subclause specifies the Decoder Configuration Information for ISO/IEC 23094-2 video content.

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 shall 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 `LCEVCProfileIndication`, `LCEVCLevelIndication`, `chroma_format_idc`, `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `bit_depth_luma_minus8`, and `bit_depth_chroma_minus8` shall be valid for all parameter sets (SC, GC, AI) that are activated when the stream described by this record is decoded (referred to as "all the parameter sets" in the reminder of this clause).

Specifically, the following restrictions apply:

— The profile indication `LCEVCProfileIndication` shall indicate a profile to which the stream associated with this configuration record conforms.

— The level indication `LCEVCLevelIndication` shall indicate a level of capability equal to or greater than the highest level indicated in all the parameter sets of this configuration record.

— The `pic_width_in_luma_samples` and `pic_height_in_luma_samples` shall contain the highest values of all parameter sets of this configuration record.

— The value of `chroma_format_idc` in all the parameter sets shall be identical.

— The value of `bit_depth_luma_minus8` in all the parameter sets shall be identical.

— The value of `bit_depth_chroma_minus8` in all the parameter sets shall be identical.

Explicit indication is provided in the LCEVC Decoder Configuration Information about the chroma format and bit depth used by the LCEVC video elementary stream. Each instance of such information shall be identical in all parameter sets, if present, in a single LCEVC configuration record. If two sequences differ in any instance of such information, two different LCEVC sample entries shall be used.

The length field is 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.

It is recommended that the parameter sets be in the order SC, GC, AI, SEI.

13.3.3.2 Syntax

```
aligned(8) class LCEVCDecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(8) LCEVCProfileIndication;
    unsigned int(8) LCEVCLevelIndication;
    unsigned int(2) chroma_format_idc;
    unsigned int(3) bit_depth_luma_minus8;
    unsigned int(3) bit_depth_chroma_minus8;
    unsigned int(2) lengthSizeMinusOne;
    bit(6) reserved = '111111'b;
    unsigned int(32) pic_width_in_luma_samples;
    unsigned int(32) pic_height_in_luma_samples;
    unsigned int(1) sc_in_stream;
    unsigned int(1) gc_in_stream;
    unsigned int(1) ai_in_stream;
    bit(5) reserved = '11111'b;
    unsigned int(8) numOfArrays;
    for (j=0; j < numOfArrays; j++) {
        bit(2) reserved = '00'b;
        unsigned int(6) NAL_unit_type;
        unsigned int(16) numOfNalus;
        for (i=0; i< numOfNalus; i++) {
            unsigned int(16) nalUnitLength;
            bit(8*nalUnitLength) nalUnit;
        }
    }
}
```

13.3.3.3 Semantics

LCEVCProfileIndication, LCEVCLevelIndication, chroma_format_idc, bit_depth_luma_minus8 and bit_depth_chroma_minus8 contain the matching values for the fields in the SC and GC for all parameter sets of the configuration record.

lengthSizeMinusOne indicates the length in bytes of the NALUnitLength field in an LCEVC video sample of the associated stream minus one. For example, a size of one byte is indicated with a value of 0. The value of this field shall be one of 0, 1, or 3 corresponding to a length encoded with 1, 2, or 4 bytes, respectively.

pic_width_in_luma_samples and pic_height_in_luma_samples contain the largest values for the fields in all SC and GC of this configuration record when the value of the sc_in_stream field is '0'. The picture dimensions of the LCEVC track, width and height in Luminance samples, are the maximum dimensions for the LCEVC track.

They will contain the largest values for the fields in all SC and GC in the stream when the value of the sc_in_stream field is '1'. The value '0' shall be used if the largest value of these fields in the SC and GC for all the parameter sets in this record is not indicated through this field when the value of the sc_in_stream field is '0' or the value of these fields in the SC and GC in the stream has the value larger than the largest value of the field in this record when the value of the sc_in_stream field is '1'.

sc_in_stream indicates that the stream may contain additional sequence_configuration, that are not included in the array of NAL units of this configuration record.

gc_in_stream indicates that the stream may contain additional global_configuration, that are not included in the array of NAL units of this configuration record.

ai_in_stream indicates that the stream may contain additional additional_info, that are not included in the array of NAL units of this configuration record.

numOfArrays indicates the number of arrays of NAL units of the indicated type(s).

NAL_unit_type indicates the type of the NAL units in the following array (which shall be all of that type); it takes a value as defined in ISO/IEC 23094-2; it is restricted to take one of the values indicating a SG, GC, AI, or SEI NAL unit.

numOfNalus indicates the number of NAL units of the indicated type included in the configuration record for the stream to which this configuration record applies.

nalUnitLength indicates the length in bytes of the NAL unit.

nalUnit contains a SC, GC, AI or a SEI NAL unit, as specified in ISO/IEC 23094-2.

13.4 Derivation from ISO base media file format

13.4.1 LCEVC video stream definition: sample entry name and format

13.4.1.1 Definition

Sample Entry and Box Types: 'lvc1','lvcC'

Container: Sample Description Box ('stsd')

Mandatory: The 'lvc1' sample entry is mandatory

Quantity: One or more sample entries may be present

A LCEVC visual sample entry shall contain a LCEVCConfigurationBox, as defined below. This includes an LCEVCDecoderConfigurationRecord, as defined in 13.3.3.

An optional `BitRateBox` may be present in the LCEVC visual sample entry to signal the bit rate information of the LCEVC video stream. Extension descriptors that should be inserted into the Elementary Stream Descriptor, when used in MPEG-4, may also be present.

The sample entry name `'lvc1'` may only be used when the stream to which this sample entry applies is a compliant LCEVC stream as viewed by a LCEVC decoder operating under the configuration (including profile and level) given in the `LCEVCConfigurationBox`.

13.4.1.2 Syntax

```
class LCEVCConfigurationBox extends Box('lvcC') {
   LCEVCDecoderConfigurationRecord()  config;
}
class LCEVCSampleEntry() extends VisualSampleEntry('lvc1'){
   LCEVCConfigurationBox  config;
   MPEG4ExtensionDescriptorsBox();  // optional
}
```

13.4.1.3 Semantics

`Compressorname` in the base class `VisualSampleEntry` indicates the name of the compressor used with the value "`\014LCEVC Coding`" being recommended (`\014` is 10, the length of the string in bytes).

`LCEVCDecoderConfigurationRecord` is defined in subclause 13.3.3.

13.4.2 LCEVC track structure

An LCEVC stream represents an enhancement to another external video track in a file.

The track containing the external base layer stream is nominated as the "base track".

For an LCEVC bitstream with its associated external base track, the base track is always assigned one track of its own.

There shall be at least one external base track that, when taken as "base" stream for the LCEVC "enhancement" stream, represents a complete set of encoded information. This group of tracks that forms the complete encoded information is called the "complete set".

The picture dimensions of the base track, width and height in Luminance samples, are specified by the relevant `DecoderConfigurationRecord` of the base track (e.g. `AVCDecoderConfigurationRecord`, `HEVCDecoderConfigurationRecord`).

The picture dimensions of the LCEVC track, width and height in Luminance samples, are specified by the specific `LCEVCDecoderConfigurationRecord`, and are the maximum dimensions for the LCEVC track.

The LCEVC track that is part of the same "complete set" shall be linked to the "base track" by means of a track reference of type `'sbas'` (scalable base), as defined in subclause 6.5.1 of this document and subclause 8.3.3 of ISO/IEC 14496-12:2022.

The LCEVC track associated to the base track shall share the same timescale as the base track.

When the base track is coded using AVC, the base track shall be constructed according to Clause 5, without using separate parameter set tracks.

When the base track is coded using HEVC, the base track shall be constructed according to Clause 8.

When the base track is coded using VVC, the base track shall be constructed according to Clause 11.

When the base track is coded using EVC, the base track shall be constructed according to Clause 12.