**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 /N18098**

**October 2018, Macao SAR, CN**

|  |  |
| --- | --- |
| **Source** | **Requirements** |
| **Title** | **Requirements for Low Complexity Video Coding Enhancements** |
| **Status** | **Approved** |

# 1. Objective

The objective for this video coding project is to develop a data stream structure defined by two component streams, a base stream decodable by a hardware decoder, and an enhancement stream suitable for software processing implementation with sustainable power consumption. The enhancement stream will provide new features such as compression capability extension to existing codecs, lower encoding and decoding complexity, for on demand and live streaming applications.

#  Use cases

The main targeted use cases are those that require live encoding and decoding, maximum device compatibility and high-quality video whilst allowing compatibility with existing ecosystems without the need to upgrade or change all existing hardware components.

Specific applications include, but are not limited to, live TV/multimedia streaming (e.g., sports, eSports, news, etc.) under constrained OTT bandwidth, live social network mobile video, live UHD broadcast at viable DTT bandwidth, live UAV/security video downlinks, SD to HD and HD to UHD improvement without the need to replace all set-top-boxes, live surveillance, live immersive video, etc.

# Requirements

The video coding project should define a codec that allows a full resolution encoded/decoded stream being formed from enhancing a stream encoded/decoded with a hardware codec and a data stream which, when added to the coded/decoded stream, would bring the video to the full resolution.

A general scheme for the decoder is illustrated in the figure in Annex 1.

The key performance requirements for the video coding project are as follows:

* when enhancing an n-th generation MPEG codec (e.g., AVC), compression efficiency for the aggregate stream is appreciably higher than that of the n-th generation MPEG codec used at full resolution and as close as possible to that of the (n+1)-th generation MPEG codec (e.g., HEVC) used at full resolution, at bandwidths and operating conditions relevant to mass market distribution; and
* encoding and decoding complexity for the aggregate full resolution video (i.e., base plus enhancement) shall be comparable with that of the base encoder or decoder, respectively, when used alone at full resolution.

The key implementation and non-technical requirements for the video coding project are:

* the video stream should be decodable without specific firmware or OS support by all devices capable to decode the base codec, with substantially same resource utilization (e.g., processing power, battery consumption, etc.) as the base decoder at full resolution decoded in hardware;
* all web browsers should be able to decode high resolution video without plug-ins and/or browser upgrade, e.g. via HTML5 javascript;
* the additional data stream should be compatible with the existing ecosystem, e.g. ad insertion, metadata management, CDNs, DRM/CA and network protocols such as DASH, HLS, MMT and SS;
* the overall processing power requirement to encode a video stream should be comparable with that of the base codec when used alone at full resolution.

The new codec shall support rectangular picture formats that will include all commonly used picture formats, ranging at least from VGA to 8Kx4K. Picture formats of arbitrary size shall also be supported, within limits specified by Levels.

The new codec shall support:

1. YCbCr colour spaces with 4:2:0 sampling, 10 bits per component;
2. High dynamic range and wide colour gamut;
3. YCbCr/RGB 4:4:4 and YCbCr 4:2:2;
4. Bit depths up to 16 bits per component;
5. progressive scan.

Fixed and variable rational frame rates shall be supported, with upper limits specified by levels.

The video coding project shall support the encoding of the full variety of characteristics of video content encountered in the envisioned applications (to the maximum extent feasible). This includes (electronic and film) camera-captured scenes, text and graphics mixed into a camera-captured video source, rendered animation content, rendered computer graphics, etc.

The video codec enhancement is not intended to be a scalable video codec. The envisioned usage does not include the conventional scalability feature. Whereas the codec includes a component stream that can be decoded by a legacy n-th generation MPEG decoder, the decoded base signal is not meant to provide backward compatibility to existing services, as the quality of the base signal is not guaranteed to satisfy that of existing services.

**Annex 1**

****