DRAFT AMENDMENT ISO/IEC 14496-10:2012/DAM 4

ISO/IEC JTC **1** Secretariat: **ANSI**

Voting begins on **2013-07-16** Voting terminates on **2013-10-16**

Information technology — Coding of audio-visual objects —

Part 10: Advanced Video Coding

AMENDMENT 4: Additional colour space and tone mapping descriptors

Technologies de l'information — Codage des objets audiovisuels —

Partie 10: Codage visuel avancé

AMENDEMENT 4

ICS 35.040

To expedite distribution, this document is circulated as received from the committee secretariat. ISO Central Secretariat work of editing and text composition will be undertaken at publication stage.

Pour accélérer la distribution, le présent document est distribué tel qu'il est parvenu du secrétariat du comité. Le travail de rédaction et de composition de texte sera effectué au Secrétariat central de l'ISO au stade de publication.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

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.

Amendment 4 to ISO/IEC 14496-10:2012 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This amendment to ITU-T Rec. H.264 | ISO/IEC 14496-10 includes additional colour space indicators and an additional tone mapping mode in the tone mapping supplemental enhancement information (SEI) message.

Information technology — Coding of audio-visual objects — Part 10: Advance Video Coding, AMENDMENT 4: Additional Colour Space and Tone Mapping Descriptors

In 2 (Normative references), add the following:

“

– ISO 12232:2006, *Photography – Digital still cameras – Determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index.*

”

Replace the text in D.1.24 with the following:

“

D.1.24 Tone mapping information SEI message syntax

tone_mapping_info(payloadSize) {	C	Descriptor
tone_map_id	5	ue(v)
tone_map_cancel_flag	5	u(1)
if(!tone_map_cancel_flag) {		
tone_map_repetition_period	5	ue(v)
coded_data_bit_depth	5	u(8)
target_bit_depth	5	u(8)
tone_map_model_id	5	ue(v)
if(tone_map_model_id == 0) {		
min_value	5	u(32)
max_value	5	u(32)
}		
if(tone_map_model_id == 1) {		
sigmoid_midpoint	5	u(32)
sigmoid_width	5	u(32)
}		
if(tone_map_model_id == 2)		
for(i = 0; i < (1 << target_bit_depth); i++)		
start_of_coded_interval[i]	5	u(v)

if(tone_map_model_id == 3) {		
num_pivots	5	u(16)
for(i=0; i < num_pivots; i++) {		
coded_pivot_value[i]	5	u(v)
target_pivot_value[i]	5	u(v)
}		
}		
if(tone_map_model_id == 4) {		
camera_iso_speed_idc	5	u(8)
if(camera_iso_speed_idc = = Extended_ISO)		
camera_iso_speed_value	5	u(32)
exposure_index_idc	5	u(8)
if(exposure_index_idc == Extended_ISO)		
exposure_index_value	5	u(32)
exposure_compensation_value_sign_flag	5	u(1)
exposure_compensation_value_numerator	5	u(16)
exposure_compensation_value_denom_idc	5	u(16)
ref_screen_luminance_white	5	u(32)
extended_range_white_level	5	u(32)
nominal_black_level_luma_code_value	5	u(16)
nominal_white_level_luma_code_value	5	u(16)
extended_white_level_luma_code_value	5	u(16)
}		
}		
}		

”

Replace D.2.24 with the following:

“

D.2.24 Tone mapping information SEI message semantics

This SEI message provides information to enable remapping of the colour samples of the output decoded pictures for customization to particular display environments. The remapping process maps coded sample values in the RGB colour space (specified in Annex E) to target sample values. The mappings are expressed either in the luma or RGB colour space domain, and should be applied to the luma component or to each RGB component produced by colour space conversion of the decoded image accordingly.

tone_map_id contains an identifying number that may be used to identify the purpose of the tone mapping model. The value of tone_map_id shall be in the range of 0 to $2^{32} - 2$, inclusive.

Values of tone_map_id from 0 to 255 and from 512 to $2^{31} - 1$ may be used as determined by the application. Values of tone_map_id from 256 to 511 and from 2^{31} to $2^{32} - 2$ are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all tone mapping information SEI messages containing a value of tone_map_id in the range of 256 to 511 or in the range of 2^{31} to $2^{32} - 2$, and bitstreams shall not contain such values.

NOTE 1 – The tone_map_id can be used to support tone mapping operations that are suitable for different display scenarios. For example, different values of tone_map_id may correspond to different display bit depths.

tone_map_cancel_flag equal to 1 indicates that the tone mapping information SEI message cancels the persistence of any previous tone mapping information SEI message in output order. tone_map_cancel_flag equal to 0 indicates that tone mapping information follows.

tone_map_repetition_period specifies the persistence of the tone mapping information SEI message and may specify a picture order count interval within which another tone mapping information SEI message with the same value of tone_map_id or the end of the coded video sequence shall be present in the bitstream. The value of tone_map_repetition_period shall be in the range of 0 to 16 384, inclusive.

tone_map_repetition_period equal to 0 specifies that the tone map information applies to the current decoded picture only.

tone_map_repetition_period equal to 1 specifies that the tone map information persists in output order until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing a tone mapping information SEI message with the same value of tone_map_id is output having PicOrderCnt( ) greater than PicOrderCnt(CurrPic).

tone_map_repetition_period equal to 0 or equal to 1 indicates that another tone mapping information SEI message with the same value of tone_map_id may or may not be present.

tone_map_repetition_period greater than 1 specifies that the tone map information persists until any of the following conditions are true:

– A new coded video sequence begins.

– A picture in an access unit containing a tone mapping information SEI message with the same value of tone_map_id is output having PicOrderCnt( ) greater than PicOrderCnt(CurrPic) and less than or equal to PicOrderCnt(CurrPic) + tone_map_repetition_period.

tone_map_repetition_period greater than 1 indicates that another tone mapping information SEI message with the same value of tone_map_id shall be present for a picture in an access unit that is output having PicOrderCnt( ) greater than PicOrderCnt(CurrPic) and less than or equal to PicOrderCnt(CurrPic) + tone_map_repetition_period; unless the bitstream ends or a new coded video sequence begins without output of such a picture.

coded_data_bit_depth specifies the $BitDepth_Y$ for interpretation of the luma component of the associated pictures for purposes of interpretation of the tone mapping information SEI message. If tone mapping information SEI messages are present that have coded_data_bit_depth that is not equal to $BitDepth_Y$, these refer to the hypothetical result of a transcoding operation performed to convert the coded video to the $BitDepth_Y$ corresponding to the value of coded_data_bit_depth.

The value of coded_data_bit_depth shall be in the range of 8 to 14, inclusive. Values of coded_data_bit_depth from 0 to 7 and from 15 to 255 are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all tone mapping SEI messages that contain a coded_data_bit_depth in the range of 0 to 7 or in the range of 15 to 255, and bitstreams shall not contain such values.

target_bit_depth specifies the bit depth of the output of the dynamic range mapping function (or tone mapping function) described by the tone mapping information SEI message. The tone mapping function specified with a particular target_bit_depth is suggested to be reasonable for all display bit depths that are less than or equal to target_bit_depth.

The value of target_bit_depth shall be in the range of 1 to 16, inclusive. Values of target_bit_depth equal to 0 and in the range of 17 to 255 are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all tone mapping SEI messages that contain a value of target_bit_depth equal to 0 or in the range of 17 to 255, and bitstreams shall not contain such values.

tone_map_model_id specifies the model utilized for mapping the coded data into the target_bit_depth range. Values greater than 4 are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all tone mapping SEI messages that contain a value of tone_map_model_id greater than 4, and bitstreams shall not contain such values. Decoders shall allow reserved values of tone_map_model_id in the range of 5 to 16 384, inclusive, to be present in the bitstream.

NOTE 2 – A tone_map_model_id value of 0 corresponds to a linear mapping with clipping; a tone_map_model_id value of 1 corresponds to a sigmoidal mapping; a tone_map_model_id value of 2 corresponds to a user-defined table mapping, and a tone_map_model_id value of 3 corresponds to a piece-wise linear mapping tone_map_model_id value of 4 corresponds to luminance dynamic range information.

min_value specifies the RGB sample value that maps to the minimum value in the bit depth indicated by target_bit_depth. It is used in combination with the max_value parameter. All sample values of the decoded picture that are less than or equal to min_value after conversion to the RGB domain (when decoded in a different domain) are mapped to this minimum value in the target_bit_depth representation.

max_value specifies the RGB sample value that maps to the maximum value in the bit depth indicated by target_bit_depth. It is used in combination with the min_value parameter. All sample values of the decoded picture that are greater than or equal to max_value after conversion to the RGB domain (when decoded in a different domain) are mapped to this maximum value in the target_bit_depth representation.

max_value shall be greater than or equal to min_value.

sigmoid_midpoint specifies the RGB sample value of the coded data that is mapped to the centre point of the target_bit_depth representation. It is used in combination with the sigmoid_width parameter.

sigmoid_width specifies the distance between two coded data values that approximately correspond to the 5% and 95% values of the target_bit_depth representation, respectively. It is used in combination with the sigmoid_midpoint parameter and is interpreted according to the following function:

$$f(i) = Round\left(\frac{2^{target\_bit\_depth} - 1}{1 + \exp\left(\frac{-6 * (i - sigmoid\_midpoint)}{sigmoid\_width}\right)}\right) \quad for\ i = 0, \cdots, (2^{coded\_bit\_depth} - 1) \qquad (D\text{-}28)$$

where *f*(i) denotes the function that maps an RGB sample value i from the coded data to a resulting RGB sample value in the target_bit_depth representation.

start_of_coded_interval[i] specifies the beginning point of an interval in the coded data such that all RGB sample values that are greater than or equal to start_of_coded_interval[i] and less than start_of_coded_interval[i + 1] are mapped to i in the target bit depth representation. The value of start_of_coded_interval[$2^{target\_bit\_depth}$] is equal to $2^{coded\_bit\_depth}$. The number of bits used for the representation of the start_of_coded_interval is ((coded_data_bit_depth + 7) >> 3) << 3.

num_pivots specifies the number of pivot points in the piece-wise linear mapping function without counting the two default end points, (0, 0) and ($2^{coded\_data\_bit\_depth}$ − 1, $2^{target\_bit\_depth}$ − 1).

coded_pivot_value[i] specifies the value in the coded_data_bit_depth corresponding to the i-th pivot point. The number of bits used for the representation of the coded_pivot_value is ((coded_data_bit_depth + 7) >> 3) << 3.

target_pivot_value[i] specifies the value in the reference target_bit_depth corresponding to the i-th pivot point. The number of bits used for the representation of the target_pivot_value is ((target_bit_depth + 7) >> 3) << 3.

camera_iso_speed_idc indicates the camera ISO speed for daylight illumination as specified by ISO 12232, interpreted as specified by Table D-8. When camera_iso_speed_idc indicates Extended_ISO, the ISO speed is represented by camera_iso_speed_value. Values identified as reserved in Table D-8 are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore (remove from the bitstream and discard) all tone mapping information SEI messages containing a reserved values of camera_iso_speed_idc, and bitstreams shall not contain such values.

camera_iso_speed_value indicates the camera ISO speed for daylight illumination as specified by ISO 12232 when camera_iso_speed_idc is set to Extended_ISO. The value of camera_iso_speed_value shall not be equal to 0.

exposure_index_idc indicates the exposure index setting of the camera as specified by ISO 12232, interpreted as specified by Table D-8. When exposure_index_idc indicates Extended_ISO, the exposure index is indicated by exposure_index_value.

The values of camera_iso_speed_idc and exposure_index_idc in the range of 31 to 254, inclusive, are reserved for future use by ITU-T | ISO/IEC, and shall not be present in bitstreams conforming to this version of this Specification. Decoders conforming to this version of this Specification shall ignore tone mapping SEI messages that contain these values.

exposure_index_value indicates the exposure index setting of the camera as specified by ISO 12232 when exposure_index_idc indicates Extended_ISO. The value of exposure_index_value shall not be equal to 0.

Table D-8 – Interpretation of camera_iso_speed_idc and exposure_index_idc

camera_iso_speed_idc or exposure_index_idc	Indicated value
0	Unspecified
1	10
2	12
3	16
4	20
5	25
6	32
7	40
8	50
9	64
10	80
11	100
12	125
13	160
14	200
15	250
16	320
17	400
18	500
19	640
20	800
21	1000
22	1250
23	1600
24	2000
25	2500
26	3200
27	4000
28	5000
29	6400
30	8000
31..254	Reserved
255	Extended_ISO

exposure_compensation_value_sign_flag, when applicable as specified below, specifies the sign of the variable ExposureCompensationValue that indicates the exposure compensation value setting used for the process of image production.

exposure_compensation_value_numerator, when applicable as specified below, specifies the numerator of the variable ExposureCompensationValue that indicates the exposure compensation value setting used for the process of image production.

exposure_compensation_value_denom_idc, when not equal to 0, specifies the denominator of the variable ExposureCompensationValue that indicates the exposure compensation value setting used for the process of image production.

When exposure_compensation_value_denom_idc is not equal to 0, the variable ExposureCompensationValue is derived from exposure_compensation_value_sign_flag, exposure_compensation_value_numerator and exposure_compensation_value_denom_idc. exposure_compensation_value_sign_flag equal to 0 indicates that the

ExposureCompensationValue is positive. exposure_compensation_value_sign_flag equal to 1 indicates that the ExposureCompensationValue is negative. When ExposureCompensationValue is positive, the image is indicated to have been further sensitized through the process of production, relative to the recommended exposure index of the camera as specified by ISO 12232. When ExposureCompensationValue is negative, the image is indicated to have been further desensitized through the process of production, relative to the recommended exposure index of the camera as specified by ISO 12232.

When exposure_compensation_value_denom_idc is not equal to 0, the variable ExposureCompensationValue is derived as follows:

$$\text{ExposureCompensationValue} = (1 - 2 * \text{exposure\_compensation\_value\_sign\_flag}) * \text{exposure\_compensation\_value\_numerator} \div \text{exposure\_compensation\_value\_denom\_idc} \qquad \text{(D-29)}$$

The value of ExposureCompensationValue is interpreted in units of exposure steps such that an increase of 1 in ExposureCompensationValue corresponds to a doubling of exposure in units of lux-seconds. For example, the exposure compensation value equal to +1/2 at the production stage may be indicated by setting exposure_compensation_value_sign_flag to 0, exposure_compensation_value_numerator to 1, and exposure_compensation_value_denom_idc to 2.

When exposure_compensation_value_denom_idc is equal to 0, the exposure compensation value is indicated as unknown or unspecified.

ref_screen_luminance_white indicates the reference screen brightness setting for the nominal white level used for image production process in units of candela per square metre.

extended_range_white_level indicates the luminance dynamic range for extended dynamic-range display of the associated pictures, after conversion to the linear light domain for display, expressed as an integer percentage relative to the nominal white level.

The value of extended_range_white_level should be greater than or equal to 100.

nominal_black_level_luma_code_value specifies the luma sample value of the associated decoded pictures to which the nominal black level is assigned. For example, when coded_data_bit_depth is equal to 8, video_full_range_flag is equal to 0, and matrix_coefficients is equal to 1, nominal_black_level_luma_code_value should be equal to 16.

nominal_white_level_luma_code_value specifies the luma sample value of the associated decoded pictures to which the nominal white level is assigned. For example, when coded_data_bit_depth is equal to 8, video_full_range_flag is equal to 0, and matrix_coefficients is equal to 1, nominal_white_level_luma_code_value should be equal to 235.

The value of nominal_white_level_luma_code_value shall be greater than nominal_black_level_luma_code_value.

extended_white_level_luma_code_value specifies the luma sample value of the associated decoded pictures to which the white level associated with an extended dynamic range is assigned.

The value of extended_white_level_luma_code_value shall be greater than or equal to nominal_white_level_luma_code_value.

”

Renumber the equations that follow the above subclause to account for the equation added above.

Replace the text in E.2.1, including Tables E-1 through E-5, up to but not including the semantics of chroma_loc_info_present_flag, with the following:

“

aspect_ratio_info_present_flag equal to 1 specifies that aspect_ratio_idc is present. aspect_ratio_info_present_flag equal to 0 specifies that aspect_ratio_idc is not present.

aspect_ratio_idc specifies the value of the sample aspect ratio of the luma samples. Table E-1 shows the meaning of the code. When aspect_ratio_idc indicates Extended_SAR, the sample aspect ratio is represented by sar_width : sar_height. When the aspect_ratio_idc syntax element is not present, aspect_ratio_idc value shall be inferred to be equal to 0.

Table E-1 – Meaning of sample aspect ratio indicator

aspect_ratio_idc	Sample aspect ratio	(informative) Examples of use
0	Unspecified	
1	1:1 ("square")	7680x4320 16:9 frame without horizontal overscan 3840x2160 16:9 frame without horizontal overscan 1920x1080 16:9 frame without horizontal overscan (cropped from 1920x1088) 1280x720 16:9 frame without horizontal overscan 640x480 4:3 frame without horizontal overscan
2	12:11	720x576 4:3 frame with horizontal overscan 352x288 4:3 frame without horizontal overscan
3	10:11	720x480 4:3 frame with horizontal overscan 352x240 4:3 frame without horizontal overscan
4	16:11	720x576 16:9 frame with horizontal overscan 528x576 4:3 frame without horizontal overscan
5	40:33	720x480 16:9 frame with horizontal overscan 528x480 4:3 frame without horizontal overscan
6	24:11	352x576 4:3 frame without horizontal overscan 480x576 16:9 frame with horizontal overscan
7	20:11	352x480 4:3 frame without horizontal overscan 480x480 16:9 frame with horizontal overscan
8	32:11	352x576 16:9 frame without horizontal overscan
9	80:33	352x480 16:9 frame without horizontal overscan
10	18:11	480x576 4:3 frame with horizontal overscan
11	15:11	480x480 4:3 frame with horizontal overscan
12	64:33	528x576 16:9 frame without horizontal overscan
13	160:99	528x480 16:9 frame without horizontal overscan
14	4:3	1440x1080 16:9 frame without horizontal overscan
15	3:2	1280x1080 16:9 frame without horizontal overscan
16	2:1	960x1080 16:9 frame without horizontal overscan
17..254	Reserved	
255	Extended_SAR	

NOTE 1 – For the examples in Table E-1, the term "without horizontal overscan" refers to display processes in which the display area matches the area of the cropped decoded pictures and the term "with horizontal overscan" refers to display processes in which some parts near the left and/or right border of the cropped decoded pictures are not visible in the display area. As an example, the entry "720x576 4:3 frame with horizontal overscan" for aspect_ratio_idc equal to 2 refers to having an area of 704x576 luma samples (which has an aspect ratio of 4:3) of the cropped decoded frame (720x576 luma samples) that is visible in the display area.

sar_width indicates the horizontal size of the sample aspect ratio (in arbitrary units).

sar_height indicates the vertical size of the sample aspect ratio (in the same arbitrary units as sar_width).

sar_width and sar_height shall be relatively prime or equal to 0. When aspect_ratio_idc is equal to 0 or sar_width is equal to 0 or sar_height is equal to 0, the sample aspect ratio shall be considered unspecified by this Recommendation | International Standard.

overscan_info_present_flag equal to 1 specifies that the overscan_appropriate_flag is present. When overscan_info_present_flag is equal to 0 or is not present, the preferred display method for the video signal is unspecified.

overscan_appropriate_flag equal to 1 indicates that the cropped decoded pictures output are suitable for display using overscan. overscan_appropriate_flag equal to 0 indicates that the cropped decoded pictures output contain visually