INTERNATIONAL STANDARD

ISO/IEC 13818-2

First edition
1996-05-15

AMENDMENT 3
1998-12-15

Information technology — Generic coding of moving pictures and associated audio information: Video

AMENDMENT 3

Technologies de l'information — Codage générique des images animées et du son associé: Données vidéo

AMENDEMENT 3

Reference number
ISO/IEC 13818-2:1996/Amd.3:1998(E)

Contents

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. 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.

Amendment 3 to ISO/IEC 13818-2:1996 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. The identical text is published as ITU-T Rec. H.262/Amd.3.

INTERNATIONAL STANDARD

ITU-T RECOMMENDATION

INFORMATION TECHNOLOGY – GENERIC CODING OF MOVING PICTURES AND ASSOCIATED AUDIO INFORMATION: VIDEO

AMENDMENT 3

1) Subclause 6.2.2.2.1

Replace 6.2.2.2.1 by the following:

6.2.2.2.1 Extension data

extension_data(i) {	No. of bits	Mnemonic
while (nextbits()== extension_start_code) {		
extension_start_code	32	bslbf
if (i == 0) { /* follows sequence_extension() */		
if (nextbits()== "Sequence Display Extension ID")		
sequence_display_extension()		
else if (nextbits()		
== "Sequence Scalable Extension ID")		
sequence_scalable_extension()		
}		
/* NOTE – i never takes the value 1 because extension_data()		
never follows a group_of_pictures_header() */		
if (i == 2) { /* follows picture_coding_extension() */		
if (nextbits() == "Quant Matrix Extension ID")		
quant_matrix_extension()		
else if (nextbits() == "Copyright Extension ID")		
copyright_extension()		
else if (nextbits() == "Picture Display Extension ID")		
picture_display_extension()		
else if (nextbits()		
== "Picture Spatial Scalable Extension ID")		
picture_spatial_scalable_extension()		
else if (nextbits()		
== "Picture Temporal Scalable Extension ID")		
picture_temporal_scalable_extension()		
else if (nextbits()		
== "Camera Parameters Extension ID")		
camera_parameters_extension()		
}		
}		
}		

2) Subclause 6.3.1

Replace Table 6-2 by:

Table 6-2 – extension_start_code_identifier codes

extension_start_code_identifier	Name
0000	Reserved
0001	Sequence Extension ID
0010	Sequence Display Extension ID
0011	Quant Matrix Extension ID
0100	Copyright Extension ID
0101	Sequence Scalable Extension ID
0110	Reserved
0111	Picture Display Extension ID
1000	Picture Coding Extension ID
1001	Picture Spatial Scalable Extension ID
1010	Picture Temporal Scalable Extension ID
1011	Camera Parameters Extension ID
1100	Reserved
…	…
1111	Reserved

3) New subclause 6.2.3.7.1

Insert new subclause 6.2.3.7.1:

6.2.3.7.1 Camera parameters extension

camera_parameters_extension() {	No. of bits	Mnemonic
extension_start_code_identifier	4	uimsbf
reserved	1	uimsbf
camera_id	7	simsbf
marker_bit	1	bslbf
height_of_image_device	22	uimsbf
marker_bit	1	bslbf
focal_length	22	uimsbf
marker_bit	1	bslbf
f_number	22	uimsbf
marker_bit	1	bslbf
vertical_angle_of_view	22	uimsbf
marker_bit	1	bslbf
camera_position_x_upper	16	simsbf
marker_bit	1	bslbf
camera_position_x_lower	16	
marker_bit	1	bslbf
camera_position_y_upper	16	simsbf
marker_bit	1	bslbf
camera_position_y_lower	16	
marker_bit	1	bslbf
camera_position_z_upper	16	simsbf
marker_bit	1	bslbf
camera_position_z_lower	16	
marker_bit	1	bslbf
camera_direction_x	22	simsbf
marker_bit	1	bslbf
camera_direction_y	22	simsbf
marker_bit	1	bslbf
camera_direction_z	22	simsbf
marker_bit	1	bslbf
image_plane_vertical_x	22	simsbf
marker_bit	1	bslbf
image_plane_vertical_y	22	simsbf
marker_bit	1	bslbf
image_plane_vertical_z	22	simsbf
marker_bit	1	bslbf
reserved	32	bslbf
next_start_code()		
}		

4) New subclause 6.3.19

Insert new subclause 6.3.19:

6.3.19 Camera parameters extension

camera_id – The number in camera_id identifies a camera.

height_of_image_device – This is a 22-bit unsigned integer which specifies the height of image device. Its value shall be measured to a resolution of 0.001 millimeter and having a range of zero to 4,194.303 mm.

focal_length – This is a 22-bit unsigned integer which specifies the focal length. Its value shall be measured to a resolution of 0.001 millimeter and having a range of zero to 4,194.303 mm.

f_number – This is a 22-bit unsigned integer which specifies the F-number. F-number is defined by (focal_length)/(effective aperture of lens). Its value shall be measured to a resolution of 0.001 and having a range of zero to 4,194.303.

vertical_angle_of_view – This is a 22-bit unsigned integer which specifies the vertical angle of the field of view as determined between the top and bottom edges of the image device. Its value shall be measured to a resolution of 0.0001 degree and having a range of zero to 180 degrees.

camera_position_x_upper, camera_position_y_upper, camera_position_z_upper – These words constitute the 16 most significant bits of camera_position_x, camera_position_y and camera_position_z respectively.

camera_position_x_lower, camera_position_y_lower, camera_position_z_lower – These words constitute the 16 least significant bits of camera_position_x, camera_position_y and camera_position_z respectively.

camera_position_x, camera_position_y, camera_position_z – A set of these values specifies the position of the optical principal point of the camera in a user-specified world coordinate system. Each of these values shall be measured to a resolution of 0.001 millimeter and having a range of +2,147,483.647 mm to –2,147,483.648 mm. The camera_position_x is a 32-bit signed (two's complement) integer, the 16 least significant bits are defined in camera_position_x_lower, the 16 most significant bits are defined in camera_position_x_upper. The camera_position_y is a 32-bit signed (two's complement) integer, the 16 least significant bits are defined in camera_position_y_lower, the 16 most significant bits are defined in camera_position_y_upper. The camera_position_z is a 32-bit signed (two's complement) integer, the 16 least significant bits are defined in camera_position_z_lower, the 16 most significant bits are defined in camera_position_z_upper.

camera_direction_x, camera_direction_y, camera_direction_z – A set of these values specifies the direction of the camera. The direction of the camera is defined by using the vector from optical principal point to a point which is in front of the camera and is on the optical axis of the camera. Each of these values is a 22-bit signed (two's complement) integer and having a range of +2,097,151 to –2,097,152.

image_plane_vertical_x, image_plane_vertical_y, image_plane_vertical_z – A set of these values specifies the upper direction of the camera. The upper direction of the camera is defined by using the vector which is parallel to the side edge of the image device and is from bottom edge to top edge. Each of these values is a 22-bit signed (two's complement) integer and is having a range of +2,097,151 to –2,097,152.

Figures in Appendix I explain these terms pictorially.

5) Clause 8

a) Replace Table 8-4 by:

Table 8-4 – Escape profile_and_level_indication identification

profile_and_level_indication	Name
10001111 to 11111111	(Reserved)
10001110	Multi-view profile @ Low level
10001101	Multi-view profile @ Main level
10001100	(Reserved)
10001011	Multi-view profile @ High1440 level
10001010	Multi-view profile @ High level
10000110 to 10001001	(Reserved)
10000101	4:2:2 profile @ Main level
10000000 to 10000100	(Reserved)

b) Add the following text as a Note after the Note on 4:2:2 Profile (as indicated in Amendment 2); accordingly, the Note on 4:2:2 Profile shall be renamed "NOTE 1 – On 4:2:2 Profile ":

NOTE 2 – On Multi-view Profile: The Multi-view Profile (MVP) is envisioned to be a profile appropriate for applications that require multiple viewpoints within the context of ITU-T Rec. H.262 | ISO/IEC 13818-2. MVP supports stereoscopic pictures as its source images for a wide range of picture resolution and quality as requested by the applications to be used. A base layer of MVP is assigned to a left view and an enhancement layer is assigned to a right view.

A monoscopic coding with the same tools as Main Profile (MP), including ISO/IEC IS 11172-2, is applied to the base layer. An enhancement layer is coded using Temporal Scalability tools and a hybrid prediction of motions and disparity can be utilized in the enhanced layer.

MVP, viewed as one of the scalable profiles in terms of multiple viewpoint layers, is expected to have the same type of compatibility features other scalable profiles have, such as compatibility with MP. For example:

1) decoders compliant to MVP at a certain Level are capable of decoding the bitstreams compliant to MP at the corresponding Level (i.e. forward compatibility);

2) decoders compliant to MP at a certain Level are capable of decoding the bitstream in the base layer of MVP (i.e. backward compatibility).

6) Subclause 8.2

Replace Table 8-5 by:

Table 8-5 – Syntactic constraints of profiles

Syntactic Element	Profile						
	Simple	Main	SNR	Spatial	High	4:2:2	Multi-view
chroma_format	4:2:0	4:2:0	4:2:0	4:2:0	4:2:2 or 4:2:0	4:2:2 or 4:2:0	4:2:0
frame_rate_extension_n	0	0	0	0	0	0	0
frame_rate_extension_d	0	0	0	0	0	0	0
aspect_ratio_information	0001, 0010, 0011	0001, 0010, 0011	0001, 0010, 0011	0001, 0010, 0011	0001, 0010, 0011	0001, 0010, 0011	0001, 0010, 0011
picture_coding_type	I, P	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B
repeat_first_field	Constrained		Unconstrained			Constrained	Unconstrained
sequence_scalable_ extension()	No	No	Yes	Yes	Yes	No	Yes
scalable_mode	–	–	SNR	SNR or Spatial	SNR or Spatial	–	Temporal
picture_spatial_scalable_ extension()	No	No	No	Yes	Yes	No	No
picture_temporal_scalable_ extension()	No	No	No	No	No	No	Yes
intra_dc_precision	8, 9, 10	8, 9, 10	8, 9, 10	8, 9, 10	8, 9, 10, 11	8, 9, 10, 11	8, 9, 10
Slice structure	Restricted 6.1.2.2						

7) Subclause 8.4

Replace the text and Table 8-9 by:

The SNR Scalable, Spatial Scalable, High and Multi-view profiles may use more than one bitstream to code the image. These different bitstreams represent layers of coding, which when combined create a higher quality image than that obtainable from one layer alone (see Annex D). The maximum number of layers for a given profile is specified in Table 8-9. The scalable layers are named according to Table 7-31. The syntactic and parameter constraints for these profile/level combinations when coded using the maximum permitted number of layers are given in Tables 8-11, 8-12, 8-13 and 8-14. When the number of layers is less than the maximum permitted, reference should also be made to Tables E-21 to E-46 as appropriate.

It should be noted that a bitstream of the base layer of SNR Scalable and Multi-view profiles can always be decoded by a Main profile decoder of equivalent level. Conversely, a Main profile bitstream shall be decodable by either SNR Scalable or Multi-view profile decoder of equivalent level.

Table 8-9 – Upper bounds for scalable layers in SNR Scalable, Spatial Scalable, High and Multi-view profiles

Level	Maximum Number of	Profile			
		SNR	Spatial	High	Multi-view
High	All layers (base + enhancement) Spatial enhancement layers SNR enhancement layers Temporal auxiliary layers			3 1 1 0	2 0 0 1
High-1440	All layers (base + enhancement) Spatial enhancement layers SNR enhancement layers Temporal auxiliary layers		3 1 1 0	3 1 1 0	2 0 0 1
Main	All layers (base + enhancement) Spatial enhancement layers SNR enhancement layers Temporal auxiliary layers	2 0 1 0		3 1 1 0	2 0 0 1
Low	All layers (base + enhancement) Spatial enhancement layers SNR enhancement layers Temporal auxiliary layers	2 0 1 0			2 0 0 1

8) Subclause 8.4.1

Replace the text and Table 8-10 by:

Table 8-10 is a summary of the permitted combinations, and is subject to the following rules:

- SNR Scalable and Multi-view profile: maximum of 2 layers; Spatial Scalable and High profile: maximum of 3 layers. (See Table 8-9.)
- Only one SNR and one Spatial scale allowed in 3-layer combinations, either SNR/Spatial or Spatial/SNR order is permitted. (See Table 8-9.)

- Adding 4:2:2 chroma format to a 4:2:0 lower layer is considered an SNR enhancement permitted for either SNR or Spatial scale.
- A 4:2:0 layer is not permitted if the lower layer is 4:2:2. (See 7.7.3.3.)

Table 8-10 – Permissible layer combinations

Profile	Scalable mode			Profile/level of simplest base layer decoder (level reference top layer) [a]
	Base layer	Enhancement layer 1	Enhancement layer 2	
SNR	4:2:0	SNR, 4:2:0	–	MP@same level
Spatial	4:2:0	SNR, 4:2:0	–	MP@same level
Spatial	4:2:0	Spatial, 4:2:0	–	MP@(level – 1)
Spatial	4:2:0	SNR, 4:2:0	Spatial, 4:2:0	MP@(level – 1)
Spatial	4:2:0	Spatial, 4:2:0	SNR, 4:2:0	MP@(level – 1)
High	4:2:0	–	–	HP@same level
High	4:2:2	–	–	HP@same level
High	4:2:0	SNR, 4:2:0	–	HP@same level
High	4:2:0	SNR, 4:2:2	–	HP@same level
High	4:2:2	SNR, 4:2:2	–	HP@same level
High	4:2:0	Spatial, 4:2:0	–	HP@(level – 1)
High	4:2:0	Spatial, 4:2:2	–	HP@(level – 1)
High	4:2:2	Spatial, 4:2:2	–	HP@(level – 1) [b]
High	4:2:0	SNR, 4:2:0	Spatial, 4:2:0	HP@(level – 1)
High	4:2:0	SNR, 4:2:0	Spatial, 4:2:2	HP@(level – 1)
High	4:2:0	SNR, 4:2:2	Spatial, 4:2:2	HP@(level – 1) [b]
High	4:2:2	SNR, 4:2:2	Spatial, 4:2:2	HP@(level – 1) [b]
High	4:2:0	Spatial, 4:2:0	SNR, 4:2:0	HP@(level – 1)
High	4:2:0	Spatial, 4:2:0	SNR, 4:2:2	HP@(level – 1)
High	4:2:0	Spatial, 4:2:2	SNR, 4:2:2	HP@(level – 1)
High	4:2:2	Spatial, 4:2:2	SNR, 4:2:2	HP@(level – 1) [b]
Multi-view	4:2:0	Temporal, 4:2:0	–	MP@same level

[a] The simplest compliant decoder to decode the base layer is specified, assuming that bitstream may contain any syntax and parameter value permitted for the stated profile @ level, except scalability. Note that for High profile @ Main level spatially scaled bitstreams, 'HP @ (level – 1)' becomes 'MP @ (level – 1)'. In the event that a base layer bitstream uses fewer syntactic elements or a reduced parameter range than permitted, profile_and_level_indication may indicate a 'simpler' profile @ level.

[b] Note that 4:2:2 chroma format is not supported as a lower spatial layer of High profile @ Main level (see Table 8-12).

9) New subclause 8.4.2

Add the following new subclause:

8.4.2 Multi-view Profile specific constraints

Both the enhancement and base layers have the same frame rate.

The picture_mux_enable, picture_mux_order and picture_mux_factor are not used in this profile and shall be ignored.

The reference_select_code should be "00" or "01" for the P-frames in the enhancement layer. The reference_select_code should be "01" for B-frames in the enhancement layer.

If the base layer coded frame is the first frame of the Group Of Pictures then the corresponding frame in the enhancement layer should be either I-frame or P-frame with the reference_select_code value of “01”.

In a P-field picture with reference_select_code = "01" and which is the first field of a frame, the following restriction applies:

- Dual prime prediction shall not be used.
- Field prediction in which motion_vertical_field_select indicates the second field of the base layer frame shall not be used.
- If base and enhancement layers do not have the same value for top_field_first, there shall be no macroblocks that are coded with macroblock_motion_forward zero and macroblock_intra zero.
- If base and enhancement layer do not have the same value for top_field_first, there shall be no skipped macroblocks.