INTERNATIONAL STANDARD

ISO/IEC 14496-10

Eighth edition
2014-09-01

AMENDMENT 3
2016-12-15

Information technology — Coding of audio-visual objects —

Part 10: Advanced Video Coding

AMENDMENT 3: Additional supplemental enhancement information

Technologies de l'information — Codage des objets audiovisuels —

Partie 10: Codage visuel avancé

AMENDEMENT 3: Information d'amélioration supplémentaire additionnelle

Reference number
ISO/IEC 14496-10:2014/Amd.3:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

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.

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).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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

For an explanation on 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 the following URL: www.iso.org/iso/foreword.html.

Amendment 3 to ISO/IEC 14496-10:2014 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — Coding of audio-visual objects —

Part 10: Advanced Video Coding

AMENDMENT 3: Additional supplemental enhancement information

Page 346, D.1

Insert the following after the row containing “mastering_display_colour_volume(payloadSize)”:

else if(payloadType = = 181)		
alternative_depth_info(payloadSize) /* specified in Annex I */	5	

Page 760, after I.13.1.6.1

Add the following subclause:

I.13.1.7 Alternative depth information SEI message syntax

alternative_depth_info(payloadSize) {	C	Descriptor
depth_type	5	ue(v)
if(depth_type = = 0) {		
num_constituent_views_gvd_minus1	5	ue(v)
depth_present_gvd_flag	5	u(1)
z_gvd_flag	5	u(1)
intrinsic_param_gvd_flag	5	u(1)
rotation_gvd_flag	5	u(1)
translation_gvd_flag	5	u(1)
if(z_gvd_flag)		
for(i = 0; i <= num_constituent_views_gvd_minus1 + 1; i++) {		
sign_gvd_z_near_flag[i]	5	u(1)
exp_gvd_z_near[i]	5	u(7)
man_len_gvd_z_near_minus1[i]	5	u(5)
man_gvd_z_near[i]	5	u(v)
sign_gvd_z_far_flag[i]	5	u(1)
exp_gvd_z_far[i]	5	u(7)
man_len_gvd_z_far_minus1[i]	5	u(5)
man_gvd_z_far[i]	5	u(v)
}		
if(intrinsic_param_gvd_flag) {		
prec_gvd_focal_length	5	ue(v)
prec_gvd_principal_point	5	ue(v)

}		
if(rotation_gvd_flag)		
prec_gvd_rotation_param	5	ue(v)
if (translation_gvd_flag)		
prec_gvd_translation_param	5	ue(v)
for(i = 0; i <= num_constituent_views_gvd_minus1 + 1; i++) {		
if(intrinsic_param_gvd_flag) {		
sign_gvd_focal_length_x[i]	5	u(1)
exp_gvd_focal_length_x[i]	5	u(6)
man_gvd_focal_length_x[i]	5	u(v)
sign_gvd_focal_length_y[i]	5	u(1)
exp_gvd_focal_length_y[i]	5	u(6)
man_gvd_focal_length_y[i]	5	u(v)
sign_gvd_principal_point_x[i]	5	u(1)
exp_gvd_principal_point_x[i]	5	u(6)
man_gvd_principal_point_x[i]	5	u(v)
sign_gvd_principal_point_y[i]	5	u(1)
exp_gvd_principal_point_y[i]	5	u(6)
man_gvd_principal_point_y[i]	5	u(v)
}		
if(rotation_gvd_flag)		
for (j = 0; j < 3; j++) /* row */		
for(k = 0; k < 3; k++) { /* column */		
sign_gvd_r[i][j][k]	5	u(1)
exp_gvd_r[i][j][k]	5	u(6)
man_gvd_r[i][j][k]	5	u(v)
}		
if (translation_gvd_flag) {		
sign_gvd_t_x[i]	5	u(1)
exp_gvd_t_x[i]	5	u(6)
man_gvd_t_x[i]	5	u(v)
}		
}		
}		
}		

Page 770, after I.13.2.6.1

Add the following subclause:

I.13.2.7 Alternative depth information SEI message semantics

The alternative depth information SEI message indicates that the view components of one output view contain samples of multiple distinct spatially packed constituent pictures. The related output view is not suitable to be displayed directly. This SEI message can be used by the application after the decoder

output to rearrange the samples to produce additional views that are appropriate for display or use by a server for other purposes (which are outside the scope of this document).

When the alternative depth information SEI message is present, the texture and depth views referred to by this SEI message shall conform to the specifications in Annexes I and J, ChromaArrayType shall be equal to 1 (4:2:0), bit_depth_luma_minus8 shall be equal to 0, bit_depth_chroma_minus8 shall be equal to 0, and the depth representation information SEI message shall be present. The depth_representation_type is defined in the depth representation information SEI message. The depth representation is only carried in the luma component.

The alternative depth information SEI message, when present, shall be associated with an IDR access unit. The information signalled in this SEI message applies to all the access units in the coded video sequence. Also, num_anchor_refs_l0[i], num_anchor_refs_l1[i], num_non_anchor_refs_l0[i] and num_non_anchor_refs_l1[i] shall be equal to 0 in the active SPS RBSP syntax structure for the coded video sequence and avc_3d_extension_flag shall be equal to 0 in all VCL NAL units of the coded video sequence.

NOTE 1 These constraints disable inter-view and inter-component prediction.

In such a coded video sequence, there shall be two output views. The base view is a complete view and the non-base view contains a packing arrangement of 1 to 4 additional views with half the width and half the height of the base view. Such lower-resolution views are referred to as constituent pictures herein.

NOTE 2 The view_id of the non-base view is not used in any particular way for this SEI message.

All constituent pictures have a width and a height equal to ((pic_width_in_mbs_minus1 + 1) * 8) and ((pic_height_in_map_unit_minus1 + 1) * 8) in luma samples, respectively. frame_mbs_only_flag shall be equal to 1. The variable i, with a value from 0 to num_constituent_views_gvd_minus1 + 1, inclusive, indicates the location of the constituent pictures in the non-base texture view as specified in Table I-4. i equal to 0 indicates the base texture view. i greater than 0 indicates the constituent texture view.

Table I-4 — Locations of the top-left luma samples of constituent pictures packed in a non-base texture view relative to the top-left luma sample of this picture

Constituent picture index i	Location of the top-left luma sample in a non-base texture view
1	(0, 0)
2	(0, (pic_height_in_map_unit_minus1 + 1) * 8)
3	((pic_width_in_mbs_minus1 + 1) * 8, 0)
4	((pic_width_in_mbs_minus1 + 1) * 8, (pic_height_in_map_unit_minus1 + 1) * 8)

The decoded depth views have the same structure as the decoded texture views, i.e., they consist of a base depth view (i equal to 0) and a non-base depth view. The non-base depth view contains up to four constituent depth pictures (i in the range of 1 to 4) for the constituent views. The constituent depth pictures are packed in the non-base depth view in the same arrangement as the constituent texture pictures as specified in Table I-4.

This arrangement of texture and depth constituent views is referred to as global view and depth (GVD) information.

depth_type shall be equal to 0. Other values are reserved for future use by ITU-T | ISO/IEC and shall not be present in bitstreams conforming to this document. Decoders shall ignore alternative depth information SEI messages in which such other values are present.

num_constituent_views_gvd_minus1 plus 1 identifies the number of constituent texture pictures packed into each texture component of the non-base view. num_constituent_views_gvd_minus1 shall be in the range of 0 to 3, inclusive.

depth_present_gvd_flag equal to 1 indicates that constituent depth pictures are packed into the depth components of the non-base view, with a packing arrangement as described above. depth_present_gvd_flag equal to 0 specifies that the depth component of the non-base view is not present.

Each constituent picture in the depth component of the non-base view is associated with a constituent picture in the texture component of the non-base view in the same relative location. The number of depth views in the coded video sequence is equal to 1 + depth_present_gvd_flag. The depth component of the base view shall always be present, independent of the value of depth_present_gvd_flag.

NOTE 3 The following SEI message parameters can be used along with the decoded depth components to project samples from the base view into the coordinates of constituent views such that reconstructed views can be generated by combining projected samples and samples from the constituent views.

The function binToFp(s, e, n, v) is specified as follows:

$$\text{binToFp}(s, e, n, v) = (-1)^{s} * (e == 0 \ ? \ (2^{-(30 + v)} * n) : (2^{e - 31} * (1 + n \div 2^{v}))) \quad \text{(I-1)}$$

NOTE 4 The above specification is similar to what is found in IEC 60559:1989, Binary floating-point arithmetic for microprocessor systems.

z_gvd_flag equal to 1 indicates the presence of the syntax elements sign_gvd_z_near_flag[i], exp_gvd_z_near[i], man_len_gvd_z_near_minus1[i], man_gvd_z_near[i], sign_gvd_z_far_flag[i], exp_gvd_z_far[i], man_len_gvd_z_far_minus1[i], and man_gvd_z_far[i], for i in the range of 0 to num_constituent_views_gvd_minus1 + 1, inclusive. z_gvd_flag equal to 0 indicates that these syntax elements are not present.

intrinsic_param_gvd_flag equal to 1 indicates the presence of intrinsic camera parameter syntax elements. intrinsic_param_gvd_flag equal to 0 indicates that these syntax elements are not present.

rotation_gvd_flag equal to 1 indicates the presence of rotation camera parameter syntax elements. rotation_gvd_flag equal to 0 indicates that these syntax elements are not present. When rotation_gvd_flag is 0, a default rotation camera parameter of a unit matrix value is inferred.

translation_gvd_flag equal to 1 indicates the presence of horizontal translation camera parameter syntax elements. translation_gvd_flag equal to 0 indicates that these syntax elements are not present.

sign_gvd_z_near_flag[i] equal to 0 indicates that the sign of the nearest depth value of the i-th camera is positive. sign_gvd_z_near[i] equal to 1 indicates that the sign of the nearest depth value of the i-th camera is negative.

exp_gvd_z_near[i] specifies the exponent part of the nearest depth value of the i-th camera. The value of exp_gvd_z_near[i] shall be in the range of 0 to 126, inclusive. The value 127 is reserved for future use by ITU-T | ISO/IEC. When exp_gvd_z_near[i] is equal to 127, the value of zNear[i] is unspecified.

man_len_gvd_z_near_minus1[i] + 1 specifies the length in bits of the mantissa of the nearest depth value of the i-th camera. The value of man_len_gvd_z_near_minus1 [i] shall be in the range of 0 to 31, inclusive.

man_gvd_z_near[i] specifies the mantissa part of the nearest depth value of the i-th camera. The length of man_gvd_z_near[i] syntax elements is man_len_gvd_z_near_minus1[i] + 1 bits.

When exp_gvd_z_near[i] is not equal to 127, zNear[i] is set equal to binToFp(sign_gvd_z_near_flag[i], exp_gvd_z_near[i], man_gvd_z_near[i], man_len_gvd_z_near_minus1[i] + 1).

sign_gvd_z_far_flag[i] equal to 0 indicates that the sign of the farthest depth value of the i-th camera is positive. sign_gvd_z_far_flag[i] equal to 1 indicates that the sign of the farthest depth value of the i-th camera is negative.

exp_gvd_z_far[i] specifies the exponent part of the farthest depth value of the i-th camera. The value of exp_gvd_z_far[i] shall be in the range of 0 to 126, inclusive. The value 127 is reserved for future use by ITU-T | ISO/IEC. When exp_gvd_z_far[i] is equal to 127, the value of zFar[i] is unspecified.

man_len_gvd_z_far_minus1[i] + 1 specifies the length in bits of the mantissa of the farthest depth value of the i-th camera. The value of man_len_gvd_z_far_minus1[i] shall be in the range of 0 to 31, inclusive.

man_gvd_z_far[i] specifies the mantissa part of the farthest depth value of the i-th camera. The length of man_gvd_z_far[i] syntax elements is man_len_gvd_z_far_minus1[i] + 1 bits.

When exp_gvd_z_far[i] is not equal to 127, zFar[i] is set equal to binToFp(sign_gvd_z_far_flag[i], exp_gvd_z_far[i], man_gvd_z_far[i], man_len_gvd_z_far_minus1[i] + 1).

prec_gvd_focal_length specifies the exponent of the maximum allowable truncation error for focalLengthX[i] and focalLengthY[i] as given by $2^{-prec\_gvd\_focal\_length}$. The value of prec_gvd_focal_length shall be in the range of 0 to 31, inclusive.

prec_gvd_principal_point specifies the exponent of the maximum allowable truncation error for principalPointX[i] and principalPointY[i] as given by $2^{-prec\_gvd\_principal\_point}$. The value of prec_gvd_principal_point shall be in the range of 0 to 31, inclusive.

prec_gvd_rotation_param specifies the exponent of the maximum allowable truncation error for r[i][j][k] as given by $2^{-pre\_gvdc\_rotation\_param}$. The value of prec_gvd_rotation_param shall be in the range of 0 to 31, inclusive.

prec_gvd_translation_param specifies the exponent of the maximum allowable truncation error for tX[i] as given by $2^{-prec\_gvd\_translation\_param}$. The value of prec_gvd_translation_param shall be in the range of 0 to 31, inclusive.

sign_gvd_focal_length_x[i] equal to 0 indicates that the sign of the focal length of the i-th camera in the horizontal direction is positive. sign_gvd_focal_length_x[i] equal to 1 indicates that the sign of the focal length of the i-th camera in the horizontal direction is negative.

exp_gvd_focal_length_x[i] specifies the exponent part of the focal length of the i-th camera in the horizontal direction. The value of exp_gvd_focal_length_x[i] shall be in the range of 0 to 62, inclusive. The value 63 is reserved for future use by ITU-T | ISO/IEC. When exp_gvd_focal_length_x[i] is equal to 63, the value of focal length of the horizontal direction for the i-th camera is unspecified.

man_gvd_focal_length_x[i] specifies the mantissa part of the focal length of the i-th camera in the horizontal direction. The length v of the man_gvd_focal_length_x[i] syntax element is determined as follows:

— if exp_gvd_focal_length_x[i] is equal to 0, the length v is set equal to Max(0, prec_gvd_focal_length − 30);

— otherwise (exp_gvd_focal_length_x[i] is in the range of 1 to 62, inclusive), the length v is Max(0, exp_gvd_focal_length_x[i] + prec_gvd_focal_length − 31).

When exp_gvd_focal_length_x[i] is not equal to 63, the variable focalLengthX[i] is set equal to binToFp(sign_gvd_focal_length_x[i], exp_gvd_focal_length_x[i], man_gvd_focal_length_x[i], v).

sign_gvd_focal_length_y[i] equal to 0 indicates that the sign of the focal length of the i-th camera in the vertical direction is positive. sign_gvd_focal_length_y[i] equal to 1 indicates that the sign of the focal length of the i-th camera in the vertical direction is negative.

exp_gvd_focal_length_y[i] specifies the exponent part of the focal length of the i-th camera in the vertical direction. The value of exp_gvd_focal_length_y[i] shall be in the range of 0 to 62, inclusive. The value 63 is reserved for future use by ITU-T | ISO/IEC. When exp_gvd_focal_length_y[i] is equal to 63, the value of focal length of the vertical direction is unspecified.

man_gvd_focal_length_y[i] specifies the mantissa part of the focal length of the i-th camera in the vertical direction.