

First edition
2009-11-15

AMENDMENT 3
2014-11-15

**Information technology — Multimedia
application format (MPEG-A) —**

Part 11:

Stereoscopic video application format

**AMENDMENT 3: Support movie fragment
for Stereoscopic Video AF**

*Technologies de l'information — Format pour application multimédia
(MPEG-A) —*

Partie 11: Format pour application vidéo stéréoscopique

*AMENDEMENT 3: Prise en charge de fragments de film pour format
d'application vidéo stéréoscopique*



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2014

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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.

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 WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword — Supplementary information.

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

IECNORM.COM : Click to view the full PDF of ISO/IEC 23000-11:2009/Amd 3:2014

Information technology — Multimedia application format (MPEG-A) —

Part 11: Stereoscopic video application format

AMENDMENT 3: Support movie fragment for Stereoscopic Video AF

In Clause 2, add the following normative reference:

ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

In 3.8, replace definition with:

maximum disparity value within successive stereoscopic samples

In 3.9, replace definition with:

set of samples which represents only monoscopic sequence

in 3.10, replace definition with:

minimum disparity value within a group of successive stereoscopic samples

In 3.18, replace text with:

stereoscopic samples

In 3.19, replace text with:

stereoscopic left samples

In 3.21, replace text with:

stereoscopic right samples

In Clause 4, add the following abbreviation:

HEVC High Efficiency Video Coding

In 6.1, replace text with:

[Table 1](#) shows a brief summary of the supported components of the Stereoscopic Video AF which consists of the ISO/IEC Standards and non-ISO/IEC Standards.

The Stereoscopic Video AF includes ISO/IEC 14496-2 Simple Profile at Level 3, ISO/IEC 14496-10 Baseline Profile at Level 1.3, and ISO/IEC 23008-2 Main/Main10 Profile for visual, ISO/IEC 14496-3 AAC and HE-AAC Profile for audio, 3GPP TS 26.071 AMR and TIA/EIA/IS-127 EVRC for voice, ISO/IEC 14496-20 LAsER for scene description, and various kinds of image such as ISO/IEC 10918-1 JPEG and ISO/IEC 15948 PNG. For this specification, ISO/IEC 14496-12 ISO base media file format is used for a base file format structure.

Table 1 — Supported components of Stereoscopic Video AF

Type	Component Name	Specification	Standard
File format	ISO base media file format	ISO/IEC 14496-12	ISO/IEC Standards
Visual	MPEG-4 Video	ISO/IEC 14496-2 Simple Profile Level 3, ISO/IEC 14496-2 Advanced Simple Profile Level 5	
	MPEG-4 AVC	ISO/IEC 14496-10 Baseline Profile Level 1.3, ISO/IEC 14496-10 High Profile Level 4.1	
	MPEG-H HEVC	ISO/IEC 23008-2 Main Profile, ISO/IEC 23008-2 Main10 Profile, ISO/IEC 23008-2 Main Still Picture Profile	
Audio	MPEG-4 Audio AAC	ISO/IEC 14496-3	
	MPEG-4 Audio HE-AAC	ISO/IEC 14496-3	
Data	MPEG-4 LSeR	ISO/IEC 14496-20	
	JPEG Image	ISO/IEC 10918-1	
	PNG Image	ISO/IEC 15948	
Voice	AMR	3GPP TS 26.071	Non-ISO/IEC Standards
	EVRC	TIA/EIA/IS-127	

In 7.1, replace

The 'mdia' box contains a 'svmi' box for the stereoscopic visual type and fragment information of the stereoscopic contents in the track.

The 'iloc' box describes the absolute offset in bytes ('extent_offset') and the size ('extent_length') of stereoscopic fragments. An item_ID is assigned to each fragment of the stereoscopic sequence for resource referencing.

with:

The 'mdia' box contains a 'svmi' box for the stereoscopic visual type and sample information of the stereoscopic contents in the track.

The 'iloc' box describes the absolute offset in bytes ('extent_offset') and the size ('extent_length') of stereoscopic samples. An item_ID is assigned to successive samples of the stereoscopic sequence for resource referencing.

In 7.2, add following text before 7.2.1:

In case of a stereoscopic content with Left/Right view sequence type, the 'stss' box which is in the track for the primary view sequence is used for random access.

In 7.2.2, replace text with:

This subclause describes the file structures for a stereo-monoscopic mixed content, which is a video sequence consisting of both stereoscopic and monoscopic samples in a single track. The stereoscopic and monoscopic samples should be stored sequentially.

Figure 9 shows an example of the file structure containing a single track for a stereo-monoscopic mixed content on the basis of the file format structure as shown in Figure 7. The item_ID under 'iloc' box is assigned to each group of stereoscopic samples sequentially. For example, when a stereoscopic contents is composed as illustrated in the below figure (S-M-S), the item_ID of the first group of samples in

the track, which is the first stereoscopic samples, is set to 1, and the `item_ID` of the third one (second stereoscopic samples) in the track is set to 2.

Figure 9 — Example of a file structure for stereoscopic and monoscopic samples in a single stereoscopic track

Figure 10 describes the file structure of a stereoscopic contents specified in 5.3.4, the composition type for storing the left and the right view sequence of stereoscopic contents in two separate tracks. Stereoscopic samples of each track have one view sequence on the basis of the file format structure as shown in Figure 8. The `item_ID` is assigned to each stereoscopic samples of only one track sequentially.

Figure 10 — Example of a file structure for stereoscopic and monoscopic samples in Left/Right view sequence type

In case of stereo-monoscopic mixed contents being shown in Figure 10, it could cause the same time stamp for monoscopic samples in the individual tracks. This ambiguity of presentation can be figured out as follows:

- Check which track is indicating a primary view sequence by the 'reference_type' and 'track_ID' of the 'tref' box in the track.
- Display each monoscopic samples of primary view sequence.

Insert following clauses after 7.2.2:

7.3 File format brands

7.3.1 The 'ss01' and 'ss02' brand

The brand 'ss01' and 'ss02' shall be used to indicate that the file is conformant with the 'stereoscopic video application format' in subclauses 7.1, 7.2, and Clause 8. If all the samples in content are stereoscopic samples, 'ss01' is used. If the content is a mixture of stereoscopic samples and monoscopic samples, 'ss02' is used.

The 'ss01' and 'ss02' brand requires support of the boxes in Table 2.

7.3.2 The 'ss03' brand

The brand 'ss03' shall be used if grouping_types for stereoscopic composition type and camera display information in Clause 8 are used.

The 'ss03' brand requires support of the 'iso2' brand. In addition, support of the following boxes is required:

					sbgp	sample-to-group
					sgpd	sample group description

Remove text from 8.1.

Remove text from 8.2.

Remove text from 8.3.

In 8.4, replace whole clause with:

8.1 Stereoscopic Video Media Information Box

8.1.1 Definition

Box Type: 'svmi'
 Container: Sample Table Box ('stbl')
 Mandatory: Yes
 Quantity: Exactly one

The 'svmi' box provides stereoscopic video media information regarding the stereoscopic visual type and also, for the care of some mixed contents, stereoscopic or monoscopic samples information. The visual type information signals the composition type of the stereoscopic video sequence and the structure of samples. The stereoscopic samples or monoscopic samples information represents the number of successive samples, the number of consecutive samples, and whether the current sample is stereoscopic or not.

8.1.2 Syntax

```
aligned(8) class StereoscopicVideoMediaInformationBox extends
  FullBox('svmi', version = 0, 0){
  // stereoscopic visual type information
  unsigned int(8) stereoscopic_composition_type;
  unsigned int(7) reserved = 0;
  unsigned int(1) is_left_first;

  // stereo_mono_change information
  unsigned int(32) stereo_mono_change_count;
  for(i=0; i<=stereo_mono_change_count; i++){
    unsigned int(32) sample_count;
    unsigned int(7) reserved = 0;
    unsigned int(1) stereo_flag;
  }
}
```

8.1.3 Semantics

stereoscopic_composition_type – the type of stereoscopic contents that are specified in [Table 4](#).

Table 4 — Stereoscopic composition type

Value	Stereoscopic_composition_type
0x00	Side-by-side (half) type
0x01	Vertical line interleaved type
0x02	Frame sequential type
0x03	Left/Right view sequence type
0x04	Top-Bottom (half) type
0x05	Side-by-side (full) type
0x06	Top-Bottom (full) type
0x07-0xFF	Reserved

is_left_first – represents positions of left and right view sequence for 3D mobile devices as being specified in [Table 5](#). When is_left_first is '1' and current stereoscopic video is composed of side-by-side type, left side and right side of the image means left view and right view, respectively. When is_left_first is '0', left side and right side means right view and left view, respectively. When is_left_first is '1' and current stereoscopic video is composed of vertical line interleaved type, odd line and even line of the image means left view and right view, respectively. When is_left_first is '0', odd line and even line means right view and left view, respectively. When is_left_first is '1' and current stereoscopic video is composed of frame sequential type, odd frame and even frame of the sequence means left view and right view, respectively. When is_left_first is '0', odd frame and even frame means right view and left view, respectively. When is_left_first is '1' and current stereoscopic video

is composed of Left/Right view sequence type, primary view sequence and secondary view sequence means left view and right view, respectively. When `is_left_first` is '0', primary view sequence and secondary view sequence means right view and left view, respectively.

Table 5 — The positions of stereoscopic Left/Right view according to the `is_left_first` value

Type	<code>is_left_first = 1</code>		<code>is_left_first = 0</code>	
	Left view	Right view	Left view	Right view
Side-by-side (half/full)	Left side	Right side	Right side	Left side
Vertical line interleaved	Odd line	Even line	Even line	Odd line
Frame sequential	Odd frame	Even frame	Even frame	Odd frame
Left/Right view sequence	Primary view sequence	Secondary view sequence	Secondary view sequence	Primary view sequence
Top-Bottom (half/full)	Top side	Bottom side	Bottom side	Top side

`stereo_mono_change_count` — is an integer that gives the number of group of successive samples when stereoscopic to/from monoscopic sample changes. If all samples are stereoscopic, `stereo_mono_change_count` is set to 0.

`sample_count` — is an integer that counts the number of consecutive samples.

`stereo_flag` — represents whether the current sample is stereoscopic or not. If this value is 1, then the current sample is stereoscopic, and if this value is 0, then the current sample is monoscopic.

In 8.5, replace whole clause with:

8.2 Stereoscopic Camera and Display Information Box

8.2.1 Definition

Box Type : 'scdi'

Container: Meta Box ('meta')

Mandatory: No

Quantity: Zero or one

The 'scdi' box, an optional box, provides primary information of the stereoscopic camera, display, and visual safety. Stereoscopic camera and display information specified in this box can be described for stereoscopic samples. Each sample including 'scdi' has a unique `item_ID` which is an identifier to be referenced by other samples.

8.2.2 Syntax

```
aligned(8) class StereoscopicCameraAndDisplayInformationBox extends
    FullBox('scdi', version = 0, 0){
    unsigned int(16) item_count;
    for( i=0; i<item_count; i++){
        unsigned int(16) item_ID;
        unsigned int(7) reserved = 0;
        unsigned int(1) is_item_ID_ref;
        if(is_item_ID_ref){
            unsigned int(16) ref_item_ID;
        }
        else{
            // stereoscopic display information
            unsigned int(4) reserved = 0;
            unsigned int(3) 3D_display_type;
            unsigned int(1) is_display_safety_info;
            if(is_display_safety_info){
                unsigned int(16) expected_display_width;
                unsigned int(16) expected_display_height;
                unsigned int(16) expected_viewing_distance;
            }
        }
    }
}
```

```

        int(16)          min_of_disparity;
        int(16)          max_of_disparity;
    }
    // stereoscopic camera information
    unsigned int(7)      reserved = 0;
    unsigned int(1)      is_cam_params;
    if(is_cam_params){
        unsigned int(32)  translation[3];
        unsigned int(32)  focal_length;
        unsigned int(32)  aspect_ratio;
        unsigned int(32)  convergence_distance;
        unsigned int(3)   reserved = 0;
        unsigned int(1)   is_camera_cross;
        unsigned int(3)   reserved = 0;
        unsigned int(1)   is_principal_point;
        if (is_camera_cross){
            unsigned int(32)  rotation[3];
        }
        if (is_principal_point){
            unsigned int(16)  primary_principal_point_x;
            unsigned int(16)  primary_principal_point_y;
            unsigned int(16)  secondary_principal_point_x;
            unsigned int(16)  secondary_principal_point_y;
        }
    }
}
}
}

```

8.2.3 Semantics

item_count — is the number of the items that indicate stereoscopic samples. In case of multiple stereoscopic samples having the same camera and display information, **item_count** is assigned to be 1.

item_ID — indicates the **item_ID** recoded in the 'iloc' box for a stereoscopic samples.

is_item_ID_ref — tells if the parameters of other item will be used.

ref_item_ID — indicates the **item_ID** of the item that is used for reference, where the item should be referenced when the camera and display information in a 'scdi' box of the item is available.

3D_display_type — specifies the types of 3D display for stereoscopic contents rendering in [Table 6](#).

Table 6 — The types of 3D display

Value	Specification
000	No limitation for 3D display type
001	Parallax Barrier
010	Lenticular
011-111	Reserved

is_display_safety_info — tells if there is display safety information.

expected_display_width — the expected display width given in *mm* (millimeter).

expected_display_height — the expected display width given in *mm*.

expected_viewing_distance — the expected viewing distance given in *mm*.

min_of_disparity — the minimum disparity in units of interger pixel resolution between left view and right view images.

`max_of_disparity` — the maximum disparity in units of interger pixel resolution between left view and right view images.

`is_cam_params` — tells if there is camera parameter information.

`translation[3]` — the relative positions of the origins between two cameras.

`focal_length` — the distance from the optical center to an image plane. Stereoscopic contents use one focal length due to that the focal length of the individual two cameras are the same for the depth effect.

`aspect_ratio` — the ratio between a horizontal focal length and a vertical focal length. (horizontal focal length/vertical focal length).

`convergence_distance` — distance from the center of baseline to convergence point.

`is_camera_cross` — specifies camera arrangement specified in [Table 7](#).

Table 7 — Camera arrangement specification

Value	Specification
0	Parallel arrangement
1	Cross arrangement

`is_principal_point` — tells if there is a principal point of individual camera. If this value is '0', a principal point is regarded as a center point of an image plane.

`rotation[3]` — relative angle values from primary view camera to secondary view camera in units of degrees.

`primary_principal_point_x` — specifies principal point in the horizontal direction of a primary view in units of quarter pixels.

`primary_principal_point_y` — specifies principal point in the vertical direction of a primary view in units of quarter pixels.

`secondary_principal_point_x` — specifies principal point in the horizontal direction of a secondary view in units of quarter pixels.

`secondary_principal_point_y` — specifies principal point in the vertical direction of a secondary view in units of quarter pixels.

Remove text from 8.6

Insert following clauses after 8.2:

8.3 Stereoscopic video sample group entry

A stereoscopic video sample group entry identifies samples of a specific type of stereoscopic content.

8.3.1 Definition

Group Types: 'svsg'

Container: Sample Group Description Box ('sgpd')

Mandatory: No

Quantity: Zero or more

8.3.2 Syntax

```
class StereoscopicVideoSampleGroupEntry() extends VisualSampleGroupEntry ('svsg')
{
    // stereoscopic visual type information
```