FINAL
DRAFT

AMENDMENT

ISO/IEC 23001-10:2015 FDAM 1

ISO/IEC JTC **1**/SC **29**

Secretariat: JISC

Voting begins on: **2016-09-09**

Voting terminates on: **2016-11-04**

Information technology — MPEG systems technologies —

Part 10: Carriage of timed metadata metrics of media in ISO base media file format

AMENDMENT 1: Carriage of special information

Technologies de l'information — Technologies des systèmes MPEG —

Partie 10: Transport de métriques de métadonnées de temporisation de supports au format de fichier de support en base ISO

AMENDEMENT 1: Transport d'information spéciale

Reference number
ISO/IEC 23001-10:2015/FDAM 1:2016(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 1 to ISO/IEC 23001-10:2015 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 — MPEG systems technologies —

Part 10:
Carriage of timed metadata metrics of media in ISO base media file format

AMENDMENT 1: Carriage of special information

Page 12

Add a new Clause 6

6 Carriage of coordinates

6.1 General

This document specifies the carriage of ROI coordinates in the ISO Base Media File Format using metadata tracks. Different coordinate types and corresponding storage formats are identified by their sample entry. This clause defines those coordinates.

The ROI metadata track shall be linked to the track it describes by means of a 'cdsc' (content describes) track reference.

The ROI described by a sample in the ROI track indicates the position of the ROI in the video track with respect to the dimensions documented by the track header (i.e. on a uniformly sampled grid, possibly upscaled to track header width and height) but before the application of the track (or movie) matrix, if any.

6.2 2D Cartesian coordinates

6.2.1 2D Cartesian coordinates Sample Entry

Sample Entry Type: `'2dcc'`

Container: Sample Description Box ('stsd')

Mandatory: No

Quantity: 0 or 1

The 2D Cartesian coordinates sample entry provides spatial information related to the referenced track expressed in a two-dimension Cartesian coordinate system.

6.2.1.1 Syntax

The 2D Cartesian coordinates sample entry shall be as follows:

```
aligned(8) class 2DCartesianCoordinatesSampleEntry
       extends MetadataSampleEntry ('2dcc') {
       unsigned int(16)    reference_width;
       unsigned int(16)    reference_height;
}
```

6.2.1.2 Semantics

reference_width and reference_height give respectively the width and height of the reference rectangular space in which all ROI coordinates (top_left_x, top_left_y, width and height) are computed.

These fields allow associating a ROI metadata track with video tracks of different resolutions but representing the same visual source.

6.2.2 2D Cartesian coordinates Sample format

6.2.2.1 Syntax

The 2D Cartesian coordinates sample shall conform to the following syntax:

```
aligned(8) class 2DCartesianCoordinatesSample(){
       unsigned int(16)    top_left_x;
       unsigned int(16)    top_left_y;
       unsigned int(16)    width;
       unsigned int(16)    height;
       unsigned int(1)     interpolate;
       unsigned int(7)     reserved;
}
```

Sync samples for ROI metadata tracks are samples for which the interpolate value is 0.

6.2.3 Semantics

top_left_x and top_left_y give respectively the horizontal and vertical coordinate of the top-left corner of the rectangle region associated with the media sample of the referenced track.

width and height give respectively the width and height of the rectangular region associated with the media sample of the referenced track.

interpolate indicates the continuity in time of the successive samples. When true, the application may linearly interpolate values of the ROI coordinates between the previous sample and the current sample. When false, there shall not be any interpolation of values between the previous and the current samples.

NOTE When using interpolation, it is expected that the interpolated samples match the presentation time of the samples in the referenced track. For instance, for each video sample of a video track, one interpolated 2D Cartesian coordinate sample is calculated.

Add a new Annex A

Annex A

(informative)

Use cases for carriage of ROI coordinates

A.1 Close-up view (video-to-video)

In this scenario, the content provider offers two videos, namely a wide angle view and a close up view. The close up view generally focuses on particularly interesting parts of the scene, e.g. most popular athletes in sport events. But to ensure a satisfying Quality of Experience for the end-user, it is utterly important to be able to describe the position of the close-up cam with respect to the wide angle cam at any point in time of the broadcast. This way, the end-user application may seamlessly switch from one video to another provide a smooth experience for the end-user. Figure A.1 illustrates this concept for a live broadcast of cycling races.

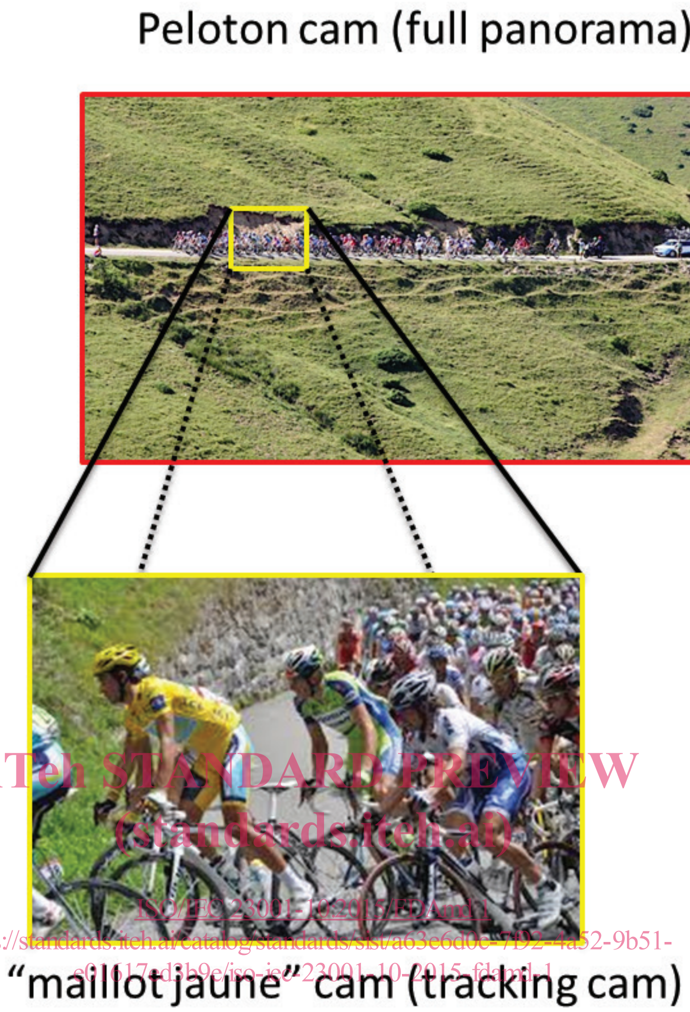

Figure A.1 — “Maillot jaune” cam use case

The file format structure for this scenario is the following: a main video track for the wide angle view, a second video track offering the close up view and a timed metadata track linked to the second video track with ‘cdsc’ track reference type, that contains ROI coordinates samples providing the position and size of the media samples of the close up view in the wide angle view.

A.2 Object annotation (metadata-to-video)

This scenario involves applications that spatially annotate dynamic object in videos. For instance, a video conference system can provide the position and size of the participants’ faces allowing the application to augment the visual experience, e.g. by displaying participant names, overlaying graphics, etc.

The file format structure for this scenario is the following: a main video offering a view of all the participants, a metadata track containing application specific metadata (e.g. participant information) and a timed metadata track containing ROI coordinates samples providing the position and size of the participant faces in the main video.

ICS 35.040
Price based on 3 pages