INTERNATIONAL STANDARD

ISO/IEC 23001-13

First edition
2019-07

Information technology — MPEG systems technologies —

Part 13: Media orchestration

Reference number
ISO/IEC 23001-13:2019(E)

© ISO/IEC 2019



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2019

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms and definitions **2**

4 Abbreviated terms, relationships between terms and mnemonics **5**
4.1 Abbreviated terms 5
4.2 Relationships between defined terms 5
4.3 Mnemonics 6

5 Architecture **6**
5.1 General 6
5.2 Architecture for temporal orchestration 9
5.2.1 General 9
5.2.2 Sink-side architecture for temporal orchestration 10
5.2.3 Source-side architecture for temporal orchestration 13
5.3 Architecture for spatial orchestration 15
5.3.1 General 15
5.3.2 M-processor side architecture for spatial orchestration 15
5.3.3 Sink-side architecture for spatial orchestration 17

6 Messaging and control **18**
6.1 Wall clock 18
6.2 MORE communication channel 18
6.2.1 General 18
6.2.2 Signalling of the MORE communication channel 19
6.3 Websocket protocol 20

7 Metadata **20**
7.1 Timed metadata 20
7.2 Position metadata 21
7.2.1 General 21
7.2.2 Global position 21
7.2.3 Altitude 21
7.2.4 Relative position 21
7.3 Orientation metadata 22
7.4 Regions of interest 23
7.4.1 General 23
7.4.2 Example implementation using DROP features 24
7.5 Quality metadata 25
7.5.1 General 25
7.5.2 Quality metadata 27
7.6 Stream monitor 27
7.6.1 General 27
7.6.2 MPEG-Audio-Sync-based stream monitor metadata 30
7.7 Capture mask 30
7.7.1 General 30
7.7.2 Capture-mask metadata 31
7.8 Camera metadata 32
7.9 Depth metadata 32

8 Transport **33**
8.1 Carriage of timed metadata in ISO base media file format 33
8.1.1 General 33
8.1.2 Carriage of position metadata in ISO base media file format 34

© ISO/IEC 2019 – All rights reserved

8.1.3 Carriage of orientation metadata in ISO base media file format 35
8.1.4 Carriage of quality metadata in ISO base media file format 36
8.1.5 Carriage of audio features in ISO base media file format 37
8.1.6 Carriage of capture-mask metadata in ISO base media file format 38
8.1.7 Carriage of dvb-css correlation timestamps in ISO base media file format 38
8.2 Carriage of timed metadata in MPEG-2 systems 41
8.2.1 General 41
8.2.2 Timed metadata access unit 41
8.2.3 Timed metadata extension descriptor 42
8.2.4 Carriage of position metadata in MPEG-2 systems 43
8.2.5 Carriage of orientation metadata in MPEG-2 systems 46
8.2.6 Carriage of capture-mask metadata in MPEG-2 systems 47
8.2.7 Carriage of audio features in MPEG-2 systems 49
8.2.8 Carriage of quality metadata in MPEG-2 systems 50
8.2.9 Carriage of correlation timestamps in MPEG-2 systems 50
8.3 Carriage of MORE information in MMT 51
8.3.1 General 51
8.3.2 Carriage of MORE information in MMT 51
8.4 Carriage of MORE information in DASH 51
8.4.1 General 51
8.4.2 Carriage of quality metadata in MPEG DASH MPD 51

9 Orchestration data 51
9.1 General 51
9.2 Timeline synchronization 52
9.3 dvb-css correlation timestamps 52

Annex A (normative) **Source-side timeline synchronization 53**

Annex B (informative) **Object model 56**

Annex C (informative) **Control architecture 59**

Bibliography 64

 © ISO/IEC 2019 – All rights reserved

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.

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) or the IEC list of patent declarations received (see http://patents.iec.ch).

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

For an explanation of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23001 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

© ISO/IEC 2019 – All rights reserved

Introduction

Audiovisual equipment is pervasive; everyone with a smartphone, a tablet or a laptop has both recording and display equipment at their disposal, usually connected to a local or a public network. This equipment is increasingly sophisticated, with higher resolutions and better lenses (for video), often multiple microphones (for audio), coupled with other sensors (e.g. for geolocation) and increasingly sophisticated processing capabilities. These devices can not only decode in real time, but usually also perform decent encoding. Sensors and displays come in the form of personal smart phones, but also include smart watches, omnidirectional cameras and a variety of virtual reality and augmented reality head-mounted devices. All these devices can either be a source of multimedia content (audio, video or audiovisual) or consume such content. Often, devices play both roles.

The proliferation of these multimedia-capable devices combined with ever-increasing bandwidth, including mobile bandwidth, necessitates better and standardized mechanisms for coordinating such devices, the associated media streams and the available resources, like media processing and transmission. This process is called “orchestration”. Orchestration includes coordination in time (synchronization) and in space, as well as the coordination of computational resources to perform the coordination functions.

This document was developed to address the need for specifying this coordination between capture devices (“sources”) and play-back devices (“sinks”) in a heterogeneous environment. In this context, “heterogeneous” refers to the fact that devices can be of different nature (e.g. a mix of consumer and professional devices) and on different networks.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of patents.

ISO and IEC take no position concerning the evidence, validity and scope of these patent rights. The holders of these patent rights have assured ISO and IEC that they are willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statements of the holders of these patent rights are registered with ISO and IEC. Information may be obtained from:

Electronics and Telecommunications Research Institute

218 Gajeong-ro, Yuseong-gu, Daejeon, 34129 Korea

Koninklijke KPN N.V.

Wilhelminakade 123, 3072 AP Rotterdam, The Netherlands

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified above. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

© ISO/IEC 2019 – All rights reserved

Information technology — MPEG systems technologies —

Part 13:
Media orchestration

1 Scope

This document specifies an architecture for media orchestration, as well as associated messaging and control, timed metadata, the carriage of that timed metadata, and orchestration data.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 13818-1:2019, *Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems*

ISO/IEC 14496-3:2009, *Information technology — Coding of audio-visual objects — Part 3: Audio*

ISO/IEC 14496-3:2009/Amd 5:2015, *Information technology — Coding of audio-visual objects — Part 3: Audio— Amendment 5: Support for Dynamic Range Control, New Levels for ALS Profile, and Audio Synchronization*

ISO/IEC 14496-12:2015, *ISO Base Media File Format*

ISO/IEC 14496-12:2015/Amd 1:2017,*ISO Base Media File Format — Amendment 1: DRC extensions*

ISO/IEC 23001-10:—[1)], *Information technology — MPEG systems technologies — Part 10: Carriage of timed metadata metrics of media in ISO base media file format*

ISO/IEC 23005-6:2019, *Information technology — Media context and control — Part 6: Common types and tools*

ISO/IEC 23008-1, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT)*

ETSI TS 103 286 02:2017, *Digital Video Broadcasting (DVB); Companion Screens and Streams; Part 2: Content Identification and Media Synchronization* [online]. Available at http://www. etsi. org/deliver/etsi_ts/103200_103299/10328602/

IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax* [online]. Edited by T. Berners-Lee. January 2005. Available at https://www.ietf.org/rfc/rfc3986.txt

IETF RFC 6455:2011, *The WebSocket Protocol* [online]. Edited by I. Fette. December 2011. Available at https://www.ietf.org/rfc/rfc6455.txt

IETF RFC 5905, *Network Time Protocol Version 4: Protocol and Algorithms Specification* [online]. Edited by D. Mills. June 2010. Available at https://tools.ietf.org/html/rfc5905

1) Under preparation. Stage at the time of publication: ISO/IEC DIS 23001-10:2019.

© ISO/IEC 2019 – All rights reserved

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
content
media data
data (3.3) that can be rendered, including audio, video, text, graphics, images, haptic and tactile information

Note 1 to entry: This data can be timed or non-timed.

3.2
controller
elementary function that controls one or more other elements, outputs *control data* (3.5) and can receive control data

3.3
data
bytes of information that are carried, processed or stored

3.4
timed data
data (3.3) that has an intrinsic *timeline* (3.25)

3.5
control data
data (3.3) exchanged for messaging and control

3.6
metadata
data (3.3) about other data that cannot be rendered independently and can affect rendering, processing or orchestration of the associated *media data* (3.1)

3.7
timed metadata
metadata (3.6) that has an intrinsic *timeline* (3.25)

3.8
orchestration data
data (3.3) that orchestrates any number of *timed data* (3.4) streams

Note 1 to entry: Orchestration data itself can also be timed data.

3.9
device
physical entity that can embody a number of elementary functions

3.10
m-processor
media processor
elementary function that can process *media data* (3.1) and/or *metadata* (3.6), that can receive media data and/or metadata, that can optionally receive *orchestration data* (3.8), that can output media data, metadata or both, and that can optionally exchange *control data* (3.5)

© ISO/IEC 2019 – All rights reserved

3.11
media orchestration
orchestration of media and *metadata* (3.6) capture, processing and presentation involving multiple *devices* (3.9)

3.12
media orchestration session
coherent application context that includes media capture and/or media processing and/or media rendering

3.13
media stream
timed *media data* (3.1)

3.14
mos more
mean opinion score of the subjective quality of a piece of timed *media data* (3.1)

3.15
sink
elementary function that can receive *media data* (3.1), and optionally *orchestration data* (3.8) and/or *metadata* (3.6), can present the media data, can optionally output metadata and can optionally exchange *control data* (3.5)

3.16
source
elementary function that can output *media data* (3.1) and/or *metadata* (3.6) and can optionally exchange *control data* (3.5)

3.17
timestamp
number that is used to identify an audio sample, video frame or *metadata* (3.6) sample

EXAMPLE MPEG-2 TS PTS (presentation timestamp) or ISOBMFF CT (composition time).

Note 1 to entry: The meaning of the term "timestamp" depends on the context. Within the MPEG context, the term is usually used as a single number to identify a point on the timeline of an MPEG-2 transport stream or within an ISOBMFF file. In the context of DVB-CSS (ETSI TS 103 286 02), the term "timestamp" is used to provide an association between a wall clock time and a point on a certain media timeline. This document uses the "dvb-css" prefix to flag where the DVB-CSS (ETSI TS 103 286 02) notion of timestamp is used.

3.18
dvb-css timestamp
pair of two values each of which represents a time value on a timeline such that the two-time values correspond to the same moment in time

[SOURCE: ETSI TS 103 286 02:2017, 3.1]

3.19
dvb-css control timestamp
dvb-css timestamp (3.18) representing the time (in relation to the wall clock) at which a *synchronization client* (3.24) is recommended to present a particular point in time (in relation to the synchronization *timeline* (3.25)) of its *timed data* (3.4)

[SOURCE: ETSI TS 103 286 02:2017, 3.1, modified — «timed content» has been replaced with «timed data» and the note has been deleted]

© ISO/IEC 2019 – All rights reserved

3.20
dvb-css correlation timestamp
time values on two or more *timelines* (3.25) that correlate with each other

[SOURCE: ETSI TS 103 286 02:2017, 3.1, modified — «ore more" has been added and the note has been deleted]

3.21
dvb-css earliest presentation timestamp
dvb-css timestamp (3.18) representing the earliest time (in relation to the wall clock) at which a *synchronization client* (3.24) believes it can present a particular point in time of its *timed data* (3.4)

[SOURCE: ETSI TS 103 286 02:2017, 3.1, modified — "(in relation to the Synchronization Timeline) of its Timed Content" has been replaced with "of its timed data" and the note has been deleted]

3.22
user
human that uses one or more *device(s)* (3.9) and/or service(s)

Note 1 to entry: A user can be a producer and consumer of media, and both at the same time.

3.23
media synchronization application server
elementary function that coordinates the process of obtaining a shared agreement on the progress of *timelines* (3.25) among all *synchronization clients* (3.24)

Note 1 to entry: This is for the purpose of enabling the synchronization clients to present timed data simultaneously with respect to each other.

[SOURCE: ETSI TS 103 286 02:2017, 3.1]

3.24
synchronization client
elementary function that aligns its presentation of *timed data* (3.4) with other synchronization clients by communicating with a *media synchronization application server* (3.23)

[SOURCE: ETSI TS 103 286 02:2017, 3.1, modified — "wishes to align" has been replaced with "aligns" and "timed content" has been replaced with "timed data"]

3.25
timeline
reference frame for describing time, represented as a linear scale against which time can be measured for a particular system

Note 1 to entry: This could manifest in various ways, such as: as a local oscillator, the progress of a broadcast or the time position within an item of media content.

[SOURCE: ETSI TS 103 286 02:2017, 3.1]

3.26
timeline correlation
correlation between two *timelines* (3.25)

[SOURCE: ETSI TS 103 286 02:2017, 3.1]

3.27
wall clock
linear monotonic clock that is not assumed to represent real date and time

Note 1 to entry: The wall clock is intended for sharing between two or more entities for the purposes of having a common synchronized time reference frame.

© ISO/IEC 2019 – All rights reserved

Note 2 to entry: DVB-CSS (ETSI TS 103 286 02) wall clock is NOT a "UTC-type date&time". It has a random offset and it has no leap seconds or leap years. It is a clock that ticks like a metronome, and it has only a local meaning in the context of media synchronization between devices. This is a different notion for a timestamp than what is commonly used across MPEG. This document uses the definition from ETSI TS 103 286 02 to maintain consistency with that specification.

[SOURCE: ETSI TS 103 286 02:2017, 3.1, modified — Note 2 to entry has been added]

3.28
wall clock synchronization
achieving a shared agreement on the progress of time of a *wall clock* (3.27)

Note 1 to entry: This is a process that takes place between devices.

[SOURCE: ETSI TS 103 286 02:2017, 3.1]

4 Abbreviated terms, relationships between terms and mnemonics

4.1 Abbreviated terms

The following abbreviated terms are used in this document:

CSA	Companion screen application
CT	Composition time
DROP	Distinctive region or pattern
GPS	Global positioning system
MORE	Media orchestration
MORE-TS	MORE timeline synchronization (interface)
MORE-WC	MORE wall clock (interface)
MSAS	Media synchronization application server
URL	Uniform resource locator
PTS	Presentation timestamp
SC	Synchronization client
SNR	Signal-to-noise ratio
WCC	Wall clock client
WCS	Wall clock server

4.2 Relationships between defined terms

Figure 1 sketches relationships between some of the defined terms, including synonyms.

© ISO/IEC 2019 – All rights reserved