INTERNATIONAL STANDARD

ISO/IEC 14496-30

Second edition
2018-11

Information technology — Coding of audio-visual objects —

Part 30: Timed text and other visual overlays in ISO base media file format

Technologies de l'information — Codage des objets audiovisuels —

Partie 30: Texte temporisé et autres recouvrements visuels dans le format ISO de base pour les fichiers médias

Reference number
ISO/IEC 14496-30:2018(E)

© ISO/IEC 2018



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2018

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 ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms, definitions and abbreviated terms ... **1**
3.1 Terms and definitions ... 1
3.2 Abbreviated terms ... 2

4 General definitions ... **2**
4.1 Layout ... 2
4.2 Timing ... 3
4.3 Language ... 3
4.4 Resources shared by multiple samples ... 4
4.5 Associating timed text tracks ... 4

5 Timed Text Markup Language (TTML) ... **4**
5.1 General ... 4
5.2 Layout ... 4
5.3 Timing ... 4
5.4 Track format ... 6
5.5 Sample entry format ... 6
5.6 Sample format ... 6
5.7 Additional considerations ... 8
5.8 Codecs parameter ... 8
5.9 Document temporal boundaries ... 8

6 Web Video Text Tracks (WebVTT) ... **9**
6.1 General ... 9
6.2 Layout ... 9
6.3 Timing ... 9
6.4 Track format ... 9
6.5 Sample entry format ... 9
6.6 Sample format ... 10
6.7 Converting to or from a WebVTT text file ... 11
6.7.1 General ... 11
6.7.2 Importing a WebVTT file into the ISO base media file format ... 11
6.7.3 Exporting a WebVTT file from the ISO base media file format ... 12
6.8 Example ... 13
6.8.1 Source file ... 13
6.8.2 Imported format ... 13

Annex A (informative) **Captioning information embedded in a media stream of another type** ... **14**

Bibliography ... **15**

© ISO/IEC 2018 – 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. 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 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 Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 14496-30:2014), which has been technically revised. It incorporates ISO/IEC 14496-30:2014/Cor.1:2015. The main changes compared to the previous edition are as follows:

— all Clauses (except Clause 1 and Clause 3) and Annex A have been technically revised.

A list of all parts in the ISO/IEC 14496 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 2018 – All rights reserved

Introduction

This document defines a storage format based on, and compatible with, the ISO base media file format (ISO/IEC 14496-12), which is used by the MP4 file format (ISO/IEC 14496-14) and the Motion JPEG 2000 file format (ISO/IEC 15444-3) among others. This document enables timed text and subtitle streams to

— be used in conjunction with other media streams, such as audio or video;

— be used in an MPEG-4 systems environment, if desired;

— be formatted for delivery by a streaming server, using hint tracks; and

— inherit all the use cases and features of the ISO base media file format on which MP4 and MJ2 are based.

© ISO/IEC 2018 – All rights reserved

Information technology — Coding of audio-visual objects —

Part 30:
Timed text and other visual overlays in ISO base media file format

1 Scope

This document describes the carriage of some forms of timed text and subtitle streams in files based on ISO/IEC 14496-12 (the ISO base media file format). The documentation of these forms does not preclude other definition of carriage of timed text or subtitles; see, for example, 3GPP Timed Text (3GPP TS 26.245), or the carriage of captioning information embedded in a media stream of another type (see Annex A).

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.

W3C Recommendation, Timed Text Markup Language 1.0, Second Edition, https://www.w3.org/TR/ttml1/

ISO/IEC 14496-12:2015, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

W3C Community Group Report, *WebVTT*, http://www.w3.org/2013/07/webvtt.html

3GPP TS 26.245:2017, *Transparent end-to-end Packet switched Streaming Service (PSS); Timed text format*

IETF RFC 2141, *URN Syntax*

IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax*

IETF RFC 6381, *MIME Codecs and Profiles*

3 Terms, definitions and abbreviated terms

3.1 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/

3.1.1
timed text document
file-based representation of textual content, possibly XML, used to produce timed text streams and possibly representing timed text track samples

© ISO/IEC 2018 – All rights reserved

3.1.2
timed text stream
stream of content, which when decoded results in textual content, possibly containing internal timing values, to be presented at a given presentation time and for a certain duration

3.1.3
subtitle stream
timed text stream potentially also presenting images

3.1.4
internal timing value
value contained in the payload of a timed text stream sample representing a time

EXAMPLE A start time, an end time, or a duration, corresponding to a timed behaviour of a part or the whole of the sample.

3.1.5
timed text track
ISOBMFF representation of a timed text stream

3.1.6
subtitle track
ISOBMFF representation of a subtitle stream

3.2 Abbreviated terms

TTML Timed Text Markup Language

WebVTT Web Video Text Tracks

ISOBMFF ISO base media file format

4 General definitions

4.1 Layout

This subclause defines common layout behaviour for processing of timed text or subtitle samples.

Unless specified by an embedding environment (e.g. an HTML page), the track header box information (i.e. `width`, `height`) shall be used to size the subtitle or timed text track content with respect to the associated track(s) as follows:

1) If the flag `track_size_is_aspect_ratio` is not set, and the track `width` and `height` are set to values different from 0, the size of the timed text track shall be the track `width` and `height`.

2) If the flag `track_size_is_aspect_ratio` is not set, and the track `width` and `height` are set to 0, the size of the timed text track shall match the reference size.

3) If the flag `track_size_is_aspect_ratio` is set, it indicates that the content of the track was authored to an aspect ratio equal to the track header `width/height`. In this case, neither `width` nor `height` shall be 0. The timed text track shall be sized to the maximum size that will fit within the reference size and should equal its width or height, while preserving the indicated aspect ratio.

If only one track is associated with the timed text track, the reference size is the size of the associated track. If multiple tracks are associated, the reference size is the size of the composition of tracks as described by the matrices in the track headers of the associated tracks.

Upon file creation, the `width` and `height` of the subtitle or timed text track should be set appropriately according to the `width` and `height` of the associated track(s), as declared in their track header. A

 © ISO/IEC 2018 – All rights reserved