**ISO/IEC 23008-12:2022/AMD 2:2022(E)**

ISO/IEC JTC1/SC 29

Secretariat: JISC

**Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 2: Support for renderable text signalling and other improvements**

Preliminary WD stage

**Copyright notice**

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to 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

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents

[6.5.3 Image spatial extents 1](#_Toc119324974)

[6.12 Text and font items 2](#_Toc119324975)

[6.12.1 Text Item 2](#_Toc119324976)

[6.12.2 Text properties 3](#_Toc119324978)

[6.12.2.1 Text layout information 3](#_Toc119324979)

[6.12.3 Font item 5](#_Toc119324983)

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 documents 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](http://www.iso.org/directives) or [www.iec.ch/members\_experts/refdocs](http://www.iec.ch/members_experts/refdocs)).

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](http://www.iso.org/patents)) or the IEC list of patent declarations received (see [patents.iec.ch](https://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 the following URL:  [www.iso.org/iso/foreword.html](http://www.iso.org/iso/foreword.html). In the IEC, see [www.iec.ch/understanding-standards](https://www.iec.ch/understanding-standards)

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

A list of all parts in the ISO/IEC 23008 series can be found on the ISO and IEC websites.

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](https://www.iso.org/members.html) and [www.iec.ch/national-committees](http://www.iec.ch/national-committees).

Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 2: Support for renderable text signalling and other improvements

*Add following definitions in clause 3*

**3.1.45**

**text item**

*item* (3.1.27) whose data is textual data

Note 1 to entry: A text item whose textual data is intended to be displayed/rendered over an image item is associated with the image item via an item reference of type 'text'.

**3.1.46**

**font item**

*item* (3.1.27) whose data is fonts

### Image spatial extents

Replace the content of subclause 6.5.3.1 (“*Definition*” for “*Image spatial extents”*) with the following

|  |  |
| --- | --- |
| Box type: | 'ispe' |
| Property type: | Descriptive item property |
| Container: | ItemPropertyContainerBox |
| Mandatory (per item): | Yes |
| Quantity (per item): | One |
|  |  |

The ImageSpatialExtentsProperty documents the width and height of the associated image item. Every image item shall be associated with one property of this type, prior to the association of all transformative properties.

The ImageSpatialExtentsProperty may be associated with items whose output can be visually rendered (e.g., text items). When ImageSpatialExtentsProperty is associated with items whose output can be visually rendered, they document the visually rendered width and height of the data which is output from the associated item.

**6.5.3.2 Syntax**

aligned(8) class ImageSpatialExtentsProperty  
extends ItemFullProperty('ispe', version = 0, flags = 0) {  
 unsigned int(32) image\_width;  
 unsigned int(32) image\_height;  
}

**6.5.3.3 Semantics**

image\_width specifies the width of the reconstructed image in pixels, as specified in 6.3.

image\_height specifies the height of the reconstructed image in pixels, as specified in 6.3.

NOTE Item properties, such as decoder configuration or layer selection, can affect the reconstructed image. As a consequence, the width and height of the reconstructed image depend on the presence and content of such properties.

When ImageSpatialExtentsProperty is associated with items whose output can be visually rendered, the image\_width and image\_height specifies the visually rendered width and height, respectively of the data which is output from the associated item.

Add the following new subclauses after subclause 6.11

### Text and font items

### Text Item

### Definition

A text item is an item with item\_type value set to 'mime' and the data in the text item is text, for example, html or plain text. The content\_type in ItemInfoEntry of the ItemInfoBox is set equal to the mime type of the data in the text item. Example values for content\_type field may include ‘text/html’ for html formatted text or ‘text/plain’ for plain text.

A text item intended to be displayed/rendered on an image item is associated with the image item using an item reference of type 'text' from the text item to the image item.

A text item may be associated with font items using an item reference of type 'font' from the text item to the font items. A font item carries the fonts used for rendering the text item.

A text item may be associated with the ImageSpatialExtentsProperty which documents the visually rendered width and height of the data which is output from the text item.

A text item may be associated with the TextLayoutProperty which documents the reference space size, the position and the language of the data which is output from the text item.

If a text item contains textual data together with possibly size, position, language, font and styling for visual rendering and is also associated with the ImageSpatialExtentsProperty and TextLayoutProperty then the values in ImageSpatialExtentsProperty and TextLayoutProperty take precedence. The association of ImageSpatialExtentsProperty and TextLayoutProperty with different mime type text items are listed in the following Table 2.

**Table 2– Mime type text item association**

|  |  |  |
| --- | --- | --- |
| **Mime type text item** | **Association with** 'ispe' **is mandatory** | **Association with** 'txlo' **is mandatory** |
| text/plain | Yes | Yes |
| text/html | No | No |
| text/vtt | Yes | Yes |

The data of a text item may be further encoded with either gzip or deflate or any other algorithm defined for content-encoding of HTTP/1.1. The encoding of data in a text item is defined by the content\_encoding parameter in ItemInfoEntry of the ItemInfoBox for the mime type text item.

If the data of text item is encoded with any of the algorithm defined for content-encoding of HTTP/1.1, the data needs to be decoded before interpreting it as the mime type text item identified by the content\_type in ItemInfoEntry of the ItemInfoBox.

If the content\_encoding parameter in ItemInfoEntry of the ItemInfoBox has an empty string, then no content encoding is applied on the text data.

If the data of text item is longer than the display/rendering size specified in the ImageSpatialExtentsProperty and TextLayoutProperty then it shall be clipped

### Text properties

### Text layout information

### Definition

Box type: 'txlo'

Property type: Descriptive item property

Container: ItemPropertyContainerBox

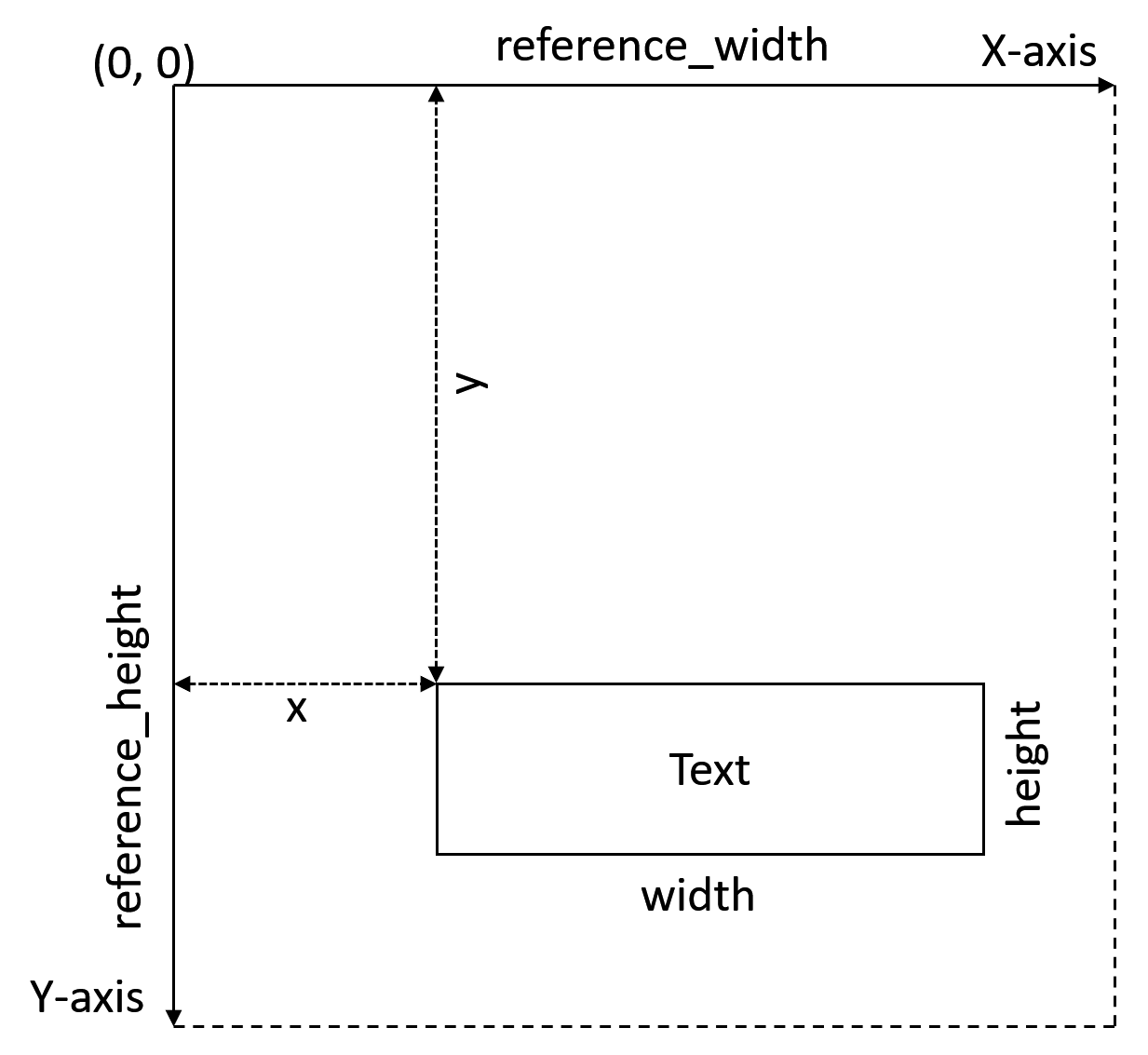
Mandatory (per item): No

Quantity (per item): One

The TextLayoutProperty documents the layout information of the associated text item. A text item shall be associated with one TextLayoutProperty prior to the association of all transformative properties.

The data in the TextLayoutProperty defines the reference space size, the position, and the language of the textual data to be displayed/rendered on the associated image item. The reference space size and the position information are used to display/render the textual data inside a reference space that is mapped to the image item with which the text item is associated after any transformative item property is applied to the image item.

The reference space is defined as a 2D coordinate system with the origin (0,0) located at the top-left corner and a maximum size defined by reference\_width and reference\_height; the x-axis is oriented from left to right and the y-axis from top to bottom. Figure 3 provides an illustration of text item displayed/rendered in the reference space. The placement of textual data inside the associated image item is obtained after applying the implicit resampling caused by the difference between the size of the reference space and the size of the associated image item. If the text item has transformative item properties, then the implicit resampling shall be performed on the text item before the first of its transformative item properties is applied.



**Figure 3: An illustration of text item in reference space.**

### Syntax

aligned(8) class TextLayoutProperty  
extends ItemFullProperty('txlo', version = 0, flags = 0) { field\_size = ((flags & 1) + 1) \* 16;  
 unsigned int(field\_size) reference\_width;  
 unsigned int(field\_size) reference\_height;   
 signed int(field\_size) x;  
 signed int(field\_size) y;  
 utf8string language;  
}

### Semantics

version shall be equal to 0.

(flags & 1) equal to 0 specifies that the length of the fields x, y, reference\_width, reference\_height is 16 bits. (flags & 1) equal to 1 specifies that the length of the fields x, y, reference\_width, reference\_height is 32 bits. The values of flags greater than 1 are reserved.

reference\_width, reference\_height specify, in pixel units, the width and height, respectively, of the reference space on which the text items are displayed/rendered.

x, y specify the top-left corner of the text item relatively to the reference space. The value (x = 0, y = 0) represents the position of the top-left pixel in the reference space.

NOTE Negative values for the x or y fields enable to specify top-left corners that are outside the image. This can be useful for updating text items during the edition of an HEIF file.

language is a character string containing an RFC 5646 compliant language tag string, such as "en-US", "fr-FR", or "zh-CN“, representing the language of the text. When language is empty, the language is unknown/undefined.

### Font item

### Definition

A font item is an item with the item\_type value set to 'mime' and the data in the font item are fonts for example ‘woff’ (Web Open Font Format) or ‘ttf’ (true type font). The content\_type in ItemInfoEntry of the ItemInfoBox is set equal to the mime type of the data in the font item. Example values for content\_type field may include ‘font/ttf’ for true type fonts or ‘font/woff’ for web open font format fonts.

A font item may be associated with text item using an item reference of type 'font' from the text item to the font item.

The data of a font item may be further encoded with either gzip or deflate or any other algorithm defined for content-encoding of HTTP/1.1. The encoding of data in a font item is defined by the content\_encoding parameter in ItemInfoEntry of the ItemInfoBox for the font item.

If the data of a font item is encoded with any of the algorithm defined for content-encoding of HTTP/1.1, the data needs to be decoded before interpreting it as the mime type font item identified by the content\_type in ItemInfoEntry of the ItemInfoBox.

If the content\_encoding parameter in ItemInfoEntry of the ItemInfoBox has an empty string, then no content encoding is applied on the font data.