**ISO/IEC 23001-xx:2024(E)**

ISO/IEC JTC 1/SC 29/WG 3

Date: 2024-MM-DD

**Information technology — MPEG systems technologies — Part xx: Carriage of green metadata**

WD stage

**Warning for WDs and CDs**

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

© ISO 2024

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

Email: copyright@iso.org

Website: www.iso.org

Published in Switzerland

Contents

*This template allows you to work with default MS Word functions and styles. You can use these if you want to maintain the Table of Contents automatically and apply auto-numbering.*

*To update the Table of Contents please select it and press "F9".*

[Foreword v](#_Toc171279077)

[Introduction vi](#_Toc171279078)

[1 Scope 1](#_Toc171279079)

[2 Normative references 1](#_Toc171279080)

[3 Terms, definitions, and abbreviated terms 1](#_Toc171279081)

[3.1 Terms and definitions 1](#_Toc171279082)

[3.2 Abbreviated terms 2](#_Toc171279083)

[4 Overview 2](#_Toc171279084)

[4.1 Overall architecture for carriage of Green metadata 2](#_Toc171279085)

[5 Carriage of Green Metadata in ISO Base Media File Format 2](#_Toc171279086)

[5.1 General 2](#_Toc171279087)

[5.2 Decoder Power Indication Metadata 2](#_Toc171279088)

[5.2.1 Definition 2](#_Toc171279089)

[5.2.2 Syntax 2](#_Toc171279090)

[5.2.3 Semantics 3](#_Toc171279091)

[5.3 Display Power Reduction Metadata 3](#_Toc171279092)

[5.3.1 Display Power Indication Metadata 3](#_Toc171279093)

[5.3.2 Display Fine Control Metadata 4](#_Toc171279094)

[5.3.3 Display Attenuation Map Metadata 5](#_Toc171279095)

[6 Encapsulation and Signalling in MPEG-DASH 10](#_Toc171279096)

[6.1 General 10](#_Toc171279097)

[6.2 Decoder Power Indication 10](#_Toc171279098)

[6.2.1 Metadata signalling in the MPD manifest file 10](#_Toc171279108)

[6.3 Display Power Indication 11](#_Toc171279111)

[6.3.1 Metadata signalling in the MPD manifest file 11](#_Toc171279115)

[6.4 Display Attenuation Map Information 12](#_Toc171279147)

[6.4.1 Metadata signalling in the MPD manifest file 12](#_Toc171279157)

[Annex A (normative) Green metadata MPEG-DASH schema 17](#_Toc171279158)

[Annex B (informative) MPEG-DASH MPD examples 19](#_Toc171279159)

[B.1 Example MPD with decoder power indication metadata 19](#_Toc171279160)

[B.2 Example MPD with display power indication metadata 20](#_Toc171279161)

[B.3 Examples of MPD with display attenuation map metadata 21](#_Toc171279162)

[B.3.1 Example 1 21](#_Toc171279163)

[B.3.2 Example 2 22](#_Toc171279164)

[B.3.3 Example 3 24](#_Toc171279165)

[Annex C (informative) Conformance and Reference Software 26](#_Toc171279166)

[C.1 Display power reduction using display adaptation 26](#_Toc171279167)

[C.1.1 Conformance test vectors 26](#_Toc171279168)

[C.1.2 Reference software 26](#_Toc171279169)

[C.2 Energy-efficient media selection 26](#_Toc171279170)

[C.2.1 Conformance test vectors 26](#_Toc171279171)

[C.2.2 Reference software 27](#_Toc171279172)

[Annex D (informative) Generation and Use of Green Metadata 28](#_Toc171279173)

[D.1 Decoder and display power indication 28](#_Toc171279174)

[D.1.1 Metadata generation at the server side 28](#_Toc171279175)

[D.1.2 Use of metadata at the client 29](#_Toc171279176)

[D.2 Display attenuation maps 32](#_Toc171279177)

[D.2.1 Metadata generation at the server side 32](#_Toc171279178)

[D.2.2 Use of metadata at the client 33](#_Toc171279179)

[Bibliography 34](#_Toc171279180)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO 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](https://www.iso.org/directives-and-policies.html)).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO *[had/had not]* received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at [www.iso.org/patents](http://www.iso.org/patents). ISO shall not be held responsible for identifying any or all such patent rights.

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](https://www.iso.org/foreword-supplementary-information.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](https://www.iso.org/members.html).

Introduction

Green metadata is used for representing information that will help to reduce the energy consumption in the video distribution chain, towards a more sustainable industry.

This document addresses technologies defining the carriage of green metadata for storage and delivery purposes. This document includes (but is not limited to):

* Storage and carriage of Green metadata using the ISO Base Media File Format (ISOBMFF) as specified in ISO/IEC 14496-12;
* Encapsulation, signalling, and streaming of Green metadata data in a media streaming system, e.g., dynamic adaptive streaming over HTTP (DASH) as specified in ISO/IEC 23009-1.

Information technology — MPEG systems technologies — Part xx: Carriage of green metadata

# Scope

This part of ISO/IEC 23001 defines a storage and delivery format for Green metadata as defined in part ISO/IEC 23001-11. The Green metadata are timed metadata which can be associated with other tracks in the ISO Base Media File Format. Timed metadata such as power consumption information and their metrics are defined in this part for carriage in files based on the ISO Base Media File Format (ISO/IEC 14496-12). These metadata can be used for multiple purposes, including optimizing power consumption during playback and supporting dynamic adaptive streaming.

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

ISO/IEC 23001-11:2023, *Information technology — MPEG Systems Technologies — Part 11: Energy-Efficient Media Consumption (Green Metadata)*

ISO/IEC 23009-1:2022, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

ISO/IEC/TR 23009-3*, Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 3: Implementation guidelines*

ISO/IEC 23001-10*, Information technology — MPEG systems technologies — Part 10: Carriage of Timed Metadata Metrics of Media in ISO Base Media File*

ISO/IEC 14496-5*, Information technology — Coding of audio-visual objects — Part 5: Reference software | Rec. ITU-T H.264.2: Reference software for ITU-T H.264 advanced video coding*

ISO/IEC 14496-15:2022*, Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

# Terms, definitions, and abbreviated terms

## Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 23001-11 and the following apply.

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

* ISO Online browsing platform: available at <https://www.iso.org/obp>
* IEC Electropedia: available at <https://www.electropedia.org/>

3.1.1

term

text of the definition

Note 1 to entry: Text of the note.

[SOURCE: …]

3.1.2

term

text of the definition

## Abbreviated terms

|  |  |
| --- | --- |
| DASH | dynamic adaptive streaming over HTTP (specified in ISO/IEC 23009-1) |
| ISOBMFF | ISO base media file format (specified in ISO/IEC 14496-12) |
|  |  |

# Overview

## Overall architecture for carriage of Green metadata

TBD

# Carriage of Green Metadata in ISO Base Media File Format

## General

If Green Metadata is carried in an ISO Base Media File Format, it shall be carried in the metadata tracks within the ISO Base Media File Format. Different Green Metadata types and corresponding storage formats are identified by their unique sample entry codes.

A metadata track carrying Green metadata is linked to the track it describes by means of a ‘cdsc’ (content describes) track reference.

## Decoder Power Indication Metadata

### Definition

Sample Entry Type: ‘depi’   
Container: Sample Description Box (‘stsd’)  
Mandatory: No  
Quantity: 0 or 1

The Decoder-Power Indication Metadata is defined in ISO/IEC 23001-11. It provides decoder complexity reduction ratios for the media track to which the metadata track refers by means of ‘cdsc’ reference.

### Syntax

The decoder power indication metadata sample entry shall be as follows.

class DecoderPowerIndicationMetaDataSampleEntry()   
 extends MetaDataSampleEntry (‘depi‘) {  
  
}

The Decoder-Power Indication sample shall conform to the following syntax:

aligned(8) class DecoderPowerIndicationMetaDataSample(){   
 unsigned int(8) Dec\_ops\_reduction\_ratio\_from\_max;  
 signed int(16) Dec\_ops\_reduction\_ratio\_from\_prev;  
}

### Semantics

Semantics are defined in ISO/IEC 23001-11.

## Display Power Reduction Metadata

The Display-Power Reduction Metadata is defined in ISO/IEC 23001-11. Display Power Reduction Metadata provides frame statistics and quality indicators for the media track that the metadata track refers to by means of ‘cdsc’ reference. These metadata allow the client to attain a specified quality level by scaling frame-buffer pixels and to reduce power correspondingly by decreasing the display backlight or OLED voltage.

Display-Power Reduction Metadata is of three types:

1. metadata that indicates power saving at different quality levels over the sample duration. This metadata shall use the 'dipi’ (display power indication) sample entry type.

2. metadata that allows fine control of the display to achieve power reduction at a specified quality level. This metadata shall use the ’dfce’ (display fine control) sample entry type.

3. metadata that conveys pixel-wise information that can be applied to frames in the original video sequence to reduce the total energy consumption resulting from rendering the frames of that video. This metadata shall use the ‘amid’ (display attenuation map) sample entry type.

Static metadata for the display fine control is stored in the sample entry. Dynamic metadata is stored in the samples.

### Display Power Indication Metadata

#### Definition

Sample Entry Type: ‘dipi’   
Container: Sample Description Box (‘stsd’)  
Mandatory: No  
Quantity: 0 or 1

This metadata indicates potential power saving at different quality levels over the sample duration.

#### Syntax

Display Power Indication Metadata shall use the following sample entry:

aligned(8) class DisplayPowerIndicationMetaDataSampleEntry() extends MetaDataSampleEntry (‘dipi‘) {  
}

The Display Power Indication sample shall use the following syntax:

class QualityLevels (num\_quality\_levels) {  
 unsigned int(8) rgb\_component\_for\_infinite\_psnr;  
 for (i = 1; i <= num\_quality\_levels; i++) {  
 unsigned int(8) max\_rgb\_component;   
 unsigned int(8) scaled\_psnr\_rgb;  
 }   
}

aligned class DisplayPowerIndicationMetaDataSample () {  
 unsigned int(4) num\_quality\_levels;  
 unsigned int(4) reserved=0;}  
 QualityLevels(num\_quality\_levels)  
}

Please note that the PSNR variables appearing in the syntax presented above are as defined in ISO/IEC 23001-11 and should not be confused with the PSNR metric defined in clause 4.2 of ISO/IEC 23001-10.

#### Semantics

Semantics are defined in ISO/IEC 23001-11.

### Display Fine Control Metadata

#### Definition

Sample Entry Type: ‘dfce’  
Container: Sample Description Box (‘stsd’)  
Mandatory: No  
Quantity: 0 or 1

The Display Fine Control Dynamic Metadata is stored in the samples and is associated with one or more video frames.

The Decoding Time to Sample box provides the decoding time for the sample so that the metadata contained therein is made available to the display with sufficient lead time relative to the video composition time. Note that the video composition time and metadata composition time are identical. The lead time is required because display settings must be adjusted in advance of presentation time for correct operation. If num\_constant\_backlight\_voltage\_time\_intervals > 1, then the lead time should be larger than the largest constant\_backlight\_voltage\_time\_interval.

#### Syntax

The Display Fine Control Metadata sample entry shall store static metadata as follows.

class DisplayFineControlMetaDataSampleEntry()   
 extends MetaDataSampleEntry (‘dfce‘) {  
 DisplayFineControlConfigurationBox();  
}  
  
aligned(8) class DisplayFineControlConfigurationBox   
 extends FullBox(‘dfcC’, version = 0, flags = 0) {  
 unsigned int(2) num\_constant\_backlight\_voltage\_time\_intervals;  
 unsigned int(6) reserved = 0;   
 unsigned int(16)constant\_backlight\_voltage\_time\_interval[   
 num\_constant\_backlight\_voltage\_time\_intervals ];  
 unsigned int(2) num\_max\_variations;  
 unsigned int(6) reserved = 0;  
 unsigned int(16) max\_variation[ num\_max\_variations ];  
}

The Display Fine Control Metadata sample shall use the following syntax:

class QualityLevels (num\_quality\_levels) {  
 unsigned int(8) rgb\_component\_for\_infinite\_psnr;  
 for (i = 1; i <= num\_quality\_levels; i++) {  
 unsigned int(8) max\_rgb\_component;   
 unsigned int(8) scaled\_psnr\_rgb;  
 }   
}

class MetadataSet (num\_quality\_levels) {  
 unsigned int(8) lower\_bound;   
 if (lower\_bound > 0)  
 unsigned int(8) upper\_bound;  
 QualityLevels(num\_quality\_levels);  
}

class DisplayPowerReductionMetaDataSample {  
 unsigned int(4) num\_quality\_levels;  
 unsigned int(4) reserved = 0;  
  
 for (k=0; k<num\_constant\_backlight\_voltage\_time\_intervals; k++)   
 for (j = 0; j < num\_max\_variations; j++)   
 MetadataSet(num\_quality\_levels);  
}

#### Semantics

Semantics are defined in ISO/IEC 23001-11.

### Display Attenuation Map Metadata

A display attenuation map video is a 2D video where each frame conveys pixel-wise information that can be applied to a corresponding frame in the original video sequence through a processing operation to reduce the total energy consumption resulting from rendering the frames of that video.

#### Attenuation Map Information Box

An AttenuationMapInformationBox contains information about the characteristics of the display attenuation map data stream carried by the track in which it is signalled. It is identified by the 4CC ‘amid’. The information in the payload of the track may include pre-processing operations that should be applied to the samples of the attenuation map sample as well as other information that may distinguish one display attenuation map track from another for the same content, which enables a player to select the most suitable track based on its energy reduction strategy.

##### Syntax

aligned(8) class AMIApproximationModel() {

bit(4) reserved = 0;

unsigned int(4) ami\_map\_approx\_model;

}

aligned(8) class AMIWindowInfo() {

unsigned int(8) ami\_window\_x;

unsigned int(8) ami\_window\_y;

unsigned int(8) ami\_window\_width;

unsigned int(8) ami\_window\_height;

}

aligned(8) class AMIVideoQualityInfo() {

bit(5) reserved = 0;

unsigned int(3) ami\_quality\_metric;

unsigned int(8) ami\_quality\_reduction;

}

aligned(8) class AMIPreprocessingInfo() {

bit(6) reserved = 0;

unsigned int(2) ami\_preprocessing\_type;

unsigned int(8) ami\_max\_value;

unsigned int(8) ami\_preprocessing\_scale;

}

aligned(8) class AttenuationMapInformationBox() extends FullBox('amid', version = 0, flags) {

bit(3) reserved = 0;

unsigned int(1) ami\_preprocessing\_info\_present\_flag;

unsigned int(1) ami\_approx\_model\_present\_flag;

unsigned int(1) ami\_window\_info\_present\_flag;

unsigned int(1) ami\_video\_quality\_info\_present\_flag;

unsigned int(5) ami\_energy\_reduction\_rate;

unsigned int(4) ami\_display\_model;

unsigned int(4) ami\_attenuation\_use\_idc;

unsigned int(4) ami\_attenuation\_component\_idc;

if (ami\_preprocessing\_info\_present) {

AMIPreprocessingInfo();

}

if (ami\_approx\_model\_present\_flag) {

AMIApproximationModel();

}

if (ami\_window\_info\_present\_flag) {

AMIWindowInfo();

}

if (ami\_video\_quality\_info\_present\_flag) {

AMIVideoQualityInfo();

}

}

##### Semantics

The semantics of the fields defined in AttenuationMapInformationBox are as follows:

ami\_processing\_info\_present\_flag is a flag indicating whether preprocessing information is present. Value 1 indicates that the box contains preprocessing information.

ami\_approx\_model\_present\_flag is a flag indicating whether approximation model information is present. Value 1 indicates that the box contains approximation model information.

ami\_window\_info\_present\_flag is a flag indicating whether window information is present. Value 1 indicates that the box contains window information.

ami\_video\_quality\_info\_present\_flag is a flag indicating whether video quality information is present. Value 1 indicates that the box contains video quality information.

ami\_energy\_reduction\_rate indicates the expected energy saving rate (percentage) when the video is displayed after applying the attenuation map sample values on the sample values of the associated video.

ami\_display\_model indicates the display models on which the attenuation map sample values may be used. The semantics of the bits of this field are as described in Table 1.

1. Semantics of the bits of the ami\_display\_model field.

|  |  |
| --- | --- |
| **Bit number** | **Display Model** |
| 0 | Transmissive pixel |
| 1 | Emissive pixel |
| 2 .. 3 | Reserved for future use |

ami\_attenuation\_use\_idc indicates which operation should be used to apply the attenuation map sample values to the corresponding frame in the associated video before rendering the frame on the display. The semantics of the values assigned to this field are described in Table 2.

1. Semantics of the values assigned to ami\_attenuation\_use\_idc.

|  |  |
| --- | --- |
| **Value** | **Description** |
| 0 | The attenuation map sample values should be added to the video frame. |
| 1 | The attenuation map sample values should be subtracted from the video frame. |
| 2 | The attenuation map sample values should be multiplied by the video frame. |
| 3 | The attenuation map sample values should be applied to the video frame according to a proprietary user-defined process. |

ami\_attenuation\_component\_idc indicates on which color component(s) of the associated video to apply the attenuation map using the operation defined by ami\_attenuation\_use\_idc. It also specifies how many components the attenuation map has. The semantics of the values assigned to this field are described in Table 3.

1. Semantics of the values assigned to ami\_attenuation\_component\_idc.

|  |  |
| --- | --- |
| **Value** | **Description** |
| 0 | Attenuation map contains only one component, and this component should be applied to the luma component of media video. |
| 1 | Attenuation map contains two components, and the first component should be applied to the luma component of the media video, and the second component should be applied to both chroma components of the media video. |
| 2 | Attenuation map contains only one component, and this component should be applied to the luma component and the chroma components of the media video. |
| 3 | Attenuation map contains only one component, and this component should be applied to the RGB components (after YUV to RGB conversion) of the media video. |
| 4 | Attenuation map contains three components and these components should be applied respectively to the luma and chroma components of the media video. |
| 5 | Attenuation map contains three components and these components should be applied, respectively, to the RGB components (after YUV to RGB conversion) of the media video. |
| 6 | User-defined |
| 7 .. 15 | Reserved for future use |

ami\_preprocessing\_type indicates which type of pre-processing interpolation model should be used to re-sample the attenuation map sample values at the same resolution as the associated video before applying it to the associated video frame. The semantics of the values assigned to this field are described in ‎Table 4.

1. Semantics of the values assigned to ami\_preprocessing\_type.

|  |  |
| --- | --- |
| **Value** | **Interpolation Type** |
| 0 | Bicubic interpolation. |
| 1 | Bilinear interpolation. |
| 2 | Lanczos interpolation. |
| 3 | User-defined interpolation process. |

ami\_max\_value indicates the maximum value of the attenuation map. This value can be optionally used to further adjust the dynamic range of the encoded attenuation map in the scaling process.

ami\_preprocessing\_scale indicates which scaling should be applied to obtain the attenuation map sample values before applying them on the sample values of the associated video.

* Value 0 specifies that a scaling of 1/255 should be applied.
* Value 1 specifies that a proprietary user-defined scaling operation is applied.

ami\_map\_approx\_model specifies the model used to extrapolate the attenuation map with individual energy reduction rate to another set of attenuation map with a different energy reduction rate. The semantics of the values assigned to this field are as follows:

* Value 0 specifies a linear scaling of the attenuation map sample values given its ami\_energy\_reduction\_rate value.
* Values 1-3 are reserved for future use.
* Value 4 specifies that a proprietary user-defined process is used to perform the approximation and infer the attenuation map sample values for different energy reduction rates.

ami\_window\_x indicates the x-coordinate of the top-left corner of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to.

ami\_window\_y indicates the y-coordinate of the top-left corner of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to.

ami\_window\_width indicates the width, in number of pixels, of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to.

ami\_window\_height indicates the height, in number of pixels, of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to.

ami\_quality\_metric indicates the type of the objective quality metric used for the measured quality reduction value resulting from applying the attenuation map to the video content and assigned to the ami\_quality\_reduction field. The semantics of the values assigned to this field are described in ‎Table 5.

1. Semantics of the values assigned to ami\_quality\_metric.

|  |  |
| --- | --- |
| **Value** | **Quality Metric** |
| 0 | PSNR |
| 1 | SSIM |
| 2 | wPSNR |
| 3 | WS-PSNR |
| 4 | User-defined |

ami\_quality\_reduction specifies the percentage of quality reduction in the media video as a result of applying the attenuation map to it.

#### Display attenuation map tracks

A display attenuation map track is a restricted video track with the sample entry type 'resv'. The original sample entry type, which is based on the video codec used for encoding the stream, is stored within the OriginalFormatBox in the RestrictedSchemeInfoBox. The scheme\_type field in SchemeTypeBox shall be set to 'gmat', indicating an *attenuation map restricted scheme*. The SchemeInformationBox shall include an AttenuationMapInformationBox, as defined in the previous section. In the track header, the track\_in\_movie flag shall be set to 0 to indicate that this track should not be presented alone.

##### Association with Video Tracks

A TrackReferenceTypeBox with the reference type 'gmam' shall be added to a TrackReferenceBox within the TrackBox of the track carrying the attenuation map data. The TrackReferenceTypeBox shall contain an array of track\_IDs designating the identifiers for the referenced video tracks.

##### Sample Format

Each sample in an attenuation map track carries a sequence of video NAL units corresponding to the encoded attenuation map for a single access unit (AU) in the associated video track(s) and are encapsulated based on the sample formats defined in ISO/IEC 14496-15:2022.

##### Signalling Alternative Attenuation Map Tracks

Multiple display attenuation map tracks may be present in a ISOBMFF file. When more than one version of an attenuation map is available for the same video track in the ISOBMFF container (e.g., different energy consumption levels, different video qualities, etc.), each version is carried in a separate display attenuation map track.

Display attenuation map tracks that are alternatives of each other shall be signalled as alternatives of each other by either setting the alternate\_group field in their respective TrackHeaderBox(es) to the same value or grouping the tracks together with an EntityToGroupBox with grouping\_type equal to 'altr', indicating that the attenuation map tracks which are mapped to this grouping are alternatives to each other, and only one of them should be processed.

# Encapsulation and Signalling in MPEG-DASH

## General

This clause explains how the green metadata can be computed at the server for adaptive streaming scenarios and how such metadata can be used at the client.

In the context of DASH delivery, a specific adaptation set within the MPD can define the available green metadata representations and their association to the available media representations, using the signalling mechanisms specified in ISO/IEC 23009‑1:2022 and ISO/IEC 23009‑3 [1].

## Decoder Power Indication

### Metadata signalling in the MPD manifest file

In the DASH context, the metadata files created for one or multiple video representations are considered as metadata representations. The available metadata representations are signalled in a specific adaptation set within the MPD. The association of a metadata representation with a media representation is signalled in the MPD through the @associationId and @associationType attributes. A metadata Segment and its associated media Segment(s) are time aligned on Segment boundaries.

The decoder-power indication metadata representation is associated with a single media representation as shown in Figure 6‑2.



**Figure 6‑2: One metadata representation for one media representation.**

## Display Power Indication

### Metadata signalling in the MPD manifest file

The display-power indication metadata representation is associated with all the available media representations as shown in Figure 6‑3.



**Figure 6‑3: One metadata representation for all media representations.**

## Display Attenuation Map Information

### Metadata signalling in the MPD manifest file

Display attenuation map Representations for a video are signalled in their own Adaptation Set within the MPD file. This Adaptation Set is henceforth referred to as a Display Attenuation Map Adaptation Set. The @codecs attribute for a Display Attenuation Map Adaptation Set, or Representations of this Adaptation Set if @codecs is not signalled for the **AdaptationSet** element, is set based on the respective codec used for encoding the attenuation map. The value of @codecs shall be set to 'resv.gmat.XXXX', where XXXX corresponds to the four-character code (4CC) of the video codec and shall be identical to the original\_format field in the RestrictedSchemeInfoBox of the sample entry of the corresponding ISOBMFF track, clause ‎5.3.3.

The association of a display attenuation map Representation with one or more video media Representations is signalled in the MPD through the @associationId and @associationType attributes. The @associationId attribute is set to a space separated list that includes the @id attribute values of the associated video Representations. The @associationType attribute shall be set to “amit”, indicating that this association is for display attenuation map information.

Multiple display attenuation maps with different characteristics (e.g., different values of ami\_energy\_reduction\_rate, different values of display model, different resolutions, etc.) can be generated for the same video Representation and each alternative attenuation map is represented by one Representation in the Display Attenuation Map Adaptation Set. As specified in subclause ‎5.3.3.2.3, these display attenuation map Representations that are alternatives of each other shall be indicated by the same value of alternate\_group in the TrackHeaderBox.

#### Descriptors

To signal the characteristics of the display attenuation map(s) carried by a Display Attenuation Map Adaptation Set, an AttenuationMap descriptor is defined. The XML elements and attributes for this descriptor are defined in a separate namespace "urn:mpeg:mpegI:green:2023”. The namespace designator “green:” is used to refer to this name space.

This descriptor is an EssentialProperty descriptor with the @schemeIdUri attribute set to the unique URI "urn:mpeg:mpegI:green:2023:ami".

At most one single AttenuationMap descriptor shall be present at the Adaptation level and/or at Representation level for each Representation in a Display Attenuation Map Adaptation Set.

The @value attribute of the AttenuationMap descriptor shall not be present. ‎Table 6 lists all the elements and attribute of the AttenuationMap descriptor.

1. Elements and attributes of the AttenuationMap descriptor.

|  |  |  |  |
| --- | --- | --- | --- |
| **Elements and attributes** | **Use** | **Data type** | **Description** |
| **AMI** | M | green:AMInfoType | An element whose attributes and sub-elements specify information for the display attenuation map present in the Representation(s) of the Adaptation Set. |
| **AMI**@displayModel | O | green:DisplayModelType | Indicates the display models on which the attenuation map may be used.  For ISO Base Media File Format Segments, the value of the @displayModel attribute shall be equal to ami\_display\_model in the AttenuationMapInformationBox in sample entries of the Initialization Segment. |
| **AMI**@attenuationUse | O | green:FourValueRangeType | Indicates which operation should be used to apply the attenuation map sample values to the corresponding frame in the associated video before rendering the frame on the display.  For ISO Base Media File Format Segments, the value of the @attenuationUse attribute shall be equal to ami\_attenuation\_use\_idc in the AttenuationMapInformationBox in sample entries of the Initialization Segment. |
| **AMI**@attenuationComponentIdc | O | green:AttenuationComponentIdcType | Indicates on which color component(s) of the associated video to apply the attenuation map using the operation defined by @attenuationUse. It also specifies how many components the attenuation map has.  For ISO Base Media File Format Segments, the value of the @attenuationComponentIdc attribute shall be equal to ami\_attenuation\_component\_idc in the AttenuationMapInformationBox in sample entries of the Initialization Segment. |
| **AMI**@energyReductionRate | O | green:EnergyReductionRateType | Indicates the expected energy saving rate (percentage) when the video is displayed after applying the attenuation map sample values on the sample values of the associated video.  For ISO Base Media File Format Segments, the value of the @energyReductionRate attribute shall be equal to ami\_energy\_reduction\_rate in the AttenuationMapInformationBox in sample entries of the Initialization Segment. |
| **AMI.QualityInfo** | O | green:QualityInfoType | An element whose attributes specify the quality reduction rate when the attenuation map is applied to the associated video representation. |
| **AMI.QualityInfo**@metric | M | xs:string | Indicates the video quality metric used to assess the quality reduction. |
| **AMI.QualityInfo**@reduction | M | xs:unsignedByte | Indicates the percentage of quality reduction in the media video as a result of applying the attenuation map to it. |
| **AMI.PreprocessingInfo** | O | green:PreprocessingInfoType | An element whose attributes and sub-elements specify information for the pre-processing interpolation model that should be used. |
| **AMI.PreprocessingInfo**@type | M | green:FourValueRangeType | Indicates which type of pre-processing interpolation model should be used to re-sample the attenuation map sample values at the same resolution as the associated video before applying it to the associated video frame. |
| **AMI.PreprocessingInfo**@max\_value | M | xs:unsignedByte | Indicates the maximum value of the attenuation map. This value can be optionally used to further adjust the dynamic range of the encoded attenuation map in the scaling process. |
| **AMI.PreprocessingInfo**@scale | M | xs:unsignedByte | Indicates the scaling that should be applied to obtain the attenuation map sample values before applying them on the sample values of the associated video. |
| **AMI.ApproxModel** | O | green:ApproxModelType | An element whose attributes specify the model used to extrapolate the attenuation map with individual energy reduction rate to another set of attenuation map with a different energy reduction rate. |
| **AMI.ApproxModel**@type | M | green:FourValueRangeType | Indicates the model used to extrapolate the attenuation map with a certain energy reduction rate to another set of attenuation map with a different energy reduction rate. |
| **AMI.WindowInfo** | O | green:WindowInfoType | An element whose attributes specify a bounding window defining a region of the associated media video to apply the attenuation map to. |
| **AMI.WindowInfo**@x | M | xs:unsignedByte | Indicates the x-coordinate of the top-left corner of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to |
| **AMI.WindowInfo**@y | M | xs:unsignedByte | Indicates the y-coordinate of the top-left corner of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to |
| **AMI.WindowInfo**@width | M | xs:unsignedByte | Indicates the width, in number of pixels, of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to |
| **AMI.WindowInfo**@height | M | xs:unsignedByte | Indicates the height, in number of pixels, of the bounding window defining a region of the associated media video to apply the attenuation map carried by the display attenuation map track to. |
| Key:  For attributes: M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory.  For elements: <minOccurs>..<maxOccurs> (N=unbounded)  Elements are **bold**; attributes are non-bold and preceded with an @. | | | |

1. (normative)  
     
   Green metadata MPEG-DASH schema

<?xml version="1.0" encoding="UTF-8"?>

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"

targetNamespace="urn:mpeg:mpegI:green:2023"

xmlns:green="urn:mpeg:mpegI:green:2023"

elementFormDefault="qualified">

<xs:element name="AMI" type="green:AMInfoType" />

<xs:complexType name="AMInfoType">

<xs:sequence>

<xs:element name="QualityInfo" type="green:QualityInfoType" minOccurs="0"

maxOccurs="1"/>

<xs:element name="PreprocessingInfo" type="green:PreprocessingInfoType"

minOccurs="0" maxOccurs="1" />

<xs:element name="WindowInfo" type="green:WindowInfoType" minOccurs="0"

maxOccurs="1" />

<xs:element name="ApproxModel" type="green:ApproxModelType" minOccurs="0"

maxOccurs="1" />

<xs:any namespace="##other" processContents="lax" minOccurs="0"

maxOccurs="unbounded" />

</xs:sequence>

<xs:attribute name="displayModel" type="green:DisplayModelType" />

<xs:attribute name="attenuationUse" type="green:FourValueRangeType" />

<xs:attribute name="attenuationComponentIdc"

type="green:AttenuationComponentIdcType" />

<xs:attribute name="energyReductionRate"

type="green:EnergyReductionRateType" />

<xs:anyAttribute namespace="##other" processContents="lax" />

</xs:complexType>

<xs:complexType name="QualityInfoType">

<xs:attribute name="metric" type="xs:string" use="required" />

<xs:attribute name="reduction" type="xs:unsignedByte" use="required" />

<xs:anyAttribute namespace="##other" processContents="lax" />

</xs:complexType>

<xs:complexType name="PreprocessingInfoType">

<xs:attribute name="type" type="green:FourValueRangeType" use="required" />

<xs:attribute name="max\_value" type="xs:unsignedByte" use="required" />

<xs:attribute name="scale" type="xs:unsignedByte" use="required" />

<xs:anyAttribute namespace="##other" processContents="lax" />

</xs:complexType>

<xs:complexType name="ApproxModelType">

<xs:attribute name="type" type="green:FourValueRangeType" use="required" />

<xs:anyAttribute namespace="##other" processContents="lax" />

</xs:complexType>

<xs:complexType name="WindowInfoType">

<xs:attribute name="x" type="xs:unsignedByte" use="required" />

<xs:attribute name="y" type="xs:unsignedByte" use="required" />

<xs:attribute name="width" type="xs:unsignedByte" use="required" />

<xs:attribute name="height" type="xs:unsignedByte" use="required" />

<xs:anyAttribute namespace="##other" processContents="lax" />

</xs:complexType>

<xs:simpleType name="DisplayModelType">

<xs:restriction base="xs:unsignedByte">

<xs:minInclusive value="0"/>

<xs:maxInclusive value="3"/>

</xs:restriction>

</xs:simpleType>

<xs:simpleType name="AttenuationComponentIdcType">

<xs:restriction base="xs:unsignedByte">

<xs:minInclusive value="0"/>

<xs:maxInclusive value="15"/>

</xs:restriction>

</xs:simpleType>

<xs:simpleType name="EnergyReductionRateType">

<xs:restriction base="xs:unsignedByte">

<xs:minInclusive value="0"/>

<xs:maxInclusive value="99"/>

</xs:restriction>

</xs:simpleType>

<xs:simpleType name="FourValueRangeType">

<xs:restriction base="xs:unsignedByte">

<xs:minInclusive value="0"/>

<xs:maxInclusive value="3"/>

</xs:restriction>

</xs:simpleType>

</xs:schema>

1. (informative)  
     
   MPEG-DASH MPD examples
   1. Example MPD with decoder power indication metadata

The following XML file provides an example of an MPD for decoder-power indication metadata.

<?xml version=”1.0” encoding=”UTF-8”?>

<MPD

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="urn:mpeg:DASH:schema:MPD:XXXX"

xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:xxxx"

type="dynamic"

minimumUpdatePeriod="PT2S"

timeShiftBufferDepth="PT30M"

availabilityStartTime="2011-12-25T12:30:00"

minBufferTime="PT4S"

profiles="urn:mpeg:dash:profile:isoff-live:2011">

<BaseURL>http://cdn1.example.com/</BaseURL>

<BaseURL>http://cdn2.example.com/</BaseURL>

<Period>

<!-- Video -->

<AdaptationSet

id="video"

mimeType="video/mp4"

codecs="avc1.4D401F"

frameRate="30000/1001"

segmentAlignment="true"

startWithSAP="1">

<BaseURL>video/</BaseURL>

<SegmentTemplate timescale="90000" media="$Bandwidth$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="v0" width="320" height="240" bandwidth="250000"/>

<Representation id="v1" width="640" height="480" bandwidth="500000"/>

<Representation id="v2" width="960" height="720" bandwidth="1000000"/>

</AdaptationSet>

<!-- English Audio -->

<AdaptationSet mimeType="audio/mp4" codecs="mp4a.0x40" lang="en" segmentAlignment="0">

<SegmentTemplate timescale="48000" media="audio/en/$Time$.mp4a">

<SegmentTimeline>

<S t="0" d="96000" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="a0" bandwidth="64000" />

</AdaptationSet>

<!-- French Audio -->

<AdaptationSet mimeType="audio/mp4" codecs="mp4a.0x40" lang="fr" segmentAlignment="0">

<SegmentTemplate timescale="48000" media="audio/fr/$Time$.mp4a">

<SegmentTimeline>

<S t="0" d="96000" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="a0" bandwidth="64000" />

</AdaptationSet>

<!-- AdaptationSet carrying Green Video Information for Video -->

<AdaptationSet id="green\_video" codecs="depi"/>

<BaseURL>video\_green\_depi/</BaseURL>

<SegmentTemplate timescale="90000" media="$id$/$Time$.mp4m">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="gv0" bandwidth="1000" associationId="v0" associationType="cdsc"/>

<Representation id="gv1" bandwidth="1000" associationId="v1" associationType="cdsc"/>

<Representation id="gv2" bandwidth="1000" associationId="v2" associationType="cdsc"/>

</AdaptationSet>

</Period>

</MPD>.

* 1. Example MPD with display power indication metadata

The following XML file provides an example of an MPD for display-power indication metadata.

<?xml version=”1.0” encoding=”UTF-8”?>

<MPD

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="urn:mpeg:DASH:schema:MPD:XXXX"

xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:xxxx"

type="dynamic"

minimumUpdatePeriod="PT2S"

timeShiftBufferDepth="PT30M"

availabilityStartTime="2011-12-25T12:30:00"

minBufferTime="PT4S"

profiles="urn:mpeg:dash:profile:isoff-live:2011">

<BaseURL>http://cdn1.example.com/</BaseURL>

<BaseURL>http://cdn2.example.com/</BaseURL>

<Period>

<!-- Video -->

<AdaptationSet

id="video"

mimeType="video/mp4"

codecs="avc1.4D401F"

frameRate="30000/1001"

segmentAlignment="true"

startWithSAP="1">

<BaseURL>video/</BaseURL>

<SegmentTemplate timescale="90000" media="$Bandwidth$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="v0" width="320" height="240" bandwidth="250000"/>

<Representation id="v1" width="640" height="480" bandwidth="500000"/>

<Representation id="v2" width="960" height="720" bandwidth="1000000"/>

</AdaptationSet>

<!-- English Audio -->

<AdaptationSet mimeType="audio/mp4" codecs="mp4a.0x40" lang="en" segmentAlignment="0">

<SegmentTemplate timescale="48000" media="audio/en/$Time$.mp4a">

<SegmentTimeline>

<S t="0" d="96000" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="a0" bandwidth="64000" />

</AdaptationSet>

<!-- French Audio -->

<AdaptationSet mimeType="audio/mp4" codecs="mp4a.0x40" lang="fr" segmentAlignment="0">

<SegmentTemplate timescale="48000" media="audio/fr/$Time$.mp4a">

<SegmentTimeline>

<S t="0" d="96000" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="a0" bandwidth="64000" />

</AdaptationSet>

<!-- AdapatationSet carrying Green Video Information for Video -->

<AdaptationSet id="green\_video" codecs="dipi"/>

<BaseURL>video\_green\_dipi/</BaseURL>

<SegmentTemplate timescale="90000" media="$id$/$Time$.mp4m">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="gv0" bandwidth="1000" associationId="v0 v1 v2"   
 associationType="cdsc"/>

</AdaptationSet>

</Period>

</MPD>

* 1. Examples of MPD with display attenuation map metadata
     1. Example 1

This example demonstrates a DASH MPD with a presentation that includes a video Adaptation Set with three Representations and a Display Attenuation Map Adaptation Set with one Representation that is associated with the first video Representation. The attenuation map carried by the Display Attenuation Map Adaptation Set results in a 20% reduction in energy consumption when applied to the associated video Representation. The application of Display Attenuation Map Representation also results in a 5% reduction of the video quality in terms of PSNR.

<?xml version=”1.0” encoding=”UTF-8”?>

<MPD

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="urn:mpeg:DASH:schema:MPD:2011"

xmlns:green="urn:mpeg:mpegI:green:2023"

xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:xxxx"

type="static"

minBufferTime="PT4S"

profiles="urn:mpeg:dash:profile:isoff-on-demand:2011">

<BaseURL>http://cdn1.example.com/</BaseURL>

<BaseURL>http://cdn2.example.com/</BaseURL>

<Period>

<!-- Video -->

<AdaptationSet id="video" mimeType="video/mp4" codecs="avc1.4D401F"

frameRate="30000/1001" segmentAlignment="true" startWithSAP="1">

<BaseURL>video/</BaseURL>

<SegmentTemplate timescale="90000" media="$Bandwidth$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="v0" width="320" height="240" bandwidth="250000"/>

<Representation id="v1" width="640" height="480" bandwidth="500000"/>

<Representation id="v2" width="960" height="720" bandwidth="1000000"/>

</AdaptationSet>

<AdaptationSet id="ami" mimeType="video/mp4" codecs="resv.gmat.avc1.4D401F">

<BaseURL>attenuation\_maps/</BaseURL>

<SegmentTemplate timescale="90000" media="$id$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="ami0" bandwidth="1000" associationId="v0"

associationType="amit" />

<EssentialProperty schemeIdUri="urn:mpeg:mpegI:green:2023:ami">

<green:AMI displayMode="3" attenuationUse="1"

energyReductionRate="20">

<green:QualityInfo metric="PSNR" reduction="5" />

<green:ApproxModel type="0” />

</green:AMI>

</EssentialProperty>

</Representation>

</AdaptationSet>

</Period>

</MPD>

* + 1. Example 2

This example demonstrates a DASH MPD with a presentation that includes a video Adaptation Set with three Representations and a Display Attenuation Map Adaptation Set with two Representations that are associated with the first video Representation. The attenuation map carried by the first Representation of the Display Attenuation Map Adaptation Set results in a 20% reduction in energy consumption when applied to the video Representation, while the second Representation of the Display Attenuation Map Adaptation Set results in a 40% reduction in energy consumption when applied to the same video Representation.

<?xml version=”1.0” encoding=”UTF-8”?>

<MPD

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="urn:mpeg:DASH:schema:MPD:XXXX"

xmlns:green="urn:mpeg:mpegI:green:2023"

xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:xxxx"

type="dynamic"

minimumUpdatePeriod="PT2S"

timeShiftBufferDepth="PT30M"

availabilityStartTime="2011-12-25T12:30:00"

minBufferTime="PT4S"

profiles="urn:mpeg:dash:profile:isoff-live:2011">

<BaseURL>http://cdn1.example.com/</BaseURL>

<BaseURL>http://cdn2.example.com/</BaseURL>

<Period>

<!-- Video -->

<AdaptationSet id="video" mimeType="video/mp4" codecs="avc1.4D401F"

frameRate="30000/1001" segmentAlignment="true" startWithSAP="1">

<BaseURL>video/</BaseURL>

<SegmentTemplate timescale="90000" media="$Bandwidth$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="v0" width="320" height="240" bandwidth="250000"/>

<Representation id="v1" width="640" height="480" bandwidth="500000"/>

<Representation id="v2" width="960" height="720" bandwidth="1000000"/>

</AdaptationSet>

<!-- Display Attenuation Map Adapatation Set for energy reduction for video -->

<AdaptationSet id="ami" mimeType="video/mp4" codecs="resv.gmat.avc1.4D401F">

<BaseURL>attenuation\_maps/</BaseURL>

<SegmentTemplate timescale="90000" media="$id$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="ami0" bandwidth="1000" associationId="v0"

associationType="amit">

<EssentialProperty

schemeIdUri="urn:mpeg:mpegI:green:2023:ami">

<green:AMI displayMode="3" attenuationUse="1"

energyReductionRate="20">

<green:ApproxModel type="1” />

</green:AMI>

</EssentialProperty>

</Representation>

<Representation id="ami1" bandwidth="1000" associationId="v0"

associationType="amit">

<EssentialProperty

schemeIdUri="urn:mpeg:mpegI:green:2023:attenuationMap">

<green:AMI displayMode="3" attenuationUse="1"

energyReductionRate="40">

<ami:ApproxModel type="1” />

</green:AMI>

</EssentialProperty>

</Representation>

</AdaptationSet>

</Period>

</MPD>

* + 1. Example 3

This example is similar to Example 2, but it uses one instance of the descriptor at the Adaptation Set level containing common information for all Representations and additional descriptor instances with representation-specific information for each Representation.

<?xml version=”1.0” encoding=”UTF-8”?>

<MPD

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="urn:mpeg:DASH:schema:MPD:XXXX"

xmlns:green="urn:mpeg:mpegI:green:2023"

xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:xxxx"

type="dynamic"

minimumUpdatePeriod="PT2S"

timeShiftBufferDepth="PT30M"

availabilityStartTime="2011-12-25T12:30:00"

minBufferTime="PT4S"

profiles="urn:mpeg:dash:profile:isoff-live:2011">

<BaseURL>http://cdn1.example.com/</BaseURL>

<BaseURL>http://cdn2.example.com/</BaseURL>

<Period>

<!-- Video -->

<AdaptationSet id="video" mimeType="video/mp4" codecs="avc1.4D401F"

frameRate="30000/1001" segmentAlignment="true" startWithSAP="1">

<BaseURL>video/</BaseURL>

<SegmentTemplate timescale="90000" media="$Bandwidth$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="v0" width="320" height="240" bandwidth="250000"/>

<Representation id="v1" width="640" height="480" bandwidth="500000"/>

<Representation id="v2" width="960" height="720" bandwidth="1000000"/>

</AdaptationSet>

<!-- Display Attenuation Map Adapatation Set for energy reduction for video -->

<AdaptationSet id="ami" mimeType="video/mp4" codecs="resv.gmat.avc1.4D401F">

<EssentialProperty schemeIdUri="urn:mpeg:mpegI:green:2023:ami">

<green:AMI displayMode="3" attenuationUse="1" />

</EssentialProperty>

<BaseURL>attenuation\_maps/</BaseURL>

<SegmentTemplate timescale="90000" media="$id$/$Time$.mp4v">

<SegmentTimeline>

<S t="0" d="180180" r="432"/>

</SegmentTimeline>

</SegmentTemplate>

<Representation id="ami0" bandwidth="1000" associationId="v0"

associationType="amit">

<EssentialProperty schemeIdUri="urn:mpeg:mpegI:green:2023:ami">

<green:AMI energyReductionRate="20">

<green:ApproxModel type="1" />

</green:AMI>

</EssentialProperty>

</Representation>

<Representation id="ami1" bandwidth="1000" associationId="v0"

associationType="amit">

<EssentialProperty schemeIdUri="urn:mpeg:mpegI:green:2023:ami">

<green:AMI energyReductionRate="40">

<green:ApproxModel type="1" />

</green:AMI>

</EssentialProperty>

</Representation>

</AdaptationSet>

</Period>

</MPD>

1. (informative)  
     
   Conformance and Reference Software
   1. Display power reduction using display adaptation
      1. Conformance test vectors

One conformance ISO BMFF file, BasketballDrill\_28\_gamma.mp4m, which contains green metadata samples of ‘dfce’ sample entry type, as specified in subclause ‎5.3.2, is available at <http://standards.iso.org/iso-iec/23001/-11/ed-3/en>.

It is composed of a sample entry which contains static metadata and samples which contain dynamic metadata.

To verify conformance of a software implementation of ‘dfce’ green metadata samples parsing in an ISO BMFF file, the conformance file shall be used to check that extracted values match expected values given in the side text file provided with the conformance file.

* + 1. Reference software

A reference software for parsing and display of ‘dfce’ green metadata samples in ISO BMFF file is available at http://standards.iso.org/iso-iec/23001/-11/ed-3/en.

It is linked with ISO BMFF reference software libraries (IsoLib), which are available in ISO/IEC 14496-5.

A readme.txt is provided to explain how to produce the executable in a Windows or Linux environment.

The reference software takes the ISO BMFF metadata file (\*.mp4m) as input and produces a text file as output, which gives a full description of the metadata stored in the samples of the input file.

To verify conformance of test metadata files, the reference software shall be used to parse the test metadata files and to check them for syntactic correctness and valid ranges.

* 1. Energy-efficient media selection
     1. Conformance test vectors

A conformance test vector for decoder-power indication metadata is available at <http://standards.iso.org/iso-iec/23001/-11/ed-3/en>.

It consists of a set of:

* ten ISO BMFF video files, which provide ten AVC video representations, with (sub)segments duration of 2 s, at the following resolutions and bitrates:
  + 1920x1080p50 @ 10Mbps,
  + 1920x1080p50 @ 8Mbps,
  + 1600x900p50 @ 8Mbps,
  + 1600x900p50 @ 6Mbps,
  + 1280x720p50 @ 6Mbps,
  + 1280x720p50 @ 5Mbps,
  + 960x540p50 @ 5Mbps,
  + 960x540p50 @ 3.5Mbps,
  + 768x432p50 @ 3.5Mbps,
  + 768x432p25 @ 2.5Mbps.
* ten ISO BMFF metadata files, which provide associated decoder-power indication (‘depi’) metadata representation of each video representation,
* a manifest file, conformant to ISO/IEC 23009-1.

The ISO BMFF metadata files contain green metadata samples of ‘depi’ sample entry type, as specified in clause ‎‎5.2.1.

To verify conformance of a software implementation of ‘depi’ green metadata samples parsing in an ISO BMFF file, the conformance metadata files shall be used to check that extracted values match expected values given in the side text files provided with the conformance files.

* + 1. Reference software

A reference software for parsing and display of decoder-power (‘depi’) or display-power (‘dipi’) indication metadata in ISO BMFF file is available at <http://standards.iso.org/iso-iec/23001/-11/ed-3/en>.

It is linked with ISO BMFF reference software libraries (IsoLib), which are available in ISO/IEC 14496-5.

A readme.txt is provided to explain how to produce the executable in Windows or Linux environment.

The reference software takes the ISO BMFF metadata file (\*.mp4m) as input and produces a text file as output, which gives a full description of the metadata stored in the samples of the input file.

To verify conformance of test metadata files, the reference software shall be used to parse the test metadata files and to check them for syntactic correctness and valid ranges.

1. (informative)  
     
   Generation and Use of Green Metadata
   1. Decoder and display power indication
      1. Metadata generation at the server side

Given *N* video representations, the decoder-power indication metadata dec\_ops\_reduction\_ratio\_from\_max(i) (DOR-Ratio-Max(i)) and dec\_ops\_reduction\_ratio\_from\_prev(i) (DOR-Ratio-Prev(i)) are computed by the encoding system and provided by the server for i = 0 to *N* – 1, as shown in ‎Figure D.1. The display-power indication metadata is computed from one representation.



1. Green metadata computation and insertion.

The DOR-Ratio-Max(i) associated with each video representation i of a Segment is computed as the power-saving ratio from the most demanding video representation produced for the Segment, as defined in subclause 8.4.1 of ISO/IEC 23001-11.

The DOR-Ratio-Prev(i) associated with each video representation i of a Segment is computed as the power-saving ratio from the previous Segment of the same representation, as defined in subclause 8.4.1 of ISO/IEC 23001-11.

To produce the normative green metadata DOR-Ratio-Max(i) and DOR-Ratio-Prev(i) for a given Segment, the encoding system needs to estimate the decoding complexity of each video representation, as a number of processing cycles. Each sample which contains the DOR-Ratio values is then stored in a specific metadata file “$id$/$Time$.mp4m” (one for each Segment) using the format specified in subclause ‎5.2.

The display-power indication metadata is a list of (ms\_num\_quality\_levels + 1) pairs of the form (ms\_max\_rgb\_component[ i ], ms\_scaled\_psnr\_rgb[ i ]) as defined in 8.4.1. This metadata is produced without considering any constraint on max\_variation, the maximal backlight variation between two successive frames. It is also assumed that the backlight can be updated on each frame so that constant\_backlight\_voltage\_time\_interval is the inter-frame interval. Therefore, the display power-indication metadata provides the maximum power saving for a given quality level.

The display-power indication metadata is stored in a specific metadata file “$id$/$Time$.mp4m” (one for each Segment) using the format specified in subclause 5.3.

* + 1. Use of metadata at the client

The client (player/decoder) can determine its remaining battery life based on the energy consumption of the current representation it is using. If it detects that its battery life is insufficient for the total duration of the video content to be consumed (given parameter in the server or requirements of duration expressed by the user), the terminal can compute the power consumption saving ratio from the current representation.

Using the following information, the terminal can determine (for the next Segment) the best power-saving allocation strategy for the decoder and for the display:

— the decoder-power saving ratio of all available video representations in the next Segment from the current (selected) representation in the previous Segment,

— the impact of RGB component scaling on video quality for the next Segment,

— for the last Segments, the decoder and display consumption as a fraction of the total consumption.

From this information, the terminal determines which representation it needs to download and what is the appropriate scaling of RGB components for this representation.

It is observed that the decoder-power saving ratio of all available video representations from the current representation in the previous Segment is not directly given by the power-indication metadata. At the server, what is given is a list of two decoder operations reduction ratios per video representation:

— the first one is the ratio of each representation from the most energy-consuming one at a given period of time *T* (dashed arrows in ‎Figure D.2),

— the second one is the ratio of each representation at a given period of time *T* from the previous period of time *T* – 1 (continuous arrow in ‎Figure D.3).

The terminal can convert this list of ratios into a list of ratios from the current representation it was using in the previous Segment.

Let us define the following terms:

— *I*refRep the index, in the current Segment, of the representation which was used by the client terminal in the previous Segment,

— dec\_ops\_reduction\_ratio\_from\_max(i) the reduction ratio from the most energy consuming representation, received from the server,

— *R*decOpsReducFromRepRef(i) the reduction ratio from representation RefRep in the current Segment,

— *R*decOpsReducFromPrevRepRef(i) the reduction ratio from representation RefRep in the previous Segment,

It is possible to express *R*decOpsReducFromRepRef(i) from dec\_ops\_reduction\_ratio\_from\_max(i), using the following formula:

(D-1)

*R*decOpsReducFromRepRef(i) are represented by dotted arrows in ‎Figure D.2. It is then possible to express *R*decOpsReducFromPrevRepRef(i) from *R*decOpsReducFromRepRef(i), using the following formula:

(D-2)

*R*decOpsReducFromPrevRepRef(i) are represented by dashed arrows in ‎Figure D.3.

NOTE 1 Floating-point numbers are used for these computations.



1. Derivation of DecOpsReductionRatios within the current Segment.



1. Derivation of DecOpsReductionRatios within the current Segment from the previous Segment.

Using the mapping between Processing frequency of processors or devices and Power Supply Voltage and the mapping between Power Supply Voltage and Power consumption, the terminal can translate this list into a list of decoder-power saving ratios from the representation which was used in the previous Segment.

In the case where the total duration of the video content to be consumed is not known (case of live content for example), the terminal can display the expected remaining usage duration based on current battery level and the energy consumption of the current representation it is using. The user can therefore act on its terminal to increase this usage duration, which are translated into a power saving ratio as in the previous case.

NOTE 2 Complexity metrics, as defined in subclause 6.2 of ISO/IEC 23001-11 can be sent with each representation to allow the client to save energy by proactively invoking C-DVFS to make the selection of a representation work at its best for energy saving.

NOTE 3 The dec\_ops\_reduction\_ratio is known to be stable across software-based platforms.

* 1. Display attenuation maps
     1. Metadata generation at the server side

Given a video representation, a set of *N* display attenuation maps is generated by the encoding system provided by the server to reduce locally and smartly the brightness of the video representation’s frames. Each display attenuation map *i* optimizes the trade-off between the resulting video quality and energy reduction, for i = 0 to *N* – 1, as shown in ‎Figure D.4.

A black screen with white text

Description automatically generated

1. Display attenuation map computation and insertion.

The display attenuation map associated with a video representation is computed to reduce the energy consumption of displaying this representation by an energy-reduction rate indicated as a reduction percentage.

A display attenuation map Media Segment and its associated video Media Segment(s) are time aligned on Segment boundaries. A display attenuation map Media Segment is an ISOBMFF file which contains samples for a restricted video track, as shown in ‎Figure D.5, where each sample is a coded display attenuation map frame. The carriage of display attenuation map data in ISOBMFF files is specified in subclause ‎5.3.3.

A black background with white text

Description automatically generated

1. Alternate group of multiple display attenuation map representations.
   * 1. Use of metadata at the client

A DASH client is guided by the information provided in the MPD. The following is an example client behaviour for streaming videos with associated display attenuation maps using the signalling presented in subclause ‎6.4.

The client first issues an HTTP request and downloads the MPD file from the content server. It then parses the MPD file to generate a corresponding in-memory representation of the XML elements in the MPD file.

To identify available display attenuation maps in a Period, the streaming client scans the AdaptationSet elements to find Adaptation Sets with an AttenuationMap descriptor whose @schemeIdUri is set to the unique URL "urn:mpeg:mpegI:green:2023:ami". For each Representation in the Display Attenuation Map Adaptation Set, the client also identifies the associated Representation in the video Adaptation Set using the @associationId.

The streaming client selects one of the Representations of the video Adaptation Set based on its capabilities and the network conditions and downloads the Initialization Segment for that Representation. It then downloads the Initialization Segment for all Representations from the Display Attenuation Map Adaptation Set that are associated with the selected video Representation. The client then starts sequentially downloading Media Segments from the video Representation. The client regularly monitors the remaining power in the device’s battery and based on the battery level, remaining playback time, and the information signalled in the AttenuationMap descriptor selects one of the display attenuation map Representations.

The client subsequently downloads with each Media Segment from the video Representation a corresponding Media Segment from the display attenuation map Representation. The downloaded display attenuation map Media Segments are decoded, and the decoded frames are applied to the corresponding decoded frames from the video before rendering.

Bibliography

[1] ISO/IEC 23009-3, Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 3: Implementation guidelines