**** **ISO/IEC JTC 1/SC 29/AG 3 N185**

**ISO/IEC JTC 1/SC 29/AG 3  
MPEG Liaison and Communication   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** White paper on Green metadata 3rd edition

**Status:** Approved

**Date of document:** 2025-01-25

**Source:** ISO/IEC JTC 1/SC 29/AG 3

**Expected action:** None

**Action due date:** None

**No. of pages:** 10 (without cover page)

**Email of Convenor:** kyuheonkim@khu.ac.kr

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29ag3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/AG 3 MPEG LIAISON AND COMMUNICATION**

**ISO/IEC JTC 1/SC 29/AG 3 N185**

**Geneva, CH, Jan. 2025**

|  |  |
| --- | --- |
| **Title** | **White paper on Green metadata 3rd edition** |
| **Source** | **MPEG Liaison and Communication** |
| **Status** | **Approved** |
| **Serial Number** | **24932** |

|  |  |  |  |
| --- | --- | --- | --- |
| *Author(s) or Contact(s):* | Edouard François Christian Herglotz Yong He Claire-Hélène Demarty Olivier Le Meur | Email: | [edouard.francois@interdigital.com](mailto:edouard.francois@interdigital.com) [christian.herglotz@fau.de](mailto:christian.herglotz@fau.de) [yonghe@qti.qualcomm.com](mailto:yonghe@qti.qualcomm.com) [claire-helene.demarty@interdigital.com](mailto:claire-helene.demarty@interdigital.com) olivier.lemeur@interdigital.com |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Introduction

The specification ISO/IEC 23001-11 “Energy-efficient consumption (green metadata)” specifies metadata that facilitate reduction of energy usage during media consumption [1]. A 3rd edition of the specification was issued in 2023, with two more amendments (amd1 in 2024 and amd2 in 2025). This paper refers to this edition and its current two amendments.

Different types of metadata are specified for the following usages:

* reduced decoder power consumption;
* reduced display power consumption;
* media selection for joint decoder and display power reduction;
* quality recovery after low-power encoding.

The metadata aim at facilitating power consumption reduction in the encoding, decoding, and display processes while preserving the user’s quality of experience (QoE). However, it is also possible to use these metadata to get larger energy savings, at the expense of some QoE degradation.

The Green Metadata specification does not impact the normative decoding processes, even if the metadata can be used in a non-normative way by an encoder, decoder or a display, to control the energy impact of the video content delivery and consumption.

# Green metadata overview

Two main types of metadata are defined in the specification. A first type consists of metadata generated by a video encoder. The encoder generates a video bitstream, plus green metadata informing the receiver about decoding complexity of the delivered bitstream, or about the quality of the decoded content, or about attenuation maps conveyed in auxiliary pictures or in a side bitstream. These metadata can be used to control the energy usage of a decoder, or of a display (presentation subsystem). A second type is made of metadata produced by a decoder as feedback commands conveyed to the encoder, in the context of point-to-point communication, in order to better adapt the decoder or display energy consumption.

Figure 1 shows the functional architecture utilizing green metadata [1]. The media pre-processor is applied to analyze (and possibly to filter) the source content. The content is encoded to generate a bitstream for delivery. The bitstream is delivered to the receiver, decoded by a video decoder, and displayed by the output presentation subsystem.

The green metadata may be generated from the media encoder and the media pre-processor. Once generated, the metadata is encapsulated in the video bitstream. At the receiver, the metadata extractor processes the packets and sends the green metadata to a power optimization module for efficient power control. For instance, the power optimization module interprets the green metadata and then applies appropriate operations to reduce the video decoder’s power consumption, or to reduce the presentation subsystem’s power consumption. In addition, the power-optimization module can collect receiver information, such as remaining battery capacity, and send requests to the transmitter as green feedback to adapt the encoder operations for power-consumption reduction.



Figure 1 – Block diagram of a green encoder and decoder system (source [1]).

The following different sets of metadata are specified, adapted to different use cases:

* Metadata for decoder-power reduction
* Metadata for display power reduction
* Metadata for energy-efficient media selection
* Metrics for quality recovery after low-power encoding

The different sets are described in more details in the following sections. The new elements added in edition 3 are also emphasized.

The metadata may be signaled in-band (inside the bitstream) in the shape of a green metadata SEI message, through request-response mechanisms as used in point-to-point video communication, or in metadata tracks within media file format (such as DASH MPD files) in the video streaming context.

# Metadata for decoder-power reduction

Two types of metadata are specified to ease the energy reduction at the decoder side. The first type relates to complexity metrics signaled in-band in the bitstream via a green metadata SEI message. The second type relates to metadata interactively exchanged between a decoder and an encoder, where the decoder may request specific decoding operation reductions or encoder configurations, and the encoder may inform the decoder about bitstream decoding complexity or characteristics.

These two types of metadata are described in the following sub-sections.

## Complexity metrics for decoder-power reduction

Complexity metrics for decoder-power reduction are specified in section 6.2 and annex A of ISO/IEC 23001-11. They may be used to inform a decoder about the decoding complexity of a bitstream. They provide information related to the relative usage of several coding tools used to encode the video and to output the consumed bitstream.

These complexity metrics can for example be used to control the clock rate of a processor, using Dynamic Voltage Frequency Scaling (DVFS), as described in [2, 3]. The most recent paper [2] reports from around 20% to 30% energy saving, using a VVC reference software decoder implementation on a RaspberryPi platform (see Table 1).

Table 1 – Maximum achievable energy savings when switching from the maximum to the minimum CPU frequency (source [2]).

A table with numbers and letters

Description automatically generated

The complexity metrics may be signaled in the bitstream via a green metadata SEI message (as a first type of metadata identified by the syntax element green\_metadata\_type equal to 0). The metrics are codec-dependent, that is, they are specific to the coding tools implemented in the corresponding video coding specification. The considered coding tools have been carefully selected based on an exhaustive coding tools decoding complexity analysis (see for example [2] for VVC).

**New elements in edition 3**

In the 3rd edition, the complexity metrics for decoder-power reduction have been amended by the addition of two new SEI messages carrying complexity metrics for VVC and EVC, respectively. The previous edition already defines two SEI messages for AVC and HEVC, respectively.

An overview of the main characteristics of the complexity metrics for decoder-power reduction is provided in Table 2.

Table 2 – Characteristics of the complexity metrics for decoder-power reduction.

|  |  |
| --- | --- |
| **Generator** | Encoder |
| **Purpose** | reduced decoder power/energy consumption |
| **Data** | complexity metrics about the bitstream (number of occurrences of coding tools)  - Spatial and temporal granularity of the information can be controlled  - Complexity information defined for AVC, HEVC, VVC, EVC (3rd edition) |
| **Carriage** | in-band green metadata SEI message |

## Interactive signaling for remote decoder-power reduction

These metadata are specified in Section 6.3 of ISO/IEC 23001-11. They target interactive communications such as **point-to-point video conferencing**. In this use case, each device contains a transmitter and a receiver. A local device sends metadata that instructs the remote device to modify the decoding complexity of the bitstream and thus reduce local decoder-power consumption. The specification specifies request and response (acknowledgement) messages allowing to adapt the bitstream complexity to the local receiving device’s capabilities.

In edition 2, the signaling mechanism was limited to one request mode (first mode), comprising one syntax element indicating the percentage reduction of local decoding operations requested by the decoder to the remote encoder.

**New elements in edition 3**

The signaling has been enhanced in edition 3 by adding new modes allowing a finer control by a decoder of the encoding operations. In a second mode, syntax elements requesting to disable some specific coding tools (e.g., loop filter, bi-prediction, intra blocks in B-pictures, fractional-pel motion compensation) have been inserted. In a third mode, requested picture resolution and frame rate may be signaled. The fourth mode defines signaling to request global or partial cancellation of the last decoding operation reduction requests.

In addition, a response/acknowledgment message has been specified to enable the encoder to respond to a received request.

[4] reports energy savings results from various experiments based on HEVC and VVC decoder implementations. For example results are reported using the different decoder-power reduction metadata modalities defined in the specification, for decoding HEVC bitstreams (cf Table 3). Playing with the picture resolution and frame rate leads to significant energy reduction (from 43% to 89% depending on the target resolution/frame rate), with a similar behavior using an hardware or software implementation. The reported energy reduction by controlling coding tools is shown to be significant on software implementations, while the gain is lower using hardware implementations. In previous works, the savings were reported to be independent from the platform, which could be observed for both software [5] and hardware decoders on mobile platforms [6].

Table 3 – Measured energy savings and BDR values for hardware and software decoding (extracted from [4]).

A table with numbers and a number of data

Description automatically generated with medium confidence

An overview of the main characteristics of the interactive signaling for remote decoder-power reduction is provided in Table 4.

Table 4 – Characteristics of interactive signalling for remote decoder-power reduction.

|  |  |
| --- | --- |
| **Generator** | decoder / encoder |
| **Purpose** | reduced decoder (and encoder) power consumption in **point-to-point video conferencing** |
| **Data** | Request message from a decoder to the remote encoder:  - requested variation of local decoding operations  - requested de-activation of coding tools  - requested spatial and temporal scaling  - requested cancelation of previous requests  Response/acknowledgment message |
| **Carriage** | can be carried interactively using request-response messages in point-to-point communications |

# Metadata for display power reduction using display adaptation

These metadata are specified in Section 7 and annex A of ISO/IEC 23001-11. They are designed to attain display energy reductions by using **display adaptation** techniques.

In edition 2, a first type of metrics consisting of RGB-component statistics and quality indicators of the video content has been defined. They can be used to perform RGB picture components **picture-level** rescaling to set the best compromise between backlight/voltage reduction and picture quality. The specified metadata differ depending on whether the use-case is point-to-point transmission or point-to-multipoint transmission. Using a picture-level scaling based on the signaled metrics, [8] reports on average 32.5% power reduction with a limited quality impact. Note that this type of metadate mainly targets liquid crystal displays (LCDs).

The metadata may be conveyed in-band in a bitstream in a green metadata SEI message, as a specific type of green metadata (identified by the syntax element green\_metadata\_type equal to 2). Alternatively, they can be carried in an MPEG Transport Stream as specified in ISO/IEC 13818‑1 or in metadata tracks within the ISO base media file format (ISO/IEC 14496‑12), as specified in ISO/IEC 23001‑10. They can also be interactively signalled using request / response messages for applications supporting a signaling mechanism from a receiver to a transmitter.

**New elements in edition 3**

The set of metrics has been enhanced in edition 3 by adding new syntax elements related to attenuation maps that allow **pixel-level** adaptation. These metadata have been defined considering LED-based displays (OLED or MicroLED) technologies. The Attenuation Map Information (AMI) metadata indicate how to interpret and use attenuation maps for display adaptation. The attenuation maps may be carried as auxiliary pictures in a scalable bitstream, or in a separate single layer bitstream.

The perceptual benefits of pixel-wise display adaptation using the Attenuation Map concept, compared to a global per-picture linear scaling approach, are emphasized in [9], with for example, for a display energy reduction of 40%, an average VMAF score of 67.1 (in a range of [0,100]) using attenuation maps, instead of 58.2 using linear scaling (see Table 5). Figure 2 illustrates on a few examples the impact of the attenuation process for a target energy reduction ratio of 40%.

Table 5 – VMAF and PSNR for different sequences and energy reduction rates (extracted from [9]). For each case, the first number in brackets indicates the gain using a pixel-wise attenuation, and the second number in brackets using a picture-level linear scaling.

A table with numbers and letters

Description automatically generated

A bird walking on the ground

Description automatically generated

A bird walking on the ground

Description automatically generated

Figure 2 – First row, original images; Second row corresponds to an energy reduction rate of 40% respectively (extracted from [9]).

An overview of the main characteristics of the display power reduction metadata is provided in Table 6.

Table 6 – Characteristics of the display power reduction metadata.

|  |  |
| --- | --- |
| **Generator** | encoder / decoder |
| **Purpose** | display energy reduction by **display adaptation** |
| **Data** | - RGB-component statistics and quality indicators (different scaling levels, and associated quality values) – for RGB components rescaling to set the best compromise between backlight/voltage reduction and quality  - attenuation map information, indicative of the way to interpret and use attenuation maps conveyed in auxiliary pictures in the bitstream, or in another video stream |
| **Carriage** | - can be carried in transport stream as specified in ISO/IEC 13818-1 (MPEG-2 TS)  - can be carried in metadata tracks within the ISO Base Media File Format (ISO/IEC 14496-12), as specified in ISO/IEC 23001-10  - can be carried using in-band green metadata SEI message  - can be carried interactively using request-response messages in point-to-point communications |

# Energy-efficient media selection

These metadata are specified in Section 8 of ISO/IEC 23001-11. They are designed for **adaptive streaming** and aim at helping the receiver to select the representation with the optimal quality for a given power consumption. Metadata are provided (for example in DASH Media Presentation Description files) for each representation of a video segment. Two types of metadata are specified in ISO/IEC 23001-11 section 8:

* decoding operations variations (compared to the most demanding representation)
* display-power indication consisting in RGB-component statistics and quality indicators

In addition, attenuation map information described in previous section are also relevant for the purpose of energy-efficient media selection.

**New elements in edition 3**

Annex B.3 provides an informative description of how to perform energy-efficient media selection in adaptive streaming using display power indication green metadata for display energy reduction. In edition 3, this annex has been amended to describe how to use the concept of attenuation maps, and related Attenuation Map Information metadata, for adapting the energy at display side.

An overview of the main characteristics of the green metadata for energy-efficient media selection is provided in Table 7.

Table 7 – Characteristics of green metadata for energy-efficient media selection.

|  |  |
| --- | --- |
| **Generator** | encoder |
| **Purpose** | reduced decoder and display consumption in **adaptive streaming**  enable a client in an adaptive streaming session, such as DASH, to determine decoder and display power-saving characteristics of available video representations and to select the representation with the optimal quality for a given power-saving target |
| **Data** | - decoder-power indication of each available representation of a video segment  - display-power indication giving the maximum potential display power saving of a video segment for a specified number of quality levels |
| **Carriage** | - can be carried in metadata tracks within the ISO Base Media File Format (ISO/IEC 14496-12), as specified in ISO/IEC 23001-10  - in DASH context, a specific Adaptation Set within the MPD can define the Green Metadata Representations and their association to the media Representations, using the signaling mechanisms specified in ISO/IEC 23009 1:2014/Amd 2 and ISO/IEC 23009 3:2014/Amd 1 |

# Metrics for quality recovery after low-power encoding

These metadata are specified in Section 9 and annex A of ISO/IEC 23001-11. They consist of quality metrics, that can be used for **encoder power reduction in segmented delivery mechanism** such as DASH. The principle is to performencoding alternating high-quality and low-quality segments, with quality indication for each segment. A decoder can therefore perform guided post-decoding picture enhancement of the lower-quality pictures, using the preceding higher-quality pictures. For picture enhancement, e.g., quality-enhancement neural networks can be used. [1, 10] reports an average of 12.1% power reduction measured on a tablet, by alternating quantization parameters of successive segments at QP and QP+5. A PSNR gain of 0.2 dB is reported, compared to not applying the cross-segment enhancement to the low-quality segments.

Even if the specification describes the use of the quality metrics in the context of quality recovery after low-power encoding, the metrics could be signaled and used for other purposes not described in the specification.

The complexity metrics may be signaled in a green metadata SEI message, when the syntax element green\_metadata\_type signaled in the SEI message is equal to 1.

**New elements in edition 3**

Before edition 3, quality metrics specified in the AVC and HEVC green metadata SEI message are limited to PSNR. In edition 3, additional quality metrics are specified for the VVC green metadata SEI message, namely, wPSNR (weighted PSNR metric dedicated to HDR PQ content), WS-PSNR (weighted PSNR metric dedicated to 360° epipolar videos), SSIM. In addition, these metrics may be signaled per subpicture.

An overview of the main characteristics of the green metadata for quality recovery after low-power encoding is provided in Table 8.

Table 8 – Characteristics of the metrics for quality recovery after low-power encoding.

|  |  |
| --- | --- |
| **Generator** | encoder |
| **Purpose** | encoder power reduction in adaptive streaming  encoding alternating high-quality and low-quality segments, with quality indication for each segment |
| **Data** | quality metric(s) of the last associated picture or subpicture of each segment |
| **Carriage** | In-band green metadata SEI message |

# New specification ISO/IEC 23001-19 on Carriage of green metadata

Some aspects of how to carry green metadata in ISO base media file format (ISOBMFF) files are defined in one clause in ISO/IEC 23001-10 “Carriage of timed metadata metrics of media in ISO base media file format”, while other aspects on how to signal and use green metadata in adaptive streaming scenarios are captured across several clauses of ISO/IEC 23001-11. This means that the information related to the carriage and signalling of green metadata and how they relate to each other is distributed across several separate clauses in different specifications.

A new specification ISO/IEC 23001-19 has been initiated, aiming at defining a consolidated specification of the carriage of green metadata in media containers and their signalling for adaptive streaming delivery based on ISO/IEC 23001-10 and 23001-11. **Green metadata are defined in ISO/IEC 23001-11 while their carriage is defined in ISO/IEC 23001-19**.ISO/IEC 23001-19 was started in July 2024 and should reach FDIS stage in July 2025.

Figure 3 illustrates the main changes resulting from the creation of this new specification.

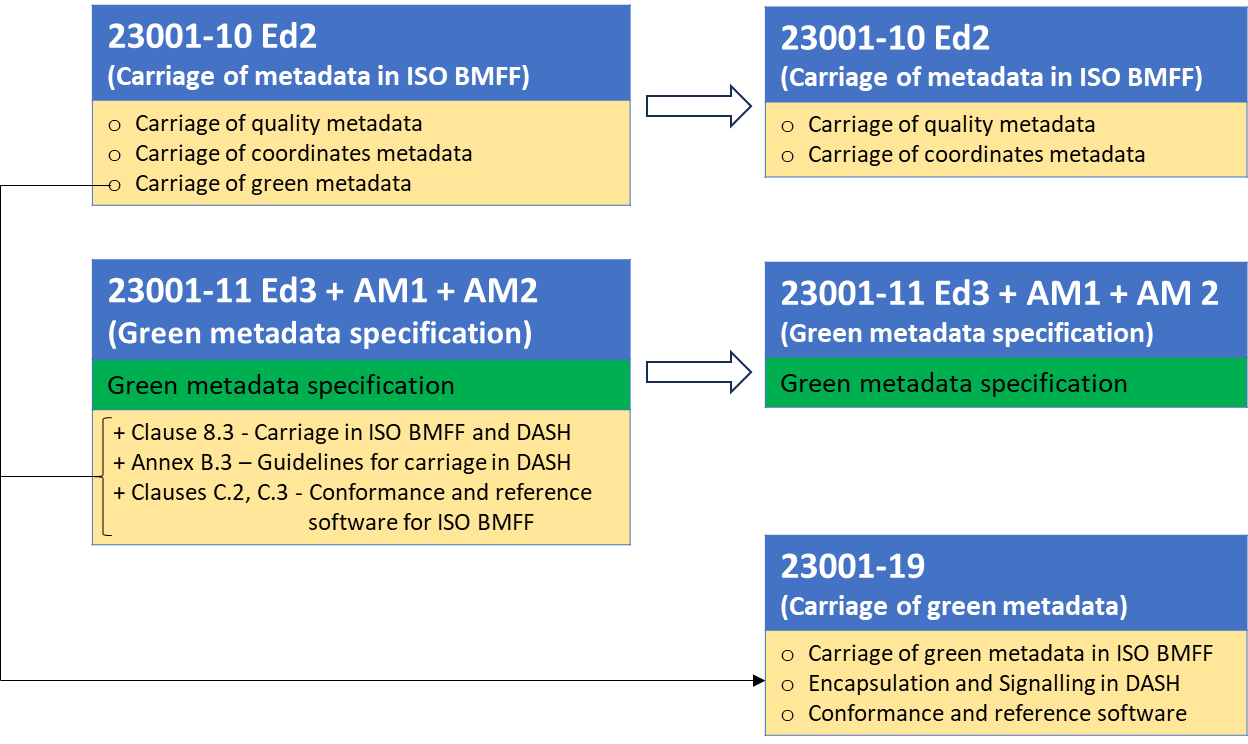


Figure 3: Consolidation of Carriage of green metadata in one single specification 23001-19.

# Conclusion

Energy impact has become a key parameter in video delivery and video consumption workflows. The Green Metadata specification aims at addressing this challenge by specifying metadata facilitating power consumption reduction in the encoding, decoding, and display processes, while preserving the user’s quality of experience.

A new edition (3rd edition) of Green Metadata specification, including two amendments, has been issued by end 2024, with new inputs related to Interactive signaling for remote decoder-power reduction, with the addition of a green metadata SEI for VVC and EVC, and with the specification of new metadata related to display adaptation mechanisms based on the concept of attenuation maps. For point-to-point video conferencing applications, the enhanced interactive signaling for remote decoder power reduction promises substantial power and energy savings for end users. The new VVC and EVC SEI messages specify corresponding ECV- and VVC-specific decoding complexity metrics, that a decoder can use to reduce its energy consumption. The VVC SEI message also defines new quality metrics that can be used for encoding energy reduction with post-decoding video quality enhancement. A new type of metadata for display adaptation, named attenuation map information (AMI), has also been added, allowing pixel-wise display adaptation while the former edition specified metrics targeting picture-level display adaptation (such as linear scaling).

# References

1. F. C. Fernandes, X. Ducloux, Z. Ma, E. Faramarzi, P. Gendron, J. Wen, “The Green Metadata Standard for Energy-Efficient Video Consumption”, IEEE MultiMedia, Vol 22, Issue 1, Jan. 2015.
2. C. Herglotz, M. Kranzler, R. Dai, A. Kaup, “Complexity Metrics for VVC Decoder Power Reduction in Green Metadata,” Proc. Picture Coding Symposium (PCS), June 2024
3. Y. Benmoussa, E. Senn, N. Derouineau, N. Tizon, J. Boukhobza, “Green metadata based adaptive DVFS for energy efficient video decoding”, In Proceedings of the IEEE Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS), Bremen, Germany, 21–23 September 2016.
4. C. Herglotz, M. Kranzler, A. Kaup, E. Francois, Y. He, “Extended Signaling Methods for Reduced Video Decoder Power Consumption Using Green Metadata,” IEEE Transactions on Circuits and Systems II: Express Briefs, Vol. 71, Issue 3, March 2024.
5. C. Herglotz; A. Heindel; and A. Kaup, “Decoding-energy-rate-distortion optimization for video coding”, IEEE Transactions on Circuits and Systems for Video Technology, 29:171–182, Nov 2017.
6. C. Herglotz; S. Coulombe; C. Vazquez; A. Vakili; A. Kaup; J.-C. Grenier, "Power modeling for video streaming applications on mobile devices", IEEE Access 8 (2020): 70234-70244.
7. RTP Control Protocol (RTCP) Messages for Temporal-Spatial Resolution, [draft-ietf-avtcore-rtcp-green-metadata-04 - RTP Control Protocol (RTCP) Messages for Temporal-Spatial Resolution](https://datatracker.ietf.org/doc/draft-ietf-avtcore-rtcp-green-metadata/), Oct. 2024.
8. F. Fernandes, E. Faramarzi, X. Li, Z. Ma, and X. Ducloux, "[Mobile Display Power Reduction for Video using Standardized Metadata](https://vision.nju.edu.cn/images/Papers/2018_TMC_da.pdf)", IEEE Trans. Mobile Computing, Vol 18, Issue 1, Jan. 2019.
9. O. Le Meur, C.-H. Demarty, E. Reinhard, F. Aumont, L. Blondé, "Energy-aware images: Quality of Experience vs Energy Reduction,", MHV’23, 2023.
10. G. Wen, W.-C. Wei, T. Mu, B. Li, S. Cheng, “Power reduction through cross segment decoding”, document ISO/IEC JTC1/SC29/WG11 M32439, Jan. 2014.