**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 N13468**

**April 2013, Incheon, Korea**

|  |  |
| --- | --- |
| **Source** | **Requirements Subgroup** |
| **Status** | **Approved** |
| **Title** | **Context, Objectives, Use Cases and Requirements for Green MPEG** |

**Abstract**

This document is submitted to discuss the technical areas, the corresponding requirements and the use cases for energy-efficient MPEG (Green MPEG).

# Context and Objectives of Green MPEG

# Background

#  MPEG has developed various technologies for multimedia processing and transport, such as MPEG-2 and MPEG-4. These technologies have been widely accepted and heavily used by various industries and applications, such as digital broadcasting, audio and video transport over the Internet and mobile phones, amongst others. Recently, power consumption associated with media processing on devices (especially on mobile devices) has become a critical issue because of the high demand for computing resources and memory transfers. In particular, media processing of high-definition and ultra-high-definition video raises serious concerns relating to energy efficiency.

This document is about the context and objectives, use cases and requirements for making MPEG technologies aware of additional information (i.e., Green MPEG metadata) that may be used to reduce energy consumption.

#

# Objectives

The main objectives of Green MPEG are:

* Efficient multimedia processing and transport: The standard will enable building interoperable solutions for energy-efficient consumption of media in this context.
* Enable efficient media encoding based on the energy resources of the encoder or the feedback from the receiver.
* Enable efficient media decoding and presentation.

# Scope

# Green MPEG will standardize metadata formats that will impact the following technical areas described in Figure 1:

* Media Encoding
* Media Decoding
* Media Presentation



Figure 1: Green MPEG Technical Areas

The transmitter will set the encoder to generate Green MPEG metadata and multiplex them in the media stream. The receiver will use the Green MPEG metadata to set the decoder and presentation subsystem in the appropriate operating mode. If a feedback channel is available, a receiver set to receive data from a transmitter will transmit appropriate metadata describing its status.



Figure 2: Example of Green MPEG technologies

Figure 2 shows an example of Green MPEG technologies. Typically, a content pre-processor is applied to analyze the content source and a video encoder is used to encode the content to a bitstream for delivery. The bitstream is delivered to the receiver and decoded by a video decoder with the output rendered on a presentation subsystem (such as a display).

“Green metadata” is extracted from either the video encoder or the content pre-processor. In the former case, the green metadata is encapsulated in a conformant bitstream from a conventional video encoder. In the latter case, the green metadata is multiplexed with the conformant bitstream. Such green metadata is used at the receiver to reduce the power consumption for video decoding and presentation. The bitstream will be packetized and delivered to the receiver for decoding and presentation. At the receiver, the bitstream parser processes the packets and sends the green metadata to a “power optimization module” for efficient “power control”. For instance, the power optimization module decodes the green metadata and then applies appropriate operations to reduce the video decoder’s power consumption when decoding the video and also to reduce the presentation subsystem’s power consumption when rendering the video. In addition, the power-optimization module could collect receiver information, such as remaining battery capacity, and send it to the transmitter through a feedback channel (i.e., “green feedback”) to adapt the encoder operations for power-consumption reduction.

* Green Metadata Extractor
	+ Extracts metadata from either the video encoder or the content pre-processor where metadata is encapsulated with a conventional conformant bitstream for delivery
* Parser
	+ Interprets the bitstream syntax information and sends it to the power optimization module.
* Power optimization module
	+ Processes the green-metadata information and applies appropriate operations for power-consumption control.
	+ Collects platform statistics (such as remaining battery capacity) and communicates with the transmitter or the receiver through a feedback channel, if available, for energy-efficient processing.
	+ Collects user preference and communicates with the transmitter or the receiver through a feedback channel, if available, for energy-efficient processing.

# Green MPEG Requirements

# General Requirements

1. Green MPEG shall provide signaling means to facilitate appropriate power consumption from the encoding, decoding and/or presentation, without loss of the QoE.
2. Green MPEG shall offer the means to choose between QoE of the presentation and energy consumption.
3. Green MPEG shall offer the means to support the communication from a receiver to a transmitter to communicate the receiver’s needs.
4. Green MPEG induced additional delay shall be minimal so as to have negligible impact on overall system performance.
5. Green MPEG shall provide compact and manageable signaling means with minimal overhead so as to have negligible impact on overall system performance.
6. Delivery of Green MPEG metadata in the chain from the source to the destination shall be guaranteed (e.g., in-band signaling, media multiplexing, re-multiplexing, statistical multiplexing, and transcoding).

# Green MPEG Encoding

1. Green MPEG may include informative guidelines to help users achieve a level of energy-efficient media encoding.
2. Green MPEG shall specify signaling to be used by a Green MPEG decoder to request an encoder to constrain encoding options (such as spatial and temporal resolution adaptation, B frames, etc.) appropriately for bandwidth/energy reduction.
3. Green MPEG encoder should be able to receive the needs from the decoder to adapt its encoding for interactive applications.
4. Green MPEG encoder shall provide means to signal Green MPEG metadata.

# Green MPEG Decoding

1. Green MPEG shall provide means for energy-efficient video decoding without affecting the compliance to the existing standards.
2. Green MPEG metadata should be agnostic to different decoder implementations.
3. The delay introduced in the decoding by the processing of Green MPEG metadata shall be minimal.

# Green MPEG Presentation

1. Green MPEG shall provide means for energy efficient presentation adaptation.
2. Green MPEG shall provide means for energy-efficient presentation by widely used display technologies such as LCD, LED and OLED for both battery-powered devices and mains-powered devices.
3. Green MPEG shall provide means for energy-efficient presentation adaptation for media other than Video.
4. Green MPEG shall enable presentation engines to provide multi-level QoE using Green MPEG metadata
5. The delay introduced in the presentation by the processing of Green MPEG metadata shall be minimal.

# Green MPEG Use Cases

# Media Decoding : Codec Dynamic Voltage/Frequency Scaling

Advanced processors and hardware chipsets provide a capability where a processor can be put into a different voltage state so as to save power/energy consumption. Usually, a lower voltage supply implies lower energy consumption while a higher voltage supply requires higher energy consumption. Also note that the frequency at which a processor operates is directly proportional to its voltage supply.

In a video stream, because of the content variation, different frames typically have quite different decoding complexities. Therefore, we can apply lower frequency (hence lower voltage) to decode lower-complexity frames and apply higher frequency (hence higher voltage) for complex frames. It is beneficial to embed the frame-complexity metadata into the bitstream so that it can guide the underlying processor voltage adaptation and thus save energy.

Please note that, depending on the application requirements, frame-level complexity metadata embedding can be extended to GOP level, time-interval level, or video-scene level complexity metadata embedding.

# Media Presentation : Display Adaptation

Large screen displays are now widely used in mobile devices such as smart phones and tablets. However, powering such large displays typically requires a significant amount of energy, especially for multimedia applications such as video presentation.

There are two major displays for mobile devices: LCDs and OLEDs. For these displays, the power consumption depends on the LCD backlight level and OLED supply voltage respectively.

Unlike conventional display subsystems, advanced displays can adapt their backlight or supply voltage according to the content statistics, and thus reduce energy consumption. This adaptation can be applied at different scales depending on the underlying applications. Moreover, if there is sufficient battery power, the normal display adaptation described previously can be applied. However in an extremely low battery state, to avoid depleting the battery before the content is consumed, aggressive display adaptation can be applied while maintaining acceptable video quality.

.

# Media Encoding: video encoding on mobile devices for interactive applications

Mobile video conferencing on mobile devices is an application for which low-power consumption is critical. Depending on the receiver battery status, the encoder can be adapted with minimal QoE degradation to extend the battery life.

One example in this category is mobile based video conferencing, such as Skype, FaceTime, Tango and etc. It is demonstrated that mobile video conferencing is a killer application for battery. Hence, it is highly desired to have a solution where energy consumption during video conferencing can be reduced. For instance, based on its battery status, a user can choose to encode smaller resolution video with low frame rate for conferencing. It can also provide user options where an energy-efficient mode can be triggered to adapt the encoding parameters for energy saving purpose.

Moreover, in a point-to-point video conference application, as each device is both an encoder and a decoder, the battery level of the device can be embedded as side information in the bitstream sent by the encoder, using the same metadata message as defined in Use-Case 1 or a separate message which can be sent at a much lower frequency than this one.

In this way, each encoder can adapt the complexity of the encoded stream as a function of the battery level of the other device communicating with it: this complexity reduction can be done through the reduction of the picture resolution or frame rate or through the filtering of pictures before coding or by limiting those encoding tools that increase decoding complexity. If the two devices are battery equipped, then the best encoding strategy can be defined by considering the battery-level evolution of the two devices.

Another interesting possibility is for each device to provide an indication of its battery-saving priority. Indeed, the nature of the energy-saving strategy, whether gradual or more aggressive, is of course dependent on the evolution of battery levels but this strategy must also be linked to the user-defined, battery-saving priority.

# Media Encoding : Dynamic Video Adaptation Considering Client Energy Consumption

State-of-the-art dynamic video adaptation mostly considers the client bandwidth. For instance, state-of-the-art video streaming over HTTP technology enables dynamic adaptive delivery of multimedia contents, where the server provides a list of available versions of multimedia contents to the clients so that the client selects one of the versions. Generally, each version provided by the server has a different configuration of encoding parameters resulting in different bitrates such as compression ratio for video or audio, spatial resolution of multimedia, temporal resolution of multimedia and so on. The list of versions also includes various audio and text with different languages. The clients select one or more of the available options according to the device capability and the network condition. During the delivery, a client can also seamlessly change to another version if it detects the change of network conditions.

Nowadays, battery life is a critical issue for battery-powered mobile devices, such as Smart Phones and tablets. More advanced solutions should consider the client battery status in addition to the network bandwidth when performing the content adaptation. Using MPEG DASH standard as an example, an extension can be amended to include additional power consumption information into the MPD (Media Presentation Description), where a client can switch to different streams with lower processing energy requirement if its battery is about to run out. With such an extension, DASH can adapt the streaming not only based on its bandwidth, but also based on the client battery status.