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

Document type:	**Press Release**
Title:	**Press Release of 140th MPEG Meeting**
Status:	**Approved**
Date of document:	**2022-10-28**
Source:	**Convenor**
Expected action:	**INFO**
No. of pages:	**5 (including this 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 N88

Mainz, Germany – October 2022

MPEG evaluates the Call for Proposals on Video Coding for Machines

Mainz, Germany – The 140th MPEG meeting was held in Mainz, Germany, 24–28 October 2022

MPEG evaluates the Call for Proposals on Video Coding for Machines

At the 140th MPEG meeting, *MPEG Technical Requirements (WG 2)* evaluated the responses to the Call for Proposals (CfP) for technologies and solutions enabling efficient video coding for machine vision tasks. A total of 17 responses to this CfP were received, with responses providing various technologies such as *(i)* learning-based video codecs, *(ii)* block-based video codecs, *(iii)* hybrid solutions combining (i) and (ii), and *(iv)* novel video coding architectures. Several proposals use a region of interest-based approach, where different areas of the frames are coded in varying qualities.

The responses to the CfP reported an improvement in compression efficiency of up to 57% on object tracking, up to 45% on instance segmentation, and up to 39% on object detection, respectively, in terms of bit rate reduction for equivalent task performance. Notably, all requirements defined by WG 2 were addressed by a variety of proposals.

Given the success of this call, MPEG will continue working on video compression methods for machine vision tasks. The work will continue in MPEG Video Coding (WG 4) within a new standardization project. A test model will be developed based on technologies from the responses to the CfP and results from the first round of core experiments in one or two meeting cycles. At the same time, the Joint Video Team with ITU-T SG 16 (WG 5) will study encoder optimization methods for machine vision tasks on top of existing MPEG video compression standards.

WG 2 thanks all proponents who submitted responses to this CfP. MPEG will continue to collect and solicit feedback to improve the test model for video coding for machines in the upcoming meetings.

MPEG evaluates Call for Evidence on Video Coding for Machines Feature Coding

At the 140th MPEG meeting, *MPEG Technical Requirements (WG 2)* evaluated the responses to the Call for Evidence (CfE) for technologies and solutions enabling efficient feature coding for machine vision tasks. A total of eight responses to this CfE were received, whereof six responses were considered valid based on the conditions described in the call:

- For the tested video dataset, increases in compression efficiency of up to 87% compared to the video anchor and over 90% compared to the feature anchor were reported.
- For the tested image dataset, the compression efficiency can be increased by over 90% compared to both image and feature anchors.

Based on the successful outcome of the CfE, WG 2 will continue working toward issuing a Call for Proposals (CfP). WG 2 thanks all proponents who submitted responses to this CfE.

MPEG reaches the First Milestone for Haptics Coding

At the 140th MPEG meeting, *MPEG Coding of 3D Graphics and Haptics (WG 7)* reached the first milestone in the approval process for the Haptics Coding (ISO/IEC CD 23090-31) standard by promoting the text to Committee Draft (CD) status. The CD comprises the MPEG-I Haptics Phase 1 codec specification which includes a JSON descriptive format based on a parametric representation of haptics and a perceptually optimized wavelet compression format addressing temporal haptic signals. These formats allow the MPEG-I Haptics Phase 1 codec to be used for the creation, editing, and interchange of haptics as well as for the efficient encoding, distribution, streaming, and storage of haptics. The JSON format is compatible with the current glTF specification allowing for future extensions of spatial and interactive haptics. The technologies selected for the CD include descriptive, human-readable representations and highly efficient psychophysical compression schemes, as well as support for both vibrotactile and kinesthetic devices. They incorporate a number of refinements and enhancements to the initial set of technologies retained after the call for proposals (termed RM0) that have passed rigorous objective and subjective perceptual tests designed to assess the quality of haptics at various bitrates.

**MPEG completes a New Standard for
Video Decoding Interface for Immersive Media**

One of the most distinctive features of immersive media compared to 2D media is that only a tiny portion of the content is presented to the user. Such a portion is interactively selected at the time of consumption. For example, a user may not see the same point cloud object's front and back sides simultaneously. Thus, for efficiency reasons and depending on the users' viewpoint, only the front or back sides need to be delivered, decoded, and presented. Similarly, parts of the scene behind the observer may not need to be accessed.

At the 140th MPEG meeting, *MPEG Systems (WG 3)* reached the final milestone of the Video Decoding Interface for Immersive Media (VDI) standard (ISO/IEC 23090-13) by promoting the text to Final Draft International Standard (FDIS). The standard defines the basic framework and specific implementation of this framework for various video coding standards, including support for application programming interface (API) standards that are widely used in practice, e.g., Vulkan by Khronos.

The VDI standard allows for dynamic adaptation of video bitstreams to provide the decoded output pictures in such a way that the number of actual video decoders can be smaller than the number of the elementary video streams to be decoded. In other cases, virtual instances of video decoders can be associated with the portions of elementary streams required to be decoded. With this standard, the resource requirements of a platform running multiple virtual video decoder instances can be further optimized by considering the specific decoded video regions to be presented to the users rather than considering only the number of video elementary streams in use. The first edition of the VDI standard

includes support for the following video coding standards: High Efficiency Video Coding (HEVC), Versatile Video Coding (VVC), and Essential Video Coding (EVC).

MPEG completes Development of
Conformance and Reference Software for Compression of Neural Networks

At the 140th MPEG meeting, MPEG *Video Coding (WG 4)* reached the final milestone for Conformance and Reference Software for Compression of Neural Networks (ISO/IEC 15938-18) by promoting the text to Final Draft International Standard (FDIS). It complements the recently published first edition of the standard for Compression of Neural Networks for Multimedia Content Description and Analysis (ISO/IEC 15938-17).

The neural network coding standard is designed as a toolbox of coding technologies. The specification contains different methods for three compression steps, *i.e.,* parameter reduction (*e.g.,* pruning, sparsification, and matrix decomposition), parameter transformation (*e.g.,* quantization), and entropy coding methods, that can be assembled into encoding pipelines combining one or more (in the case of reduction) methods from each step. The reference software is written in Python and provides a framework defining interfaces for these three steps in the coding pipeline and components implementing all supported methods. Additionally, bitstreams for testing the conformance to the neural network coding standard are provided.

MPEG White Papers

At the 140th MPEG meeting, *MPEG Liaison and Communication (AG 3)* approved the following two MPEG white papers.

MPEG-H 3D Audio

The MPEG-H 3D Audio standard specifies a universal audio coding and rendering environment that is designed to efficiently represent high-quality spatial or immersive audio content for storage and transmission. Since there is no generally accepted "one-size-fits-all" format for immersive audio, it supports *(i)* common loudspeaker setups including mono, stereo, surround, and 3D audio (*i.e.,* setups including loudspeakers above ear level and possibly below ear level) and *(ii)* rendering over a wide range of reproduction conditions (*i.e.,* various loudspeaker setups or headphones, possibly with background noise in the listening environment).

MPEG-I Scene Description

MPEG has been working on technologies and standards for immersive media under the umbrella of the MPEG immersive media coding project (MPEG-I). MPEG Systems (WG 3) recognized the need for an interoperable and distributable scene description solution as a key element to foster the emergence of immersive media services and to enable the delivery of its immersive content in the consumer market. As part of the MPEG-I project, WG 3 started investigating architectures for immersive media and possible solutions for a scene description format in 2017, which resulted in the ISO/IEC 23090-14 standard.

This white paper introduces ISO/IEC 23090-14, which provides a set of extensions under the "MPEG" prefix to Khronos glTF[1] (also available as ISO/IEC 12113), as well as extensions to the MPEG-defined ISO Base Media file format, also known as ISO/IEC 14496-12 ISOBMFF. These extensions enable the description and delivery of timed immersive media into glTF-based immersive scenes. Furthermore, the standard defines an architecture together with an application programming interface (API) that allows

---

[1] Khronos Group, glTF : https://www.khronos.org/gltf/

the application to separate the access to the immersive timed media content from the rendering of this media. The white paper concludes with an outlook and future plans for the standard.

How to contact MPEG and further information

Everyone who wish to receive MPEG Press Releases by email should contact Dr. Christian Timmerer at christian.timmerer@aau.at or subscribe at https://lists.aau.at/mailman/listinfo/mpeg-pr. Further information can be found on the MPEG Website: http://www.mpeg.org/.

Future MPEG meetings are planned as follows:
> No. 141, Online, 16 – 20 January 2023
> No. 142, Anatalya, TR, 24 – 28 April 2023
> No. 143, Geneva, CH, 17 – 21 July 2023

For further information about MPEG, please contact:
> Prof. Dr.-Ing. Jörn Ostermann (Convenor of MPEG Technical Coordination, Germany)
> Leibniz Universität Hannover
> Appelstr. 9A
> 30167 Hannover, Germany
> Tel: +49 511 762 5316
> Fax: ++49 511 762 5333
> Email: ostermann@tnt.uni-hannover.de

or
> Prof. Dr. Kyuheon Kim (Convenor of MPEG Liaison and Communication, Korea)
> Department of Electronic Engineering
> Kyung Hee University
> Seoul, South Korea
> Tel: +82 31 201 3810
> Email: kyuheonkim@khu.ac.kr

or
> Priv.-Doz. Dr. Christian Timmerer (MPEG Press Officer, Austria)
> Alpen-Adria-Universität Klagenfurt | Bitmovin Inc.
> 9020 Klagenfurt am Wörthersee, Austria, Europe
> Tel: +43 463 2700 3621
> Email: christian.timmerer@aau.at