



ISO/IEC JTC 1/SC 29/AG 03

N00027

ISO/IEC JTC 1/SC 29/AG 03

MPEG Liaison and Communication

Convenorship: KATS (Korea, Republic of)

Document type: Press Release

Title: Press Release of 134th MPEG Meeting

Status: Approved

Date of document: 2021-04-30

Source: Convenor

Expected action: INFO

No. of pages: 9 (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 03
 MPEG LIAISON AND COMMUNICATION

ISO/IEC JTC 1/SC 29/AG 03 **N00027**

Online Meeting – April 2021

Source: Convenor of ISO/IEC JTC 1/SC 29/AG 03
 Status: Approved by AG 03
 Subject: Press Release of 134th MPEG Meeting
 Date: 30 April 2021
 Serial Number: 20395



**MPEG Video ratifies the First International Standard on
 Neural Network Compression for Multimedia Applications**

The 134th MPEG meeting was held online, 26 – 30 April 2021

Table of Contents:

<i>MPEG Video ratifies the First International Standard on Neural Network Compression for Multimedia Applications.....</i>	<i>1</i>
<i>MPEG Systems completes carriage of VVC and EVC</i>	<i>1</i>
<i>MPEG Systems completes carriage of V3C in ISO/BMFF.....</i>	<i>1</i>
<i>MPEG announces a Call for Proposals on New Advanced Genomics Features and Technologies.....</i>	<i>2</i>
<i>MPEG announces Call for Proposals for MPEG-I Immersive Audio</i>	<i>2</i>
<i>MPEG issues Call for Proposals on the Coded Representation of Haptics.....</i>	<i>3</i>
<i>MPEG evaluated Responses on Incremental Compression of Neural Networks.....</i>	<i>4</i>
<i>MPEG evaluated Responses to the Call for Evidence for Video Coding for Machines</i>	<i>4</i>
<i>Progression of MPEG 3D Audio Standards.....</i>	<i>4</i>
<i>MPEG Systems reaches the first milestone of the amendment to Open Font Format</i>	<i>5</i>
<i>MPEG Video completes Low Complexity Enhancement Video Coding (LCEVC) Verification Test.....</i>	<i>5</i>
<i>Verification tests for more application cases of Versatile Video Coding (VVC)</i>	<i>6</i>
<i>Standardization work on Version 2 of VVC and VSEI started</i>	<i>6</i>
<i>How to contact MPEG and further information.....</i>	<i>7</i>

MPEG Video ratifies the First International Standard on Neural Network Compression for Multimedia Applications

Artificial neural networks have been adopted for a broad range of tasks in multimedia analysis and processing, such as visual and acoustic classification, speech recognition, and extraction of multimedia descriptors, and their potential is also being explored for other uses such as image and video coding. The trained neural networks for these applications contain many parameters (*i.e.*, weights), resulting in a considerable amount of data necessary to represent the neural network. Thus, transferring a neural network representation to several clients (*e.g.*, mobile phones, smart cameras) can benefit from using a compressed representation of the neural network.

At the 134th MPEG meeting, MPEG Video completed the first international standards on *Neural Network Compression for Multimedia Applications (ISO/IEC 15938-17)*, designed as a toolbox of compression technologies. The specification contains different (*i*) parameter reduction (*e.g.*, pruning, sparsification, matrix decomposition), (*ii*) parameter transformation (*e.g.*, quantization), and (*iii*) entropy coding methods that can be assembled to encoding pipelines combining one or more (in the case of reduction) methods from each group. The results show that trained neural networks for many multimedia problems such as image or audio classification or image compression can be compressed by a factor of 10-20 with no performance loss and even by more than a factor of 30 with some performance trade-off. The new standard is not limited to a particular neural network architecture and is independent of choice of the neural network exchange format. Interoperability with common neural network exchange formats is described in the annexes of the standard.

MPEG Systems completes carriage of VVC and EVC

At the 134th MPEG meeting, MPEG Systems completed two standards for carriage of the recently developed video coding standards Versatile Video Coding (VVC) and Essential Video Coding (EVC).

The *8th edition of ISO/IEC 13818-1 (MPEG-2 Systems)* was promoted to Final Draft International Standard (FDIS) status including support for the carriage of VVC and EVC in MPEG-2 Transport Streams (TS). The standard defines constraints on elementary streams of VVC and EVC to enable them to be carried in packetized elementary stream (PES) packets. Buffer management mechanisms and transport system target decoder (T-STD) model extensions are also defined to deliver the video bitstreams under the constraints of MPEG-2 Systems.

Amendment 2 of ISO/IEC 14496-15 (Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format) has been completed with the support for the carriage of VVC and EVC promoted to Final Draft Amendment (FDAM) status. The carriage of codec initialization information in the ISOBMFF for VVC and EVC are defined in this amendment. The standard also defines samples and sub-samples reflecting the high-level bitstream structure and independently decodable units for both video codecs. For VVC, signalling and extraction of a specific operating point with a combination of video layers is also supported.

MPEG Systems completes carriage of V3C in ISOBMFF

At the 134th MPEG meeting, MPEG Systems completed the standard for the carriage of Visual Volumetric Video-based Coding (V3C) data using the ISO Base Media File Format (ISOBMFF) by promoting it to Final Draft International Standard (FDIS), the last milestone of the ISO/IEC standards development process.

The standard introduces support for media comprising multiple independent component bitstreams. Considering that only some portions of immersive media assets need to be rendered according to the users' position and viewport, the metadata indicating the relationship between the region in the 3D space data to be rendered and its location in the bitstream are defined. In addition, the delivery of the ISOBMFF file containing a V3C content over DASH and MMT is also specified in this standard.

MPEG announces a Call for Proposals on New Advanced Genomics Features and Technologies

The extensive usage of high-throughput DNA sequencing technologies enables a new approach to healthcare known as “precision medicine”. DNA sequencing technologies produce extremely large amounts of raw data which are stored in different repositories worldwide. One challenge is to efficiently handle the increasing flood of sequencing data. Another challenge is the ability to process such a deluge of data to 1) increase the scientific knowledge of genome sequence information and 2) search genome databases for diagnosis and therapy purposes. High-performance compression of genomic data is required to reduce the storage size and increase transmission speed of large data sets.

The current MPEG-G standard series (ISO/IEC 23092) addresses the representation, compression, and transport of genome sequencing data with support for annotation data under development. They provide a file and transport format, compression technology, metadata specifications, protection support, and standard APIs for the access of genomic data in the native compressed format.

In line with the traditional MPEG practice of continuous improvement of the quality and performance of its standards, after having assessed the evidence of the availability of new technology at its 134th meeting, MPEG has issued a Call for Proposals (CfP). This CfP aims to collect submissions of new technologies that can (i) provide improvements to the current compression, transport and indexing capabilities of the ISO/IEC 23092 standards suite, particularly applied to data consisting of very long reads generated by 3rd generation sequencing devices, (ii) provide the support for representation and usage of graph genome references, (iii) include coding modes relying on machine learning processes, satisfying data access modalities required by machine learning and providing higher compression, and (iv) support of interfaces with existing standards for the interchange of clinical data.

Companies and organizations are invited to submit proposals in response to this call. Parties do not need to be MPEG members to respond. Responses should be submitted by January 10th, 2022 and will be evaluated during the 137th MPEG meeting in January. Detailed information, including instructions on how to respond to the call for proposals, the requirements that must be considered, the test data to be used, and the submission and evaluation procedures for proponents are available at <http://www.mpeg.org/>, submenu Emerging Technologies. For further information about the call, please contact Dr. Igor Curcio, WG 02 MPEG Technical Requirements Convenor, at igor.curcio@nokia.com and Dr. Marco Mattavelli, WG 08 MPEG Genomic Coding Convenor, at marco.mattavelli@epfl.ch.

MPEG announces Call for Proposals for MPEG-I Immersive Audio

The MPEG-I Immersive Audio Call for Proposals (CfP) is for technology to be standardized in MPEG-I Part 4, “Immersive Audio.” Along with other parts of MPEG-I (*i.e.*, Part 3, “Immersive Video” and Part 2, “Systems Support”), the suite of standards will support a Virtual Reality (VR) or an Augmented Reality (AR) presentation in which the user can navigate and interact with the environment using 6 degrees of freedom (6 DoF), that being spatial navigation (x, y, z) and user head orientation (yaw, pitch, roll).

The goal in MPEG-I presentations is to impart the feeling that the user is actually present in the virtual world. Audio in the virtual world (or scene) is perceived as in the real world, with sounds coming from an associated visual figure. That is, perceived with the correct location and distance. Physical movement of the user in the real world is perceived as having matching movement in the virtual world. Furthermore, and importantly, the user can interact with the virtual scene and cause sounds that are perceived as realistic and matching the users' experience in the real world.

The audio compression engine used in MPEG-I Immersive Audio is MPEG-H 3D Audio (ISO/IEC 23008-3), specifically the Low Complexity (LC) Profile. The new technology that is to be standardized in MPEG-I Immersive Audio is as follows:

- Technology for rendering the audio presentation while permitting the user to have 6 DoF movement.

- Metadata to support this rendering.
- A bitstream syntax that enables efficient storage and streaming of the MPEG-I Immersive Audio content.

The evaluation of proponent technology submitted in response to the CfP will be done using a real-time audio-visual presentation engine. This requires proponents to submit audio renderers that will run in real-time and permits evaluation subjects to experience a compelling virtual or augmented reality audio-visual presentation that is key to experiencing truly immersive audio content.

Companies and organizations are invited to submit proposals in response to this call. Parties do not need to be MPEG members to respond. Respondents need to register their intent to participate by September 27th, 2021. Responses should be submitted by November 10th, 2021 and will be evaluated during the 137th MPEG meeting in January 2022. Detailed information, including instructions on how to respond to the call for proposals, the requirements that must be considered, the test data to be used, and the submission and evaluation procedures for proponents are available at <http://www.mpeg.org/>, submenu Emerging Technologies. For further information about the call, please contact Dr. Igor Curcio, WG 02 MPEG Technical Requirements Convenor, at igor.curcio@nokia.com and Dr. Schuyler Quackenbush, WG 06 MPEG Audio Coding Convenor, at srq@audioresearchlabs.com.

MPEG issues Call for Proposals on the Coded Representation of Haptics

Haptics provide an additional layer of entertainment and sensory immersion beyond audio and visual media. The user experience and enjoyment of media content can be significantly enhanced by the addition of haptics to audio/video content, be it in ISO Base Media File Format (ISOBMFF) files or media streams such as ATSC 3.0 broadcasts, streaming games, and mobile advertisements. To that end, haptics has been proposed as a first-order media type, co-equal with audio and video, in ISOBMFF. Furthermore, haptics has been proposed as an addition to the MPEG-DASH standard, which would make DASH streaming clients aware of the presence of haptics within MP4 segments. Lastly, the MPEG-I use cases have been extended to include haptics, resulting in a set of haptic-specific requirements for MPEG-I. All these proposals are in various stages of the MPEG standardization process.

These ongoing efforts to standardize haptics in media highlight the need for a standard coded representation of haptics. A standard haptics coding format, and associated standardized decoder, will facilitate the incorporation of haptics into the ISOBMFF, MPEG-DASH, and MPEG-I standards, making it easier for content creators and streaming content providers to incorporate haptics and benefit from its effect on the user experience and for consumer device makers to support haptic playback.

At the 134th MPEG meeting, MPEG Requirements issued a Call for Proposals (CfP) on the Coded Representation of Haptics. This CfP seeks submissions of technologies that can provide efficient representation and compression of time-dependent haptic signals and are suitable for coding of timed haptic tracks that may be synchronized with audio and/or video media.

Companies and organizations are invited to submit proposals in response to this call. Parties do not need to be MPEG members to respond. Respondents need to register their intent to participate by May 22nd, 2021. Responses should be submitted by July 5th, 2021 and will be evaluated during the 136th MPEG meeting in October 2021. Detailed information, including instructions on how to respond to the call for proposals, the requirements that must be considered, the test data to be used, and the submission and evaluation procedures for proponents are available at <http://www.mpeg.org/>, submenu Emerging Technologies. For further information about the call, please contact Dr. Igor Curcio, WG 02 MPEG Technical Requirements Convenor, at igor.curcio@nokia.com.

MPEG evaluated Responses on Incremental Compression of Neural Networks

An increasing number of applications of artificial networks for multimedia analysis and processing (*e.g.*, visual and acoustic classification, extraction of multimedia descriptors or image and video coding) utilize edge-based content processing or federated training. The trained neural networks for these applications contain many parameters (*i.e.*, weights), resulting in a considerable size. The MPEG standard for compressed representation of neural networks for multimedia content and analysis, currently being finalized, addresses these requirements, and provides technologies for parameter reduction and quantization to compress entire neural networks.

In scenarios like edge-based content processing or federated training, updates of neural networks (*e.g.*, after training on additional data) need to be exchanged. Such updates include changes of the network parameters but may also involve structural changes of the network (*e.g.*, when extending a classification method with a new class). In scenarios like federated training, these updates are more frequent than initial deployments of trained networks and, thus, require much more bandwidth over time.

After issuing a Call for Proposals (CfP) in fall 2020, MPEG has evaluated the responses to the CfP at its 134th meeting. Two responses describing compression toolchains have been received. The results show that updates relative to a base model can be compressed very efficiently. In a distributed training scenario, a model update after a training iteration can be represented at 1% or less of the base model size on average, without sacrificing the classification performance of the neural networks.

MPEG evaluated Responses to the Call for Evidence for Video Coding for Machines

MPEG is studying the standardization of coding technologies of unprocessed or processed video for machine intelligence use cases. Those technologies are expected to be with a compression capability that exceeds that of state-of-the-art video coding standards to fulfill a single or multiple machine intelligent task, or optionally to support hybrid machine/human vision at sufficient quality.

At the 133rd MPEG meeting, MPEG Requirements issued a Call for Evidence (CfE) for Video Coding for Machines (VCM) addressing two specific technologies:

- Efficient compression of processed or unprocessed videos.
- The shared backbone of feature extraction.

At the 134th MPEG meeting, MPEG Requirements evaluated five responses to the CfE from which two were accepted per the MPEG VCM evaluation framework. Based on the evaluation of these responses, it has been demonstrated that for object detection tasks (*i*) an end-to-end compression network with a BD-rate (mAP) of 22.80% saving can be achieved compared to the OpenImages-v6 anchors and (*ii*) a BD-rate (Pareto mAP) with a 30.76% saving can be attained compared to the FLIR anchors.

MPEG will continue accepting and evaluating additional evidence through a supplemental CfE with a scope beyond the object detection tasks at the 135th MPEG meeting.

Progression of MPEG 3D Audio Standards

At its 134th meeting, MPEG promoted the 3rd Edition of the ISO/IEC 23008-3 MPEG-H 3D Audio standard to Committee Draft (CD). This 3rd Edition incorporates two amendments to the 2nd Edition which introduced several clarifications and corrections as well as metadata enhancements and the 3D audio baseline profile. As MPEG-H 3D Audio is one of the most advanced audio systems for broadcast and streaming applications, it has been revised with many improvements after its first publication. Now the 3rd Edition contains the full specification of the widely adopted low complexity and baseline profiles and offers essential support to implementers around the world. This edition of MPEG-H 3D Audio will contain additional improvements and will reach its final milestone by mid of 2022.

Additionally, MPEG has finalized its work on revising the references software and conformance testing for MPEG-H 3D Audio and is pleased to announce the promotion of ISO/IEC 23008-6, MPEG-H 3D

Audio Reference Software 3rd Edition and ISO/IEC 23008-9 MPEG-H 3D Audio Conformance Testing 2nd Edition to Final Draft International Standard (FDIS) status. With its 3rd Edition, the MPEG-H 3D Audio Reference Software incorporates fixes for all identified bugs reported in the past years and is now better aligned to the standard text. The 2nd Edition of the MPEG-H 3D Audio Conformance Testing incorporates additional conformance bitstreams for the baseline profile and offers a better coverage of the MPEG-H 3D Audio features.

MPEG Systems reaches the first milestone of the amendment to Open Font Format

At the 134th MPEG meeting, MPEG Systems reached the first milestone of the development of a standard extending the Open Font Format by promoting ISO/IEC 14496-22 Amendment 2 to Committee Draft Amendment (CDAM) status.

This amendment extends the colour font capabilities to implement feature-rich variable colour fonts with multiple layers supporting complex design using advanced graphical primitives. It adds support for multi-coloured glyphs in a manner that integrates with the rasterizers of existing text engines and that is designed to be easy to support with current OpenType font files. One version of it allows for a simple composition of coloured elements and another version supports additional graphic capabilities such as gradient fills, affine transformations, and various blending modes.

This standard is expected to reach its final milestone, Final Draft Amendment (FDAM), in early 2022.

MPEG Video completes Low Complexity Enhancement Video Coding (LCEVC) Verification Test

At the 134th MPEG meeting, MPEG Video completed the verification test of the Low Complexity Enhancement Video Coding standard (ISO/IEC 23094-2) by measuring the benefits of enhancing four existing codecs of different generations (*i.e.*, AVC, HEVC, EVC, VVC) and validating the project's requirements.

A first set of tests compared LCEVC-enhanced encoding with full-resolution single-layer anchors. The average bit rate savings produced by LCEVC when enhancing AVC were determined to be approximately 46% for UHD and 28% for HD. When enhancing HEVC approximately 31% for UHD and 24% for HD. Test results tend to indicate an overall benefit also when using LCEVC to enhance EVC and VVC.

A second set of tests confirmed that LCEVC provided a more efficient means of resolution enhancement of half resolution anchors than unguided up-sampling. Comparing LCEVC full-resolution encoding with the up-sampled half-resolution anchors, the average bit-rate savings when using LCEVC with AVC, HEVC, EVC and VVC were calculated to be approximately 28%, 34%, 38%, and 33% for UHD and 27%, 26%, 21%, and 21% for HD, respectively.

LCEVC adds an enhancement data stream that can appreciably improve the resolution and visual quality of reconstructed video with effective compression efficiency of limited encoding/decoding complexity by building on top of existing and future video codecs. It is designed to be compatible with existing video workflows (*e.g.*, CDNs, metadata management, DRM/CA) and streaming/media formats (*e.g.*, HLS, DASH, CMAF) to facilitate the rapid deployment of enhanced video services. LCEVC can be used to deliver higher video quality in limited bandwidth scenarios, especially when the available bit rate is low for high-resolution video delivery and encoding or decoding complexity is a challenge.

Verification tests for more application cases of Versatile Video Coding (VVC)

Despite the difficulties caused by the pandemic situation, the second round of verification testing for VVC (ISO/IEC 23090-3 and Rec. ITU-T H.266) has been completed. This includes

- 360° video for equirectangular and cubemap formats, where VVC shows on average more than 50% bit rate reduction compared to the previous major generation of MPEG video coding standard known as High Efficiency Video Coding (HEVC), developed in 2013.
- Low-delay applications such as compression of conversational (teleconferencing) and gaming content, where the compression benefit is about 40% on average,
- HD video streaming, with an average bit rate reduction of close to 50%.

A previous set of tests for 4K UHD content completed in October 2020 had showed similar gains. These verification tests used formal subjective visual quality assessment testing with “naïve” human viewers. The tests were performed under a strict hygienic regime in two test laboratories to ensure safe conditions for the viewers and test managers.

In addition to demonstrating the effectiveness of the VVC standard for high compression capability, the test confirmed the feasibility of practical implementation with subjective perceptual optimization, as the test included a VVC encoder with much higher running speed than the reference software example implementation that has been developed along with the VVC standard as an experiment platform and demonstration of the VVC syntax capability. The subjectively optimized encoder showed compression performance that was subjectively generally comparable or better than the current version of the reference software.

Further subjective testing of VVC capability is also planned for additional use cases including the coding of high dynamic range (HDR) video content.

Standardization work on Version 2 of VVC and VSEI started

Work on the second version of both Versatile Video Coding (VVC) and Versatile SEI (VSEI) standards has entered the formal standardization phase by issuing CDAMs on both.

Profiles in version 2 of VVC will support video with bit depth larger than 10 bits, and new tools have been added to further increase compression performance at extremely high bit rates.

Version 2 of VSEI adds more SEI messages supporting scalable and multiview, depth data and alpha channels, extending the dependent random access concept and mapping functions for HDR, and carries over some more SEI messages from HEVC to be used in the VVC context.

While the new SEI messages have initially been developed for use in the VVC context, it is also planned that the use of at least two of these SEI messages will be supported for the HEVC standard as well, taking advantage of the generality of the VSEI specification (ISO/IEC 23002-7 and Rec. ITU-T H.274).

How to contact MPEG and further information

Journalists that wish to receive MPEG Press Releases by email should contact Dr. Christian Timmerer at christian.timmerer@aau.at or subscribe via <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. 135, Online, 12 – 16 July 2021
- No. 136, Antalya, TR, 11 – 15 October 2021
- No. 137, Geneva, CH, 17– 21 January 2022
- No. 138, TBD, April 2022

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)
Alpen-Adria-Universität Klagenfurt | Bitmovin Inc.
9020 Klagenfurt am Wörthersee, Austria, Europe
Tel: +43 463 2700 3621
Email: christian.timmerer@aau.at