ISO/IEC JTC 1 "Information technology"
Secretariat: **ANSI**
Committee manager: **Rajchel Lisa Mrs.**

JTC 1/SC 29 Business Plan 2022

Document type	Related content	Document date	Expected action
Project / Other	Meeting: Tokyo (Japan) 14 Nov 2022	2022-09-27	**COMMENT/REPLY** by 2022-10-17

Description

This document is circulated for review and consideration at the November 2022 JTC1 Plenary.

JTC 1/SC 29 Business Plan

Coding of Audio, Picture, Multimedia and Hypermedia Information

Period covered: September 2021 – August 2022

Produced and submitted by

Gary Sullivan, Chair

(garysull@microsoft.com)

1.0 Executive summary

SC 29 is a very active SC that develops international standards that are basis of digital media services and applications, particularly those for efficient coding of images, audio, moving pictures, graphics, and some other types of digital data (such as genomics and haptics signals) and digital information support for media technology in systems. These standards are designed to represent, package, record, preserve and convey digital media information. The standards support functionalities such as the coding, composition, transport, and file storage of media and their control, multiplexing, interfaces, and middleware for general and/or specific applications.

Since its restructuring in mid-2020, the work of SC 29 has been organized into 8 working groups and 5 advisory groups. One of the working groups is known as JPEG, the Joint Photographic Experts Group, and five of the working groups, together with three of the advisory groups, are collectively known as MPEG, the Moving Picture Experts Group. The SC 29 structure is illustrated in Figure 1.

SC 29 and its WGs and AGs comprise a community of a thousand or more people, with about 400 people attending each of the quarterly sets of meetings of its WGs and roughly a thousand total registered experts. As with most of the rest of JTC 1, of course, the meetings of SC 29 and its WGs and AGs have all been held in virtual mode since April 2000.

A highlight for SC 29 during this period is the receipt of *two* Technology & Engineering Emmy® Awards presented in April 2022 in recognition of the MPEG Dynamic Adaptive Streaming over HTTP (DASH, ISO/IEC 23009-1) and Open Font Format (ISO/IEC 14496-22) standards. This comes in the wake of the recent Primetime Engineering Emmy® received for the JPEG standard (ISO/IEC 10918-1) in 2019 and Technology & Engineering Emmy® for the ISO Base Media File Format standard (ISOBMFF, ISO/IEC 14496-12) in 2021, bringing the total number of Emmy awards received by SC 29 to nine.

SC 29 has 28 P-member national bodies and 17 O-members, more than 600 published standards under its direct responsibility, and more than 100 standards currently under development.

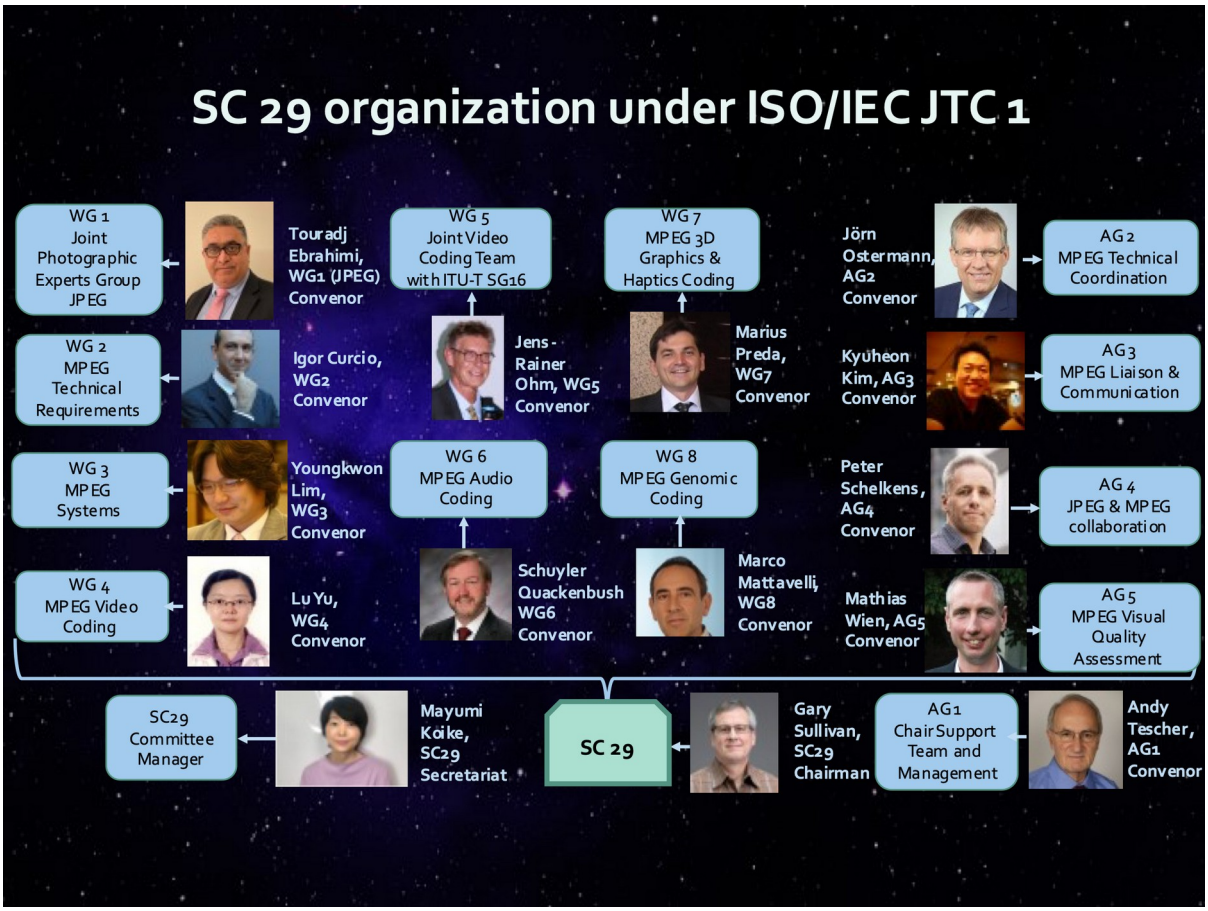

Figure 1: Structure of ISO/IEC JTC 1/SC 29

2.0 Chair's Remarks

2.1 Market requirements, innovation

For WG 1 ("JPEG"), a perennial example of major innovation, the legacy JPEG format has remained one of the leading still image formats for the last three decades, enabling the world to use and share billions of images every day. The WG 1 ISO/IEC Joint Photographic Expert Group of ISO/IEC JTC 1/SC 29 and ITU-T SG 16 was recognized with a Primetime Engineering Emmy® award in 2019 for its contribution to engineering developments. In receiving the award, the ISO, IEC and ITU JPEG group was given "special recognition for developing the image compression standard that has been universally adopted to preserve high-quality imagery in television production and its tangential yet essential workflows". More recently, JPEG has been developing an advanced image compression specifications known as JPEG XL with attractive features for applications that rely on image content. These features range from better compression efficiency to the possibility of transcoding legacy JPEG images to JPEG XL losslessly, while maintaining a very low complexity. JPEG XL has been now adopted for inclusion in a number of popular web browsers as an optional format. Another JPEG specification for the low-latency lightweight image coding system known as JPEG XS, which has been increasingly used in applications where until now uncompressed formats were preferred. More recently, JPEG has started collaborative work with ITU-T SG16 on neural-network based image coding known as JPEG AI. Additional potential new projects are also being explored such as JPEG Pleno Point Cloud Coding based in AI, new addition to JPEG AIC to cope with ranges of quality that are not yet addressed in image quality assessment, as well as system related specifications for provenance to cope with fake media in both good faith (e.g. content creation) and ill intended situations (e.g. misinformation). JPEG also continues to explore additional areas such as coding for archival on DNA support and support mechanisms to cope with NFTs represented in JPEG formats.

A roadmap of standards development in JPEG is shown in in Figures 2 and 3, and public-facing information about JPEG can be found at http://www.jpeg.org/.

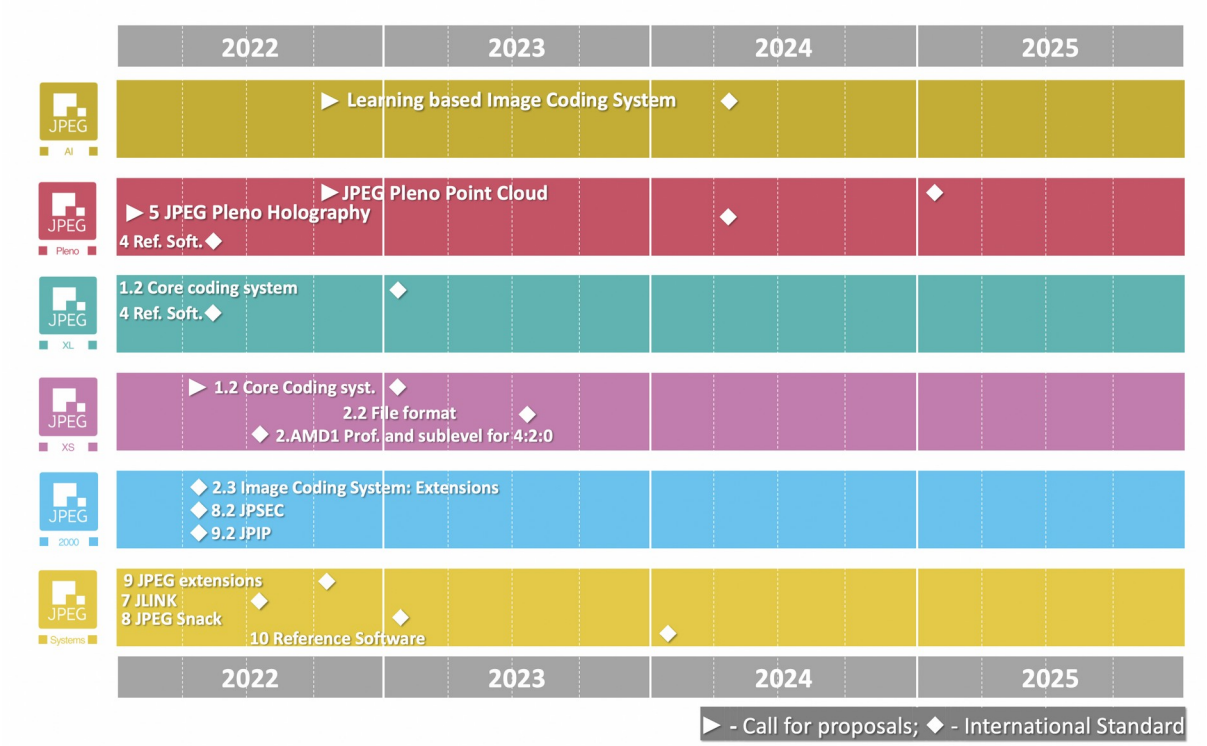

Figure 2: Roadmap of standards development in JPEG (WG1), as of July 2022 (part 1 of 2)

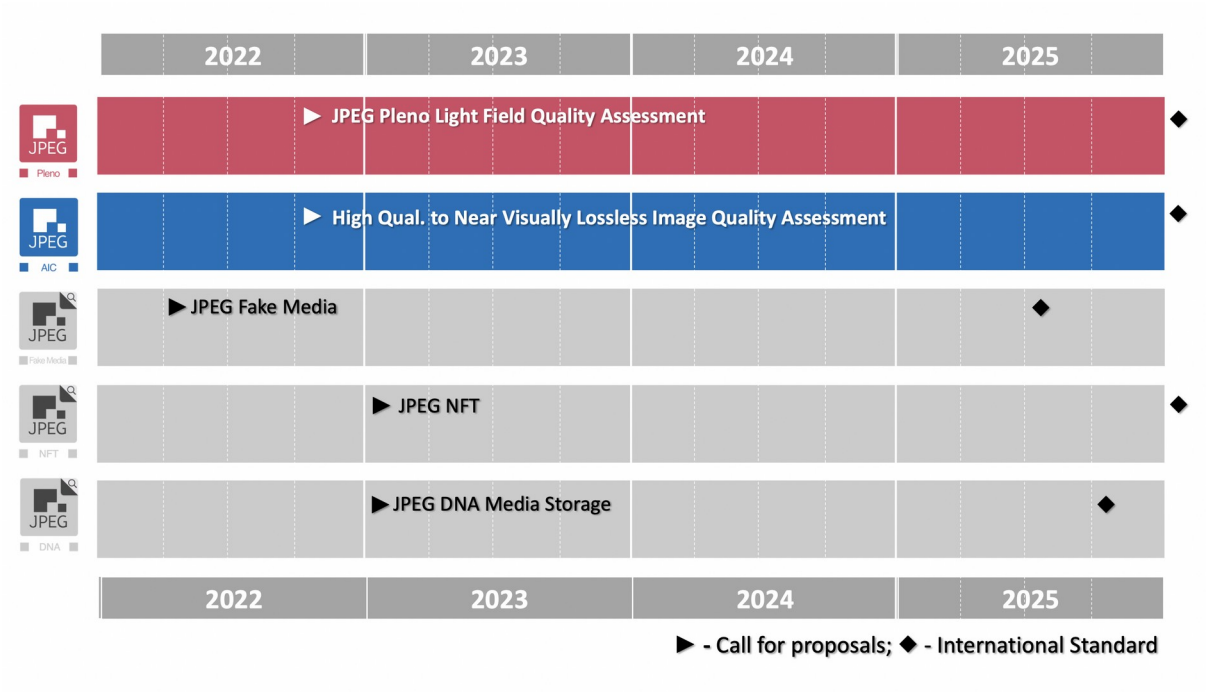

Figure 3: Roadmap of standards development in JPEG (WG1), as of July 2022 (part 2 of 2)

Concerning WGs 2-8 (the "MPEG" WGs, previously WG 11), over a span of 34 years since its founding in 1988, MPEG has become a center of expertise for the development of successful standards that have changed the industry and improved the lives of billions of people. Since its reorganization in 2020, the coherence and coordination of the MPEG community was preserved in its character and coherence by a particular modus-operandi including a common meeting calendar, regular joint meetings between the WGs, an advisory group (SC 29/AG 2) for technical coordination among all the MPEG WG convenors and meeting planning, a common advisory group (SC 29/AG 3) for liaison and communication coordination, a common advisory group (SC 29/AG 5) for visual quality assessment of the capabilities of MPEG visual coding standards and technology proposals, common work on Requirements coordinated by SC 29/WG 2, and plans to hold consistently collocated future meetings once it again becomes feasible to hold meetings with face-to-face participation. The first of such meetings is planned to be held in Mainz, Germany, in October 2022 to include most of the MPEG WGs and AGs.

The coordinated set of affiliated groups known as MPEG includes WGs 2 through 8, together with AGs 2, 3 and 5, and have held their meetings together in a coordinated fashion with routinely held joint sessions for sharing of information and coordination across the groups. Their meetings of October 2021 are collectively referred to as the 136th MPEG meeting; their meetings of January 2022 are collectively referred to as the 137th MPEG meeting; their meetings of April 2022 are collectively referred to as the 138th MPEG meeting; and their meetings of July 2022 are collectively referred to as the 139th MPEG meeting. At each of these meetings, more than 300 participants have continued to work efficiently on standards for the future needs of the industry. As a group, MPEG has also continued exploring new application areas that will benefit from standardization of efficient coding methods and related systems technologies in the future.

A roadmap of projects under recent and current development by the MPEG WGs is shown in Figure 4.

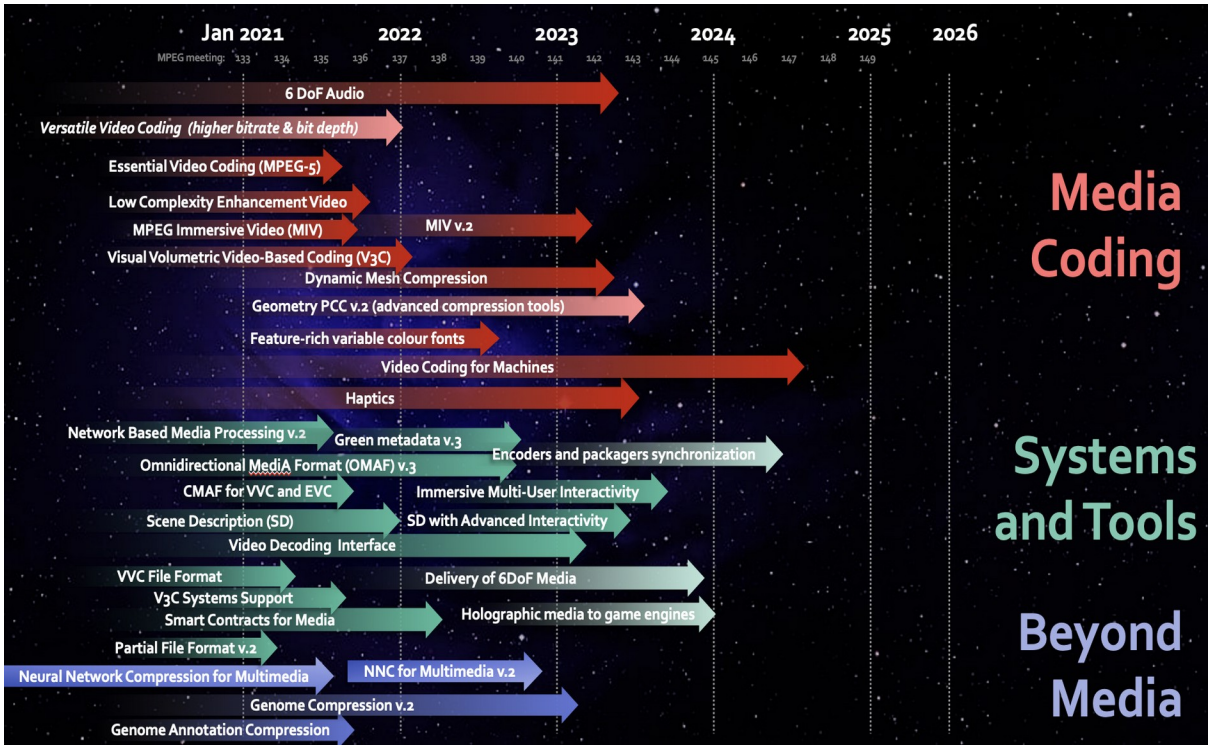

Figure 4: Roadmap of standards development in the MPEG WGs (WGs 2-8), as of July 2022

The MPEG community consists of AG 2 MPEG Technical Coordination (for overall MPEG work coordination), WG 2 MPEG Technical Requirements, WG 3 MPEG Systems, WG 4 MPEG Video Coding, WG 5 MPEG Joint Video Coding Team with ITU-T SG 16, WG 6 MPEG Audio Coding, WG 7 MPEG 3D Graphics and Haptics Coding, WG 8 MPEG Genomic Coding, AG 3 MPEG Liaison and Communication, and AG 5 MPEG Visual Quality Assessment. More than 300 participants continued to work efficiently on standards for the future needs of the industry.

Public-facing information about MPEG can be found at http://www.mpeg.org/.

Two additional advisory groups assist in SC 29 coordination. AG 1 assists the chair on various SC 29 matters and particularly provides advice on management issues, and AG 4 is constituted for coordination between the JPEG and MPEG communities. AG 4 has particularly been active in organizing workshops for exchange of information and outreach on JPEG and MPEG projects in related areas.

For WGs 1 to 7**,** immersive applications, including virtual and augmented reality, 360° image/video, 3D audio and 3D representations (e.g. point clouds), are continuing to emerge. To avoid fragmentation in the market and to ensure interoperability, standards are required for such applications.

Media data is expected to be more machine- or system-friendly as rich sensory data with the development of Cyber Physical Systems and IoT-type applications. There is also a rapid expansion of multimedia content generation for security and monitoring applications, and increased use of machine-based consumption and analysis of multimedia content. The coordinated use of content from multiple sources is also rapidly expanding, as contrasted with the traditional model of linear content generation, storage and playout. Having standardized mechanisms to support such systems and applications is needed.

Moreover, there is an ongoing rapid expansion of the use of artificial intelligence / machine learning / neural network technology across a broad range of applications, including media coding, coordination and analysis. Such developments have already begun to appear in multiple projects within SC 29, and a large amount of further future impact of such technology is anticipated.

SC 29's WG 3 MPEG Systems wins two more Technology & Engineering Emmy® Awards

SC 29 is pleased to report its standards again being recognized this year by the National Academy for Television Arts and Sciences (NATAS), this time with *two* Technology & Engineering Emmy® Awards presented to its WG 3 *MPEG Systems* working group, for *(i)* "standardization of font technology for custom downloadable fonts and typography for Web and TV devices and for *(ii)* "standardization of HTTP encapsulated protocols", respectively.

The first of these Emmys is related to MPEG's Open Font Format (ISO/IEC 14496-22). Fonts are the critical components of any written communication. Text carries a meaning, but it is the font that makes text readable, *i.e.*, fonts give the written word its voice. The standardization of the Open Font Format technology by MPEG over the last 20 years has significantly influenced the capabilities of all classes of consumer electronic devices, bringing advanced font technology for digital TV, streaming media environments, and the Web. It also inspired many open-source projects that enabled mass adoption of high-quality font rendering and advanced text support, making high quality font support easy and cost-effective for OEMs, service providers and content authors to deploy new features and applications supporting all of the world's languages and writing systems.

The second of these Emmys is related to MPEG Dynamic Adaptive Streaming over HTTP (*i.e.*, MPEG DASH, ISO/IEC 23009). The MPEG DASH standard is the only commercially deployed international standard technology for media streaming over HTTP and it is widely used in many products. MPEG developed the first edition of the DASH standard in 2012 in collaboration with 3GPP and since then has produced four more editions amending the core specification by adding new features and extended functionality. Furthermore, MPEG has developed six other standards as additional "parts" of ISO/IEC 23009 enabling the effective use of the MPEG DASH standards with reference software and conformance testing tools, guidelines, and enhancements for additional deployment scenarios. MPEG DASH has dramatically changed the streaming industry by providing a standard that is widely adopted by various consortia such as 3GPP, ATSC, DVB, and HbbTV, and across different sectors. The success of this standard is due to its technical excellence, large participation of the industry in its development, addressing the market needs, and working with all sectors of industry all under ISO/IEC JTC 1/SC 29 MPEG Systems' standard development practices and leadership.

Members of the leadership of JTC 1 and SC 29 lauded the accomplishment:

> "Thank you to the National Academy of Television Arts & Sciences (NATAS) for recognizing the outstanding contributions of media coding experts in ISO/IEC JTC 1. Their work has significantly expanded the benefits of MPEG standardization." — *Philip C. Wennblom, Chair of ISO/IEC JTC 1*
>
> "It's a tremendous recognition for the MPEG Systems work in SC 29 to receive two Technology & Engineering Emmy® awards in 2022. The MPEG Open Font Format and DASH standards have been very widely adopted in industry and have become fundamental to the interoperability of today's television and web-based multimedia services." — *Gary J. Sullivan, Chair of ISO/IEC JTC 1/SC 29*

> "MPEG Systems has been persistently working to respond to industry needs in collaboration with the key industry players and other standards development organizations. The two Technology & Engineering Emmy® awards this year for the MPEG Open Font Format and DASH standards show the great success and industry recognition of such efforts." — *Youngkwon Lim, Convenor of MPEG Systems*

These are SC 29's eighth and ninth Emmy® Awards. They are the fifth and sixth Technology & Engineering Emmy® Awards (after MPEG-1 and MPEG-2 together with JPEG in 1996, Advanced Video Coding (AVC) in 2008, MPEG-2 Transport Stream in 2014, and ISO Base Media File Format in 2021) and SC 29's eighth and ninth overall Emmy® Awards when also including the three Primetime Engineering Emmy® Awards for Advanced Video Coding (AVC) High Profile in 2008, High Efficiency Video Coding (HEVC) in 2017 and JPEG in 2019.

See also the ISO press release at https://www.iso.org/contents/news/2022/04/record-breaking-mpeg-takes-two-e.html and the IEC press release at https://iec.ch/blog/mpeg-wins-two-new-technology-and-engineering-emmys.

White papers for educational outreach

An example of the outreach and educational activities of SC 29 is the production, at the January 2022, April 2022 and July 2022 meetings, by SC 29/AG 3 *MPEG Liaison and Communication*, of informal publicly released white papers on the following five topics.

- **Neural Network Coding (NNC)**: Artificial neural networks have been adopted for a broad range of tasks in almost every technical field, such as medical applications, transportation, network optimization, big data analysis, surveillance, speech, audio, image and video classification, image and video compression, and many more. Their recent success is based on the feasibility of processing much larger and more complex neural networks (deep neural networks, DNNs) than in the past, and the availability of large-scale training data sets. An additional factor for the exponential growth in the use of such neural networks is the appearance of new use cases, such as federated learning with continuous communication between many devices. Accordingly, this requires the highest capability for compression of neural networks to minimize the overall communication traffic and to reduce the size of the networks when used for inference. Thus, a standard for neural network coding (NNC) has been defined in ISO/IEC 15938-17 "Compression of Neural Networks for Multimedia Description and Analysis".
- **Low Complexity Enhancement Video Coding (LCEVC)**: This white paper provides an overview of ISO/IEC 23094-2 Low Complexity Enhancement Video Coding (LCEVC). The coding format is designed for use in conjunction with existing video coding formats, leveraging encoder-driven upsampling and specific tools for the encoding of "residuals", i.e., the difference between the original video and a predicted rendition. LCEVC can improve the compression efficiency and reduce the overall computational complexity for coding a given resolution and bit depth by using a small number of specialized enhancement tools. This white paper provides an outline of the architecture, coding tools, and an overview of the compression efficiency of LCEVC.
- **MPEG Immersive vide**o (MIV): The MPEG Immersive video (MIV) standard produced by MPEG was completed with its Final Draft International Standard (FDIS) issued in October 2021 for final approval and publication. The goal of the MIV standard is to provide efficient coding of immersive, six degree of freedom (6DoF) volumetric visual scenes. An immersive 6DoF representation, unlike a three degree of freedom (3DoF) representation, provides a larger viewing space, where viewers have both translational and rotational freedom of movement at their disposition. 6DoF video also enables the perception of motion parallax, where the relative positions of scene geometry change with the pose of the viewer. Without the 6DoF capabilities of MIV, the absence of motion parallax in 3DoF video is inconsistent with the workings of the human visual system and often leads to visual discomfort. This white paper provides a brief overview of the important new MIV standard.
- **MPEG Smart Contracts for Media**: In the last few years, MPEG has developed a set of standardized RDF ontologies and XML schemas for the codification of intellectual property (IP) rights information related to music and media. ISO/IEC 21000-19 Media Value Chain Ontology (MVCO) facilitates rights tracking for fair, timely, and transparent transactions by capturing user roles and their permissible

actions on a particular IP entity. ISO/IEC 21000-19/AMD1 Audio Value Chain Ontology (AVCO) extends MVCO functionality related to the description of IP entities in the audio domain, e.g., multitrack audio and time segments. ISO/IEC 21000-21 (2nd Ed) Media Contract Ontology (MCO) facilitates the conversion of narrative contracts to digital ones related to the exploitation of IP rights, payments, and notifications. With respect to the latter, XML schemas have been developed as ISO/IEC 21000-20 (2nd Ed) Contract Expression Language (CEL). Furthermore, the axioms in these XML schemas and RDF ontologies can drive the execution of rights-related workflows in controlled environments, e.g., Distributed Ledger Technologies (DLTs), where transparency and interoperability are favoured toward fair trade of music and media. Thus, the aim of ISO/IEC 21000-23 Smart Contracts for Media is to provide the means (e.g., application programming interfaces) for converting these XML and RDF media contracts to smart contracts that can be executed on existing DLT environments. By performing this conversion in a standard way for several smart contract languages, MPEG-21 CEL/MCO schemas and ontologies will prevail as the interlingua for transferring verified contractual data from one DLT to another. Another important feature of this standard is that it offers the possibility to bind the clauses of a smart contract with those of a narrative contract and vice versa. In this way, each party signing a smart contract knows exactly what the clauses stored in the smart contract express.

- **MPEG Common Media Application Format (CMAF)**: The Common Media Application Format (CMAF), as defined in ISO/IEC 23000-19, has found significant market adoption as the media streaming format used by industry, in particular for a convergence between ISO/IEC 23009 MPEG Dynamic Adaptive Streaming over HTTP (DASH) and other streaming technology. Based on these developments and new market needs, the CMAF standard has been continuously developed, and bug fixes and maintenance have been addressed taking into account feedback from deployments. The most recent result of these efforts is MPEG's completion of the third edition of CMAF which incorporates the changes introduced by three approved amendments. The third edition provides several additional capabilities, along with general bug fixes, clarifications and improvements. The new features include the definition of new logical structures, namely CMAF random access chunk and CMAF Principal Header; the addition of timed metadata tracks; the addition of new media profiles for HEVC for High Frame Rate, interlaced video content, and 8K-UHD; adding a listing of common source formats based on display resolution operating point sets and frame rate operating point sets; and the addition of new media profiles for MPEG-H audio and the Versatile Video Coding and Essential Video Coding formats.

Verification testing

Completion of Subjective Verification Tests for Video-based Point Cloud Compression (V-PCC)

At the 136th MPEG meeting, a subjective test to assess the performance of the new Video-based Point Cloud Compression (V-PCC) standard (ISO/IEC 23090-5) was completed by the *MPEG Visual Quality Assessment advisory group* (AG 5). The results confirm that excellent visual quality can be achieved using bit rates between 15 and 25 Mbps. All tested V-PCC profiles show considerable compression gain (*i.e.*, 8-9 times better) compared to the anchor Point Cloud Library (PCL) codec.

Subjective quality was assessed using rigorous methods proven in 2D video evaluation wherein a dynamic 3D point cloud is rendered to a 2D video by following a preselected camera path.

Four profiles have been subjectively tested: *(i)* one profile with a very simple reconstruction in 3D space and *(ii)* three profiles with a more sophisticated reconstruction in the 3D space. The first profile used High Efficiency Video Coding (HEVC) as the underlying video codec, while the others used either HEVC or Versatile Video Coding (VVC) as the underlying video codec. The subjective quality assessment further confirms that compelling gains can be achieved by replacing of the underlying video codec (*i.e.*, VVC instead of HEVC) and by keeping the point cloud codec unchanged.

V-PCC is specified in ISO/IEC 23090-5:2021 — Information technology — Coded representation of immersive media — Part 5: *Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC)*.

Future Project Explorations

Explorations toward potential additional new projects within SC 29 include various topics such as:

- Fake media countermeasures such as media authentication
- Best practices and specifications for non-fungible tokens (NFTs)
- DNA-based storage for archival of media information
- Point cloud coding based on neural network technology
- Coding of point clouds generated by advanced LiDAR and volumetric scanners
- Video coding for machines
- Neural-network based video coding
- Video coding with compression capability beyond that of the Versatile Video Coding (VVC) standard
- Incremental compression of neural networks for multimedia content description and analysis
- Lenslet-view video coding for dense light fields
- AI-based technologies for graphics compression
- Coding for advanced genomics features and technologies
- Implicit neural visual representation

2.2 Accomplishments

The status of all standards published by SC 29, together with those under development, may be viewed in the SC 29 Standards Catalogue.

The accomplishments of SC 29 are reported below, categorized within Clause 3 below according to the Working Group assignments.

2.3 Resources

SC 29 has sufficient participation of experts to proceed with its work on the JPEG, JPEG 2000, AIC, JPEG XR, JPEG XT, JPEG Systems, JPEG XS, JPEG XL, JPEG Pleno, JPEG AI, MPEG-2, MPEG-4, MPEG-5, MPEG-7, MPEG-21, MPEG-A to MPEG-E, MPEG-G, MPEG-H, MPEG-I, MPEG-M, MPEG-V, MPEG-U, MPEG-CICP, MPEG-DASH and MPEG-IoMT projects in addition to other standards being developed and additional topics under investigation. SC 29 and its WGs comprise a community of a thousand or more people, with about 400 people attending each quarterly set of the meetings of its WGs and roughly a thousand total registered experts.

2.4 Competition and cooperation (including consortia)

SC 29 maintains many liaisons with other organizations to meet the requirements and expectations of the standards users from other communities. SC 29 has, in particular, been continuing its productive collaboration with ITU-T. The Joint Photographic Experts Group (JPEG) and the Joint Video Exploration Team (JVET) are continued joint activities that SC 29 shares with ITU-T SG 16 (the ITU-T Study Group on Multimedia), and the MPEG-2 Systems standard (ISO/IEC 13818-1, a.k.a. ITU-T H.222.0) is also maintained as common text.

JVET, reporting as WG 5 to SC 29, is a major ongoing joint activity with ITU-T. JVET was originally formed in October 2017, and it completed the development of the first edition of Versatile Video Coding (VVC, ISO/IEC 23090-3, a.k.a. ITU-T H.266) in 2020, to provide a significant improvement in compression performance over the prior High Efficiency Video Coding standard (HEVC, ISO/IEC 23008-2, a.k.a. ITU-T H.265). Together with VVC, JVET also produced a standard on *Versatile supplemental enhancement information messages for coded video bitstreams* (VSEI, ISO/IEC 23002-7 | ITU-T H.274). JVET has more recently been standardizing an operation range extension for VVC, additional supplemental enhancement information for VSEI, a conformance testing standard, and reference software. JVET is also responsible for maintenance and minor enhancement of other jointly developed standards of SC 29 and ITU-T that are related to video coding, and is conducting further exploration studies of the potential for additional future joint developments in this area.

A number of standards in SC 29/WG 1 (JPEG) are also developed jointly with SG 16. Most recently, an agreement was reached for joint standardization of the upcoming JPEG AI standard (ISO/IEC 29170). Other WG 1 standards that were jointly developed and have been subsequently maintained and extended with ITU-T SG 16 include the original 1992 JPEG standard, the JPEG 2000 family, and JPEG XR.

SC 29 and its WGs are open to welcoming new members and seeking opportunities to collaborate with new liaison partners, and it is well recognized that this field of work requires rapid progress and has the potential to bring about new initiatives and new competitors. SC 29 and its WGs strive to always communicate with the stakeholders in the industry and recruit appropriate members in order to maintain relevance and competitiveness.

In the area of video coding, other organizations have also been recently developed their own video coding specifications. The AOM (Alliance for Open Media) consortium produced the AV1 video coding specification in 2018 and has promoted its use for video streaming services. AVS (Audio Video coding Standard Workgroup of China), working recently together with a working group in IEEE, have developed AVS3 video coding technology. The Society of Motion Picture and Television Engineers (SMPTE) has also published various video coding standards. Although recent studies indicate that the coding efficiency of VVC (Versatile Video Coding) is substantially higher than that of its competitors, competitive efforts have continued to evolve, and SC29 will continue monitor and respond to such outside developments.

The Khronos Group consortium is also highly relevant to SC 29 in several areas, recently especially in regard to SC 29/WG 7 (MPEG 3D Graphics and Haptics Coding) and SC 29/WG 3 (MPEG Systems). In mid-2021, Khronos was approved as a PAS submitter for JTC 1 and is expected to submit specifications as candidate standards for graphics representation. SC 29 members have a keen interest in these developments, and liaison communication has been held between SC 29 and Khronos.

Two other subcommittees within JTC 1 have recently been considered especially relevant to the work of SC 29:

- SC 24 (*Computer graphics, image processing and environmental data representation*) works on closely related technology, and is especially relevant to SC 29 WG 7 (MPEG 3D Graphics and Haptics Coding). SC 29 has recently been communicating with SC 24 in its effort to propose a revised description of its scope for approval by JTC 1 and has also worked to improve communication and strengthen ties between the two SCs on other matters.
- SC 42 (*Artificial intelligence*) is chartered with work on technology that is emerging very rapidly with many applications, particularly including the coding and analysis of audio-visual media.

3.0 Discussion of SC program of work

3.1 SC Opportunities and Strategies

SC 29 has delivered several generations of major media coding standards. The capabilities and characteristics of these standards, such as compression capability and the functionalities supporting media composition, transport and storage, have been evolutionally enhanced. There are still evident needs for support of high quality, usability and advanced features. Further compression of media is expected to deliver media over the Internet and mobile channels. Ultra HD, 3D 360-degree, holographic and immersive media emerging capabilities requiring support for existing and new applications. Regarding these requirements, the industry needs further efficiency and adaptability in compression, composition, description and control of digital media.

In the era of machine learning based processing of massive quantities of data, SC 29 has identified a high potential but still challenging opportunity in developing artificial intelligence based media coding.

The recipients or consumers of digital media information are not necessarily always human but can rather be machines or systems in the era of Internet of Things. Thus, SC 29 has many opportunities and challenges in its future work.

SC 29 continues to provide information on the progress of standardization work to the public through SC's and WGs' web sites as well as press releases, white papers and awareness events in order for attention to be paid to its recent developments the area of the multimedia information technology and to solicit new input and participation.

SC 29 continues the practice of having standards contain conformance testing bitstreams and reference software accessible as freely available standards from ISO/IEC. SC 29 is continually soliciting information about application requirements from industry, so that appropriate WGs under SC 29 will study and take advantage of the opportunities to deliver the standards that meet such requirements. The WGs may hold public seminars to introduce their activities and to align future standardization with users and industry needs.

3.2 SC and WG risks

Risks noted in the work of SC 29 include the following:

— Scope, structure, and personnel transitions: A revised structure for SC 29 became effective on July 17, 2020, and a new Chair was introduced with a term that began in January 2021. Minor refinement of the terms of reference for some groups within SC 29 has since been conducted, and the new Convenors and Chair have settled into their fields of responsibility. Further study and consideration of organizational matters is likely, although no major problems have been evident so far, following these transitions.

— Virtual meeting practices: The COVID-19 pandemic has been a major disruption to the conduct and planning of meetings, forcing an abrupt switch to all-virtual meetings that began for SC 29 with its WG meetings of April 2020. As of the time of preparation of this report, no hybrid or face-to-face meetings of SC 29 or its WGs have been held since then, although there are plans for face-to-face meetings for some of the SC 29 WGs to begin to be held in October 2022 and onwards. While SC 29 has worked hard to maintain momentum and to progress its work programme through this crisis, the loss of the ability to conduct face-to-face work has interfered with many aspects of standards development. One example is the difficulty of conducting tests of perceptual effects of technology that require human test subjects, special equipment and on-site expertise. A deterioration in group engagement and coherence is evident, and virtual meetings require difficult choices between attendee fatigue, abbreviated meeting times, and more frequent ad-hoc meetings. The camaraderie and convivial social environment that accompanies face-to-face gatherings has diminished, together with the associated opportunities for informal discussions that help identify and resolve key issues. While SC 29 is proud that some of its own standards that it has developed have been key tools for assisting with effective telecommunication, there is a strong yearning for resumption of some amount of in-person standards development activity. On the other hand, some of the lessons learned during this difficult period are likely to have a persistent and positive impact, such as greater use of online work coordination tools. Some people who in the past were not able to participate due to difficulty with travel have been able to participate more, and certainly some travel time and expenses have been saved. Looking forward, the future plans within SC 29 that include a face-to-face element are being planned as “face-to-face mode” with provisions for remote participation on a best-effort basis rather than as “hybrid mode” meetings in the sense defined in (the recently updated) SD 19, as it seems difficult to provide the full assurances needed for the hybrid mode.

— Risk related to patents: In order to meet the requirements of industry and to cope with competition, SC 29’s standards should incorporate new technologies and ideas, which may be associated with patent rights that require licensing. The number of relevant patents and patent holders corresponding to the advanced features of complex technology standards can grow so large as to make one-by-one licensing negotiations difficult and can include parties that have not participated in the development process of the standard. Issues can arise that delay or cause suspension of the standards development. Hesitance to adopt some recent standards produced by SC 29 has been evident due to perceived high costs, complexity and business risk associated with patent rights licensing (e.g., for High Efficiency Video Coding ISO/IEC 23008-2 and Versatile Video Coding ISO/IEC 23090-3), and some large and prolonged lawsuits over patent rights for past SC 29 standards have occurred in the past (e.g., for Advanced Video Coding, ISO/IEC 14496-10). One recent standard from SC 29 (High-Throughput JPEG 2000, ISO/IEC 15444-15) had a “type 3” assertion of a lack of willingness to license that was filed a year after the FDIS approval ballot and eight months after publication of the standard by a company that had experts participating in the WG and was then later withdrawn after about seven months, introducing potential confusion into the deployment plans of the targeted usage community. SC 29 and its WGs continue to encourage their members to submit patent rights declarations. Meanwhile, some competitive efforts in industry consortia have developed specifications and software implementations under different policies which are promoted as “open and royalty free” or with “one-stop shopping” licensing availability. Further study may be needed to consider how ISO/IEC can better respond to these developments.

— Free public availability of the work: Industry practices are changing, and standard developments have become more open public activities. Several SDOs developing standards for similar areas, such as 3GPP, IETF, ITU, W3C, and Khronos, make their published standards and other publications (e.g., technical reports) available to the public for free. Some such SDOs also do this for non-final drafts. Potential participants may choose not to do their work in SDOs that do not follow this increasingly common practice. Restrictions on the ability to exchange information, invite participation and solicit input and feedback from non-members can also discourage participation and put at risk the quality and relevance of the standards that are produced. Greater flexibility for making published texts and the work under development externally available (and available without cost) would improve alignment with this new industry practice.

— Delay of final publication of standards: As the technology development cycle has continued to shorten, industry demands access to standards in a timely fashion. However, there have been cases in which the publication of standards has been delayed for quite some time, i.e. more than one year, to resolve minor editorial issues. Requirements for extensive editorial changes to be made to unchanged parts of existing standards that are being revised have aggravated the problem and risk the inadvertent introduction of technical errors. House Style and publication workflow requirements that are rigid and different from conventions followed elsewhere can be a source of significant frustration without necessarily contributing to the quality or relevance of the published result. Some parts of the publication workflow do harm to the ability to use the result as the basis of further work – e.g., by removing automatic numbering and automatically updated cross-reference links. SC 29 has tried to improve the communication between the ISO CS Editors and the Project Editors to resolve such issues quickly and efficiently, but such issues have continued to be an ongoing problem. Accelerating the standards publication process and making technically mature texts rapidly available to industry would serve the user of standards and strengthen the relevance of the standards to the real-world applications where standards are needed.

3.3 WG 1 JPEG Coding of Digital Representations of Images

3.3.1 WG 1 accomplishments

WG 1 Terms of Reference

WG 1 develops of international standards that support broad interoperability targeting user experiences based on digital representations of light.

WG 1 summary

Besides the well-known legacy JPEG image coding standard ISO/IEC 10918, also known as the “JPEG” standard, the WG1 continuously develops new image coding standards triggered by market needs and new technologies. Examples of the latter are the increased availability of multicore CPU or GPU processors, artificial intelligence, new image modalities like point cloud, light field or holograms or new storage and transport mechanisms. This progress leads to the development of new coding algorithms, image representations, and support mechanisms or functionality extensions.

The typical process for standardization within WG1 is as follows: Based on request from stakeholders new ideas on standardization are brought into the Committee, investigations are launched to define relevant use cases and potential solutions, and if sufficient evidence is found, either an extension of a specific standard or a new work item for a new standard is launched with often a call for proposals issued, followed by evaluation, core experiments and consensus finding inside the Committee for defining specifications. Where relevant, WG1 also takes into account backward and forward compatibility and the cross standardization issues.

Currently, WG1 has several ongoing standardization activities shown in the timelines of Figures 2 and 3. It can be observed that in 2022 several activities have launched their call for proposals and are following their first phase of standardization, notably the new learning-based image coding system JPEG AI, the JPEG Pleno Holography and Point Cloud coding activities as well as a number of JPEG Systems extensions. Many activities are soon to become new or revised international standards as seen in the standardization projects JPEG Pleno, JPEG XL, JPEG XS, JPEG 2000 and JPEG Systems. JPEG AI is being developed as a joint standard between ISO/IEC and ITU-T.

JPEG 2000 (ISO/IEC 15444)

Work continues on maintaining and improving the JPEG 2000 family of standards, with an objective of adapting to evolving industry requirements and resolving issues identified through implementation experience.

- Part 17 (Extensions for coding of discontinuous media) has reached Final Draft International Standard and is expected to be finalized in 2023. Part 17 adds breakpoint-dependent spatial wavelet transforms and scalable breakpoint coding to JPEG 2000, and was designed for discontinuous media such as optical flow data and stereo disparity maps.
- Part 15 (HTJ2K), which brings an order of magnitude increase in throughput to JPEG 2000, is now supported by multiple commercial and open source implementations.
- Extensive maintenance work has been completed on Part 1 and Part 2, which form the core of the JPEG 2000 family of standards. This work fixed outstanding bugs, merged the many amendments that had accumulated over the years and brought the documents to modern editorial standards. The objective is to improve interoperability and facilitate implementation.
- Part 4, 5 and 16, which define JPEG 2000 conformance testing, reference software and wrapping in HEIF, respectively, have also benefited from general improvements.

JPEG XS (ISO/IEC 21122)

JPEG XS is a high quality, low complexity, low latency image codec for mezzanine compression. Typical use cases include compression ratios from 4:1 to 16:1 and transport mechanisms like RTP streaming and file wrapping with MXF. Work continues on extending the parameter set and performance of the coding system.

- The second editions of Part 1 (Core coding system) and Part 2 (Profiles and buffer models) add support for 420 subsampling, CFA RAW Bayer pattern coding and mathematically lossless coding, that were published recently in 2022. The CFA RAW Bayer pattern coding was improved significantly by using a so-called Star-Tetrix transformation and adoption of specific tools like pre- and post-emphasis.
- The second editions of Part 3 (Transport and Container Formats), Part 4 (Conformance Testing) and Part 5 (Reference Software) reflect the changes of Part 1 and 2.
- In 2022, the work on the third edition of JPEG XS parts has started. The main goal is to improve the coding efficiency especially for screen content by relaxing the buffer size requirements but keeping the latency. It is expected that the inclusion of a buffer with data from the previous frame will be allowed to increase the prediction performance.
- It should be mentioned that outside JPEG also additional standards have been created to enable the use of JPEG XS in different applications. These include an IETF standard for RTP payload of JPEG XS codestreams called RFC 9134, a SMPTE standard for file wrapping of JPEG XS codestreams with the MXF file format in ST2124, Technical Recommendations of the Video Services Forum (VSF) for transport of codestreams in MPEG 2 transport streams as TR-07 or in ST2110 as TR-08 [6] or a specification within the Networked Media Open Specification (NMOS) from AMWA for enabling registration, discovery and connection management of JPEG XS endpoints.

JPEG Systems (ISO/IEC19566)

JPEG Systems was introduced to describe additional tools and features of the context and to allow for better interoperability across JPEG standards. Some of these newer additional features contain extensions for privacy and security, universal metadata, 360-degree images, linked images or multimedia shows. Main activities since 2020 have been:

- JPEG Systems Part 7 - JPEG Linked Media Format (JLINK): Multiple images (a collection) can be stored in one standard JPEG file format and linked together with action points. This allows for example virtual tours in properties or zooming in to details of a specific image area. The user can navigate through the image collection by clicking these action points.
- JPEG Systems Part 8 - JPEG Snack: Multiple images or other media types are composed together as animated multi-media show to tell a short story.

- JPEG Systems Part 9 - JPEG extensions mechanisms to facilitate forward and backward compatibility. This technical report describes the mechanism and syntax to enable forward and backward compatibility.

Important to know is that these features can be used in both, the legacy JPEG file formats (JPEG, JPEG XT) and the more recent box based file formats (JPEG 2000, JPEG XS, JPEG XL). To enable this compatibility of features in legacy and box-based file architecture, new features are defined in boxes similar to those used in JPEG 2000 and they are split to App11 marker segments in the legacy JPEG formats. This allows a backward compatible use of such images, meaning that an old JPEG decoder could still decode the main image, whereas a new decoder can benefit from additional features.

JPEG AIC (ISO/IEC 29170)

The purpose of JPEG's Advanced Image Coding and Evaluations (JPEG AIC) is to produce best practices and new specifications for performance assessment of image coding and in particular quality evaluation methodologies and procedures. The standard has already produced two specifications divided in two parts. Part 1 of JPEG AIC is a technical report and recommends best practices for coding system evaluation of images and image sequences for use cases where distortions are visible. Part 2, defines a set of subjective evaluation methodologies and procedures for use cases in which the quality of the content is either totally transparent when compared to uncompressed, or nearly lossless. WG1 is now investigating creating additional specifications for use cases not efficiently covered in Parts 1 and 2, such as subjective quality assessment methods and objective metrics related to web imaging where the quality of images are higher than those covered in Part 1 and less than those in Part 2. During the 95th JPEG meeting, the Committee released a Draft Call for Contributions on Subjective Image Quality Assessment. The new JPEG AIC standard will be developed considering all the submissions to the Call for Contributions in a collaborative process. The deadline for the submission is set for 14 October 2022. Multiple types of contributions are accepted, notably subjective assessment methods including supporting evidence and detailed description, test material, interchange format, software implementation, criteria and protocols for evaluation, additional relevant use cases and requirements, and any relevant evidence or literature.

JPEG AIC AHG also organized a workshop on subjective assessment methods for the investigated quality range. The workshop targeted on obtaining different views on the problem from experts and stakeholders both inside and outside of WG1.

JPEG Pleno (ISO/IEC 21794)

General

JPEG Pleno supports the representation of new imaging modalities, such as light field, point cloud, and holographic imaging. Such imaging should be understood as representations of light inspired by the plenoptic function, which describes the radiance in time and in space obtained by positioning a pinhole camera at every viewpoint in 3D spatial coordinates, every viewing angle and every wavelength, resulting in a 7D function.

JPEG Pleno standard tools are designed in a manner to consider their synergies and dependencies for the whole to be effectively greater than the sum of its parts. To fully exploit this holistic approach, JPEG Pleno is not just a set of efficient coding tools addressing compression efficiency. It is also a representation framework understood as a fully integrated system for providing advanced functionalities support for image manipulation, metadata, random access and interaction as well as various file formats. In addition, it should offer privacy protection, ownership rights, and security.

JPEG Pleno Part 1 describes the overall framework and the generic JPEG Pleno file format, JPL. This file format is built on the box-based format specified in JPEG 2000 Part 1 and 2 and allows for carrying light field, point cloud and holographic data. Since plenoptic data can represent huge amounts of data in several hundreds of Gigabytes and more, specific directory functionality has been integrated to easily retrieve particular segments of the carried data. Moreover, a system has been integrated that enables carrying for example several separately encoded point clouds (or other modalities) and signaling their position and orientation in a 3D reference grid. Also, light fields can be positioned in this grid and it is possible to signal in addition intrinsic and extrinsic camera calibration data for fine-tuning the individual views. The latter is particularly important when sending data from more complex camera array configurations. The offered functionality is particularly

interesting for more advanced 3D media production and 3D rendering environments. Also, AR/VR services and devices will benefit from this when displaying complex 3D scenes.

Light Field Coding

JPEG Pleno Part 2 specifies the codestream syntax for **light field data** and associated metadata. Two decoding modes are supported: the 4D-Prediction mode and the 4D-Transform mode. The 4D-Prediction mode exploits the inter-view redundancy of the 4D light field via depth-based warping. The 4D-Transform mode employs 4D block partitioning, 4D Discrete Cosine Transform (DCT) and 4D bit-plane hexadeca-tree decoding. For each mode a specific profile is defined containing all coding tools necessary for that mode. For each profile, four levels are defined depending on the: Maximum number of samples and Maximum block dimension for the Baseline block-based profile; and only on the Maximum number of samples for the Baseline view-based profile (a sample corresponds to a single component value per channel, texture or depth).

JPEG Pleno Part 3 establishes the conformance testing methodology to guarantee that an application complies with the JPEG Pleno plenoptic image coding system standard. It provides a framework to enable conformity testing for JPEG Pleno encoders and decoders.

JPEG Pleno Part 4 provides the reference decoder software capable of decoding codestreams and a sample encoder software capable of producing codestreams both conformant to JPEG Pleno Part 1 and JPEG Pleno Part 2.

In the context of the JPEG Pleno standardization process with respect to light field coding (Part 2), various subjective visual quality assessment procedures have been designed and used. To further improve this task, WG1 has launched a new standardization effort known as JPEG Pleno Quality Assessment. The first phase of this effort will address the light field modality and should build on the light field quality assessment tools developed by WG1 in recent years. WG1 recently issued a call for contributions (CfC) to collect new methodologies and best practices for light field subjective quality assessment methods to assess artifacts induced by coding algorithms. Subjectively-annotated light field databases will support the evaluation of the objective quality assessment methods in a second call.

Point Cloud Coding

The JPEG Pleno activities also focused on **Point Cloud coding** where several explorations were conducted to prepare a Call for Proposals for point cloud coding technology. Some of those explorations took the form of Exploratory Studies (ES) and dealt mostly with choosing point cloud objective quality metrics both for geometry-only point clouds and geometry plus color attributes point clouds. Other ES focused on the design and test of protocols for subjective evaluation of encoded point clouds, on the choice of adequate test content (point clouds with different characteristics) and anchor point cloud codecs suitable to be used in performance comparison with codecs submitted. After analyzing the results of these ESs and those of a Call for Evidence (CfE), JPEG Pleno PCC AHG prepared a Call for Proposals (CfP) which was issued in January 2022. One important feature of this Call was that it specified deep-learning based coding technology only, excluding other types of processing models. It is expected that after the evaluation of the proposed codec, scheduled in Summer 2022, a series of decisions will be taken concerning the course of action towards a JPEG standard for Point Cloud coding, with the release of a working draft (WD) in the Fall 2022.

Holography

Work on **JPEG Pleno Holography** started recently with a document for use cases and requirements and a Call for Proposals. The target date for this standard is set to April 2024.

JPEG XL (ISO/IEC 18181)

JPEG XL has a rich set of features and is designed to meet the needs of image delivery on the web, as well as professional photography, digital art, printing, medical and scientific applications. It supports wide color gamut as well as high dynamic range and high bit depth images. JPEG XL further includes features such as animation, alpha channels, layers, thumbnails, lossless and progressive coding to support a wide range of use cases. It is particularly optimized for responsive web environments, so that content renders well on a wide range of devices.

JPEG XL offers significantly better image quality and compression ratios compared to legacy JPEG. Compared to contemporary alternatives (in particular, video-based image formats like WebP, HEIC and AVIF), JPEG XL has a shorter specification and allows a simpler implementation, while still offering similar or better compression performance. It is designed for computationally efficient encoding and decoding using software implementations without the need for additional hardware acceleration, even on mobile devices.

Moreover, JPEG XL includes several features that help transition from the legacy JPEG coding format. Existing JPEG files can be losslessly transcoded to JPEG XL files, significantly reducing their size. These can be reconstructed to the exact same JPEG file, ensuring backward compatibility with legacy applications. This provides a smooth transition path from legacy JPEG platforms to the modern JPEG XL.

Parts 1 (core coding system) and 2 (file format) were published in early 2022; parts 3 (conformance testing) and 4 (reference implementation) are planned for publication at the end of 2022. Work has started on a second edition of Parts 1 and 2. The reference software libjxl is available royalty free under a permissive free and open-source license (BSD-3). JPEG XL is already supported in many popular applications and libraries including ImageMagick, FFmpeg, libvips, Imlib2, Qt, GDK, Gimp, Krita, and there is preliminary browser support (behind a configuration flag) in Chrome, Firefox, Edge, and Opera.

JPEG AI (ISO/IEC 6048)

The scope of JPEG AI is the creation of a learning-based image coding standard offering a single-stream, compact compressed domain representation, targeting both human visualization, with significant compression efficiency improvement over image coding standards in common use at equivalent subjective quality, and effective performance for image processing and computer vision tasks, with the goal of supporting a royalty-free baseline.

JPEG AI targets a wide range of applications such as cloud storage, visual surveillance, autonomous vehicles and devices, image collection storage and management, live monitoring of visual data and media distribution. The objective is to design a coding solution with significant compression efficiency improvements over coding standards in common use at equivalent subjective quality as well as an effective compressed domain processing for machine learning-based image processing and computer vision tasks. Other key requirements include hardware/software implementation-friendly encoding and decoding, support for 8- and 10-bit depth, efficient coding of images with text and graphics, and progressive decoding.

JPEG AI learning-based image coding solution framework serves a triple purpose: decoding for human visualization, image processing and computer vision analysis. First, as usual, the encoder receives a digital image and outputs a bitstream which can be used to obtain standard decoded images for human visualization (by performing entropy decoding and standard reconstruction). However, standard reconstruction can be skipped since the latent representation produced by the encoder contains the necessary information to perform image processing and computer vision tasks at the decoder side (after entropy decoding). This offers compelling advantages for applications where an image needs to be enhanced or modified before human consumption or where semantic (or higher-level) information needs to be extracted. Some examples of image processing tasks are super-resolution, denoising and exposure correction while relevant computer vision tasks are image classification, object detection and identification, and face recognition.

Following the release of the JPEG AI joint ISO/IEC and ITU-T Call for Proposals at the 94th JPEG meeting, twelve deep-learning based image codecs were submitted for the standard reconstruction task. For computer vision and image processing tasks, several compressed domain decoders were submitted, notably six for image classification. The ongoing work focuses on the evaluation of the Call for Proposals submissions, which is expected to be finalized at the 96th JPEG meeting, July 2022. A first Working Draft is expected to be produced by October 2022 and an International Standard to be available in 2024.

Exploration on DNA-based storage

The scope of JPEG DNA is the creation of a standard for efficient coding of images that considers biochemical constraints and offers robustness to noise introduced by the different stages of the storage process that is based on DNA synthetic polymers.

DNA is a chain of macromolecules present in life forms made of simple units called nucleotides which line up in a particular order. There are basically four types of units in DNA, referred to as A, C, G and T. The order of the

units carries the genetic information for a specific living organism, similar to how the order of letters in a text carries information about its content. With advances in technology, DNA molecules can now be sequenced to read the information that they carry, and also synthesized artificially to produce new information. This opens doors to the use of DNA as a means to store and retrieve all sorts of information. In digital media, notably images, the relevant representation symbols, e.g. quantized DCT coefficients, are expressed in bits (binary units) but they could be expressed in any other form as well. This includes, in DNA units which follow a quaternary (4-ary) representation basis. As a result, artificial DNA molecules may be created with a specific DNA configuration for storage of media including images.

Use of DNA as a storage mechanism brings several added values, such as archival at an extremely high density while being energy-friendly as well as durable. In this context, as the cost of sequencing and synthesis of DNA becomes affordable, storage of all sorts of information also becomes feasible in many applications ranging from long term preservation to content tracing.

As several JPEG standards are widely used in archival, notably the legacy JPEG and JPEG 2000, the JPEG Committee started an exploration to set the groundwork for creation of solutions to efficiently code images represented in DNA representation, while coping with various constraints in such an approach. After the organization of five workshops where stakeholders presented various technologies, applications, challenges and solutions, the Committee produced an overview document that not only motivates creation of such a standard in terms of potential applications but also has identified their requirements, and presents a survey of solutions under study in the state of the art in DNA storage. Furthermore, an experimentation software for image coding in DNA code has been produced by the JPEG Committee and made available to experts in order to carry out experimentations so as to better identify advantages, current challenges and potential solutions for image coding in DNA.

WG1 is on its way to launch a Draft Call for Proposals in October 2022, with a final call in January 2023 with the aim of publishing a JPEG image coding standard for DNA storage in 2025.

Exploration on JPEG Fake Media

Recent advances in media creation and modification technologies allow the production of near realistic media assets that are often, to the human eye, indistinguishable from original assets. These developments open opportunities for creative production of new media in the entertainment and art industry. However, the intentional or unintentional spread of manipulated media, such as modified media with the intention to induce misinterpretation, also imposes risks such as social unrest, spread of rumors for political gain or encouraging hate crimes.

Clear and transparent annotation of media asset creation and modifications is a crucial element in many usage scenarios bringing trust to the users. This has already triggered various organizations to develop a wide range of mechanisms that can detect and/or annotate modified media assets when they are shared. However, these annotations should be attached to the media in a secure way to deter them from being altered. In addition, to achieve a wide adoption of such an annotation ecosystem, interoperability is essential.

Hence, the scope of JPEG Fake Media is the creation of a standard that can facilitate a secure and reliable annotation of media asset creation and modifications. The standard shall address use cases that are in good faith as well as those with malicious intent.

At the 95th JPEG meeting in April 2022, the Committee issued a Final Call for Proposals (CfP) on JPEG Fake Media. The call for proposals seeks for contributions that address at least one of the extensive list of requirements specified in the associated "Use Cases and Requirements for JPEG Fake Media" document. The deadline for proposal submissions is the 19th of October 2022. According to the current timeline, the resulting International Standard is expected to be published in 2025.

Exploration on JPEG NFT

Non-Fungible Tokens (NFTs) have garnered considerable interest and debate in terms of usefulness and potential applications. Several digital assets that NFTs point to are either in existing JPEG formats or can be represented in current and emerging formats under development by the JPEG Committee. In particular, various trust and security concerns have been raised about NFTs and the digital assets on which they rely. To better understand user requirements for media formats in NFT, the JPEG Committee launched the JPEG NFT

exploration initiative. The scope of JPEG NFT is the creation of effective specifications that support a wide range of applications relying on NFTs applied to media assets. The standard shall be secure, trustworthy and environmentally aware, allowing for an interoperable ecosystem relying on NFT within a single application or across applications. The JPEG Committee strives to engage stakeholders from diverse backgrounds, including technical, legal, artistic, and end-user, to establish use cases and requirements.

It is the aim of this exploration to investigate a Call for Proposals to be issued in the near future when the needs and challenges in terms of imaging are better identified.

3.3.2 WG 1 current work items

WG 1 is progressing the development of the following standards.

- ISO/IEC AWI 6048 JPEG AI Learning-based Image Coding System
- ISO/IEC AWI 10918-4 Information technology — Digital compression and coding of continuous-tone still images— Part 4: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF colour spaces, APPn markers, SPIFF compression types and Registration Authorities (REGAUT)
- ISO/IEC AWI 10918-7 Information technology — Digital compression and coding of continuous-tone still images — Part 7: Reference software
- ISO/IEC DIS 15444-2 Information technology — JPEG 2000 image coding system — Part 2: Extensions
- ISO/IEC AWI 15444-4 Information technology — JPEG 2000 image coding system — Part 4: Conformance Testing
- ISO/IEC DIS 15444-8 Information technology — JPEG 2000 image coding system — Part 8: Secure JPEG 2000
- ISO/IEC DIS 15444-9 Information technology — JPEG 2000 image coding system — Part 9: Interactivity tools, APIs and protocols
- ISO/IEC FDIS 15444-17 Information technology — JPEG 2000 image coding system — Part 17: Extensions for coding of discontinuous media
- ISO/IEC WD 18181-1 Information technology — JPEG XL Image Coding System — Part 1: Core coding system
- ISO/IEC WD 18181-2 Information technology — JPEG XL image coding system — Part 2: File format
- ISO/IEC AWI 18477-3 Information technology — Scalable compression and coding of continuous-tone still images — Part 3: Box file format
- ISO/IEC CD 19566-5 Information technologies — JPEG systems — Part 5: JPEG universal metadata box format (JUMBF)
- ISO/IEC FDIS 19566-7 Information technologies — JPEG systems — Part 7: JPEG linked media format (JLINK)
- ISO/IEC DIS 19566-8 Information technologies — JPEG systems — Part 8: JPEG Snack
- ISO/IEC AWI TR 19566-9 Information technology — JPEG Systems — Part 9: JPEG extensions mechanisms to facilitate forwards and backwards compatibility
- ISO/IEC AWI 19566-10 Information technologies — JPEG systems — Part 10: Reference Software
- ISO/IEC AWI 21122-1 Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system
- ISO/IEC AWI 21122-2 Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models
- ISO/IEC FDIS 21122-3 Information technology — JPEG XS low-latency lightweight image coding system — Part 3: Transport and container formats

- ISO/IEC CD 21122-4 Information technology — JPEG XS low-latency lightweight image coding system — Part 4: Conformance testing
- ISO/IEC CD 21122-5 Information technology — JPEG XS low-latency lightweight image coding system — Part 5: Reference software
- ISO?IEC AWI 21122-1 Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system
- ISO/IEC AWI 21122-2 Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models
- IISO/IEC 21122-2:2022/DAmd 1 Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models — Amendment 1: Profile and sublevel for 4:2:0 content
- ISO/IEC AWI 21794-5 Information technology — Plenoptic image coding system (JPEG Pleno) — Part 5: Holography
- ISO/IEC AWI 21794-6 Information technology — Plenoptic image coding system (JPEG Pleno) — Learning-based Point Cloud Coding

3.3.3 WG 1 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 1:

- ISO.IEC 10918-7:2021 Information technology — Digital compression and coding of continuous-tone still images — Part 7: Reference software
- ISO/IEC 15444-2:2021 Information technology — JPEG 2000 image coding system — Part 2: Extensions
- ISO/IEC 15444-4:2021 Information technology — JPEG 2000 image coding system — Part 4: Conformance Testing
- ISO/IEC 15444-5:2021 Information technology — JPEG 2000 image coding system — Part 5: Reference software
- ISO/IEC 15444-16:2021 Information technology — JPEG 2000 image coding system — Part 16: Encapsulation of JPEG 2000 images into ISO/IEC 23008-12
- ISO/IEC 18181-1:2022 Information technology — JPEG XL Image Coding System — Part 1: Core coding system
- ISO/IEC 18181-1:2022/DAmd 1:2022 Information technology — JPEG XL Image Coding System — Part 1: Core coding system — Amendment 1: Profiles and levels for JPEG XL image coding system
- ISO/IEC 18181-2:2021 Information technology — JPEG XL image coding system — Part 2: File format
- ISO/IEC 18181-3:2022 Information technology — JPEG XL Image Coding System — Part 3: Conformance testing
- ISO/IEC 18181-4:2022 Information technology — JPEG XL Image Coding System — Part 4: Reference software
- ISO/IEC 21122-1:2022 Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system
- ISO/IEC 21122-2:2022 Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models
- ISO/IEC 21122-3:2022 Information technology — JPEG XS low-latency lightweight image coding system — Part 3: Transport and container formats
- ISO/IEC 21122-4:2022 Information technology — JPEG XS low-latency lightweight image coding system — Part 4: Conformance testing

- ISO/IEC 21122-5:2022 Information technology — JPEG XS low-latency lightweight image coding system — Part 5: Reference software
- ISO/IEC 21794-3:2021 Information technology — Plenoptic image coding system (JPEG Pleno) — Part 3: Conformance testing
- ISO/IEC 21794-4:2022 Information technology — Plenoptic image coding system (JPEG Pleno) — Part 4: Reference software

3.4 WG 2 MPEG Technical Requirements

3.4.1 WG 2 accomplishments

WG 2 Terms of Reference

WG 2 conducts requirements studies for SC 29 MPEG standards with the purpose of ensuring that requirements coming from prospective application areas of generic coding algorithms developed by SC 29 WG2 to WG8 are taken into consideration in all phases of work.

Call for Proposals on Dynamic Mesh Coding

At the 136th MPEG meeting in October 2021, *MPEG Technical Requirements* jointly with *MPEG Coding of 3D Graphics* issued a Call for Proposals (CfP) on Dynamic Mesh Coding. A dynamic mesh sequence may require a large amount of data since it may consist of a significant amount of information changing in time. Therefore, efficient coding technologies are required to store and transmit such content. MPEG had previously developed mesh coding standards; however, these standards do not consider time-varying attribute maps and connectivity information, which are considered in the new work.

Evaluation of Responses to the Haptics Phase 1 Call for Proposals

At the 134th MPEG meeting, *MPEG Technical Requirements* issued a Call for Proposals (CfP) on the Coded Representation of Haptics. This CfP sought submissions of technologies that could provide efficient representation and compression of time-dependent haptic signals and were suitable for coding of timed haptic tracks that may be synchronized with audio and/or video media. This CfP was an essential part of Phase 1 of the multi-year MPEG roadmap for haptics.

At the 136th MPEG meeting, *MPEG Technical Requirements* concluded its evaluation of four responses to the Haptics CfP. The technologies evaluated included descriptive, human-readable representations, highly efficient psychophysical compression schemes, and support for both vibrotactile and kinesthetic devices. Based on the evaluation of these responses, a subset of the technologies has been selected to form a combined codec (termed RM0) that leverages the best of these technologies and ensures long term relevance for future use cases, including Extended Reality (XR). One of the technologies selected had excellent scores using the subjective MUSHRA tests; a second one had the best PSNR scores and is based on optimized perceptual wavelet coding. A third included technology provides a lightweight JSON framework compatible with the Khronos glTF framework, which is already extended by the MPEG-I scene description standardization efforts, in which haptics is expected to play a significant role.

After the evaluation of the responses, MPEG will move into the collaborative phase of this effort where its members will work together to improve and fine-tune the combined RM0 technology over the next several months. It was expected that phase 1 of Haptics would reach the first milestone of standardization (*i.e.*, issuing a Committee Draft) by the end of 2022.

Call for Proposals for Encoder and Packager Synchronization

At the 137th MPEG meeting, *MPEG Technical Requirements (WG 2)* issued a Call for Proposals (CfP) for technologies enabling encoder and packager synchronization and related distributed media asset storage.

The encoder and packager synchronization framework will define preferred ways of generating content from distributed sources based on existing MPEG standards such as Common Media Application Format (CMAF), MPEG Dynamic Adaptive Streaming over HTTP (DASH), and potentially other standards and specifications. This will enable redundant live and (Video on Demand) VoD content generation setups producing synchronized content that is robust to failures and loss of input in one or more of the components of the setup. Additionally,

solutions for storing media assets at scale are solicited. The asset storage solution will enable live-to-VoD and VoD-to-live use cases and take advantage of the synchronized encoder/packager framework.

This CfP requested proposals from companies and other organizations. Registration was required by 10 April 2022 and the submission by 17 April 2022. Evaluation of the submissions in response to the CfP was performed at the 138th MPEG meeting in April 2022.

Call for Proposals for Video Coding for Machines

At the 138th MPEG meeting, *MPEG Technical Requirements* (WG 2) issued a Call for Proposals (CfP) for technologies and solutions enabling efficient video coding for machine vision tasks.

This work on "Video Coding for Machines" aims at compressing input video and images or feature maps for machine tasks. As machines consume and understand visual data differently to human vision, coding technologies and solutions could be different from conventional ones, even for coding video and images, in order to achieve optimized performance for machine usage. With the rise of machine learning technologies and machine vision applications, the amount of video and images consumed by machines has been rapidly growing. Typical use cases include intelligent transportation, smart city, intelligent content management, etc., which incorporate machine vision tasks such as object detection, instance segmentation, and object tracking. Due to the large volume of video data, it is essential to compress video for efficient transmission, analysis and storage. Besides the highly demanded compression benefit, VCM can also be helpful in some other regards, such as computational offloading and privacy protection.

Over the last three years, MPEG has investigated potential technologies for efficient compression of visual data for machine vision tasks and established an evaluation mechanism that includes common test conditions (CTC), rate-distortion based metrics, and evaluation pipelines. In addition to existing image datasets, some new datasets, and video datasets in particular, have been built and donated to MPEG which are valuable for future research and standardization.

Evaluation of the submissions in response to the CfP will be performed at the 140th MPEG meeting in October 2022.

Call for Evidence for Video Coding for Machines

At the 139th MPEG meeting of 18-22 July 2022, *MPEG Technical Requirements* (WG 2) issued a Call for Evidence (CfE) for technologies and solutions enabling efficient feature coding for machine vision tasks.

MPEG's exploration work on "Video Coding for Machines" aims at compressing features for machine-performed tasks such as video object detection and event analysis. As neural networks increase in complexity, architectures such as 'Collaborative Intelligence', whereby a network is distributed across an edge device and the cloud, become advantageous. With the rise of newer network architectures being deployed amongst a heterogenous population of edge devices, such architectures bring flexibility to systems implementers. Due to such architectures, there is a need to efficiently compress intermediate feature information for transport over wide area networks (WANs). As feature information differs substantially from conventional image or video data, coding technologies and solutions for machine usage could differ from conventional human-viewing-oriented applications to achieve optimized performance. With the rise of machine learning technologies and machine vision applications, the amount of video and images consumed by machines has rapidly grown. Typical use cases include intelligent transportation, smart city technology, intelligent content management, etc., which incorporate machine vision tasks such as object detection, instance segmentation, and object tracking. Due to the large volume of video data, extracting and compressing the feature from a video is essential for efficient transmission and storage. Feature compression technology solicited in this CfE can also be helpful in other regards, such as computational offloading and privacy protection.

Over the last three years, MPEG has investigated potential technologies for efficiently compressing feature data for machine vision tasks and established an evaluation mechanism that includes feature anchors, rate-distortion-based metrics, and evaluation pipelines.

Discussion of the submissions in response to the CfE was planned for the 140th MPEG meeting in October 2022.

At the 139th MPEG meeting, MPEG Technical Requirements (WG 2) also issued an updated version of its previously-issued Call for Proposals (CfP) on Video Coding for Machines. The changes include a clarification that responses to this call need to support the coding of video data for machine tasks and include a questionnaire that summarizes the requirements each proposal fulfills. Furthermore, the registration deadline was extended.

3.4.2 WG 2 current work items

WG 2 is progressing the development of the following standards.

- ISO/IEC CD TR 23090-1 Information technology - Coded representation of immersive media — Part 1: Immersive media
- ISO/IEC WD TR 23090-27 Information technology - Coded representation of immersive media — Part 27: Media, renderers, and game engines for render- based systems and applications.

3.5 WG 3 MPEG Systems

3.5.1 WG 3 accomplishments

WG 3 Terms of Reference

WG 3 develops coded representation of encapsulation formats, delivery protocols, multimedia presentation information; and schemes for declaring, and description of multimedia content and related information.

Promotion of Third Edition of Energy-Efficient Media Consumption (Green Metadata) to Committee Draft

At the 136th MPEG meeting, *MPEG Systems* promoted the third edition of ISO/IEC 23001-11 Energy-Efficient Media Consumption (Green Metadata) to Committee Draft (CD) stage. ISO/IEC 23001-11 specifies metadata (Green Metadata) that facilitates the reduction of energy usage during media consumption. The metadata enable means for *(a)* reducing decoder power consumption, *(b)* reducing display power consumption, *(c)* media selection for the joint decoder and display power reduction, and *(d)* quality recovery after low-power encoding. The third edition adds three main aspects to the standard: *(1)* the existing green metadata for interactive signalling for remote decoder power reduction is enhanced by adding new syntax elements for a finer control of decoder complexity on a high variety of platforms and processor cores; *(2)* a Versatile Video Coding (VVC) Supplemental Enhancement Information (SEI) message carrying green metadata related to complexity metrics for decoder power reduction; and *(3)* metrics for quality recovery after low-power encoding. The development of the standard is expected to be completed by the end of 2022.

FDIS for Event Messages in Tracks of the ISO Base Media File Format

At the 136th MPEG meeting, *MPEG Systems* has reached the last milestone of the development of a standard to carry event messages in tracks of the ISO Base Media File Format (ISOBMFF) by promoting ISO/IEC 23001-18 to Final Draft International Standard (FDIS) stage.

This specification provides a method for the carriage of event messages in tracks of the ISO Base Media File Format (ISO/IEC 14496-12). This event message track format associates the timeline of the DashEventMessageBox messages to the track timeline. The specified track format enables common ISOBMFF processing such as multiplexing and de-fragmentation. In addition, multiplexing and de-multiplexing operations using top-level DashEventMessageBox based on this event message track format are defined. The carriage in the event message track format will also make this information more easily accessible to devices that can seek through ISOBMFF formatted media files.

Promotion of MPEG-I Scene Description to the Final Stage

At the 137th MPEG meeting, *MPEG Systems (WG 3)* completed the MPEG-I Scene Description standard, a key technology in enabling immersive 3D user experiences, by promoting it to the final approval stage as a Final Draft International Standard (FDIS). The specification describes the composition of a 3D scene by referencing and positioning different media assets in the scene. The information provided in the scene description is then used by an application to render the 3D scene accordingly. To address the needs of immersive applications, the specification has developed MPEG extensions for Khronos glTF 2.0, a scene description solution widely used by the industry. glTF 2.0 provides a solid and efficient baseline for exchangeable and interoperable scene

descriptions and can enable realistic rendering of immersive content such as by using Physically Based Rendering (PBR). However, glTF 2.0 has primarily been designed for static scenes and assets, which does not fully address the requirements and needs of dynamic and rich 3D scenes in immersive environments.

Based on this analysis, MPEG has developed extensions to Khronos's glTF 2.0 to integrate real-time media, *i.e.*, support for dynamic visual objects, audio, timed updates of the scene, and media access related functions. The standard also defines an architecture that decouples media rendering from media access by specifying the relevant APIs to access media referenced by the scene description. This work has continuously been coordinated with Khronos and 3GPP. Further details can be found at http://mpeg-sd.org/.

Promotion of Smart Contracts for Media to the Final Stage

In the last few years, MPEG has developed a set of standardized Resource Description Framework (RDF) ontologies and XML schemas for the codification of intellectual property (IP) rights information related to music and media. The ISO/IEC 21000-19 Media Value Chain Ontology (MVCO) facilitates rights tracking for fair, timely, and transparent marketplace transactions by capturing user roles and their permissible actions on a particular IP entity. The ISO/IEC 21000-21 Media Contract Ontology (MCO) facilitates the conversion of narrative contracts to digital ones related to the management of IP rights, payments, and notifications. With respect to the latter, XML schemas have been developed as the ISO/IEC 21000-20 Contract Expression Language (CEL).

At the 137th MPEG meeting, *MPEG Systems (WG 3)* completed the development of ISO/IEC 21000-23 Smart Contracts for Media by promoting the standard to the Final Draft International Standard (FDIS) stage, which is the final approval milestone in the development of a standard. The standard specifies the means (*e.g.*, application programming interfaces) for converting the above-mentioned RDF ontologies and XML schemas to smart contracts that can be executed on existing Distributed Ledger Technology (DLT) environments. This important standard will greatly assist the music and media industry and its stakeholders in achieving effective interoperability for the exchange of verified contractual data between different DLT environments. In this way, it will increase trust among the stakeholders for sharing high-value data (*e.g.*, music rights) in the ecosystem. Another important feature of this standard is that it offers the possibility to bind, through persistent links, the clauses of a smart contract to their corresponding ones of a human-readable contract. In this way, each party signing an ISO/IEC 21000-23 conforming smart contract will be able to know exactly what its clauses express.

Further enhancement of the ISOBMFF Standard

At the 137th MPEG meeting, *MPEG Systems (WG 3)* reached the first formal milestone in the approval process for an amendment of its recently Emmy Award-winning standard ISO/IEC 14496-12 ISO Base Media File Format (ISOBMFF) comprising improved brand documentation and other improvements. The Committee Draft Amendment (CDAM) of ISO/IEC 14496-12:2021 Amendment 1 includes the enhancement of random access of the media stored in ISOBMFF by using an external elementary stream providing inter-frame prediction references for the decoding of an elementary stream that is stored as the main stream, supporting so-called extended dependent random access points (EDRAPs). Additionally, the amendment provides a way to list the sets of (one or multiple) media components representing one version of the media presentation that may be selected by a user for simultaneous decoding and presentation. The amendment also clarifies the use of major brands and compatible brands in the file type box to provide guidelines for defining a new brand identifier for ISOBMFF. This amendment is expected to reach its final approval milestone as a Final Draft Amendment (FDAM) in early 2023.

Acceptance of two Emmy® Awards

As discussed above in section 2.1, in its 73rd annual ceremony on April 26, 2022, the National Academy of Television Arts & Sciences (NATAS) awarded ISO/IEC MPEG Systems Working Group (WG 3) with two Technology & Engineering Emmy® Awards for its MPEG-DASH and Open Font Format standards.

Since 1948, the Technology & Engineering Emmy® Awards honor development and innovation in TV-related technology and recognize companies, organizations, and individuals for breakthroughs in this field.

This year, MPEG Systems received a Technology & Engineering Emmy® for its work on "Standardization of HTTP Encapsulated Protocols". 3GPP, which partnered with MPEG on the collaborative development of the MPEG-DASH standard, was also recognized with an Emmy.

MPEG Systems started the project on HTTP streaming in 2009. MPEG published the first edition of the Dynamic Adaptive Streaming over HTTP (DASH) standard (ISO/IEC 23009) in 2012, followed by three more editions since then. The 5th edition of the MPEG-DASH standard is expected to be published in 2022.

The MPEG-DASH standard defines a delivery format for streaming multimedia content with the highest quality possible over networks with variable bandwidth using CDNs. Its features support many services and applications such as on-demand video, live streaming, low latency streaming, and targeted ad insertion. MPEG-DASH is the first true open international standard for video streaming over the internet that enabled multi-vendor interoperable solutions and has been widely adopted by industry and various consortia.

The standard was developed with the strong participation of experts from around the world. More than 90 experts from over 60 companies participated in the development of the state-of-art technology defined by the standard for more than a decade. The following experts, in particular, are recognized by MPEG for contributing to the development of the standards in various ways: Ali C. Begen, Romain Bouqueau, Imed Bouzazi, Zachary Cava, Mary-Luc Champel, Cyril Concolato, Igor Curcio, Franck Denoual, Mike Dolan, Wang Fang Gerard Fernando, Per Fröjdh, Alexander Giladi, Jeff Goldberg, Miska Hannuksela, Mitsuhiro Hirabayashi, Paul Higgs, Ingo Hoffmann, Kilroy Hughes, Will Law, Jin Young Lee, Jean Le Feuvre, Brendan Long, Sylvain Kervadec, Yongliang Liu, Brenden Long, Shivakumar Mahadevappa, Frederic Maze, Nhut Nguyen, Harry Pile (late), Yuriy Reznik, Sungryeul Rhyu, Yago Sanchez, Thomas Schierl, Iraj Sodagar, Thomas Stockhammer, Kevin Streeter, Yasser Syed, Viswanathan (Vishy) Swaminathan, Emmanuel Thomas, Christian Timmerer, Ye Xiaoyang, Xin Wang, Ye-Kui Wang, Mark Watson, Yang Yanzi, Shaobo Zhang, Zhijie Zhao and Waqar Zia. Many companies and institutions participated by providing experts and resources for the development of the DASH standard, including Adobe, Akamai, Bitmovin, Brightcove, Bytedance, CableLabs, Canon, Cisco, Comcast, Ericsson, ETRI, Fraunhofer, Google, Huawei, Hulu, InterDigital Communications, LG Electronics, KPN, Microsoft, Netflix, Orange, Nokia, Ozyegin University, Panasonic, Qualcomm, Samsung, Sony, Telecom Paris, Tencent America, Xiaomi, Leibniz University Hannover, University of Klagenfurt, Vubiquity, and ZTE Corporation.

The chair of the MPEG-DASH subgroup, Iraj Sodagar said, “The success of the DASH standard shows that when the industry as a whole participates in a standardization project under a reputable SDO such as ISO/IEC JTC 1, the resulting standard can make a direct impact on the entire industry by enabling interoperability between different parts of the ecosystem.”

The second Technology & Engineering Emmy® Award presented to MPEG Systems this year is for the development of the ISO/IEC 14496-22 “Open Font Format”, with its 4th edition published in 2019 and work currently underway toward 5th edition. The Emmy was presented for “Standardization of Font Technology for Custom Downloadable Fonts and Typography for Web and TV Devices”.

Font technology standardization work as part of the MPEG-4 family of standards was initiated in 2004 with the publication of the ISO/IEC 14496-18 “Font compression and streaming”. This work continued with the development of the new part 22 (ISO/IEC 14496-22), “Open Font Format”, which was based on the contribution of OpenType technology by Microsoft and Adobe in 2004. Since then, the standard has been revised and extended many times, with the participation of a wide community of experts who offer combined expertise in typography, linguistics, font design, and computer science. Experts contributing to this work include representatives of a number of large companies, including Adobe, Apple, Google, Microsoft, Monotype, and many individual contributors and subject matter experts including John Hudson, Laurence Penney, and Adam Twardoch (among many others) who made valuable contributions. Vladimir Levantovsky (Type Standards LLC) has served as a chair of the Font subgroup and the Project Editor for both standards.

“Fonts are the critical components of any written communication. Text carries a meaning, but it’s a font that makes text readable – fonts give the written word a voice! Standardization of the Open Font Format technology by ISO/IEC JTC 1’s MPEG Systems Working Group (SC 29/WG 3) significantly influenced the capabilities of all classes of consumer electronic devices bringing advanced font technology for digital TV, streaming media environments, and the Web. It also inspired many open-source projects that enabled mass adoption of high-quality font rendering and advanced text support, making it easy and cost-effective for OEMs, service providers, and content authors to deploy new features and applications supporting all world languages and writing systems,” the chair of the Fonts subgroup, Vladimir Levantovsky, said.

Selection of Technologies for Encoder and Packager Synchronization and Asset Storage

At the 137th MPEG meeting of January 2022, *MPEG Technical Requirements* (WG 2) issued a Call for Proposals (CfP) on technologies for encoder and packager synchronization and asset storage. At the 138th meeting, MPEG concluded the evaluation of the responses to the CfP. The CfP responses fulfilled and supplemented the CfP requirements, with one proposal dealing with encoder/packager synchronization and a second proposal dealing with asset storage and recording. The combination of the responses already covers 85 percent of the expressed synchronization requirements and 75 percent of the storage requirements. Furthermore, the proposals are well aligned with existing MPEG specifications, in particular MPEG-DASH and CMAF. As a result, MPEG Systems (WG 3) produced a draft text of a Working Draft at the 138th meeting for the upcoming specification. The final standard is expected to be completed at the beginning of 2023.

This standard will greatly assist the industry in achieving effective interoperability for the production and storage of 24×7 live content. It meets specific requirements for 24×7 live media production and distribution in cloud-based workflows, such as the use of object-based cloud storage. The proposal addresses the use cases of redundant encoder synchronization for failover handling, distributed encoding of (multi-codec) bit-rate ladders, and A/B watermarking. As the proposals are based on MPEG-DASH and CMAF, this standard also introduces best practices for using these MPEG technologies for the key use cases targeted by broadcasters and content owners.

Completion of CMAF support of EVC and VVC

As a part of efforts to integrate recently developed video coding standards into storage and delivery standards, at the 138th MPEG meeting, MPEG Systems (WG 3) completed the standardization process for carriage of Versatile Video Coding (VVC) and Essential Video Coding (EVC) by the Common Media Application Format (CMAF).

The third edition of ISO/IEC 23000-19 *Common media application format (CMAF) for segmented media* has reached the final milestone of standard development, Final Draft International Standard (FDIS) with support for these new video coding technologies. The specification defines CMAF track and media profiles for VVC and EVC. For both video codecs, constraints on elementary streams for use in CMAF and constraints on permitted values of some fields of parameter sets are specified. Additionally, restrictions on the usage of Supplemental Enhancement Information (SEI) and Video Usability Information (VUI) are also specified. The method to indicate the conformance point of the video codecs, such as profile, level, toolset indication, etc., are also defined by this specification. Concerning VVC, the specification supports the carriage of both single- and multi-layer bitstreams.

As CMAF has already been adopted by other Standards Development Organizations (SDOs) and many of them have been anticipating the integration of new video codecs into CMAF with various application requirements and constraints, and the development process has been conducted with active collaboration with them. Many valuable inputs have been received, carefully studied, and reflected as much as possible.

Completion of Third Edition of Green Metadata, a Standard for Energy-Efficient Media Consumption

At the 139th MPEG meeting, *MPEG Systems* (WG 3) issued the Final Draft International Standard (FDIS) of the third edition of ISO/IEC 23001-11 *Energy-Efficient Media Consumption (Green Metadata)*. FDIS is the final milestone of standard development.

MPEG Systems has been working on Green Metadata for the last 10 years to enable the adaptation of the client's power consumption according to the complexity of the bitstream. Many modern implementations of video decoders can adjust their operating voltage or clock speed to adjust the power consumption level according to the required computational power. Thus, if the decoder implementation knows the variation in the complexity of the incoming bitstream, then the decoder can adjust its power consumption level to the complexity of the bitstream. This will allow less energy use in general and extended video playback for the battery-powered devices.

The third edition enables support for *Versatile Video Coding* (VVC, ISO/IEC 23090-3, a.k.a. ITU-T H.266) encoded bitstreams and enhances the capability of this standard for real-time communication applications and services. While finalizing the support of VVC, MPEG Systems has also started the development of a new

amendment to the Green Metadata standard adding the support of *Essential Video Coding* (EVC, ISO/IEC 23094-1) encoded bitstreams.

Completion of Third Edition of the Common Media Application Format by adding Support for 8K and High Frame Rate for High Efficiency Video Coding

At the 139th MPEG meeting, *MPEG Systems* (WG 3) issued the Final Draft International Standards (FDIS) of the third edition of the ISO/IEC 23000-19 *Common Media Application Format* (CMAF) for segmented media. FDIS is the final milestone of standard development.

The third edition of CMAF adds two new media profiles for *High Efficiency Video Coding* (HEVC, ISO/IEC 23008-2, a.k.a. ITU-T H.265), namely for *(i)* 8K and *(ii)* High Frame Rate (HFR). Regarding the former, the media profile supporting 8K resolution video encoded with HEVC (Main 10 profile, Main Tier with 10 bits per colour component) has been added to the list of CMAF media profiles for HEVC. The profile will be branded as 'c8k0' and will support video with up to 7680×4320 pixels (8K) and up to 60 frames per second. Regarding the latter, another media profile has been added to the list of CMAF media profiles, branded as 'c8k1' and supports HEVC encoded video with up to 8K resolution and up to 120 frames per second.

Finally, chroma location indication support has been added to the 3rd edition of CMAF.

MPEG Scene Descriptions adds Support for Immersive Media Codecs

At the 139th MPEG meeting, *MPEG Systems* (WG 3) issued the Committee Draft Amendment (CDAM) for ISO/IEC 23090-14 *Scene Description* support for immersive media codecs. CDAM is the first formal milestone of the amendment development approval process.

ISO/IEC 23090-14 specifies the coded representation of scene descriptions, and this new amendment facilitates the integration of *Video-based Point Cloud Compression* (V-PCC, specified in ISO/IEC 23090-5) and *MPEG Immersive Video* (MIV, ISO/IEC 23090-12) into a scene. Immersive media codecs such as V-PCC and MIV encode immersive media data with multiple conventional 2D video codecs. The decoded bitstreams from each video decoder are then combined to reconstruct the immersive media presentation. The scene description of ISO/IEC 23090-14 is based on the GL Transmission Format (glTF), which lacks support for integrating multiple video bitstreams into a single 3D object. This amendment enables the association of multiple buffers with a single mesh object. Additionally, it enables these buffers to deliver decoded components of V-PCC or MIV objects to the rendering engine so that the rendering engine can reconstruct the 3D point cloud or immersive video content. As the components of Visual Volumetric Video-based Coded (V3C) objects are encoded as 2D video, and support for YCbCr formats is also being added. The amendment is planned to be completed, i.e., to reach the status of Final Draft Amendment (FDAM), by the end of 2023.

3.5.2 WG 3 current work items

WG 3 is progressing the development of the following standards.

- ISO/IEC 13818-1/FDAmd 1 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems — Amendment 1: Carriage of LCEVC in MPEG-2 TS and other improvements
- ISO/IEC 13818-1/CD Cor 1 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems — Technical Corrigendum 1
- ISO/IEC WD 14496-1 Information technology — Coding of audio-visual objects — Part 1: Systems
- ISO/IEC 14496-12:2022/CD Amd 1 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format — Amendment 1: Improved brand documentation and other improvements
- ISO/IEC DIS 14496-12 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format
- ISO/IEC FDIS 14496-15 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

- ISO/IEC FDIS 14496-15/FDAmd 1 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 — Amendment 1: Support for LCEVC
- ISO/IEC FDIS 14496-15/CDAmd 2 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 — Amendment 2: Picture-in-picture support and other extensions
- ISO/IEC 14496-22:2019/FDAmd 2 Information technology — Coding of audio-visual objects — Part 22: Open Font Format — Amendment 2: Extending colour font functionality and other updates
- ISO/IEC WD 14496-34 Information technology — Coding of audio-visual objects — Part 34: Syntactic description language
- ISO/IEC FDIS 21000-23 Information technology — Multimedia framework (MPEG-21) — Part 23: Smart Contracts for Media
- ISO/IEC FDIS 23000-19 Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media
- ISO/IEC FDIS 23000-19:2020/CDAmd 4 Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media — Amendment 4: LCEVC and other technologies
- ISO/IEC 23000-22/CD Amd 3 Information technology — Multimedia application format (MPEG-A) — Part 22: Multi-image application format (MIAF) - Amendment 3: Chroma Subsampling and others for MIAF
- ISO/IEC FDIS 23001-7 Information technology — MPEG systems technologies — Part 7: Common encryption in ISO base media file format files
- ISO/IEC FDIS 23001-11 MPEG systems technologies — Part 11: Energy-efficient media consumption (green metadata)
- ISO/IEC FDIS 23001-11/WD Amd 1 MPEG systems technologies — Part 11: Energy-efficient media consumption (green metadata) — Amendment 1: Energy-efficient media consumption (green metadata) for EVC
- ISO/IEC CD 23001-17 Information technology — MPEG systems technologies — Part 17: Carriage of Uncompressed video in the ISO base media file format
- ISO/IEC FDIS 23008-1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT)
- ISO/IEC DIS 23008-1/WD Amd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT) — Amendment 1: Improvement of MMT message transaction
- ISO/IEC 23008-4:2020/CD Amd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 4: MMT reference software — Amendment 2: Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 4: MMT reference and conformance software — Amendment 2: Support for MMTP extensions
- ISO/IEC FDIS 23008-12 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format
- ISO/IEC FDIS 23008-12/CDAmd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format— Amendment 1: Support for progressive rendering signalling and other improvements
- ISO/IEC DIS 23009-1/DAmd 1 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats — Amendment 1: Preroll, nonlinear playback and other extensions

- ISO/IEC DIS 23009-1/WD Amd 2 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats — Amendment 2: EDRAP streaming and other extensions
- ISO/IEC WD TR 23009-3 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 3: Implementation guidelines
- ISO/IEC WD TR 23009-7 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 7: Delivery of CMAF contents with DASH
- ISO/IEC FDIS 23009-8:2022/DAmd 1 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 8: Session-based DASH operations — Amendment 1: URL customization and other extensions
- ISO/IEC WD 23009-9 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 9: Encoder and packager synchronization
- ISO/IEC FDIS 23090-2 Information technology — Coded representation of immersive media — Part 2: Omnidirectional media format
- ISO/IEC 23090-6:2021/DAmd 1 Information technology — Coded representation of immersive media — Part 6: Immersive media metrics — Amendment 1: Immersive media metrics for V3C Data and OMAF
- ISO/IEC FDIS 23090-7 Information technology — Coded representation of immersive media — Part 7: Immersive media metadata
- ISO/IEC FDIS 23090-7/WD Amd 1 Information technology — Coded representation of immersive media — Part 7: Immersive media metadata — Amendment 1: Common metadata for immersive media
- ISO/IEC FDIS 23090-8 Information technology — Coded representation of immersive media — Part 8: Network based media processing
- ISO/IEC DIS 23090-10/FDAmd 1 Information technology — Coded representation of immersive media — Part 10: Carriage of visual volumetric video-based coding data — Amendment 1: Support of packed video data
- ISO/IEC DTR 23090-11 Information technology — Coded representation of immersive media — Part 11: Network-based media processing implementation guidelines
- ISO/IEC DIS 23090-13 Information technology — Coded representation of immersive media — Part 13: Video Decoding Interface for Immersive Media
- ISO/IEC FDIS 23090-14 Information technology — Coded representation of immersive media — Part 14: Scene Description for MPEG Media
- ISO/IEC FDIS 23090-14/CDAmd 1 Information technology — Coded representation of immersive media — Part 14: Scene Description for MPEG Media — Amendment 1: Support for immersive media codecs in scene description
- ISO/IEC FDIS 23090-18 Information technology — Coded representation of immersive media — Part 18: Carriage of geometry-based point cloud compression data (pending assignment)
- ISO/IEC FDIS 23090-18/DAmd 1 Information technology — Coded representation of immersive media — Part 18: Carriage of geometry-based point cloud compression data — Amendment 1: Support for temporal scalability
- ISO/IEC WD 23090-24 Information technology — Coded representation of immersive media — Part 24: Conformance and Reference Software for Scene Description for MPEG Media
- ISO/IEC WD 23090-25 Information technology — Coded representation of immersive media — Part 25: Conformance and Reference Software for Carriage of Visual Volumetric Video-based Coding Data
- ISO/IEC WD 23090-26 Information technology — Coded representation of immersive media — Part 26: Conformance and Reference Software for Carriage of Geometry-based Point Cloud Compression Data

3.5.3 WG 3 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 3:

- ISO/IEC 13818-1:2022 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems
- ISO/IEC 13818-1:2019/Cor 1:2022 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems — Technical Corrigendum 1
- ISO/IEC 14496-30:2018/Amd 1:2022 Information technology — Coding of audio-visual objects — Part 30: Timed text and other visual overlays in ISO base media file format — Amendment 1: Timing improvements
- ISO/IEC 21000-22:2022 Information technology — Multimedia framework (MPEG-21) — Part 22: User Description
- ISO/IEC 23000-22:2019/Amd 1:2021 Information technology — Multimedia application format (MPEG-A) — Part 22: Multi-image application format (MIAF) — Amendment 1: Reference software and conformance for multi-image application format
- ISO/IEC 23000-22:2019/Amd 2:2021 Information technology — Multimedia application format (MPEG-A) — Part 22: Multi-image application format (MIAF) — Amendment 1: HEVC Advanced HDR profile and other clarifications
- ISO/IEC 23001-10:2020/Amd 1:2021 Information technology — MPEG systems technologies — Part 10: Carriage of timed metadata metrics of media in ISO base media file format — Amendment 1: Support for content-guided transcoding and spatial relationship of immersive media
- ISO/IEC 23001-14:2019/Amd 1:2021 Information technology — MPEG systems technologies — Part 14: Partial file format — Amendment 1: Support for HTTP entities, enhanced file type and byte-range priorities (pending assignment)
- ISO/IEC 23001-16:2021 Information technology — MPEG systems technologies — Part 16: Derived visual tracks in the ISO base media file format
- ISO/IEC DIS 23001-18:2022 Information technology — MPEG systems technologies — Part 18: Event message tracks in the ISO base media file format
- ISO/IEC 23008-10:2015/Amd 1:2021 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 10: MPEG media transport forward error correction (FEC) codes — Amendment 1: Window-based FEC code
- ISO/IEC 23009-1:2022 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats
- ISO/IEC 23009-8:2022 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 8: Session-based DASH operations
- ISO/IEC DIS 23090-10:2022 Information technology — Coded representation of immersive media — Part 10: Carriage of visual volumetric video-based coding data
- ISO/IEC PRF 23090-17:2021 Information technology — Coded representation of immersive media — Part 17: Reference software and conformance for omnidirectional media format (OMAF)

3.6 WG 4 MPEG Video

3.6.1 WG 4 accomplishments

WG 4 Terms of Reference

WG 4 develops coding and compression of digital representations and descriptions of moving pictures generated or captured by any technology; efficiency evaluation of coding systems and content description

methods related to visual information; implementation feasibility analysis of coding systems and content description methods.

Second edition of MPEG Immersive Video

MPEG immersive video (MIV) standard (ISO/IEC 23090-12) was developed to support compression of immersive video content in which multiple real or virtual cameras capture a real or virtual 3D scene. The standard enables storage and distribution of immersive video content over existing and future networks for playback with 6 Degrees of Freedom (6DoF) of view position and orientation. MIV is a flexible standard for multiview video with depth (MVD) that leverages the strong hardware support for commonly used video codecs to code volumetric video. Views may use equirectangular, perspective, or orthographic projection. By packing and pruning views, MIV can achieve affordable bit rates at significantly reduced coding pixel rate. The MIV standard defines multiple profiles: (i) the MIV Main profile for MVD, (ii) the MIV Geometry Absent profile targeting cloud-based and decoder-side depth estimation, and (iii) the MIV Extended profile enabling the coding of multi-plane images (MPI).

The MIV standard is designed as a set of extensions and profile restrictions for the Visual Volumetric Video-based Coding (V3C) standard (ISO/IEC 23090-5). The main body of this standard is shared between MIV and the Video-based Point Cloud Coding (V-PCC) standard (ISO/IEC 23090-5 Annex H). It may potentially be used by other MPEG-I volumetric codecs under development. The carriage of MIV is specified through the Carriage of V3C Data standard (ISO/IEC 23090-10).

MPEG Video Coding (WG 4) is working on the second edition of ISO/IEC 23090-12, of which a preliminary working draft was issued from the 139th MPEG meeting meeting in July 2022. The second edition extends MIV to support use cases such as full 6DoF immersive content experience, natural scenes with non-Lambertian characteristics and captured with moving cameras and adjustable focus, bullet time, object-based content composition and compression, etc.

Second edition of 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, video and 3D graphics 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. The FDIS of the first international standards on neural network compression for multimedia applications (ISO/IEC 15938-17 ed.1) has been approved in July 2022, and is thus soon to be published. This specification defines a set of tools for the compressed representation of neural networks for more efficient transmission and inference. The results show that trained neural networks for many multimedia applications 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 Video Coding (WG 4) is also already working on a second edition of ISO/IEC 15938-17, of which a Draft International Standard (DIS) has been completed in the 8th WG 4 meeting in July 2022. The second edition adds the functionality to compress incremental updates of neural networks, which can e.g. be applied to sending updates of neural networks or to federated learning scenarios.

Promotion of Reference Software and Conformance for Neural Network Coding to Committee Draft

MPEG recently ratified the first international standard on Neural Network Compression for Multimedia Applications (ISO/IEC 15938-17). At the 136th MPEG meeting, *MPEG Video* promoted ISO/IEC 15938-18 providing *(i)* reference software, *(ii)* guidelines, and *(iii)* assets for conformance testing to Committee Draft (CD), the first milestone of its formal approval process.

The neural network coding standard is designed as a toolbox of coding technologies. The specification contains different parameter reduction (*e.g.*, pruning, sparsification, matrix decomposition), parameter transformation (*e.g.*, quantization), and entropy coding methods, that can be assembled to encoding pipelines combining one or more (in the case of reduction) methods from each group. The reference software is written in Python and provides a framework defining interfaces for each of the three steps in the coding pipeline and components implementing all supported methods. Additionally, the new standard provides bitstreams for testing the conformance of decoders against the neural network coding specification (ISO/IEC 15938-17).

Completion of Conformance and Reference Software for LCEVC

At the 137th MPEG meeting, *MPEG Video Coding (WG 4)* completed the development of the Conformance and Reference software standard for Low Complexity Enhancement Video Coding (LCEVC) (ISO/IEC 23094-3) with its promotion to the Final Draft International Standard (FDIS) stage for final approval and publication.

This standardization will assist implementers of LCEVC in checking the proper functioning of their implementations during the development of their products. The tests will also help the users and potential users of LCEVC products by providing a way to verify claims of the conformance of such products. Thus, such tests will assist the community in achieving interoperability of encoder and decoder products and will encourage the adoption and use of LCEVC.

LCEVC adds an enhancement data stream that can appreciably improve the resolution, bit depth and visual quality of reconstructed video with effective compression efficiency of limited encoding and decoding complexity by building on top of existing and future video coding formats. It is designed to be compatible with existing video workflows (*e.g.*, with CDNs, metadata management, and DRM/CA) and streaming/media formats (*e.g.*, DASH, and 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 relatively low for high-resolution video encoding and delivery or when decoding complexity is a challenge.

Committee Draft of Conformance and Reference Software for MPEG Immersive Video

At the 137th MPEG meeting, *MPEG Video Coding (WG 4)* promoted its MPEG Immersive Video (MIV) Conformance and Reference software standard (ISO/IEC 23090-23) to the Committee Draft (CD) stage, the first formal milestone of its approval process. The document specifies how to conduct conformance tests and provides reference encoder and decoder software for ISO/IEC 23090-12 MPEG Immersive video. This draft includes 18 verified and validated conformance bitstreams and encoding and decoding reference software based on version 12.0 of the Test model for MPEG Immersive Video (TMIV).

MIV was developed to support compression of immersive video content, in which a real or virtual 3D scene is captured by multiple real or virtual cameras. The standard enables the storage and distribution of immersive video content over existing and future networks, for playback with 6 degrees of freedom (6DoF) of view position and orientation. MIV is a flexible standard for multiview video with depth (MVD) that leverages the strong hardware support for commonly used video formats to code volumetric video. Views may use equirectangular, perspective, or orthographic projection. By pruning and packing views, MIV can achieve bit rates around 15 to 30 Mb/s using High Efficiency Video Coding (HEVC) and a pixel rate equivalent to HEVC Level 5.2. Besides the MIV Main profile for MVD, there is the MIV Geometry Absent profile, which is suitable for use with cloud-based and decoder-side depth estimation, and the MIV Extended profile, which enables coding of multi-plane images (MPI). The MIV standard is designed as a set of extensions and profile restrictions on the Visual Volumetric Video-based Coding and Video-based Point Cloud Coding (ISO/IEC 23090-5) standard, and its conformance bitstreams cover all of the specified profiles.

In addition to conformance testing, work on the verification testing of MIV is ongoing, and the carriage of MIV is specified through the Carriage of V3C Data standard (ISO/IEC 23090-10). *MPEG Requirements (WG 2)* produced final use cases and requirements for the MIV 2nd edition, which will be an evolution of the MIV standard, and *MPEG Liaison and Communication (AG 3)* produced a white paper on MIV.

3.6.2 WG 4 current work items

WG 4 is progressing the development of the following standards.

- ISO/IEC DIS 15938-18 Information technology — Multimedia content description interface — Part 18: conformance and reference software for compression of neural networks
- ISO/IEC FDIS 23090-12 Information technology — Coded representation of immersive media — Part 12: MPEG immersive video
- ISO/IEC DIS 23090-12/DAmd 1 Information technology — Coded representation of immersive media — Part 12: MPEG immersive video — Amendment 1: V3C extension mechanism
- ISO/IEC CD 23090-23 Information technology — Coded representation of immersive media — Part 23: Conformance and reference software for MPEG Immersive video
- ISO/IEC 23094-1:2020/DAmd 1 Information technology – General video coding — Part 1: Essential video coding — Amendment 1: Green metadata supplemental enhancement information

3.6.3 WG 4 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 4:

- ISO/IEC 15938-17:2022 Information technology — Multimedia content description interface — Part 17: Compression of neural networks for multimedia content description and analysis
- ISO/IEC 23094-2:2021 Information technology — General video coding — Part 2: Low complexity enhancement video coding
- ISO/IEC 23094-3:2022 Information technology — General video coding — Part 3: Conformance and Reference Software for Low Complexity Enhancement Video Coding
- ISO/IEC 23094-4:2022 Information technology — General video coding — Part 2: Conformance and reference software of essential video coding

3.7 WG 5 MPEG Joint Video Coding Team with ITU-T SG 16

3.7.1 WG 5 accomplishments

WG 5 Terms of Reference

WG 5 corresponds to the Joint Video Experts Team (JVET) established between SC 29 MPEG and ITU-T SG 16. JVET is established to collaboratively develop technically aligned twin text or common text for Recommendations | International Standards for video coding technology.

Progression of VVC Conformance Testing

At the 136th MPEG meeting, two important milestones were achieved by the *MPEG Joint Video Experts Team* with ITU-T SG16 in developing conformance testing for ISO/IEC 23090-3 Versatile Video Coding (VVC): *(1)* Conformance Testing for Versatile Video Coding (ISO/IEC 23090-15) reached Final Draft International Standard (FDIS) which provides means to test implementations of VVC for their conformance to the profile specified in VVC version 1. *(2)* VVC version 2 (incl. operation range extensions for higher bit depth and high bit rates) is in the last phase of its development and is planned to be finalized in January 2022. Therefore, a related conformance testing specification has been promoted to Committee Draft Amendment (CDAM) for testing the features of the new profiles in VVC version 2.

The standardization of such tests will assist implementers of VVC in checking the proper functioning of their implementations during the development of their products. Such tests will also help the users and potential users of VVC products by providing a way to verify claims of the conformance of such products. Thus, such tests will assist the community in achieving interoperability of encoder and decoder products and will encourage the adoption and use of VVC.

Second Editions of VVC & VSEI and finalization of VVC Reference Software

At the 137th MPEG meeting, *MPEG Joint Video Coding Team(s) with ITU-T SG 16 (WG 5; JVET)* completed the development of the second editions of Versatile Video Coding (VVC, ISO/IEC 23090-3 | ITU-T H.266) and Versatile supplemental enhancement information messages for coded video bitstreams (VSEI, ISO/IEC 23002-7 | ITU-T H.274) by promoting them to Final Draft International Standard (FDIS) status for final approval and publication. The new VVC version defines profiles and levels supporting larger bit depths (up to 16 bits), including some low-level coding tool modifications to obtain improved compression efficiency with high bit-depth video at high bit rates. VSEI version 2 adds SEI messages giving additional support for scalability, multi-view, display adaptation, improved stream access, and other use cases. Furthermore, a Committee Draft Amendment (CDAM) for the next amendment of VVC was issued to begin the formal approval process to enable linking VVC with the Green Metadata (ISO/IEC 23001-11) and Video Decoding Interface (ISO/IEC 23090-13) standards and add a new unconstrained level for exceptionally high capability applications such as certain uses in professional, scientific, and medical application scenarios. Finally, the reference software package for VVC (ISO/IEC 23090-16) also been completed with its achievement of FDIS status. Reference software is extremely helpful for developers of VVC devices, helping them in testing their implementations for conformance to the video coding specification.

Promotion of Tenth Edition of AVC to Final Draft International Standard

At the 137th MPEG meeting, *MPEG Joint Video Coding Team(s) with ITU-T SG 16 (WG 5; JVET)* finalized the 10th edition of Advanced Video Coding (ISO/IEC 14496-10 | ITU-T H.264) by issuing it as a Final Draft International Standard (FDIS) for final approval and publication. Beyond various text improvements, this specifies a new SEI message for describing the shutter interval applied during video capture. This can be variable in video cameras, and conveying this information can be valuable for analysis and post-processing of the decoded video.

Extension of HEVC for High-Capability Applications up to 16K and Beyond

At the 137th MPEG meeting, *MPEG Joint Video Coding Team(s) with ITU-T SG 16 (WG 5; JVET)* started the approval process of a new second amendment of High Efficiency Video Coding (HEVC, ISO/IEC 23008-2 | ITU-T H.265) by issuing a Committee Draft Amendment (CDAM) that defines new levels and tiers providing support for very high bit rates and video resolutions up to 16K, as well as defining an unconstrained level. This will enable the usage of HEVC in new application domains, including professional, scientific, and medical video sectors.

New Amendment of VSEI containing Technology for Neural Network-based Post Filtering

At the 139th MPEG meeting, the *MPEG Joint Video Experts Team with ITU-T SG 16* (WG 5; JVET) issued a Committee Draft Amendment (CDAM) text for the *Versatile Supplemental Enhancement Information* (VSEI) standard (ISO/IEC 23002-7, a.k.a. ITU-T H.274). Beyond the SEI message for shutter interval indication which is already known from its specification in *Advanced Video Coding* (AVC, ISO/IEC 14496-10, a.k.a. ITU-T H.264) and *High Efficiency Video Coding* (HEVC, ISO/IEC 23008-2, a.k.a. ITU-T H.265), and a new indicator for subsampling phase indication which is relevant for variable-resolution video streaming, this new amendment contains two Supplemental Enhancement Information (SEI) messages for describing and activating post filters using neural network technology in video bitstreams. This could for example be used for reducing coding noise, upsampling, colour improvement, or denoising. The description of the neural network architecture itself is based on MPEG's neural network coding standard (ISO/IEC 15938-17). Results from an exploration experiment have shown that neural network-based post filters can deliver better performance than conventional filtering methods. Processes for invoking these new post-processing filters have already been tested in a software framework, and will be made available in an upcoming version of the *Versatile Video Coding* (VVC, ISO/IEC 23090-3, a.k.a. ITU-T H.266) reference software (ISO/IEC 23090-16, a.k.a. ITU-T H.266.2).

New Edition of Video Coding-Independent Code Points Standard

At the 139th MPEG meeting, the *MPEG Joint Video Experts Team with ITU-T SG 16* (WG 5; JVET) issued a Committee Draft (CD) text for the third edition for the video part of the Coding-Independent Code Points (CICP) standard (ISO/IEC 23091-2, a.k.a. ITU-T H.273). This new edition will include two new YCgCo-R colour type identifiers enabling efficient luma-chroma representations with improved support of lossless conversion to and from RGB colour spaces without requiring support of different bit depths for luma and chroma.

3.7.2 WG 5 current work items

WG 5 is progressing the development of the following standards.

- ISO/IEC DIS 14496-10 Information technology — Coding of audio-visual objects — Part 10: Advanced video coding
- ISO/IEC FDIS 23002-7 Information technology — MPEG video technologies — Part 7: Versatile supplemental enhancement information messages for coded video bitstreams
- ISO/IEC FDIS 23002-7/WD Amd 1 Information technology — MPEG video technologies — Part 7: Versatile supplemental enhancement information messages for coded video bitstreams — Amendment 1: Additional SEI messages
- ISO/IEC WD TR 23002-9 Information technology — MPEG video technologies — Part 9: Film grain synthesis technologies for video applications
- ISO/IEC 23008-2:2020/CD Amd 2 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding — Amendment 2: High-range levels
- ISO/IEC FDIS 23090-3 Information technology — Coded representation of immersive media — Part 3: Versatile video coding
- ISO/IEC FDIS 23090-3/CD Amd 1 Information technology — Coded representation of immersive media — Part 3: Versatile video coding — Amendment 1: New level and systems-related supplemental enhancement information
- ISO/IEC FDIS 23090-15 Information technology — Coded representation of immersive media — Part 15: Conformance testing for versatile video coding
- ISO/IEC FDIS 23090-15/DAmd 1 Information technology — Coded representation of immersive media — Part 15: Conformance testing for versatile video coding
- ISO/IEC FDIS 23090-16 Information technology — Coded representation of immersive media — Part 16: Reference software for versatile video coding
- ISO/IEC WD 23091-2 Information technology — Coding-independent code points — Part 2: Video (pending assignment)

3.8 WG 6 MPEG Audio Coding

3.8.1 WG 6 accomplishments

WG 6 Terms of Reference

WG 6 develops standards for coded digital representation of audio and associated metadata.

6DoF Technology Selected for MPEG-I Immersive Audio

MPEG had previously issued a Call for Proposals (CfP) for 6DoF immersive audio technology in April 2021 for convincing Virtual Reality (VR) and Augmented Reality (AR) experiences. At the 137th MPEG meeting, submissions to this CfP were reviewed and technology was selected. Fourteen submissions were evaluated in three subjective VR/AR listening tests, conducted at 12 test sites around the world, as a basis for the selection.

The technology selected in the CfP permits the user to have a VR or AR experience in which the user can freely navigate and interact with the virtual environment using 6 degrees of freedom (6DoF), thus enabling unconstrained spatial movement and user rotation. The rendered audio signals can be in multiple formats using audio objects, channels, and Higher Order Ambisonics (HOA). The VR/AR environment (scene) is encoded in the bitstream and the MPEG-I Immersive Audio technology renders a binaural audio signal to the user's headphone based on the transmitted and decoded scene and implements a rich set of audio effects, such as directionality, localization, extent, occlusion, diffraction and Doppler shift of sound sources, and sophisticated modelling of the acoustic environment. A wide range of user interactions with the VR/AR environment are supported.

It is expected that the standard will progress to the first formal stage of its approval process with a Committee Draft (CD) in January 2023, Draft International Standard (DIS) in April 2023, and International Standard (IS) in October 2023.

3.8.2 WG 6 current work items

WG 6 is progressing the development of the following standards.

- ISO/IEC DTR 14496-24 Information technology — Coding of audio-visual objects — Part 24: Audio and systems interaction
- ISO/IEC AWI 14496-26 Information technology — Coding of audio-visual objects — Part 26: Audio conformance
- ISO/IEC 23003-4:2020/CD Amd 2 Information technology — MPEG audio technologies — Part 4: Dynamic range control — Amendment 2: Loudness Leveling
- ISO/IEC 23008-9:2022/DAmd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 9: 3D Audio conformance testing — Amendment 1: Sample rate conversion
- ISO/IEC AWI 23090-4 Information technology — Coded representation of immersive media — Part 4: MPEG-I immersive audio

3.8.3 WG 6 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 6:

- ISO/IEC 23003-3:2020/Amd 1:2021 Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding — Amendment 1: Reference software and conformance
- ISO/IEC 23003-4:2020/Amd 1:2022 Information technology — MPEG audio technologies — Part 4: Dynamic range control — Amendment 1: Side chain normalization
- ISO/IEC 23003-6:2022 Information technology — MPEG audio technologies — Part 6: Unified speech and audio coding reference software
- ISO/IEC 23003-7:2022 Information technology — MPEG audio technologies — Part 7: Unified speech and audio coding conformance testing
- ISO/IEC 23008-3:2022 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio
- ISO/IEC 23008-6:2021 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 6: 3D audio reference software
- ISO/IEC 23008-9:2022 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 9: 3D Audio conformance testing
- ISO/IEC 23091-3:2018/Amd 1:2022 Information technology — Coding-independent code points — Part 3: Audio — Amendment 1: Headphone support

3.9 WG 7 MPEG 3D Graphics and Haptics Coding

3.9.1 WG 7 accomplishments

WG 7 Terms of Reference

WG 7 develops standards for the coding of 3D data representations for graphics objects and environments, of sensing and actuating data including all modalities of haptics, computer generated or captured from the physical world, and integration of natural and synthetic audio - video, haptics, and interaction information.

WG 7 summary

Recent advances in computer graphics have made it possible to create realistic digital representations of three-dimensional objects and surroundings allowing real-time interactions with users. Such representations use various mathematical models to describe both geometries and attributes (colors, materials, and lighting behavior) of the objects. Among them, 3D meshes and point-clouds have become popular and extensively used in the gaming and 3D movie industries. Nevertheless, the creation of a high-fidelity content requires enormous computational resources for storage, transmission, processing, and visualization purposes. This problem penalizes their deployment on light mobile devices. Such graphics representations are well-suited for various industrial applications, including 3D scanning and modeling, environmental monitoring, agriculture and forestry, and autonomous driving, being able to fulfill different requirements in terms of precision, resolution, and related attributes. Therefore, graphics representation with high density (up to millions of 3D points) are required for hyperrealistic visualization applications, which raise enormous demands in terms of computational and memory requirements. Sparse representations are used for autonomous navigation applications, but high precision is required, which needs to be preserved. Moreover, for applications involving dynamic content, the supplementary temporal dimension explodes the related bandwidth and needs to be appropriately considered.

Video-based compression for graphics content (ISO/IEC 23090 Parts 5 and 29)

With the publication of the first edition of MPEG-I Part 5, WG 7 initiated a compression technology track by considering that any 3D graphics representation can be transformed into a set of 2D projections that can further on be compressed by 2D (image and video) solutions.

This strategy was first used to compress dense point clouds and the corresponding standard is MPEG-I Part 5, Video-based Point Cloud Coding (V-PCC). The main goal for V-PCC is to enable rapid deployment of point cloud representations on current devices through the reuse of existing hardware. This is achieved by first transforming the 3D geometry and attribute data of the point cloud into a set of 2D patches. These patches are then mapped to a predefined set of 2D planes through orthogonal projections, without self-occlusions and with limited distortion. A mapping between the point cloud and a regular 2D grid is then obtained by packing the projected patches. This mapping is then used to generate 2D images representing the point cloud geometry and its attributes. Since the mapping between the point cloud and the 2D grid is not bijective, an extra binary image, referred to as the occupancy map, is needed in order to distinguish between the filled (i.e., associated with a point) and the empty (i.e., not associated with any point) cells of the grid. Similar to the geometry and attribute video sequences, the occupancy map video sequence is compressed using 2D based image or video codecs. In order to be able to reconstruct the 3D point cloud from the 2D geometry, attribute, and occupancy video, additional information specifying per patch information (e.g., 3D/2D patch location and projection plane index) is required and V-PCC introduces a new codec specifically optimized to handle the patch information sub-stream. This sub-stream occupies a relatively small amount (less than 5%) of the overall bitstream. Additional information needed to synchronize and link the video and patch sub-streams is also signaled in the bitstream. All these components can be multiplexed into a single bitstream or encoded as separate tracks of an ISOBMFF container.

A second instance of the same 3D to 2D projection strategy is the compression of immersive video content, a standard developed by WG 4 and published under MPEG-I Part 12. WG 4 reused the same framework built by WG 7 and enriched it support volumetric video content.

A third instance of the framework is currently under development in WG 7 in MPEG-I Part 29 Video-based dynamic mesh coding (V-Mesh). A dynamic mesh sequence may require a large amount of data since it may consist of a significant amount of information changing in time. Therefore, efficient coding technologies are required to store and transmit such content and Part 29 key concept is the representation of the mesh as a set of refinements of a coarse mesh version and the encoding of these refinements as video frames. The standardization process is ongoing but the current results show significant improvements compared with state of the art mesh coding technologies.

Selection of Technology for MPEG-I Video-based Dynamic Mesh Compression

As noted above, *MPEG Technical Requirements* issued a Call for Proposals (CfP) for dynamic mesh compression technology in October 2021. At the 138th MPEG meeting, *MPEG 3D Graphics Coding* (WG 7) reviewed submissions to this CfP and selected technology for the subsequent standardization process. Five submissions

covering full codecs were evaluated based on (i) objective measurements for geometry and attributes and (ii) subjective visualization tests conducted at two test sites.

The technology selected in the CfP already enables the efficient representation of realistic dynamic objects (including humans), and the MPEG 3D Graphics Coding will conduct additional experiments to further improve this basis technology and to produce a standard specification document and test model for the full codec (encoder and decoder). The final standard will make it possible to include realistic 3D dynamic objects in games and VR or AR experiences, transmit them between virtual worlds, and create immersive experiences for artistic performances and sports events.

It is expected that the standard will progress to the first formal stage of its approval process with a Committee Draft (CD) in April 2023, followed by a Draft International Standard (DIS) in January 2024 and completed International Standard (IS) by October 2024.

Geometry-based compression for graphics content (ISO/IEC 23090 Part 9)

The projective based method of the Video-based compression for graphics content introduced previously provides good results when the content is dense. For sparse 3D content such method would conduct to unoccupied regions in the 2D planes and therefore an efficient video coding. In order to address these cases, WG 7 developed in parallel to V-PCC a geometry-based coding approach able to accommodate any point cloud representation. The geometry-based compression of 3D graphics content is the traditional path that WG 7 (and previously the 3DG sub-group of WG 11) followed when developing standards for graphics. Such approach was first implemented in the MPEG-4 parts related to graphics (Part 2, 16 and 25) and continued in MPEG-I Part 9 (G-PCC) when point cloud compression was addressed. G-PCC is a compression standard dedicated for generic point clouds, covering sparse and dense, static and dynamic objects and scenes. The G-PCC coding process is made up of two different paths: one for positions (or geometry) and the other for attributes (e.g., color). The two-path approach cannot be conducted in parallel because the attribute compression path can only begin after geometry is compressed. If the compression of the point coordinates is lossy, it means that the locations of points may be different from the original point coordinates and also that the number of points after compression may be different. For this reason, attributes need to be recomputed based on the reconstructed point coordinates. G-PCC includes several compression tools for geometry (octree, context modeling, triangle soup, predictive tree) and attributes (Level of Details decomposition, RAHT transform) that are combined together in several profiles to efficiently address specific applications. WG 7 finalized the FDIS of G-PCC in 2021 and is now working on the second edition addressing better inter-prediction coding and additional functionalities.

Low latency, low complexity LiDAR coding (ISO/IEC 23090 Part 30)

While evaluating technology that improves its existing point cloud compression standards for LiDAR acquired point clouds, WG 7 has realized that there is significant difference in the acquisition characteristics of available test material and direct sensing methods available to industry. The acquisition technology used to capture point cloud geometry has fundamental characteristics that may be exploited by compression technologies and standards. There are various acquisition protocols such as spinning LiDAR (for example, 360° rotating sensor with one or more scanning lasers, each at constant elevation with scanning at a constant or variable rate), scanning LiDAR (for example, mirror-based scanning according to a raster or parametric scan pattern producing a non-linear sweep over a limited field of view with scanning that may be horizontally or vertically oriented), array LiDAR / Radar (for example, flash LiDAR, or other 2D sensor arrays), etc. WG 7 is currently developing Part 30 of MPEG-I addressing compression of point clouds obtained by LiDAR technologies while exploiting the characteristics of the acquisition protocol. By doing so, applications with low latency and low complexity constraints may be addressed.

Call for LiDAR and Volumetric Point Cloud Content

MPEG Coding of 3D Graphics has been evaluating technologies that improve its existing point cloud compression standards for volumetric scanners. To ensure that future compression standards are tailored to the needs and requirements of the industry, *MPEG 3D Graphics Coding* issued a Call for Point Cloud Content at the 136th MPEG meeting. Industry partners were invited to propose point cloud content acquired using novel acquisition methods such as advanced LiDAR or volumetric scanners.

AI-based coding for graphics

WG 7 is currently conducting an exploration on investigating the advantages and limitations of using AI-based technologies for graphics coding. Currently, machine learning technologies are game changers in the majority of activity domains involving ICT and a new generation of technical standards addressing aspects related to machine learning are needed. The objectives of the activity are first to investigate the potential value of AI based technologies compared with existent (MPEG) standards for 3D graphics compression and, secondly, to prepare a methodology for a potential Call for Proposal, including the generation of several draft documents (requirements, overall framework, common test conditions) needed to initiate the standardization process. While the compression results of the AI based methods are not yet better than the ones obtained with the traditional methods, it is highly expected that the situation be reversed in the near future due to the advancements in the AI and of models that are becoming more and more mature. WG 7 is progressing the exploration by focusing the activities on three types of activities: creating the guidelines for conducting AI exploration experiments for point cloud compression, conducting the performance analysis of currently AI-based available solutions and collecting a preliminary data set collection for AI experiments.

Efficient 3D graphics media representation for render-based systems and applications (ISO/IEC 23090-28)

Display technologies capable of creating fully formed holograms, e.g. 3D images floating in free space comprised of waveforms emitted from a display panel surface, require input source media that are sufficiently dense in information such that each of the objects in a holographic visual scene is described in terms of its complete and explicit geometry, and surface properties to characterize how the surface responds to the presence of light. Such visual information is furthermore desired to enable display technologies to produce visual scenes of objects that cannot be distinguished from the same objects when viewed in the natural world; i.e., the visual information of the media should be sufficiently dense such that the display is enabled to produce a photorealistic result. Commonly used raster media formats such as those specified in ITU-R BT.601, ITU-R BT.709, ITU-R BT.2020, ITU-R BT.2100 were originally developed to efficiently provide information sufficient for 2D displays to produce 2D visual presentations, albeit with varying degrees of spatial resolutions, frame rates, and pixel depth. Such commonly used raster formats, however, neither provide sufficient information in a practical nor efficient manner, to enable a holographic display to produce holograms. However, 3D graphics formats used in digital content creation tools and production workflows offer a solution to provide sufficiently dense information to enable photorealistic results where raster formats do not do so. The use of 3D graphics technologies hence serves as a fundamental technology in enabling immersive displays to create their optimal visual presentations.

By initiating a new part of ISO/IEC 23090 WG 7's objective is to specify a canonical scene graph that is used to translate a concrete scene graph format (or scene graph representation) to another concrete scene graph format (or scene graph representation). The canonical scene graph itself is not intended to be implemented nor used as a concrete format. Rather, the canonical scene graph facilitates the translation from one concrete format to another concrete format by specifying terms and associated definitions that are commonly employed in existing concrete scene representations. Such terms and definitions describe nodes, media formats, media attributes, and render processing instructions of commonly-used concrete formats.

Haptics Coding (To be requested as a new part of ISO/IEC 23090)

Haptics digital encoding is the storing of tactile data in digital format. As per music and video, digital encoding is of fundamental importance to allow digital haptics devices to function. Haptics encoding gained relevance in the last years with the increased market importance of wideband haptics in consumer peripherals. The prior generation of haptics peripherals was based on less expressive haptics actuators usually based on state machine control processes. The current market and technological landscape are highly fragmented with a large spectrum of encoding formats developed and used by various technology manufacturers. This brings to several cross-compatibility problems blocking the haptics market to scale.

In ISO/IEC 23090-31, the WG 7 objective is to specify technology which supports the efficient transmission and rendering of haptic signals for the playback of immersive experiences in a wide variety of scenarios. The standard aims at specifying a robust coded representation of haptic media targeting the two most popular haptics perception leveraged by devices today: vibrotactile and kinaesthetic. The coded representation allows to encode both descriptive and quantized data in a human readable JSON format used for exchange purposes and a compressed bitstream version, optimized for memory usage, for distribution purposes. This approach

allows to meet the expectations of compatibility with both descriptive and quantized formats on the market and offers an exchange format allowing interoperability between devices for both 3D immersive experiences and distribution purposes.

Internet of Media Things (ISO/IEC 23093)

The MPEG IoMT standard family (formally known as ISO/IEC 23093 Parts 1, 2 and 3) provides an architecture and specifies application programming interfaces (APIs) and compressed representation of data flowing among “media things” (MThings). The APIs for the media things facilitate discovering other media things in the network, connecting, and efficiently exchanging data between media things. The APIs also provide means for supporting transaction tokens to access valuable functionalities, resources, and data from media things. This way, the media things can be represented in industrial applications as abstract entities, providing their functionalities in a standard way. While this process entered a mature phase, various additional MThings and/or MThing characteristics are required for fulfilling new applicative needs. The results relating to these aspects are expected to contribute to 3rd editions of ISO/IEC 23093 Parts 1, 2 and 3.

In parallel, WG 7 further explores an incremental industrial need, namely the possibility of defining autonomous services over media things. In this respect, exploratory work was carried out with a two-folded aim: (1) to describe in a formal (standard) way the mission to be performed by various media things that collaboratively achieve a high-level task, and (2) to identify how such descriptions can be exchanged with external processing resources (like DLT) that can ensure the controlled processing of such solution. The results related to these aspects are expected to contribute to the new versions of the existing parts as well as to a new part of the standard – IoMT Part 5 – IoMT Autonomous Collaboration. This preliminary work brings to light the need for extending the activities to specify the interfaces for a new component (Mission Management and Control) as well as for the interfaces required for the data exchange among media things/system manager and existing transaction-related technologies.

3.9.2 WG 7 current work items

WG 7 is progressing the development of the following standards.

- ISO/IEC DIS 23090-5 Information technology — Coded representation of immersive media — Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC)
- ISO/IEC DIS 23090-5/DAmd 1 Information technology — Coded representation of immersive media — Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC) — Amendment 1: V3C extension mechanism and payload type
- ISO/IEC DIS 23090-9 Information technology — Coded representation of immersive media — Part 9: Geometry-based point cloud compression
- ISO/IEC DIS 23090-19 Information technology — Coded representation of immersive media — Part 19: Reference Software for V-PCC
- ISO/IEC DIS 23090-20 Information technology — Coded representation of immersive media — Part 20: Conformance for V-PCC
- ISO/IEC CD 23090-21 Information technology — Coded representation of immersive media — Part 21: Reference Software for G-PCC
- ISO/IEC CD 23090-22 Information technology — Coded representation of immersive media — Part 22: Conformance for G-PCC
- ISO/IEC AWI 23090-28 Information technology — Coded representation of immersive media — Part 28: Efficient 3D graphics media representation for render-based systems and applications
- ISO/IEC AWI 23090-29 Information technology — Coded representation of immersive media — Part 29: Video-based dynamic mesh coding
- ISO/IEC AWI 23090-30 Information technology — Coded representation of immersive media — Part 30: Low latency, low complexity LiDAR coding

- ISO/IEC DIS 23093-4 Information technology — Internet of media things — Part 4: IoMT Reference Software and Conformance
- ISO/IEC WD 23093-5 Information technology — Internet of media things — Part 5: IoMT autonomous collaboration

3.9.3 WG 7 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 7:

- ISO/IEC 23093-1:2022 Information technology — Internet of media things — Part 1: Architecture
- ISO/IEC 23093-2:2022 Information technology — Internet of media things — Part 2: Discovery and communication API
- ISO/IEC 23093-3:2022 Information technology — Internet of media things — Part 3: Media data formats and APIs
-

3.10 WG 8 MPEG Genomic Coding

3.10.1 WG 8 accomplishments

WG 8 Terms of Reference

WG 8 develops standards for various types of genomic data.

Evaluation of Responses on New Advanced Genomics Features and Technologies

The extensive usage of high-throughput DNA sequencing technologies enables new approaches to healthcare known as “precision medicine” as well as many other emerging applications such as monitoring the evolution of outbreaks and pathogen surveillance in agriculture and the food industry. DNA sequencing technologies produce extremely large amounts of heterogeneous data including raw sequence reads, analysis results, annotations, and associated metadata which are stored in different repositories worldwide and the use of the data needs to be enabled by standardized and interoperable formats. Structuring and high-performance compression of such genomic data is required to reduce the storage size, increase the transmission speed, and improve the browsing and searching performance of these large data sets as required by a wide variety of applications and use cases. The current MPEG-G standard series (ISO/IEC 23092) addresses the representation, indexing, compression, and transport of genome sequencing data with support for annotation data and searching capability. The ISO/IEC 23092 (MPEG-G) standard series provides a file and transport format, compression technology, metadata specifications, protection support, and standard APIs for the access of genomic data and annotation data in a native compressed format.

At the 134th MPEG meeting, *MPEG Genomic Coding* (WG 8) had previously issued a Call for Proposals (CfP) to collect submissions of new technologies improving the current compression, transport and indexing capabilities of the ISO/IEC 23092 standard series.

At the 137th MPEG meeting, *MPEG Genomic Coding* (WG 8) evaluated submitted responses to the CfP, addressing the representation and usage of graph genome references and supporting interfaces with existing standards for the interchange of clinical data. The initial evaluation results show that the incorporation of native representations of graph genome references into the MPEG-G standard is desirable and will provide new, advanced, and efficient representation capabilities for extended support of use cases of genomic sequencing data. Concerning the support of interfaces with existing standards for the interchange of clinical data (HL7 and FHIR) the response received shows that such an extension is desirable for the efficient integration of MPEG-G data into clinical workflows and shows a standardization path to achieve such integration. With such purpose 5 core experiments have been defined: CE 1 High level Syntax and Graph Genome Representation, CE 2 High Level Syntax of Sequence-to-graph Alignment Mappings, CE 3 Indexing and Querying of Graph Genomes and Sequence-to-graph Alignment Mappings, CE 4 Graph Partitioning for Efficient Compression and Selective Access, CE 5 Entropy Coders for the Compression of Graph Genomes and Sequence-to-graph Alignment Mappings. The results of such core experiments will define how to extend the different parts of ISO/IEC 23092 to support the validated advanced functionalities. The plan is to be able to represent in

compressed form pan-genomes in the form of graphs and be able to use them as “reference” for the compression of sequencing (raw and aligned) data. Such new feature will make possible to provide efficient compressed representation of sequencing data also for species for which reference genomes do not exists or are not general enough to represent all types of data.

Coding of genomic annotations

After the standardization of the compressed representation of sequencing data, WG 8 has intensively worked to extend the 23092 standard family to include the indexed compression of genomic annotations and linkage of them to the corresponding sequencing data. This work is about to be completed for the compressed representation part with the promotion to DIS of ISO-IEC 23092-6 Coding of Genomic Annotations at the 139th MPEG meeting.

Entropy coding for Part 6 Genomic Annotations and Low-Complexity entropy coding

With the objective of specifying efficient entropy coding algorithms for the heterogeneous data generated by annotation ISO/IEC 23092-2 was amended and is about to be published as 3rd edition. Such standard extension supports also the entropy coding of sequencing data with new low-complexity codecs that provide relevant speed-ups and performance improvements.

Application program Interfaces and metadata

With the development of Part 6 Coding of genomic Annotation it has been necessary to work on the extension of APIs and metadata to support the compressed representation of annotations. The study of a new amendment of Part 3 has started collecting all input contributions received in the last 18 months including the answers to the Call for new advanced technologies received in January 2022. The current activity of revision and extension of part 3 includes not only the necessary extensions for Part 6, but also the new proposed support for the interchange of medical data in the form of Hl7 and FIRH standards.

Collaborations and synergy with other ISO TCs: TC 276/WG 5 and TC34/SC 16

Discussions and technical work have started in collaboration with ISO TC 276/WG 5 on the mutual support of “ISO/IEC 23092 Information technology – Representation of Genomic Information” (developed by SC29/WG 8) and “ISO 23494 Biotechnology – Provenance Information Model for Biological Material and Data”. The objective is that the representation of the metadata associated to sequencing and annotation data supported by ISO 23092 standard family will be included by ISO 23492 in the wider unique standard definition of provenance information associated to each biological samples. Liaisons and common work items have been established to achieve the identified goals.

A liaison with TC 34/SC 16 “Horizontal methods for molecular biomarker analysis” has been established to proceed with possible common standardization activities. The Convenor of TC 34/SC 16 has provided use cases and information of relevant standards requiring usage/support of ISO/IEC 23092 standard series. Exchange of data and the validation of the standard support of ISO/IEC 23092 for the generation of compressed and indexed databases of genomic information according to the TC 34/SC 16 requirements are planned to define the common standardization goals and activities.

3.10.2 WG 8 current work items

WG 8 is progressing the development of the following standards.

- ISO/IEC CD 23092-1 Information technology — Genomic information representation — Part 1: Transport and storage of genomic information
- ISO/IEC 23092-1:2020/CD Amd 1 ed.2 Information technology — Genomic information representation — Part 1: Transport and storage of genomic information — Amendment 1: Support for Part 6
- ISO/IEC DIS 23092-2 Information technology — Genomic information representation — Part 2: Coding of genomic information
- ISO/IEC 23092-4:2020/CD Amd 1 ed.1 Information technology — Genomic information representation — Part 4: Reference software — Amendment 1: Version 2 and Part 6
- ISO/IEC 23092-5:2020/WD Amd 1- Information technology — Genomic information representation — Part 5: Conformance — Amendment 1: Version 2 and Part 6 support
- ISO/IEC FDIS 23092-6- Information technology — Genomic information representation — Part 6: Coding of genomic annotations

3.10.3 WG 8 published deliverables

Since September 2021, the following deliverables have been published that are allocated to SC 29/WG 8:

- ISO/IEC 23092-3:2022- Information technology — Genomic information representation — Part 3: Metadata and application programming interfaces (APIs)