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

JTC 1/SC 29 Business Plan 2024

Document type	Related content	Document date	Expected action
General / Other		2024-09-10	**COMMENT/REPLY** by 2024-10-07

Description

This document is circulated for review and consideration at the November 2024 JTC 1 Plenary.

JTC 1/SC 29 Business Plan

Coding of Audio, Picture, Multimedia and Hypermedia Information

Period covered: September 2023 – August 2024

Produced and submitted by
Gary J. Sullivan, Chair
(gary.sullivan@dolby.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.

SC 29 was established in 1991. 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 (WG 1) is known as **JPEG**, the Joint Photographic Experts Group, and the other seven working groups (WGs 2–8), together with three of the advisory groups (AGs 2, 3 and 5), are collectively known as **MPEG**, the Moving Picture Experts Group.

SC 29 and its WGs and AGs comprise a community of more than a thousand people, with about 500 people attending each of the quarterly sets of meetings of its JPEG and MPEG WGs, and roughly a thousand total registered experts. Following the COVID-19 emergency that caused all meetings to be held in virtual mode for more than three years (as with most of the rest of JTC 1), the meetings of the WGs and AGs of SC 29 have returned to primarily being held at physical locations since mid-2023, although they are generally held in hybrid mode (with remote access provided on a best-effort basis). The shift back from exclusive use of virtual mode has been accomplished without major difficulty, while an interest in holding some meetings in fully virtual mode has persisted, with the WGs typically planning to continue holding one or two of their meetings per year (out of a total of four) in virtual mode.

SC 29's work is well known and widely used in high-volume consumer technology as well as professional settings. SC 29 has continually served the ever-growing need for advances in digital media technology, including leading the way in the introduction of consumer technologies including digital photography, digital music, digital video broadcast and consumer disc storage in the early 1990s and through many new developments such as high-definition and ultra-high-definition television, multichannel and immersive sound experiences, high-dynamic-range video, 3D and free-viewpoint media experiences, and online streaming media. The work of SC 29 has been recognized by nine Emmy Awards. Within the last five years, this included one in 2019 for the JPEG standard

(ISO/IEC 10918-1), one in 2021 for the ISO Base Media File Format standard (ISOBMFF, ISO/IEC 14496-12), and two in 2022 for the MPEG Dynamic Adaptive Streaming over HTTP (DASH, ISO/IEC 23009-1) and Open Font Format (ISO/IEC 14496-22) standards.

A key area of recent focus in the work of SC 29 that is important to consumer use of media is media authentication and provenance, a timely need with the emerging prevalence and societal recognition of deep fake media content such as that recently demonstrated by generative AI technology. The WG 1 JPEG working group has led the way in recent efforts with its JPEG Trust family of standards (ISO/IEC 21617). The DIS ballot for the core foundation standard for JPEG Trust closed in July 2024, and development of the standard will be completed by the time of the current (November 2024) meeting of JTC 1. A statement issued by Prof. Touradj Ebrahimi, the Convenor of the JPEG working group said "JPEG Trust will bring trustworthiness back to imaging with specifications under the governance of the entire International community and stakeholders as opposed to a small number of companies or countries." JPEG Trust has been developed in collaboration with the Coalition for Content Provenance and Authenticity (C2PA). Closely related exploration studies are under way in other working groups of SC 29 as well, specifically including WG 2 MPEG Requirements, WG 3 MPEG Systems, and the WG 5 Joint Video Experts Team with ITU-T SG 16, and further dialog on the subject is under way with ITU-T SG 16. These recent efforts add to related prior technologies such as those specified in MPEG 21 (ISO/IEC 21000), MPEG DASH (ISO/IEC 23009), MPEG-4 (ISO/IEC 14496), JPSec (ISO/IEC 15444-8), and JPEG Systems (ISO/IEC 19566).

Other recent areas of new innovation in SC 29 for emerging consumer applications include recent work on 1) "coding for machines" to address machine analysis requirements for media coding including imaging, audio and video content, 2) point cloud coding, e.g. for LiDAR and volumetric scanners, and 3) event-based vision technology for new imaging applications with very-high time resolution for rapid detection and machine analysis of visual events.

As of August 2024, SC 29 has 31 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.

2.0 Environment

Recent advances in technology that have affected SC 29 include the emerging uses of neural network technology for many applications including media content production and many areas of signal processing. Neural network technology advances have driven increased interest in media authentication and media provenance tracking as well as new government regulatory initiatives. Immersive media applications and technologies have continued to expand, including more types of immersive video and audio, virtual, augmented and extended reality, and "metaverse" use cases.

See section 4 for a discussion of collaboration relationships and joint projects in SC 29.

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 (a project of the Joint Development Foundation, which became a PAS submitter to JTC 1 in 2020) produced the AV1 video coding specification in 2018 and has been working on future video coding technology. AVS (Audio Video coding Standard Workgroup of China), working recently together with a working group in IEEE, has developed AVS3 video coding technology. The Society of Motion Picture and Television Engineers (SMPTE) has also published various video coding standards. SC29 will continue to 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 completed the PAS transposition process for the glTF Specification (ISO/IEC 12113:2022). SC 29 members have had a keen interest in these developments, and liaison communication has been conducted 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 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.

See also the discussion of cooperation and collaboration relationships in section 4.

3.0 Achievements and benefits

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.

SC 29's work is well known and widely used in high-volume consumer technology as well as professional settings. SC 29 has continually served the ever-growing need for advances in digital media technology, including leading the way in the introduction of digital photography, digital music, digital video broadcast and consumer disc storage in the early 1990s and through many new developments such as high-definition and ultra-high-definition television, multichannel and immersive sound experiences, high-dynamic-range video, 3D and free-viewpoint media experiences, and online streaming media. The work of SC 29 has been recognized by nine Emmy Awards. Within the last five years, this included one in 2019 for the JPEG standard (ISO/IEC 10918-1), one in 2021 for the ISO Base Media File Format standard (ISOBMFF, ISO/IEC 14496-12), and two in 2022 for the MPEG Dynamic Adaptive Streaming over HTTP (DASH, ISO/IEC 23009-1) and Open Font Format (ISO/IEC 14496-22) standards.

Most of SC 29's standards contribute to Sustainable Development Goals 9 (Industry, Innovation and Infrastructure) and/or 12 (Responsible Consumption and Production).

4.0 Participation and cooperation/collaboration.

As of August 2024, SC 29 has 31 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.

Approximately 60 people attended each of the last two plenary meetings of SC 29, including representatives of about 25 P-members. SC 29 typically meets twice per year, and its WGs meet four times per year. Most meetings are held in hybrid mode (with remote access provided on a best-effort basis), and some meetings are also held in virtual-only mode. Altogether about 500 people attend each set of quarterly meetings of the WGs.

Internal and external liaison relationships with SC 29 are listed at https://www.iso.org/committee/45316.html.

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 collaborations with ITU-T. The Joint Photographic Experts Group (JPEG) and the Joint Video Experts Team (JVET) have 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) maintained by SC 29 WG 3 (MPEG Systems) is also maintained as common text with ITU-T SG 16.

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 also been standardizing minor extensions 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. This particularly includes exploration studies of the potential for neural-network based video coding and enhanced compression capabilities using more classical signal processing technology.

A number of standards developed 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 the JPEG XR standard. More recently, WG 1 multiple offers to ITU-T SG 16 for joint standardization projects around JPEG XS, JPEG XL, and JPEG Trust, to mention a few, have been unsuccessful, with unenthusiastic reactions for which it may be interesting to explore the reasons.

On the other hand, WG 1 has successfully coordinated and collaborated with the Coalition for Content Provenance and Authenticity (C2PA), resulting in a fully interoperable JPEG Trust standard, currently at the DIS stage, to address the growing challenge of authenticity in multimedia.

Several WGs of SC 29 are actively collaborating with 3GPP SA4 through liaison relationships and the participation of key experts in both organizations. MPEG Systems (WG 3), in particular successfully developed the MPEG DASH standard through this collaboration and continuously working together to develop further improvements. Recently WG 3 has also received requests from 3GPP SA4 to develop standards for their use cases and started related projects, e.g. for the Messaging Media Application Format (MeMAF).

See also the discussion of related organizations in the environment description in section 2.

5.0 Objectives and strategies to achieve them

SC 29 is acknowledged worldwide as the primary center of excellence in media coding and media communication systems for a very broad range of applications, continually serving the needs of the industry at large and cooperating with other SDOs and consortia to apply its expertise to many application-specific communities.

The standards produced in SC 29 are designed to represent, package, record, preserve and convey digital media information. These 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.

Compression technology, expressed in the SC 29 scope description as “efficient coding” is a major focus in SC 29, spanning across images, audio, moving pictures / video, graphics, haptics, genomic data, etc.

SC 29 has a structure with 8 working groups and 5 advisory groups to cover the specializations across this domain and to coordinate between them. Digital information support for media technology in systems is also standardized in SC 29 and is one of its most active areas (especially in WG 3, which is the most active WG in SC 29 in terms of the number of work items under development).

The structure of SC 29 has been stable since a restructuring in 2020 that increased the number of working groups to strengthen the topic-specific expertise by establishing corresponding groups while also establishing an advisory structure to ensure cooperation between them. No significant changes to that structure have been proposed by NBs since this restructuring.

6.0 Factors affecting completion and adoption of the work program

SC 29 has sufficient participation of experts to proceed with the development of its work programme, including the JPEG, JPEG 2000, AIC, JPEG XR, JPEG XT, JPEG Systems, JPEG AIC, JPEG XS, JPEG XL, JPEG Pleno, JPEG AI, JPEG Systems, JPEG Trust, 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 500 people attending each quarterly set of the meetings of its WGs and roughly a thousand total registered experts.

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

- **Succession planning**: As with the rest of JTC 1, SC 29 has been considering succession planning and how that planning can be conducted. Having been recently renewed, the terms of most SC 29 convenorships as well as the term of the chair will end at the end of 2026.
- **Identifying new opportunities for standardization and attracting stakeholders**: For ensuring the development of relevant standards to address industry needs, a strong collaborative atmosphere must be maintained and continuing outreach is needed to attract the widest possible range of stakeholders. Studies of requirements for the development of standards are critical, and SC 29 has established both a requirements subgroup within its WG 1 and also an MPEG Technical Requirements working group (WG 2) where such activities take place. Coordination on these matters is further facilitated by the JPEG and MPEG coordination advisory group (AG 4) and outreach is assisted by the JPEG liaison subgroup and the MPEG liaison and communications advisory group (AG 3). Stakeholder engagement, e.g., including consumer stakeholder involvement, is needed.
- **Keeping the work plan aligned with industry and consumer needs**: Any standardization effort must be aligned with the actual needs of industry and consumers. Stakeholder requirements must be fulfilled. This includes technical capabilities, the timing of the development process, and the quality of the provided solution, along with sufficient outreach, evangelism and educational material. These require a critical mass of competent experts and continual attention from the management of the groups involved.
- **Vulnerability to key adoption decisions**: While we may strive to develop standards that are appropriate for a broad application environment, in some cases the breadth of adoption of a standard is vulnerable to the decisions of key high-volume product makers. In such cases, the early signs of momentum toward deployment that arose during development of a standard can dissipate, limiting the prospects for widespread adoption. While this phenomenon may be partly unavoidable, it can potentially be minimized by ensuring that the key stakeholders are engaged in the process of developing the standards.
- **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 process. 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, and substantial disputes have arisen over SC 29 projects and standards in the past. SC 29 and its WGs continue to encourage their members to submit patent rights declarations and abide by the

ISO/IEC patent rights policy. 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. Such developments present challenges to the successful development and adoption of SC 29 standards.

- **Lack of free public availability of publications**: 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 the published texts and the work under development externally available (and available without cost) would improve alignment with this new industry practice. SC 29 welcomes the recent developments reported to the May JTC 1 meeting that improve the ability to make some deliverables available to the public without cost and to make this availability more clear and discoverable to the public.
- **Delays and problems in preparation for 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 welcomes the recent efforts in JTC 1 and ISO to address this 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.

7.0 Structure, current projects, and publications

SC 29 was established in 1991. 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 (WG 1) is known as **JPEG**, the Joint Photographic Experts Group, and the other seven working groups (WGs 2–8), together with three of the advisory groups (AGs 2, 3 and 5), are collectively known as **MPEG**, the Moving Picture Experts Group. The SC 29 structure is illustrated in Figure 1, and the current work

programme of the more than 100 standards and amendments under development in SC 29 is found at https://www.iso.org/committee/45316/x/catalogue/p/0/u/1/w/0/d/0.

The terms of service for the Chair of SC 29 and all but one of the convenorships within it expired at the end of 2023. The Convenor of WG 6 MPEG Audio, Schuyler Quackenbush of the United States, retired with distinction at the end of his term and was replaced by Thomas Sporer of Germany. The other Convenors and the Chair of SC 29 were appointed to an additional three-year term starting at the beginning of 2024, except for the AG 1 Convenor, whose term will expire at the end of this year. A call for volunteers for that position was issued by SC 29 in March 2024, and the NB of Japan responded in May 2024 with a nomination of Teruhiko Suzuki, a former Chair of SC 29, as a candidate for that position.

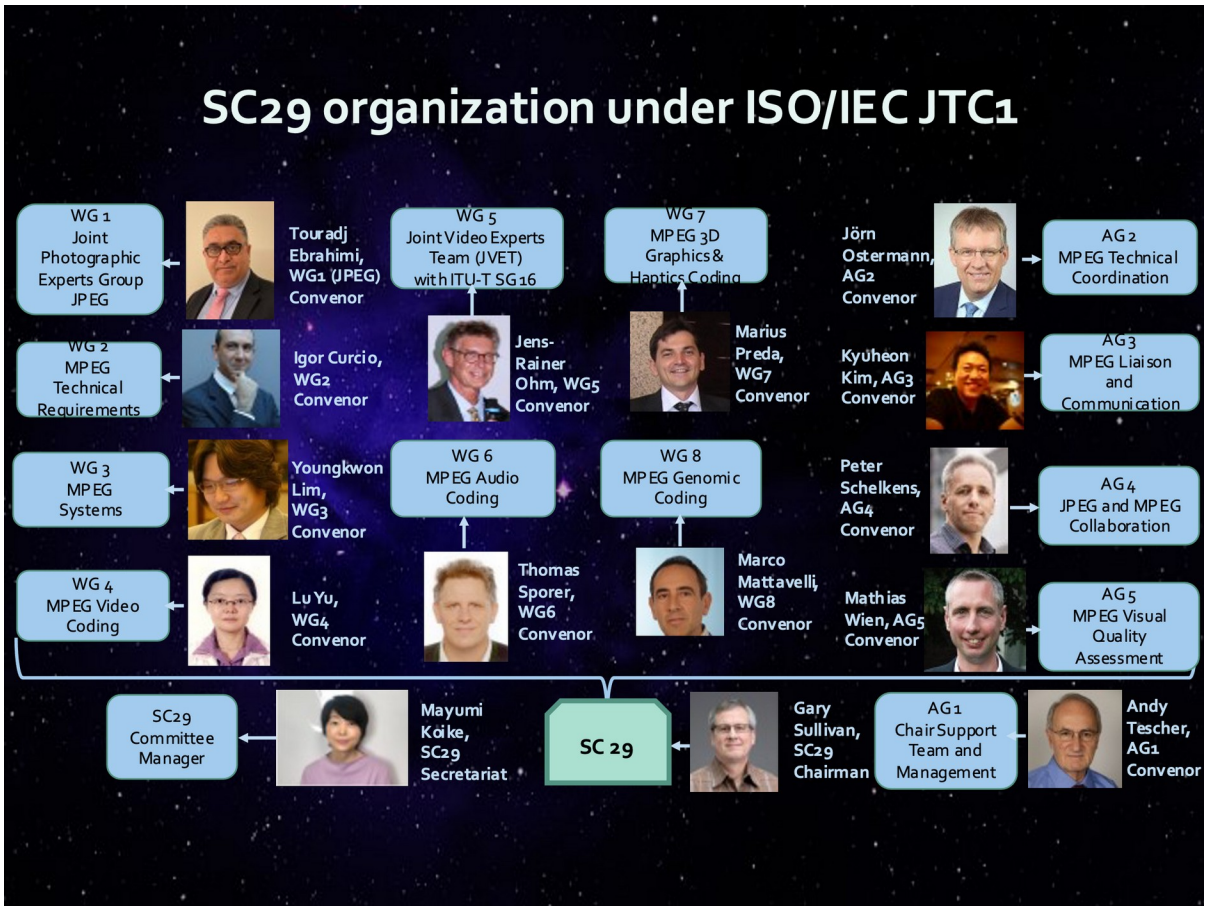

Figure 1: Structure of ISO/IEC JTC 1/SC 29

WG 1 (JPEG) has a long tradition of creating still image coding standards, dating back nearly four decades. WG 1 (JPEG) is responsible for the popular JPEG, JPEG 2000, JPEG XR, JPSearch, JPEG XT, and, more recently, the JPEG XS, JPEG Systems, JPEG AIC, JPEG Pleno, JPEG XL, JPEG AI, and JPEG Trust families of imaging standards.

WG 1 (JPEG) is organized around six subgroups, Communication, Requirements, Systems & Integration, Plenoptic Coding & Quality, Image Coding & Quality and Coding and Performance for Machines. It is furthermore coordinated by a JPEG Executive Team (JET), composed of Subgroup chairs and the webmaster, and is chaired by the WG 1 Convenor.

JPEG meets usually four times a year, with as much as possible alternation between virtual and hybrid modes. It also alternates the time zones and locations of its meetings to avoid penalizing the same experts worldwide. WG 1 (JPEG) celebrated its 100th meeting in Covilha, Portugal, in July 2023 with around 100 in-person participants. Typical WG 1 meetings have about 100 registered experts, of which around 80 actually participate, among which about 50 are in-person when meetings are in hybrid mode.

The roadmap of current projects in the JPEG WG is presented in Figure 2, excluding those regularly under maintenance.

Public-facing information about JPEG can be found at http://www.jpeg.org/.

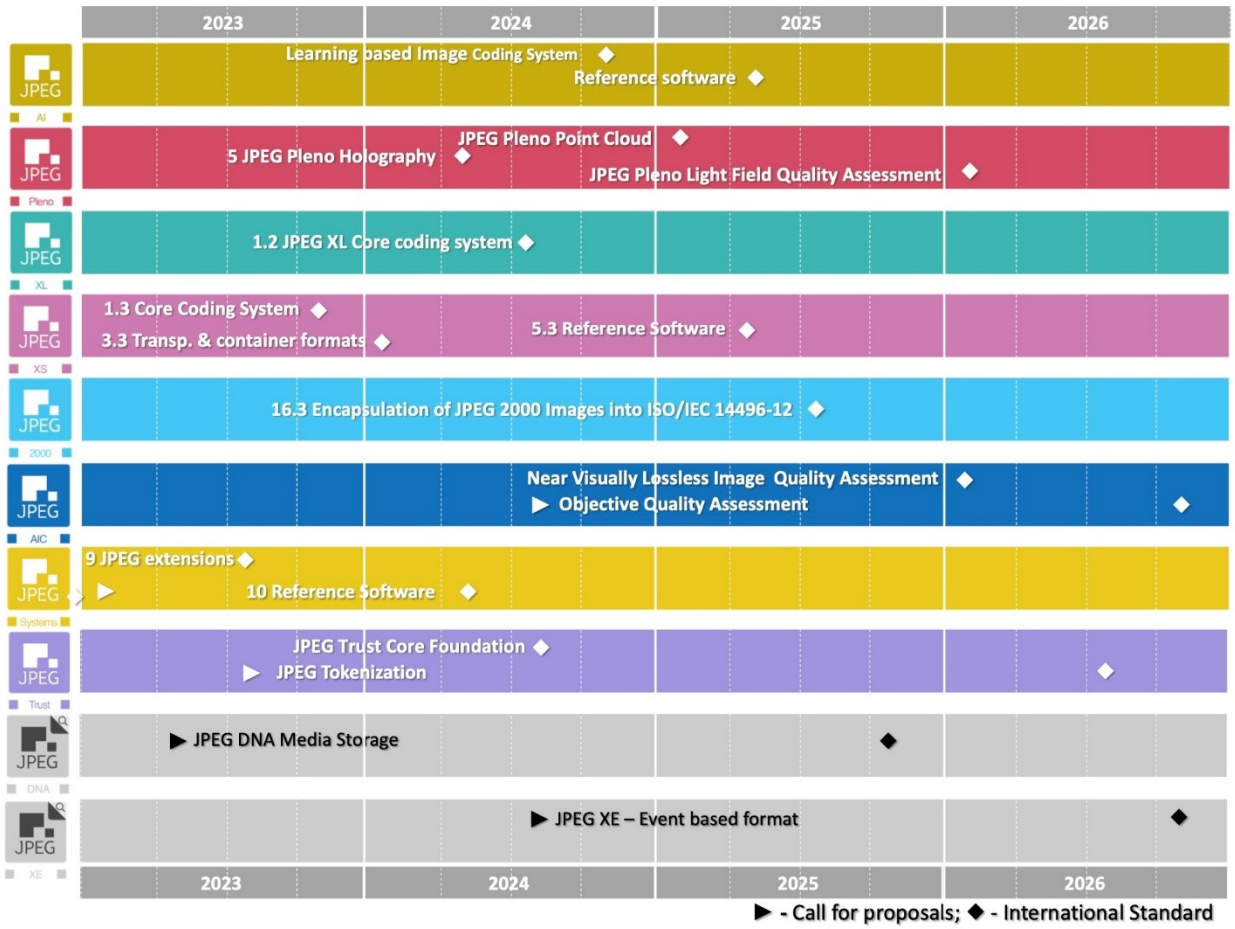

Figure 2: Roadmap of standards development in the JPEG WG (WG 1), as of August 2024

WGs 2-8 and AGs 2, 3 and 5 constitute the "MPEG" working groups. Over a span of 35 years since its founding in 1988, MPEG has become a center of expertise for the development of successful digital media 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, coordination of meeting plans including collocation of physical meeting locations, 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, and common work on Requirements coordinated by SC 29/WG 2.

The meetings of the MPEG WGs are collectively referred to as MPEG meetings, with the October 2023 meetings called the 144th MPEG meeting, the January 2024 meetings called the 145th MPEG meeting, etc. At each of these meetings, more than 400 participants have continued to work efficiently on standards for the future needs of the industry. The MPEG groups have also collaborated with each other on continued exploration of 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 3, and explorations for potential future standardization are shown in Figure 4 along with major maintenance items for enhancements of existing standards.

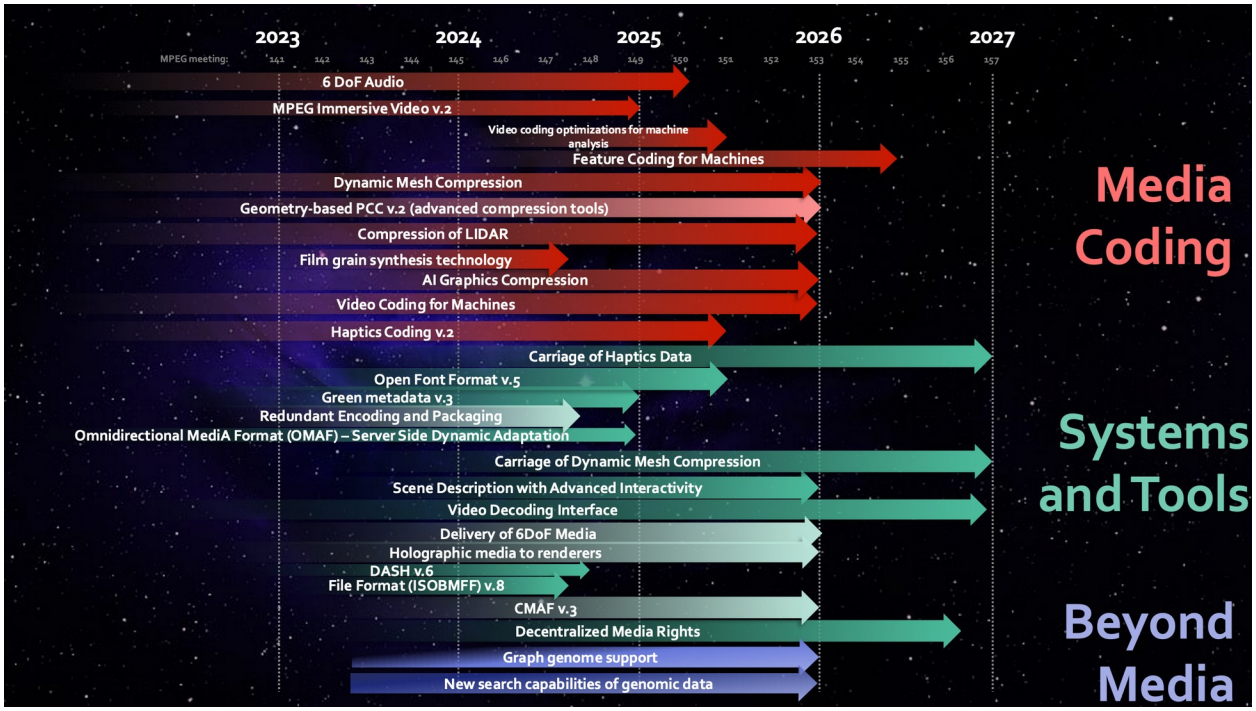

Figure 3: Roadmap of standards development in the MPEG WGs (WGs 2-8), as of August 2024

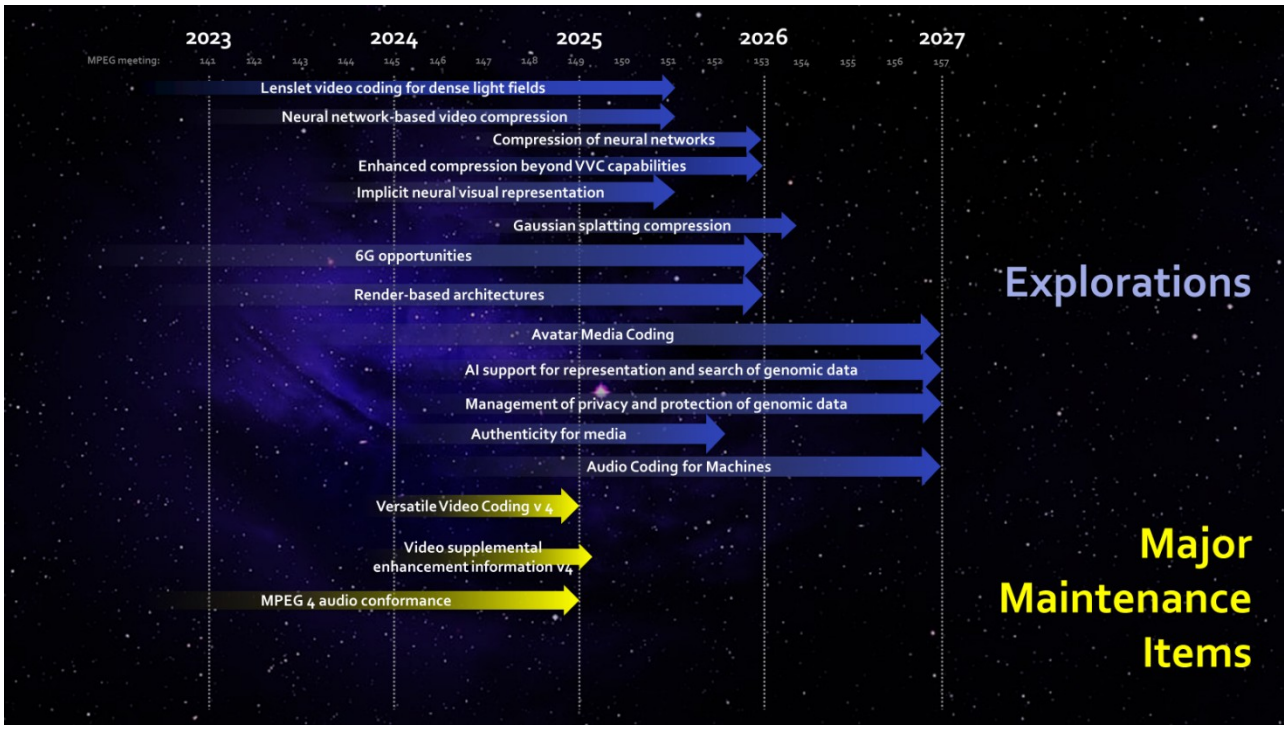

Figure 4: Exploration activities and major maintenance items in the MPEG WGs (WGs 2-8), as of August 2024

More specifically, 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 Experts 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. Within these WGs, than 400 participants have 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.

There has been an ongoing and 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 increasingly appeared in multiple projects within SC 29, and a large amount of further future impact of such technology is anticipated.

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.

Major accomplishments in the WGs are described in subsequent sections of this document, organized chronologically on a per-meeting basis.

Workshop organization and participation

In this reporting period, for outreach, coordination and education, several workshops were held in SC 29 as follows:

- A Light Field workshop on 22 November 2023 for discussion of standardization of light field coding technology.
- A workshop on event-based vision technology on 24 October 2023 toward "JPEG XE" exploration activity.
- A workshop on MPEG and JPEG Systems Activities on 3 April 2024.
- A workshop on JPEG and MPEG Emerging Activities on 27 February 2024.

The SC 29 Chair also presented and participated at the 6 December 2023 ISO and IEC 4th ISO/IEC AI Workshop Organized by SC 42 on "AI Applications Applications Roundtable – Streams, screens and synthetic dreams: how AI is transforming media and photography". The workshop material can be found at https://jtc1info.org/technology/subcommittees/ai/workshops/.

White papers for educational outreach

An example of the outreach and educational activities of SC 29 is the production of the following informally publicly released white papers on the following topics.

- The JPEG Trust framework
- Neural Network Coding (NNC) – Efficient Storage and Inference of Neural Networks for Multimedia Applications
- MPEG-I Immersive Audio

Verification testing

In this reporting period, the following verification test was completed:

- VVC multi-layer coding: Content layering

Preparations for the following verification and subjective evaluation test were a further focus of work:

- VVC multi-layer coding: Spatial scalability
- Subjective quality testing of film grain characteristics supplemental enhancement information message

Future project explorations

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

- Authorship and Ownership Signalling Mechanisms in Imaging Ecosystems
- Watermarking of Images for Authenticity and Copyright Protection
- Multimedia Content Tokenization
- Audio Coding for Machines (ACoM) to enable machine analysis of audio signal fields
- Coding of Immersive Contents represented using Neural Radiance Fields (NeRFs) and Gaussian Splatting
- DNA-based image coding for storage and archival
- Coding of point clouds generated by advanced LiDAR and volumetric scanners
- Implicit neural visual representation
- Neural-network based video coding
- Video coding with compression capability beyond that of the Versatile Video Coding (VVC) standard
- Video coding using generative face rendering technology with neural networks
- Avatar coding for visual communication
- Lenslet-view video coding for dense light fields
- AI-based technologies for graphics compression
- AI support for representation and searching of genomic data

7.1 Developments of the 101st Meeting of JPEG (October-November 2023)

The 101st JPEG meeting was held online, from the 30th of October to the 3rd of November 2023.

At this meeting, JPEG Trust became a Committee Draft. JPEG AI also reached Committee Draft status, reflecting the large number of core experiments held after the 100th JPEG meeting. In addition, JPEG analyzed the responses to its Calls for Proposals for JPEG DNA.

The following sections summarize the main highlights of the 101st JPEG meeting.

7.1.1 JPEG Trust

The 101st meeting marked an important milestone for JPEG Trust project with its Committee Draft (CD) for Part 1 “Core Foundation” (ISO/IEC 21617-1) of the standard approved for consultation. It was expected that a Draft International Standard (DIS) of the Core Foundation would be approved at the 102nd JPEG meeting in January 2024. This rapid schedule is necessitated by the speed at which fake media and misinformation are proliferating especially in respect of generative AI.

Aligned with JPEG Trust, the NFT Call for Proposals (CfP) yielded two expressions of interest, and submission of proposals was kept open through the 15th of January 2024.

Additionally, the Use Cases and Requirements document for JPEG Fake Media (the JPEG Fake Media exploration preceded the initiation of the JPEG Trust international standard) was updated to reflect the change to JPEG Trust as well as incorporate additional use cases that have arisen since the previous JPEG meeting, namely in respect of composited images. This document is available on the JPEG website.

"The release of the first Committee Draft of JPEG Trust is a strong signal that the JPEG working group is reacting with a timely response to demands for solutions that inform users when digital media assets are created or modified, in particular through Generative AI, hence contributing to bringing back trust into media-centric ecosystems.” said Prof. Touradj Ebrahimi, the Convenor of the JPEG working group.

7.1.2 JPEG AI

At the 101st meeting, the JPEG working group issued a request for re-establishing the JPEG AI (6048-1) project, along with a Committee Draft (CD) of its version 1. A new JPEG AI timeline has also been approved, where a Draft International Standard (DIS) of the Core Coding Engine of JPEG AI version 1 is foreseen at the 103rd JPEG meeting (April 2024), a rather important milestone for JPEG AI. The JPEG working group also established that JPEG AI version 2 will address requirements not yet fulfilled (especially regarding machine consumption tasks) but also significant improvements on requirements already addressed in version 1, e.g. compression efficiency. A final Call for Proposals for JPEG AI version 2 was expected to occur in January 2025 and the presentation and evaluation of JPEG AI version 2 proposals to occur in July 2025. During 2023, the JPEG AI Verification Model (VM) has evolved from a complex system (800kMAC/pxl) to two acceptable complexity-efficiency operation points, providing 11% compression efficiency gains at 20 kMAC/pxl and 25% compression efficiency gains at 200 kMAC/pxl. The decoder for the lower-end operating point has now been implemented on mobile devices and demonstrated during the 100th and 101st JPEG meetings. A presentation with the JPEG AI architecture, networks, and tools is now available. To avoid project

delays in the future, the promising input contributions from the 101st meeting were to be combined in JPEG AI Core Experiment 6.1 (CE6.1) to study interaction and resolve potential issues during the next meeting cycle. After this integration, a model was planned to be trained and cross-checked to be approved for release (JPEG AI VM5 release candidate) along with the study DIS text. Among promising technologies included in CE6.1 are high quality and variable rate improvements, with a smaller number of models (from 5 to 4), a multi-branch decoder that allows up to three reconstructions with different levels of quality from the same latent representation, but with synthesis transform networks with different complexity along with several post-filter and arithmetic coder simplifications.

7.1.3 JPEG Pleno Learning-based Point Cloud coding

The JPEG Pleno Learning-based Point Cloud coding activity progressed at the 101st meeting with a major investigation into point cloud quality metrics. The JPEG working group decided to continue this investigation into point cloud quality metrics as well as explore possible advancements to the VM in the areas of parameter tuning and support for residual lossless coding. The JPEG working group is targeting a release of the Committee Draft of Part 6 of the JPEG Pleno standard relating to Learning-based point cloud coding at the 102nd JPEG meeting in San Francisco, USA in January 2024.

7.1.4 JPEG Pleno Light Field

The JPEG working group has been creating several standards to provision the dynamic demands of the market, with its royalty-free patent licensing commitments. A light field coding standard has recently been developed, and JPEG Pleno is constantly exploring novel light field coding architectures.

The JPEG working group is also preparing standardization activities – among others – in the domains of objective and subjective quality assessment for light fields, improved light field coding modes, and learning-based light field coding.

A Light Field Industry Workshop was held on November 22nd, 2023, aiming at providing a forum for industrial actors to exchange information on their needs and expectations with respect to standardization activities in this domain.

7.1.5 JPEG AIC

During the 101st JPEG meeting, the AIC activity continued its efforts on the evaluation of the contributions received in April 2023 in response to the Call for Contributions on Subjective Image Quality Assessment. Notably, the activity is currently investigating three different subjective image quality assessment methodologies. The results of the newly established Core Experiments were planned be considered during the design of the AIC-3 standard, which has been carried out in a collaborative way since its beginning.

The AIC activity also initiated the discussion on Part 4 of the standard on Objective Image Quality Metrics (AIC-4) by refining the Use Cases and Requirements document. During the 102nd JPEG meeting in January 2024, the activity is planning to work on the Draft Call for Proposals on Objective Image Quality Assessment.

7.1.6 JPEG XE

The JPEG working group continued its activity on Event-based Vision. This activity revolves around a new and emerging image modality created by event-based visual sensors. JPEG XE aims at the creation and development of a standard to represent events in an efficient way allowing interoperability between sensing, storage, and processing, targeting machine vision and other relevant applications. For better dissemination and raising external interest, a workshop around Event-based Vision was organized and took place on Oct 24th, 2023. The workshop triggered the attention of various stakeholders in the field of Event-based Vision, who were expected to start contributing to JPEG XE. The workshop proceedings are available on jpeg.org. In addition, the JPEG working group created a minor revision for the Use cases and Requirements as v1.0, adding an extra use case on scientific and engineering measurements. Finally, a first draft of the Common Test Conditions for JPEG XE was produced, along with the first Exploration Experiments to start practical experiments in the coming 3-month period until the next JPEG meeting. The Ad-hoc Group on Event-based Vision was re-established to continue the work towards the next 102nd JPEG meeting in January of 2024.

7.1.7 JPEG DNA

As a result of the Call for Proposals issued by the JPEG working group for contributions to JPEG DNA standard, 5 proposals were submitted under three distinct codecs by three organizations. Two codecs were submitted to both coding and transcoding categories, and one was submitted to the coding category only. All proposals showed improved compression efficiency when compared to three selected anchors by the JPEG working group. After a rigorous analysis of the proposals and their cross checking by independent parties, it was decided to create a first Verification Model (VM) based on V- DNA, the best performing proposal. In addition, a number of core experiments were designed to improve the JPEG DNA VM with elements from other proposals submitted by quantifying their added value when integrated in the VM.

7.1.8 JPEG XS

The JPEG working group continued its work on JPEG XS 3rd edition. The primary goal of the 3rd edition is to deliver the same image quality as the 2nd edition, but with half of the required bandwidth. The Final Draft International Standard for Part 1 of the standard – Core coding tools – was produced at this meeting. With this FDIS version, all technical features are now fixed and completed. Part 2 – Profiles and buffer models – and Part 3 – Transport and container formats – of the standard were still in DIS ballot, with ballot results to be known by the end of January 2024. The JPEG working group is now working on Part 4 – Conformance testing, to provide the necessary test streams of the 3rd edition for potential implementors. A first Working Draft for Part 4 was issued. Completion of the JPEG XS 3rd edition was scheduled for April 2024 (Parts 1, 2, and 3) and Parts 4 and 5 were planned to follow shortly after that. Finally, the new Use cases and Requirements for JPEG XS document was created containing a new use case to use JPEG XS for transport of 4K/8K video over 5G mobile networks. It is expected that the new use case can already be covered by the 3rd edition, meaning that no further updates to the standard would be needed. However, more investigations and experimentations were to be conducted on this subject.

7.1.9 JPEG XL

The second editions of JPEG XL Part 1 (Core coding system) and Part 2 (File format) have proceeded to the FDIS stage, and the second edition of JPEG XL Part 3 (Conformance testing) has proceeded to the CD stage. These second editions provide clarifications, corrections and editorial improvements to facilitate independent implementations. At the same time, the development of hardware implementation solutions continues.

7.2 Developments of the 144th Meeting of MPEG (October 2023)

The 144th meeting of MPEG took place in Hanover, Germany, from 2023-10-16 until 2023-10-20. More information can be found at https://www.mpeg.org/meetings/mpeg-144/.

7.2.1 MPEG Visual Quality Assessment (AG 5) Issues Call for Learning-Based Video Codecs for Study of Quality Assessment

At the 144th MPEG meeting, MPEG Visual Quality Assessment (AG 5) issued a call for learning-based video codecs for study of quality assessment. AG 5 has been conducting subjective quality evaluations for coded video content and studying their correlation with objective quality metrics. Most of these studies focused on the High Efficiency Video Coding (HEVC) and Versatile Video Coding (VVC) standards. MPEG maintains the Compressed Video for study of Quality Metrics (CVQM) dataset for the purpose of this study.

Given the recent advancements in the development of learning-based video compression algorithms, MPEG studies compression using learning-based codecs. MPEG anticipates that different types of distortion would be present in a reconstructed video that has been compressed using learning-based codecs compared to those induced by traditional block-based motion-compensated video coding designs. In order to facilitate a deeper understanding of these distortions and their impact on visual quality, MPEG issued a call for learning-based video codecs for study of quality assessment. MPEG welcomes inputs in response to the call. Upon evaluating the responses, MPEG planned to invite those responses that meet the call's requirements to submit compressed bitstreams for further study of their subjective quality and potential inclusion into the CVQM dataset.

Given the continued rapid advancements in the development of learning-based video compression algorithms, MPEG planned to keep this call open and anticipated future updates to the call.

Interested parties are requested to contact the MPEG AG 5 Convenor Mathias Wien (wien@lfb.rwth-aachen.de) and submit responses for review at the 145th MPEG meeting in January 2024. Further details are given in the call, issued as AG 5 document N 104 and available from the mpeg.org website.

7.2.2 MPEG Technical Requirements (WG 2) Evaluates Call for Proposals on Feature Compression for Video Coding for Machines

At the 144th MPEG meeting, MPEG Technical Requirements (WG 2) evaluated the responses to the Call for Proposals (CfP) on Feature Compression for Video Coding for Machines (FCVCM). Feature Compression for Video Coding for Machines investigates technology directed towards compression of intermediate 'features' encountered within neural networks, enabling use cases such as

distributed execution of neural networks. This stands in contrast to Video Coding for Machines, which compresses conventional video data but with optimizations targeting machine consumption of the decoded video, rather than human consumption.

According to the 12 responses received to this CfP, the overall pipeline of FCVCM can be divided into two stages: (1) feature reduction and (2) feature coding. Technologies related to feature reduction include – but are not limited to – neural network-based feature fusion, temporal and spatial resampling, and adaptive feature truncation. Technologies related to feature coding include learning-based codecs, block-based exiting video codecs, and hybrid codecs.

All responses were evaluated on three tasks across four datasets. The results provide an overall gain, measured in average Bjøntegaard-Delta (BD) rate, of up to 94% against the feature anchors and 69% against the visual anchors. All requirements that were defined by WG 2 were addressed by different proposals and a test model has been defined.

Given the success of this call, MPEG planned to continue working on video feature compression methods for machine vision purposes. The work was to continue in MPEG Video Coding (WG 4) where a new standardization project was expected to be started and was planned to be completed and reach the status of Final Draft International Standard (FDIS) by July 2025.

WG 2 thanked the proponents who submitted responses to the CfP and the test administrator. MPEG planned to continue to collect and solicit feedback to improve the test model in the upcoming meetings.

7.2.3 MPEG Systems (WG 3) Progresses ISOBMFF-related Standards for the Carriage of Network Abstraction Layer Video Data

At the 144th MPEG meeting, MPEG Systems (WG 3) progressed the development of various ISO Base Media File Format (ISOBMFF) related standards.

As a part of the family of ISOBMFF-related standards, ISO/IEC 14496-15 defines the carriage of Network Abstract Layer (NAL) unit structured video data such as Advanced Video Coding (AVC), High Efficiency Video Coding (HEVC), Versatile Video Coding (VVC), Essential Video Coding (EVC), and Low Complexity Enhancement Video Coding (LCEVC). ISO/IEC 14496-15 has been further improved by adding support for enhanced features such as Picture-in-Picture (PiP) use cases particularly enabled by VVC, which resulted in the approval of the Final Draft Amendment (FDAM). Additionally, separately developed amendments have been consolidated in the 7th edition of ISO/IEC 14496-15, which has been promoted to Final Draft International Standard (FDIS), the final milestone of the standard development.

At the same time, the 2nd edition of ISO/IEC14496-32 (file format reference software and conformance) has been promoted to Committee Draft (CD), the first stage of standard development, and is planned to be completed and reach the status of Final Draft International Standard (FDIS) by the beginning of 2025. This standard was expected to be essential for industry professionals who require a reliable and standardized method of verifying the conformance of their implementation.

7.2.4 MPEG Systems (WG 3) Enhances the Support of Energy-Efficient Media Consumption

At the 144th MPEG meeting, MPEG Systems (WG 3) promoted the ISO/IEC 23001-11 Amendment 1 (energy-efficient media consumption (green metadata) for Essential Video Coding (EVC)) to Final Draft Amendment (FDAM), the final milestone of the standard development. This latest amendment defines metadata that enables a reduction in decoder power consumption for ISO/IEC 23094-1 (Essential Video Coding (EVC)).

At the same time, ISO/IEC 23001-11 Amendment 2 (energy-efficient media consumption for new display power reduction metadata) has been promoted to Committee Draft Amendment (CDAM), the first stage of standard development. This amendment introduces a novel way to carry metadata about display power reduction encoded as a video elementary stream interleaved with the video it describes. The amendment is expected to be completed and reach the status of Final Draft Amendment (FDAM) by the beginning of 2025. These developments represent a significant step towards more energy-efficient media consumption and a more sustainable future.

7.2.5 MPEG Systems (WG 3) Ratifies the Support of Temporal Scalability for Geometry-based Point Cloud Compression

At the 144th MPEG meeting, MPEG Systems (WG 3) promoted ISO/IEC 23090-18 Amendment 1 (support of temporal scalability) to Final Draft Amendment (FDAM), the final stage of standard development. The amendment enables the compression of a single elementary stream of point cloud data using ISO/IEC 23090-9 and storing it in more than one track of ISO Base Media File Format (ISOBMFF)-based files, thereby enabling support for applications that require multiple frame rates within a single file. The amendment introduces a track grouping mechanism to indicate multiple tracks carrying a specific temporal layer of a single elementary stream separately. The standard also provides information about reconstructing a single elementary stream from the data stored in more than one track, taking into consideration the frame rate suitable for specific applications.

7.2.6 MPEG Coding of 3D Graphics and Haptics (WG 7) Reaches the First Milestone for the Interchange of 3D Graphics Formats

At the 144th MPEG meeting, MPEG Coding of 3D Graphics and Haptics (WG 7) promoted ISO/IEC 23090-28 (efficient 3D graphics media representation for render-based systems and applications) to Committee Draft (CD), the first stage of standard development. This standard aims to streamline the interchange of 3D graphics formats. It primarily tackles the challenge of consistent asset interchange among prevalent 3D formats such as glTF, USD, ITMF, and others across multiple rendering platforms. For instance, a glTF scene might not render similarly on different renderers or players due to existing interchange limitations. ISO/IEC 23090-28 addresses this by introducing a comprehensive metadata vocabulary designed to ensure compatibility between popular 3D model formats on platforms like Unity Technologies and Unreal Engine. This standard delineates the initial mappings for the standard, starting with aligning ISO/IEC 23090-28 metadata to the glTF2.0 specification, most recently recognized as ISO/IEC 12113. This standard is planned to be completed, i.e., to reach the status of Final Draft International Standard (FDIS), by the beginning of 2025.

7.2.7 MPEG Genomic Coding (WG 8) Announces Completion of Coding of Genomic Annotations

At the 144th MPEG meeting, MPEG Genomic Coding (WG 8) announced the completion of ISO/IEC 23092-6 (coding of genomic annotations). This standard addresses the need to provide compressed representations of genomic annotations linked to the compressed representation of raw sequencing data and metadata.

ISO/IEC 23092-6 complements existing MPEG genomics standards to incorporate not only the primary (raw sequencing data) and secondary (aligned sequencing data) but also tertiary genomic data, including variant calls, gene expressions, mapping statistics, contact matrices (e.g., Hi-C), genomic tracks information, and functional annotations, which are collectively referred to as annotation data in the ISO/IEC 23092 series of standards, with efficient compression, indexing, and search capabilities. The formats specified in ISO/IEC 23092-6 also include advanced features such as selective encryption and signing of the data, auditing support, data provenance information, traceability, and support for direct linkage to external clinical data repositories expressed in common standard formats.

7.3 Developments of the 102nd Meeting of JPEG (January 2024)

The 102nd JPEG meeting was held in San Francisco, California, USA, from 22 to 26 January 2024. At this meeting, JPEG Trust became a Draft International Standard. Moreover, the responses to the Call for Proposals of JPEG NFT were received and analysed. As a consequence, relevant steps were taken towards the definition of standardized tools for certification of the provenance and authenticity of media content in a time where tools for effective media manipulation should be made available to the general public.

The following sections summarize the main highlights of the 102nd JPEG meeting.

7.1.1 JPEG Trust

At its 102nd meeting the JPEG working group produced the DIS (Draft International Standard) of JPEG Trust Part 1 “Core Foundation” (21617-1). It was expected that the standard would be published as an International Standard during the Summer of 2024. This rapid standardization schedule has been necessary because of the speed at which fake media and misinformation are proliferating especially with respect to Generative AI.

The JPEG Trust Core Foundation specifies a comprehensive framework for individuals, organizations, and governing institutions interested in establishing an environment of trust for the media that they use, and for supporting trust in the media they share online. This framework addresses aspects of provenance, authenticity, integrity, copyright, and identification of assets and stakeholders. To complement Part 1, a plan for a proposed new Part 2 “Trust Profiles Catalogue” was established. This new Part was planned to specify a catalogue of Trust Profiles, targeting common usage scenarios.

During the meeting, the WG also evaluated responses received to the JPEG NFT Final Call for Proposals (CfP). Certain portions of the submissions were planned to be incorporated in the JPEG

Trust suite of standards to improve interoperability with respect to media tokenization. As a first step, the WG planned to focus on standardization of declarations of authorship and ownership.

Finally, the Use Cases and Requirements document for JPEG Trust was updated to incorporate additional requirements in respect of composited media. This document is available on the JPEG website.

A white paper describing the JPEG Trust framework is also available on the JPEG website.

"In its efforts to provide standardized solutions to ascertain authenticity and provenance of the visual information, the JPEG working group has released the Draft international Standard of the JPEG Trust. JPEG Trust will bring trustworthiness back to imaging with specifications under the governance of the entire International community and stakeholders as opposed to a small number of companies or countries." said Prof. Touradj Ebrahimi, the Convenor of the JPEG working group.

7.1.2 JPEG AI

At the 102nd JPEG meeting, the JPEG AI Verification Model was improved by integrating nearly all the contributions adopted at the 101st JPEG meeting. The major change is a multi-branch JPEG AI decoding architecture with two encoders and three decoders (6 possible compatible combinations) that have been jointly trained, which allows the coverage of encoder and decoder complexity-efficiency tradeoffs. The entropy decoding and latent prediction portion is common for all possible combinations and thus differences reside at the analysis/synthesis networks. Moreover, the number of models has been reduced to 4, both 4:4:4 and 4:2:0 coding is supported, and JPEG AI can now achieve better rate-distortion performance in some relevant use cases. A new training dataset has also been adopted with difficult/high-contrast/versatile images to reduce the number of artifacts and to achieve better generalization and color reproducibility for a wide range of situations. Other enhancements have also been adopted, namely feature clipping for decoding artifacts reduction, improved variable bit-rate training strategy and post-synthesis transform filtering speedups.

The resulting performance and complexity characterization show compression efficiency (BD-rate) gains of 12.5% to 27.9% over the VVC Intra anchor, for relevant encoder and decoder configurations with a wide range of complexity-efficiency tradeoffs (7 to 216 kMAC/px at the decoder side). For the CPU platform, the decoder complexity is 1.6x/3.1x times higher compared to VVC Intra (reference implementation) for the simplest/base operating point. At the 102nd meeting, 12 core experiments were established to further continue work related to different topics, namely about the JPEG AI high-level syntax, progressive decoding, training dataset, hierarchical dependent tiling, spatial random access, to mention the most relevant. Finally, two demonstrations were shown where JPEG AI decoder implementations were run on two smartphone devices, Huawei Mate50 Pro and iPhone14 Pro.

7.1.3 JPEG Pleno Learning-based Point Cloud coding

The 102nd JPEG meeting marked an important milestone for JPEG Pleno Point Cloud with the release of its Committee Draft (CD) for ISO/IEC 21794-Part 6 "Learning-based point cloud coding" (21794-6). Part 6 of the JPEG Pleno framework brings an innovative Learning-based Point Cloud Coding technology adding value to existing Parts focused on Light field and Holography coding. It is expected that a Draft International Standard (DIS) of Part 6 would be approved at the 104th JPEG

meeting in July 2024 and the International Standard to be published during 2025. The 102nd meeting also marked the release of version 4 of the JPEG Pleno Point Cloud Verification Model updated to be robust to different hardware and software operating environments.

7.1.4 JPEG Pleno Light Field

The JPEG working group has recently produced a light field coding standard, and JPEG Pleno is constantly exploring novel light field coding architectures. The JPEG working group is also preparing standardization activities – among others – in the domains of objective and subjective quality assessment for light fields, improved light field coding modes, and learning-based light field coding.

As the JPEG working group seeks continuous improvement of its use case and requirements specifications, it organized a Light Field Industry Workshop. The presentations and video recording of the workshop that took place on November 22nd, 2023 are available on the JPEG website.

7.1.5 JPEG AIC

During the 102nd JPEG meeting, work on Image Quality Assessment continued with a focus on JPEG AIC-3, targeting standardizing a subjective visual quality assessment methodology for images in the range from high to nearly visually lossless qualities. The activity is currently investigating three different subjective image quality assessment methodologies.

The JPEG working group also launched the activities on Part 4 of the standard (AIC-4), by initiating work on the Draft Call for Proposals on Objective Image Quality Assessment. The Final Call for Proposals on Objective Image Quality Assessment is planned to be released in July 2024, while the submission of the proposals is planned for October 2024.

7.1.6 JPEG XE

The JPEG working group continued its activity on JPEG XE and event-based vision. This activity revolves around a new and emerging image modality created by event-based visual sensors. JPEG XE is about the creation and development of a standard to represent events in an efficient way allowing interoperability between sensing, storage, and processing, targeting machine vision and other relevant applications. The JPEG working group is preparing a Common Test Conditions document that provides the means to perform an evaluation of candidate technology for the efficient coding of event sequences. The Common Test Conditions provide a definition of a reference format, a dataset, a set of key performance metrics and an evaluation methodology. In addition, the WG is preparing a Draft Call for Proposals on lossless coding, with the intent to issue it publicly in April of 2024. Standardization will first start with lossless coding of event sequences as this seems to have the higher application urgency in industry. However, the WG acknowledges that lossy coding of event sequences is also a valuable feature, which will be addressed at a later stage. The Ad-hoc Group on Event-based Vision was reestablished to continue the work towards the next 103rd JPEG meeting in April of 2024.

7.1.7 JPEG DNA

During the 102nd JPEG meeting, the JPEG DNA Verification Model description and software were approved along with continued efforts to evaluate its rate-distortion characteristics. Notably, during the 102nd meeting, a subjective quality assessment was carried out by expert viewing using a new

approach under development in the framework of AIC-3. The robustness of the Verification Model to errors generated in a biochemical process was also analysed using a simple noise simulator. After meticulous analysis of the results, it was decided to create a number of core experiments to improve the Verification Model rate-distortion performance and the robustness to the errors by adding an error correction technique to the latter. In parallel, efforts are underway to improve the rate-distortion performance of the JPEG DNA Verification Model by exploring learning-based coding solutions. In addition, further efforts are defined to improve the noise simulator so as to allow assessment of the resilience to noise in the Verification Model in more realistic conditions, laying the groundwork for a JPEG DNA robust to insertion, deletion and substitution errors.

7.1.8 JPEG XS

The JPEG working group is happy to announce that the core parts of JPEG XS 3rd edition are ready for publication as International standards. The Final Draft International Standard for Part 1 of the standard – Core coding tools – was created at the last meeting in November 2023, and is scheduled for publication. DIS ballot results for Part 2 – Profiles and buffer models – and Part 3 – Transport and container formats – of the standard came back, allowing the JPEG working group to produce and deliver the proposed IS texts to ISO. This means that Part 2 and Part 3 3rd edition are also scheduled for publication.

At this meeting, the JPEG working group continued the work on Part 4 – Conformance testing, to provide the necessary test streams of the 3rd edition for potential implementors. A Committee Draft for Part 4 was issued. With Parts 1, 2, and 3 now ready, and Part 4 ongoing, the JPEG working group initiated the 3rd edition of Part 5 – Reference software. A first Working Draft was prepared and work on the reference software will start.

Finally, experimental results were presented on how to use JPEG XS over 5G mobile networks for the transmission of low-latency and high quality 4K/8K 360 degree views with mobile devices. This use case was added at the previous JPEG meeting. It is expected that the new use case can already be covered by the 3rd edition, meaning that no further updates to the standard would be necessary. However, investigations and experimentation on this subject continue.

7.1.9 JPEG XL

The second edition of JPEG XL Part 3 (Conformance testing) has proceeded to the DIS stage. Work on a hardware implementation continues. Experiments are planned to investigate HDR compression performance of JPEG XL.

7.2 Developments of the 145th Meeting of MPEG (January 2024)

The 145th meeting of MPEG took place online from 2024-01-22 until 2024-01-26. More information can be found at https://www.mpeg.org/meetings/mpeg-145/.

7.2.1 MPEG Systems (WG 3) Issues Latest Edition of the High Efficiency Image Format Standard to Unveil Cutting-Edge Features for Enhanced Image Decoding and Annotation

At the 145th MPEG meeting, MPEG Systems (WG 3) ratified the third edition of its High Efficiency Image Format (HEIF; ISO/IEC 23008-12: Image file format). HEIF has solidified its position as one of the most rapidly and widely adopted standards in the imaging industry. The newest edition

represents a significant leap forward, introducing progressive decoding capabilities that elevate image quality through a sequential, single-decoder instance process. This enhancement empowers users to decode a bitstream in successive steps, with each phase delivering perceptible improvements in image quality compared to the preceding step.

Additionally, this edition introduces a sophisticated data structure that meticulously describes the spatial configuration of the camera and outlines the distinctive characteristics of the camera responsible for generating the image content.

Furthermore, the updated HEIF specification encompasses innovative tools for annotating specific areas in diverse shapes, enhancing the versatility of image content manipulation. The inclusion of these annotation features adds a layer of creativity and customization, catering to the diverse needs of users across various industries.

Not stopping at these remarkable upgrades, the HEIF standard is actively advancing its technology portfolio. The ongoing development efforts promise to introduce support for renderable text items, providing a comprehensive solution for incorporating textual elements seamlessly into images. Additionally, the inclusion of slim versions of image files addresses the demand for efficient use cases, particularly for smaller image sizes such as icons.

7.2.2 MPEG Systems (WG 3) finalizes Standards supporting Interoperability Testing

At the 145th MPEG meeting, MPEG Systems (WG 3) finalized two standards comprising conformance and reference software by promoting it to the Final Draft International Standard (FDIS), the final stage of standards development. This pivotal milestone represents the culmination of rigorous standards development and underscores MPEG Systems' commitment to innovation and excellence in the field.

The finalized standards, ISO/IEC 23090-24 and ISO/IEC 23090-25, showcase the pinnacle of conformance and reference software for scene description and visual volumetric video-based coding data, respectively. These standards offer not only reference implementations but also essential bitstreams for conformance testing, ensuring robustness and reliability in real-world applications.

ISO/IEC 23090-24 focuses on conformance and reference software for scene description, providing a comprehensive reference implementation and bitstream tailored for conformance testing related to ISO/IEC 23090-14, scene description. This standard opens new avenues for advancements in scene depiction technologies, setting a new standard for conformance and software reference in this domain.

Similarly, ISO/IEC 23090-25 targets conformance and reference software for the carriage of visual volumetric video-based coding data. With a dedicated reference implementation and bitstream, this standard is poised to elevate the conformance testing standards for ISO/IEC 23090-10, the carriage of visual volumetric video-based coding data. The introduction of this standard is expected to have a transformative impact on the visualization of volumetric video data.

Both standards, ISO/IEC 23090-24 and ISO/IEC 23090-25, will be made freely accessible for download on the official ISO website, ensuring widespread availability for industry professionals,

researchers, and enthusiasts alike. This commitment to openness and accessibility aligns with MPEG Systems' mission to contribute to the broader technological community and foster collaboration.

7.2.3 MPEG Audio Coding (WG6) finalizes the Third Edition of MPEG-D Dynamic Range Control

At the 145th MPEG meeting, MPEG Audio Coding (WG6) completed the work on the third edition of ISO/IEC 23003-4, Dynamic range control, promoting it to the Final Draft International Standard (FDIS) stage. This update incorporates two amendments into the second edition, originally published in 2020.

The third edition includes the specification of dynamic range control (DRC) side chain information and metadata-based real-time loudness leveling for live workflows. The technologies enable producers of live content, such as sports broadcasts and concerts, to seamlessly integrate MPEG-D DRC-based loudness leveling into their existing workflows. The metadata-based approach offers highest possible quality of loudness processing and dynamic range control while maintaining full flexibility and control in playback devices. The technology can be tightly integrated with existing audio codecs such as MPEG-D USAC, MPEG-H Audio or any other audio codec supporting MPEG-D DRC.

7.2.4 MPEG Audio Coding (WG6) finalizes the Second Edition of MPEG-4 Audio Conformance

At the 145th MPEG meeting, MPEG Audio Coding (WG6) celebrated the completion of the second edition of ISO/IEC 14496-26, audio conformance, elevating it to the Final Draft International Standard (FDIS) stage. This significant update incorporates seven corrigenda and five amendments into the initial edition, originally published in 2010.

ISO/IEC 14496-26 serves as a pivotal standard, providing a framework for designing tests to ensure the compliance of compressed data and decoders with the requirements outlined in ISO/IEC 14496-3 (MPEG-4 Audio). The second edition reflects an evolution of the original, addressing key updates and enhancements through diligent amendments and corrigenda.

This latest edition, now at the FDIS stage, marks a notable stride in MPEG Audio Coding's commitment to refining audio conformance standards and ensuring the seamless integration of compressed data within the MPEG-4 Audio framework.

7.2.5 MPEG Genomic Coding (WG 8) extended MPEG Genomic Coding to support Transport and File Format for Genomic Annotations

At the 145th MPEG meeting, the MPEG Genomic Coding (WG 8) working group extended the support of transport and file format to the coding of any common type of annotations obtained by the analysis results of DNA sequencing data.

The ISO/IEC 23092-1 (3rd edition) – Transport and file format, supporting a joint coding of sequencing and annotation data, has been promoted to Final Draft International Standard (FDIS). The current MPEG-G standard series (ISO/IEC 23092) can now support full application pipelines, covering data representation and compression from the output of the sequencing up to the results of tertiary analysis support in a single structured transport and file format. The extended structured and compressed representation provides the basis for standard APIs implementing advanced standard browsing and searching features. They include standard APIs for exact and approximate

string-matching capabilities directly in the compressed domain for sequencing data metadata and annotations. These new standard functionalities are fundamental for searching large databases of compressed sequencing and annotation data resulting from the massive amounts of sequencing data that are generated by next generation sequencing technologies.

In addition, the MPEG Genomic Coding working group also reached the first milestone for the 2nd edition of ISO/IEC 23092-5 (MPEG-G Part 5, Conformance) by promoting the text to Committee Draft (CD) status. This new edition incorporates support for the newly issued Part 6 of the MPEG Genomics family of standards: the coding of genomic annotations. The conformance standard supports detailed diagnostic assessment of decoder implementations so that conformant implementations can be certified and can provide functional guarantees as required by regulations of diagnostic devices.

7.2.6 MPEG Liaison and Communication (AG 3) issues White Paper on Neural Network Coding (NNC)

At the 145th MPEG meeting, MPEG Liaison and Communication (AG 3) approved an MPEG white paper called *Neural Network Coding (NNC) – Efficient Storage and Inference of Neural Networks for Multimedia Applications*, which is available at https://www.mpeg.org/whitepapers/.

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. An additional factor for the exponential growth is the appearance of new use cases, such as federated learning with continuous communication between many devices. To effectively reduce bandwidth usage in communication and reduce the size of networks for inference, achieving an optimal compression ratio must be prioritized. 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), with the second edition adding new compression tools and support for coding incremental updates of neural networks.

Incremental coding, one of the main extensions in the second edition, targets neural network updates as a difference signal between a base neural network (i.e., an instance of a trained neural network for the particular use case) and an updated neural network. The updated neural network is typically the result of one of the following operations (this list is considered non-exhaustive), for example:

- The base neural network is retrained with other data or hyper-parameters.
- The base neural network and the updated neural network are compressed versions of the same network with different compression ratio.
- The updated neural network is the result of applying transfer learning, starting from the base neural network.
- The updated neural network uses the base neural network in its structure (possibly retrained end-to-end).

7.3 Developments of the 103nd Meeting of JPEG (April 2024)

The 103rd JPEG meeting was held online from April 8 to 12, 2024. During this JPEG meeting, the first learning-based standard, JPEG AI, reached the Draft International Standard (DIS) and was sent for balloting after a very successful development stage that led to performance improvements above 25% against its best-performing anchor, VVC. This high performance, combined with implementation in current mobile phones or the possibilities given by the latent representation to be used in image processing applications, leads to new opportunities and will certainly launch a new era of compression technology.

The following sections summarize the main highlights of the 103rd JPEG meeting.

7.3.1 JPEG AI

At its 103rd meeting the JPEG working group produced the Draft International Standard (DIS) of the JPEG AI Part 1 Core Coding Engine which is expected to be published as an International Standard in October 2024. JPEG AI offers a coding solution for standard reconstruction with significant improvements in compression efficiency over previous image coding standards at equivalent subjective quality. The JPEG AI coding design allows for hardware/software implementation encoding and decoding, in terms of memory and computational complexity, efficient coding of images with text and graphics, support for 8- and 10-bit depth, region of interest coding, and progressive coding. To cover multiple encoder and decoder complexity-efficiency tradeoffs, JPEG AI supports a multi-branch coding architecture with two encoders and three decoders (6 possible compatible combinations) that have been jointly trained. Compression efficiency (BD-rate) gains of 12.5% to 27.9% over the VVC Intra coding anchor, for relevant encoder and decoder configurations, can be achieved with a wide range of complexity tradeoffs (7 to 216 kMAC/px at the decoder side).

The work regarding JPEG AI profiles and levels (part 2), reference software (part 3) and conformance (part 4) has started and a request for sub-division has been issued in this meeting to establish a new part on the file format (part 5). At this meeting, most of the work focused on the JPEG AI high-level syntax and improvement of several normative and non-normative tools, such as hyper-decoder activations, training dataset, progressive decoding, training methodology and enhancement filters. There are now two smartphone implementations of JPEG AI available. In this meeting, a JPEG AI demo was shown running on a Huawei Mate50 Pro with a Qualcomm Snapdragon 8+ Gen1 with high resolution (4K) image decoding, tiling, full base operating point support and arbitrary image resolution decoding.

"The JPEG AI Draft International Standard is a yet another important milestone in an age where AI is rapidly replacing previous technologies. With this achievement, the JPEG working group has demonstrated its ability to reinvent itself and adapt to new technological paradigms, offering standardized solutions based on latest state-of-the-art technologies.” said Prof. Touradj Ebrahimi, the Convenor of the JPEG working group.

7.3.2 JPEG Trust

At the 103rd meeting, the JPEG working group produced an updated version of the Use Cases and Requirements for JPEG Trust (v2.0). This document integrates the use cases and requirements of the JPEG NFT exploration with the use cases and requirements of JPEG Trust. In addition, a new document with Terms and Definitions for JPEG Trust (v1.0) was published which incorporates all

terms and concepts as they are used in the context of the JPEG Trust activities. Finally, an updated version of the JPEG Trust White Paper v1.1 has been released. These documents are available on the JPEG Trust/Documentation page.

7.3.3 JPEG Pleno Learning-based Point Cloud coding

The JPEG working group continued its activity on Learning-based Point Cloud Coding under the JPEG Pleno family of standards. During the 103rd JPEG meeting, comments on the Committee Draft of IS0/IEC 21794 Part 6: “Learning-based point cloud coding” were received and the activity is on track for the release of a Draft International Standard for balloting at the 104th JPEG meeting in Sapporo, Japan in July 2024. A new version of the Verification Model (Version 4.1) was released during the 103rd JPEG meeting containing an updated entropy coding module. In addition, version 2.1 of the Common Training and Test Conditions was released.

7.3.4 JPEG Pleno Light Field

The JPEG Pleno Light Field activity progressed at this meeting with a number of technical submissions for improvements to the JPEG PLeno Model (JPLM). The JPLM provides reference implementations for the standardized technologies within the JPEG Pleno framework. The JPEG Pleno Light Field activity has an ongoing standardization activity concerning a novel light field coding architecture that delivers a single coding mode to efficiently code all types of light fields. This novel coding mode does not need any depth information resulting in significant improvement in compression efficiency.

The JPEG Pleno Light Field is also preparing standardization activities in the domains of objective and subjective quality assessment for light fields, aiming to address other plenoptic modalities in the future. During the meeting, important decisions were made regarding the execution of multiple collaborative subjective experiments aiming at exploring various aspects of subjective light field quality assessments. Additionally, a specialized tool for subjective quality evaluation has been developed to support these experiments. The outcomes of these experiments will guide the decisions during the subjective quality assessment standardization process. They will also be utilized in evaluating proposals for the upcoming objective quality assessment standardization activities.

7.3.5 JPEG AIC

During the 103rd JPEG meeting, the work on visual image quality assessment continued with a focus on JPEG AIC-3, targeting a standard for a subjective quality assessment methodology for images in the range from high to nearly visually lossless quality. The activity is currently investigating three kinds of subjective image quality assessment methodologies, notably the Boosted Triplet Comparison (BTC), the In-place Double Stimulus Quality Scale (IDSQS), and the In-place Plain Triplet Comparison (IPTC), as well as a unified framework capable of merging the results of two among them.

The JPEG working group has also worked on the preparation of the Part 4 of the standard (JPEG AIC-4) by initiating work on the Draft Call for Proposals on Objective Image Quality Assessment. The Final Call for Proposals on Objective Image Quality Assessment is planned to be released in January 2025, while the submission of the proposals is planned for April 2025.

7.3.6 JPEG XE

The JPEG working group continued its activity on JPEG XE and event-based vision. This activity revolves around a new and emerging image modality created by event-based visual sensors. JPEG XE is about the creation and development of a standard to represent events in an efficient way allowing interoperability between sensing, storage, and processing, targeting machine vision and other relevant applications. The JPEG working group finished the Common Test Conditions v1.0 document that provides the means to perform an evaluation of candidate technologies for efficient coding of event sequences. The Common Test Conditions define a canonical raw event format, a reference dataset, a set of key performance metrics and an evaluation methodology. In addition, the JPEG working group also finalized the Draft Call for Proposals on lossless coding for event-based data. This call will be finalized at the next JPEG meeting in July 2024. Both the Common Test Conditions v1.0 and the Draft Call for Proposals are available on jpeg.org. Standardization will start with lossless coding of event sequences as this has the most imminent application urgency in industry. However, the JPEG working group acknowledges that lossy coding of event sequences is also a valuable feature, which will be addressed at a later stage. The Ad-hoc Group on Event-based Vision was reestablished to continue the work towards the 104th JPEG meeting.

7.3.7 JPEG DNA

JPEG DNA is an exploration aiming at developing a standard that provides technical solutions that are capable of representing bi-level, continuous-tone grey-scale, continuous-tone colour, or multichannel digital samples in a format representing nucleotide sequences for supporting DNA storage. A Call for Proposals was published at the 99th JPEG meeting and based on performance assessment and a descriptive analysis of the solutions that had been submitted, the JPEG DNA Verification Model was created during the 102nd JPEG meeting. A number of core experiments were conducted to validate the Verification Model, and notably, the first Working Draft of JPEG DNA was produced during the 103rd JPEG meeting. Work towards the creation of the specification will start with newly defined core experiments to improve the rate-distortion performance of Verification Model and the robustness to insertion, deletion, and substitution errors. In parallel, efforts are underway to improve the noise simulator produced at the 102nd JPEG meeting to allow the assessment of the resilience to noise in the Verification Model in more realistic conditions and to explore learning-based coding solutions.

7.3.8 JPEG XS

The JPEG working group is happy to announce that the core parts of JPEG XS 3rd edition are ready for publication as International Standards. The Final Draft International Standard for Part 1 of the standard – Core coding tools – is ready, and Part 2 – Profiles and buffer models – and Part 3 – Transport and container formats – are both being prepared by ISO for immediate publication. At this meeting, the JPEG working group continued the work on Part 4 – Conformance testing, to provide the necessary test streams and test protocols to implementers of the 3rd edition. Consultation of the Committee Draft for Part 4 took place and a DIS version was issued. The development of the reference software, contained in Part 5, continued and the reference decoder is now feature-complete and fully compliant with the 3rd edition. A Committee Draft for Part 5 was issued at this meeting. Development of a fully compliant reference encoder is scheduled to be completed by July.

Finally, new experimental results were presented on how to use JPEG XS over 5G mobile networks for the wireless transmission of low-latency and high quality 4K/8K 360 degree views with mobile devices and VR headsets. More experiments will be conducted, but first results show that JPEG XS is capable of providing immersive and excellent quality of experience in VR use cases, mainly thanks to its native low-latency and low-complexity properties.

7.3.9 JPEG XL

The performance of JPEG XL on HDR images was investigated and the experiments will continue. Work on a hardware implementation continues, and further improvements are made to the libjxl reference software. The second editions of Parts 1 and 2 are in the final stages of the ISO process and will be published soon.

7.4 Developments of the 146th Meeting of MPEG (April 2024)

The 146th meeting of MPEG took place in Rennes, France, from 2024-04-22 until 2024-04-26. More information can be found at https://www.mpeg.org/meetings/mpeg-146/.

7.4.1 MPEG Technical Requirements (WG 2) issues Call for Proposals for AI-based Point Cloud Coding

At the 146th MPEG meeting, MPEG Technical Requirements (WG 2) issued a Call for Proposals (CfP) focusing on AI-based point cloud coding technologies. This initiative stems from ongoing explorations by MPEG into potential use cases, requirements, and the capabilities of AI-driven point cloud encoding, particularly for dynamic point clouds.

With recent significant progress in AI-based point cloud compression technologies, MPEG is keen on studying and adopting AI methodologies. MPEG is specifically looking for learning-based codecs capable of handling a broad spectrum of dynamic point clouds, which are crucial for applications ranging from immersive experiences to autonomous driving and navigation.

As the field evolves rapidly, MPEG expects to receive multiple innovative proposals. These may include a unified codec, capable of addressing multiple types of point clouds, or specialized codecs tailored to meet specific requirements, contingent upon demonstrating clear advantages. MPEG has therefore publicly called for submissions of AI-based point cloud codecs, aimed at deepening the understanding of the various options available and their respective impacts. Submissions that meet the requirements outlined in the call will be invited to provide source code for further analysis, potentially laying the groundwork for a new standard in AI-based point cloud coding. MPEG welcomes all relevant contributions and looks forward to evaluating the responses.

Interested parties are requested to contact the MPEG WG 2 Convenor Igor Curcio (igor.curcio@nokia.com) and MPEG WG 7 Convenor Marius Preda (marius.preda@it-sudparis.eu) to register their participation to the CfP and to submit responses for review at the 148th MPEG meeting in November 2024. Further details are given in the CfP, issued as WG 2 document N 365 and available from https://www.mpeg.org/meetings/mpeg-146/.

7.4.2 MPEG Technical Requirements (WG 2) issues Call for Interest in Object Wave Compression

At the 146th MPEG meeting, MPEG Technical Requirements (WG 2) issued a Call for Interest (CfI) in object wave compression. Computer holography, a 3D display technology, utilizes a digital fringe pattern called a computer-generated hologram (CGH) to reconstruct 3D images from input 3D models. Holographic near-eye displays (HNEDs) reduce the need for extensive pixel counts due to their wearable design, positioning the display near the eye. This positions HNEDs as frontrunners for the early commercialization of computer holography, with significant research underway for product development.

Innovative approaches facilitate the transmission of object wave data, crucial for CGH calculations, over networks. Object wave transmission offers several advantages, including independent treatment from playback device optics, lower computational complexity, and compatibility with video coding technology. These advancements open doors for diverse applications, ranging from entertainment experiences to real-time two-way spatial transmissions, revolutionizing fields such as remote surgery and virtual collaboration. As MPEG explores object wave compression for computer holography transmission, a Call for Interest seeks contributions to address market needs in this field.

Interested parties are requested to contact the MPEG WG 2 Convenor Igor Curcio (igor.curcio@nokia.com) or to submit inputs for review at the 147th MPEG meeting in July 2024. Further details are given in the Call for Interest, issued as WG 2 document N 377 and available from https://www.mpeg.org/meetings/mpeg-146/.

7.4.3 MPEG Systems (WG 3) reaches First Milestone for Fifth Edition of Open Font Format

At the 146th MPEG meeting, MPEG Systems (WG 3) promoted the 5th edition of ISO/IEC 14496-22 Open font format to Committee Draft (CD), marking the initial stage of standard development.

The importance of textual representation within multimedia content cannot be understated. In recognition of this, MPEG Systems has diligently pursued the standardization of interoperable font formats. With the commencement of its 5th edition, a pivotal milestone has been achieved. This latest iteration not only enhances the legibility of the specification but also transcends previous limitations, notably the 64K glyph encoding constraint in a single font file. By surpassing this barrier, the new edition facilitates the comprehensive coverage of the entire Unicode repertoire, accommodating diverse world languages and writing systems, including multiple glyph variants, within a singular font file.

Moreover, the latest edition introduces more space-efficient composite glyph representations, along with a myriad of novel features and capabilities tailored for variable fonts. This innovation culminates in substantial reductions in font file sizes and empowers the creation of parametric variable fonts utilizing higher order interpolations.

The development trajectory of this standard is projected for completion, culminating in the attainment of the Final Draft International Standard (FDIS) status by the conclusion of 2025.

7.4.4 MPEG Systems (WG 3) ratifies Second Edition of Scene Description

At the 146th MPEG meeting, MPEG Systems (WG 3) promoted the 2nd edition of ISO/IEC 23090-14 Scene description to Final Draft International Standard (FDIS), the final stage of standard development.

Since the inaugural release of the standard on immersive media scene description in 2022, the momentum in extending its capabilities has remained unwavering. The latest iteration seamlessly integrates two amendments into its predecessor, prioritizing enhanced user readability. Noteworthy advancements include the seamless integration of MPEG-developed immersive media objects, such as Video-based Point Cloud Compression (V-PCC, as specified in ISO/IEC 23090-5), and MPEG Immersive Video (MIV, as delineated in ISO/IEC 23090-12), within a scene framework. Furthermore, this edition fortifies support for a myriad of data types essential for immersive scenes, encompassing haptics, augmented reality, avatars, interactivity, and lighting, among others.

Looking ahead, MPEG Systems is steadfast in its commitment to advancing the standard's development, with plans to expand support to encompass MPEG-I immersive audio and beyond.

7.4.5 MPEG Video Coding (WG 4) reaches First Milestone for Second Edition of MPEG Immersive Video (MIV)

At the 146th MPEG meeting, MPEG Video Coding (WG 4) reached the Committee Draft (CD) stage of the 2nd edition of ISO/IEC 23090-12 MPEG immersive video (MIV), the first stage of standard development.

MIV was developed to support the compression of immersive video content, in which multiple real or virtual cameras capture a real or virtual 3D scene. 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 multi-view video plus depth (MVD) and multi-planar video (MPI) that leverages strong hardware support for commonly used video formats to compress volumetric video.

New features in the 2nd edition are coloured depth, capture device information, patch margins, background views, static background atlases, support for decoder-side depth estimation, chroma dynamic range modification, piecewise linear normalized disparity quantization, and linear depth quantization. These features provide additional functionality and improved performance.

The first edition of the standard included the MIV Main profile for MVD, the MIV Extended profile, which enables MPI, and the MIV Geometry Absent profile, which is suitable for use with cloud-based and decoder-side depth estimation. In the 2nd edition, the MIV 2 profile is being added, which is a superset of the existing profiles and covers all new functionality. In addition, a document entitled "profiles under consideration" was started to study the inclusion of narrower profiles in this edition.

Finally, it was expected that issuing a 2nd edition of ISO/IEC 23090-23 Conformance and reference software for MPEG immersive video would be requested at the next MPEG meeting.

7.4.6 MPEG Joint Video Experts Team with ITU-T SG 16 (WG 5, JVET) releases New Editions of AVC, HEVC, and Video CICP

At the 146th MPEG meeting, the MPEG Joint Video Experts Team with ITU-T SG 16 (WG 5), also known as JVET, promoted (i) the 11th edition of ISO/IEC 14496-10 Advanced Video Coding (AVC), (ii) the 5th edition of ISO/IEC 23008-2 High Efficiency Video Coding (HEVC), and (iii) the 3rd edition of ISO/IEC 23091-2 Video Coding-independent Code Points (Video CICP) to Final Draft International Standard (FDIS), the final stage of standard development.

The latest editions of AVC and HEVC now incorporate support for additional supplemental enhancemen information (SEI) messages, drawing from ISO/IEC 23002-7 Versatile Supplemental Enhancement Information (VSEI) Messages for Coded Video Bitstreams. Specifically, this includes the integration of (a) neural network post-filtering SEI message and (b) phase indication SEI message with these standards. HEVC has been expanded to include extended multiview profiles for 8-bit and 10-bit, as well as monochrome multiview profiles supporting standalone depth map coding with up to 16 bits. Additionally, the new version of Video CICP introduces additional color code points and implements text improvements and clarifications.

These advancements demonstrate a commitment to maintaining support for legacy standards developed jointly with ITU-T, ensuring their relevance to current market needs.

7.4.7 MPEG Joint Video Experts Team with ITU-T SG 16 (WG 5, JVET) Promotes Standard Development for Machine-Optimized Video Compression

At the 146th MPEG meeting, the MPEG Joint Video Experts Team with ITU-T SG16 (WG 5), also known as JVET, advanced ISO/IEC 23888-3 “Optimization of Encoders and Receiving Systems for Machine Analysis of Coded Video Content” as part 3 of MPEG AI to Committee Draft Technical Report (CDTR), marking the initial stage of standard development.

In recent years, the efficacy of machine learning-based algorithms in video content analysis has steadily improved. However, an encoder designed for human consumption does not always produce compressed video conducive to effective machine analysis. This challenge lies not in the compression standard but in optimizing the encoder or receiving system. The forthcoming technical report addresses this gap by showcasing technologies and methods that optimize encoders or receiving systems to enhance machine analysis performance.

Developed collaboratively with ITU-T SG16, this technical report will be published as a technically aligned twin text, corresponding to a forthcoming supplement or technical paper of ITU-T. It is available at https://www.mpeg.org/meetings/mpeg-146/.

7.4.8 MPEG Audio Coding (WG 6) reaches First Milestone for MPEG-I Immersive Audio

At the 146th MPEG meeting, MPEG Audio Coding (WG 6) promoted ISO/IEC 23090-4 MPEG-I immersive audio and ISO/IEC 23090-34 Immersive audio reference software to Committee Draft (CD) stage, the first stage of standard development. The MPEG-I immersive audio standard sets a new benchmark for compact and lifelike audio representation in virtual and physical spaces, catering to Virtual, Augmented, and Mixed Reality (VR/AR/MR) applications. By enabling high-quality, real-

time interactive rendering of audio content with six degrees of freedom (6DoF), users can experience immersion, freely exploring 3D environments while enjoying dynamic audio.

Designed in accordance with MPEG's rigorous standards, MPEG-I immersive audio ensures efficient distribution across bandwidth-constrained networks without compromising on quality. Unlike proprietary frameworks, this standard prioritizes interoperability, stability, and versatility, supporting both streaming and downloadable content while seamlessly integrating with MPEG-H 3D audio compression.

MPEG-I's comprehensive modeling of real-world acoustic effects, including sound source properties and environmental characteristics, guarantees an authentic auditory experience. Moreover, its efficient rendering algorithms balance computational complexity with accuracy, empowering users to finely tune scene characteristics for desired outcomes.

The release of the CD for ISO/IEC 23090-34 Immersive Audio Reference Software, which encompasses all aspects of the standard, facilitates real-time evaluation and adoption in industry and consumer applications. Interested parties can access both the text specification and reference software at https://www.mpeg.org/meetings/mpeg-146/, with additional insights available through a dedicated white paper released during this meeting.

7.4.9 MPEG Coding of 3D Graphics and Haptics (WG 7) reaches First Milestone for Video-based Dynamic Mesh Coding (V-DMC)

At the 146th MPEG meeting, MPEG Coding of 3D Graphics and Haptics (WG 7) reached the Committee Draft (CD) stage of ISO/IEC 23090-29 Video-based Dynamic Mesh Compression (V-DMC), the first stage of standard development. This standard represents a significant advancement in 3D content compression, catering to the ever-increasing complexity of dynamic meshes used across various applications, including real-time communications, storage, free-viewpoint video, augmented reality (AR), and virtual reality (VR). The standard addresses the challenges associated with dynamic meshes that exhibit time-varying connectivity and attribute maps, which were not sufficiently supported by previous standards.

Video-based Dynamic Mesh Compression promises to revolutionize how dynamic 3D content is stored and transmitted, allowing more efficient and realistic interactions with 3D content globally. The Committee Draft follows an extensive call for proposals issued by MPEG, which invited technology developers to submit innovations that could contribute to the new standard. Proposals were evaluated based on various objective and subjective metrics to ensure the selected technologies meet and exceed the current market and technical demands. MPEG extends its gratitude to all contributors who have submitted proposals and participated in the rigorous testing and evaluation process. The results of these evaluations have shaped the draft of the standard, ensuring it meets the high expectations and needs of the industry.

The Committee Draft of the Video-based Dynamic Mesh Compression standard is now available for further comments and evaluation by national bodies. It is available at https://www.mpeg.org/meetings/mpeg-146/. MPEG encourages continued participation from the community to finalize the standard for publication.

7.4.10 MPEG Coding of 3D Graphics and Haptics (WG 7) reaches First Milestone for Low Latency, Low Complexity LIDAR coding

At the 146th MPEG meeting, MPEG Coding of 3D Graphics and Haptics (WG 7) reached the Committee Draft (CD) stage of ISO/IEC 23090-30 Low Latency, Low Complexity LiDAR Coding, the first stage of standard development. This milestone underscores MPEG's commitment to advancing coding technologies required by modern LiDAR applications across diverse sectors. The new standard addresses critical needs in the processing and compression of LiDAR-acquired point clouds, which are integral to applications ranging from automated driving to smart city management. It provides an optimized solution for scenarios requiring high efficiency in both compression and real-time delivery, responding to the increasingly complex demands of LiDAR data handling.

LiDAR technology has become essential for various applications that require detailed environmental scanning, from autonomous vehicles navigating roads to robots mapping indoor spaces. The Low Latency, Low Complexity LiDAR Coding standard will facilitate a new level of efficiency and responsiveness in LiDAR data processing, which is critical for the real-time decision-making capabilities needed in these applications.

This Committee Draft builds on comprehensive analysis and industry feedback to address specific challenges such as noise reduction, temporal data redundancy, and the need for region-based quality of compression. The standard also emphasizes the importance of low latency coding to support real-time applications, essential for operational safety and efficiency in dynamic environments.

Key applications highlighted for the new standard include:

- Automotive Industry: enhancing driver assistance systems and self-driving functionalities through efficient and rapid processing of road and environmental data.
- Robotics: optimizing navigation and operational efficiency in automated robots.
- Surveillance: Supporting advanced security systems with combined video and LiDAR data processing capabilities.
- Aerial Drones: enabling safer and more effective use of drones in professional and emergency scenarios through improved obstacle detection and environmental mapping.
- Industrial Automation: enhancing precision and safety in industrial applications through better tracking and positioning of machinery.

The Committee Draft is available at https://www.mpeg.org/meetings/mpeg-146/.

7.3.11 MPEG Liaison and Communication (AG 3) White Paper on MPEG-I Immersive Audio

At the 146th MPEG meeting, MPEG Liaison and Communication (AG 3) approved a White paper on MPEG-I Immersive Audio, available at https://www.mpeg.org/whitepapers/.

The MPEG-I immersive audio standard aims at providing a convincing solution for compact representation and for high-quality real-time interactive rendering of virtual audio content with six

degrees of freedom (6DoF), i.e., the user can not only turn his/her head in all directions (pitch/yaw/roll) but also move around freely in 3D space.

By exploring the 6DoF virtual world, many acoustic effects of the real world must be modeled accurately to provide a realistic user experience, including properties of sound sources (e.g., level, size, radiation/directivity characteristics, Doppler processing) as well as effects of the acoustic environment (e.g., sound reflections and reverberation, diffraction, total- and partial occlusion). MPEG-I immersive audio features a plethora of technology components that support computationally efficient rendering of such aspects. Distinguishing from many existing technologies, it offers scene descriptions using physics-inspired metadata (for easier scene authoring from CAD scenes and material databases) and possibilities for artistic tuning of the scene characteristics to achieve the desired results.

During the standardization process, extensive listening test comparisons and evaluations were conducted.

7.5 Developments of the 104th Meeting of JPEG (July 2024)

The 104th JPEG meeting was held in Sapporo, Japan from July 15 to 19, 2024.

7.5.1 JPEG XE

During this JPEG meeting, a Call for Proposals on event-based vision representation was launched for the creation of the first standardized representation of this type of data. This CfP addresses lossless coding, and aims to provide the first standard representation for event-based data that ensures interoperability between systems and devices.

JPEG XE is an activity focused on event-based vision. This activity revolves around a new and emerging image modality created by event-based visual sensors. JPEG XE is about the creation and development of a standard to represent events in an efficient way allowing interoperability between sensing, storage, and processing, targeting machine vision and other relevant applications. The JPEG working group completed the Common Test Conditions (CTC) v2.0 document that provides the means to perform an evaluation of candidate technologies for efficient coding of events. The Common Test Conditions document also defines a canonical raw event format, a reference dataset, a set of key performance metrics and an evaluation methodology.

The JPEG working group furthermore issued a Final Call for Proposals (CfP) on lossless coding for event-based data. This call marks an important milestone in the standardization process and the JPEG working group is eager to receive proposals. The deadline for submission of proposals is set to March 31st of 2025. Standardization will start with lossless coding of events as this has the most imminent application urgency in industry. However, the JPEG working group acknowledges that lossy coding of events is also a valuable feature, which will be addressed at a later stage.

Accompanying these two new public documents, a revised Use Cases and Requirements v2.0 document was also released to provide a formal definition for lossless coding of events that is used in the CTC and the CfP. These documents are publicly available on jpeg.org. The Ad-hoc Group on event-based vision was re-established to continue work towards the 105th JPEG meeting.

"The JPEG working group has reached a new milestone by releasing a new Call for Proposals to code events. This call is aimed at creating the first International Standard to efficiently represent events, enabling interoperability between devices and systems that rely on event sensing." said Prof. Touradj Ebrahimi, the Convenor of the JPEG working group.

7.5.2 JPEG Trust

JPEG Trust provides a comprehensive framework for individuals, organizations, and governing institutions interested in establishing an environment of trust for the media that they use, and supports trust in the media they share. At the 104th meeting, the JPEG working group produced an updated version of the Use Cases and Requirements for JPEG Trust (v3.0). This document integrates additional use cases and requirements related to authorship, ownership, and rights declaration. The JPEG working group also requested a new Part to JPEG Trust, entitled "Media asset watermarking". This new Part will define the use of watermarking as one of the available components of the JPEG Trust framework to support usage scenarios for content authenticity, provenance, integrity, labeling, and binding between JPEG Trust metadata and corresponding media assets. This work will focus on various types of watermarking, including explicit or visible watermarking, invisible watermarking, and implicit watermarking of the media assets with relevant metadata.

7.5.3 JPEG AI

At the 104th meeting, the JPEG working group reviewed recent integration efforts, following the adoption of the changes in the past meeting and the creation of a new version of the JPEG AI verification model. This version reflects the JPEG AI DIS text and was thoroughly evaluated for performance and functionalities, including bitrate matching, 4:2:0 coding, region adaptive quantization maps, and other key features. JPEG AI supports a multi-branch coding architecture with two encoders and three decoders, allowing for six compatible combinations that have been jointly trained. The compression efficiency improvements range from 12% to 27% over the VVC Intra coding anchor, with decoding complexities between 8 to 215 kMAC/px.

The meeting also focused on Part 2: Profiles and Levels, which is moving to Committee Draft consultation. Two main concepts have been established: 1) the stream profile, defining a specific subset of the code stream syntax along with permissible parameter values, and 2) the decoder profile, specifying a subset of the full JPEG AI decoder toolset required to obtain the decoded image. Additionally, Part 3: Reference Software and Part 5: File Format will also proceed to Committee Draft consultation. Part 4 is significant as it sets the conformance points for JPEG AI compliance, and some preliminary experiments have been conducted in this area.

7.5.4 JPEG Pleno Learning-based Point Cloud coding

Learning-based solutions are the state of the art for several computer vision tasks, such as those requiring high-level understanding of image semantics, e.g., image classification, face recognition and object segmentation, but also 3D processing tasks, e.g. visual enhancement and super-resolution. Learning-based point cloud coding solutions have demonstrated the ability to achieve competitive compression efficiency compared to available conventional point cloud coding solutions at equivalent subjective quality. At the 104th meeting, the JPEG working group instigated balloting for the Draft International Standard (DIS) of ISO/IEC 21794 Information technology — Plenoptic

image coding system (JPEG Pleno) — Part 6: Learning-based point cloud coding. This activity is on track for the publication of an International Standard in January 2025. The 104th JPEG meeting also began an exploration into advanced point cloud coding functionality, in particular the potential for progressive decoding of point clouds.

7.5.5 JPEG Pleno Light Field

The JPEG Pleno Light Field effort has an ongoing standardization activity concerning a novel light field coding architecture that delivers a single coding mode to efficiently code light fields spanning from narrow to wide baselines. This novel coding mode is depth information agnostic resulting in significant improvement in compression efficiency. The first version of the Working Draft of the JPEG Pleno Part 2: Light Field Coding second edition (ISO/IEC 21794-2 2ED), including this novel coding mode, was issued during the 104th JPEG meeting in Sapporo, Japan.

The JPEG PLeno Model (JPLM) provides reference implementations for the standardized technologies within the JPEG Pleno framework, including the JPEG Pleno Part 2 (ISO/IEC 21794-2). Improvements to the JPLM have been implemented and tested, including the design of a more user-friendly platform.

The JPEG Pleno Light Field effort is also preparing standardization activities in the domains of objective and subjective quality assessment for light fields, aiming to address other plenoptic modalities in the future. During the 104th JPEG meeting in Sapporo, Japan, the collaborative subjective experiments aiming at exploring various aspects of subjective light field quality assessments were presented and discussed. The outcomes of these experiments will guide the decisions during the subjective quality assessment standardization process, which has issued its third Working Draft. A new version of a specialized tool for subjective quality evaluation, that supports these experiments, has also been released.

7.5.6 JPEG AIC

At its 104th meeting, the JPEG working group reviewed results from previous Core Experiments that collected subjective data for fine-grained quality assessments of compressed images ranging from high to near-lossless visual quality. These crowdsourcing experiments used triplet comparisons with and without boosted distortions, as well as double stimulus ratings on a visual analog scale. Analysis revealed that boosting increased the precision of reconstructed scale values by nearly a factor of two. Consequently, the JPEG working group has decided to use triplet comparisons in the upcoming AIC-3.

The JPEG working group also discussed JPEG AIC Part 4, which focuses on objective image quality assessments for compressed images in the high to near-lossless quality range. This includes developing methods to evaluate the performance of such objective image quality metrics. A draft call for contributions is planned for January 2025.

7.5.7 JPEG Systems

At the 104th JPEG meeting, Part 10 of JPEG Systems (ISO/IEC 19566-10), the JPEG Systems Reference Software, reached the IS stage. This first version of the reference software provides a reference implementation and reference dataset for the JPEG Universal Metadata Box Format

(JUMBF, ISO/IEC 19566-5). Meanwhile, work is in progress to extend the reference software implementations of additional Parts, including JPEG Privacy and Security and JPEG 360.

7.5.8 JPEG DNA

JPEG DNA is an initiative aimed at developing a standard capable of representing bi-level, continuous-tone grey-scale, continuous-tone colour, or multichannel digital samples in a format using nucleotide sequences to support DNA storage. A Call for Proposals was published at the 99th JPEG meeting. Based on the performance assessments and descriptive analyses of the submitted solutions, the JPEG DNA Verification Model was created during the 102nd JPEG meeting. Several core experiments were conducted to validate this Verification Model, leading to the creation of the first Working Draft of JPEG DNA during the 103rd JPEG meeting.

The next phase of this work involves newly defined core experiments to enhance the rate-distortion performance of the Verification Model and its robustness to insertion, deletion, and substitution errors. Additionally, core experiments to test robustness against substitution and indel noise are conducted. A core experiment was also performed to integrate JPEG AI into the JPEG DNA VM, and quality comparisons have been carried out. A study on visual quality assessment of JPEG AI as an alternative to JPEG XL in the VM will be carried out.

In parallel, efforts are underway to improve the noise simulator developed at the 102nd JPEG meeting, enabling a more realistic assessment of the Verification Model's resilience to noise. There is also ongoing exploration of the performance of different clustering and consensus algorithms to further enhance the VM's capabilities.

7.5.9 JPEG XS

The core parts of JPEG XS 3rd edition were prepared for immediate publication as International Standards. This means that Part 1 of the standard – Core coding tools, Part 2 – Profiles and buffer models, and Part 3 – Transport and container formats, will be available before the end of 2024. Part 4 – Conformance testing was currently still under DIS ballot and it will be finalized in October 2024. At the 104th meeting, the JPEG working group continued the work on Part 5 – Reference software. This part was currently at Committee Draft stage and the DIS is planned for October 2024. The reference software has a feature-complete decoder that is fully compliant with the 3rd edition. Work on the encoder is ongoing.

Finally, additional experimental results were presented on how JPEG XS can be used over 5G mobile networks for wireless transmission of low-latency and high quality 6K/8K 360 degree views with mobile devices and VR headsets. This work will be continued.

7.5.10 JPEG XL

Objective metrics results for HDR images were investigated (using among others the ColorVideoVDP metric), indicating very promising compression performance of JPEG XL compared to other codecs like AVIF and JPEG 2000. Both the libjxl reference software encoder and a simulated candidate hardware encoder were tested. Subjective experiments for HDR images are planned.

The second editions of JPEG XL Part 1 (Core coding system) and Part 2 (File format) became ready for publication. The second edition of JPEG XL Part 3 (Conformance testing) moved to the FDIS stage.

7.6 Developments of the 147th Meeting of MPEG (July 2024)

The 147th meeting of MPEG took place in Sapporo, Japan, from 2024-07-15 until 2024-07-19. More information can be found at https://www.mpeg.org/meetings/mpeg-147/.

7.6.1 *MPEG Systems (WG 3) produces Eight Edition of ISO Base Media File Format*

At the 147th MPEG meeting, MPEG Systems (WG 3) promoted the 8th edition of ISO/IEC 14496-12 ISO base media file format (ISOBMFF) to Final Draft International Standard (FDIS), the final stage of standard development.

The ever-growing expansion of the ISO/IEC 14496-12 ISO base media file format (ISOBMFF) application area has continuously brought new technologies to the standards. During the last couple of years, MPEG Systems (WG 3) has received new technologies on ISOBMFF for more seamless support of ISO/IEC 23009 Dynamic Adaptive Streaming over HTTP (DASH) and ISO/IEC 23000-19 Common Media Application Format (CMAF) leading to the development of the 8th edition of ISO/IEC14496-12.

The new edition of the standard includes new technologies to explicitly indicate the set of tracks representing various versions of the media presentation of a single media for seamless switching and continuous presentation. Such technologies will enable more efficient processing of the ISOBMFF formatted files for DASH manifest or CMAF Fragments.

7.6.2 MPEG Systems (WG 3) ratifies Syntactic Description Language

At the 147th MPEG meeting, MPEG Systems (WG 3) promoted ISO/IEC 14496-34 Syntactic description language to Final Draft International Standard (FDIS), the final stage of standard development.

The Syntax Description Language (SDL) has been developed as part of ISO/IEC 14496-1 MPEG-4 Systems to define the syntax of various standards within MPEG-4 and beyond. To enable independent development of and the ability to reference SDL by other standards, MPEG started the development of ISO/IEC 14496-34 Syntactic description language, which reached its final milestone at the 147th MPEG meeting.

The new standard comes with various improvements, such as clarifications of parsable and non-parsable variable definitions and value coercion, among others. Finally, it includes various examples for a better understanding of the standard.

7.6.3 MPEG Systems (WG 3) reaches First Milestone for Low-Overhead Image File Format

At the 147th MPEG meeting, MPEG Systems (WG 3) promoted Amendment 3 of ISO/IEC 23008-12 low-overhead image file format to Committee Draft Amendment (CDAM), marking the initial stage of standard development.

The ISO/IEC 23008-12 image format specification defines generic structures for storing image items and sequences based on ISO/IEC 14496-12 ISO base media file format (ISOBMFF). As it allows the use of various high-performance video compression standards for a single image or a series of images, it has been adopted by the market quickly. However, it was challenging to use it for very small-sized images such as icons or emojis. While the initial design of the standard was versatile and

useful for a wide range of applications, the size of headers becomes an overhead for applications with tiny images. Thus, Amendment 3 of ISO/IEC 23008-12 low-overhead image file format aims to address this use case by adding a new compact box for storing metadata instead of the ‘Meta’ box to lower the size of the overhead.

The standard is planned to be completed, i.e., to reach the status of Final Draft Amendment (FDAM), by the end of 2025.

7.6.4 MPEG Video Coding (WG 4)ratifies Second Edition of Conformance and Reference Software for Compression of Neural Networks

At the 147th MPEG meeting, MPEG Video Coding (WG 4) promoted the 2nd edition of ISO/IEC 15938-18 conformance and reference software for compression of neural networks to Final Draft International Standard (FDIS), the final stage of standard development.

An increasing number of artificial intelligence applications based on artificial neural networks, such as edge-based multimedia content processing, content-adaptive video post-processing filters, or federated training, need to exchange updates of neural networks (e.g., after training on additional data or fine-tuning to specific content). For this purpose, MPEG developed a second edition of the standard for coding of neural networks for multimedia content description and analysis (NNC, ISO/IEC 15938-17, published in 2024), adding syntax for differential coding of neural network parameters as well as new coding tools. Trained models can be compressed to at least 10-20% for several architectures, even below 3%, of their original size without performance loss. Higher compression rates are possible at moderate performance degradation. In a distributed training scenario, a model update after a training iteration can be represented at 1% or less of the base model size on average without sacrificing the classification performance of the neural network.

In order to facilitate the implementation of the standard, the accompanying standard ISO/IEC 15938-18 has been updated to cover the second edition of ISO/IEC 15938-17. This standard provides a reference software for encoding and decoding NNC bitstreams, as well as a set of conformance guidelines and reference bitstreams for testing of decoder implementations. The software covers the functionalities of both editions of the standard, and can be configured to test different combinations of coding tools specified by the standard.

7.6.5 MPEG Coding of 3D Graphics and Haptics (WG 7) reaches First Milestone for Third Edition of Reference Software and Conformance for the Internet of Media Things

At the 147th MPEG meeting, MPEG Coding of 3D Graphics and Haptics (WG 7) reached the Committee Draft (CD) stage of reference software and conformance for the Internet of Media Things (IoMT), the first stage of standard development.

This new edition comprises a comprehensive reference implementation of IoMT functionalities as specified in the MPEG series of Media Things-related standards, outlined across various parts of ISO/IEC 23093. These standards define data formats and APIs for Media Things, enabling robust integration into diverse applications and services. The software implementation features an extensive suite of Media Things, designed to facilitate the collaborative execution of complex artificial intelligence tasks within distributed environments. This standard targets a wide range of applications, such as crowd monitoring with distributed camera networks and forest fire prevention

and surveillance. The advanced capabilities of IoMT will significantly enhance the efficiency and effectiveness of such operations, demonstrating the potential of Media Things in real-world scenarios.

7.7 SC 29 Current projects

SC 29 is progressing the development of the following projects:

7.7.1 WG 1

- ISO/IEC 10918-4:2024 Information technology — Digital compression and coding of continuous-tone still images — Part 4: APPn markers
- ISO/IEC 10918-7:2023 Information technology — Digital compression and coding of continuous-tone still images — Part 7: Reference software
- ISO/IEC 15444-2:2023 Information technology — JPEG 2000 image coding system — Part 2: Extensions
- ISO/IEC 15444-8:2023 Information technology — JPEG 2000 image coding system — Part 8: Secure JPEG 2000
- ISO/IEC 18181-1:2024 Information technology — JPEG XL image coding system — Part 1: Core coding system
- ISO/IEC 18181-2:2024 Information technology — JPEG XL image coding system — Part 2: File format
- ISO/IEC 18477-1:2024 Information technology — Scalable compression and coding of continuous-tone still images — Part 1: Core coding system specification
- ISO/IEC 18477-3:2023 Information technology — Scalable compression and coding of continuous-tone still images — Part 3: Box file format
- ISO/IEC TR 19566-9:2024 Information technology — JPEG Systems — Part 9: JPEG extensions mechanisms to facilitate forwards and backwards compatibility
- ISO/IEC 21122-1:2024 Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system
- ISO/IEC 21122-2:2024 Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models
- ISO/IEC 21122-3:2024 Information technology — JPEG XS low-latency lightweight image coding system — Part 3: Transport and container formats

7.7.2 WG 3

• ISO/IEC 13818-1:2023 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems

• ISO/IEC 14496-15:2022/Amd 1:2023 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 23000-19:2024 Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media

• ISO/IEC 23000-19:2024/Amd 1:2024 Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media — Amendment 1: Low complexity enhancement video Coding (LCEVC) and other technologies

• ISO/IEC 23001-11:2023/Amd 1:2024 Information technology — MPEG systems technologies — Part 11: Energy-efficient media consumption (green metadata) — Amendment 1: Energy-efficient media consumption (green metadata) for EVC

• ISO/IEC 23001-17:2024 Information technology — MPEG systems technologies — Part 17: Carriage of uncompressed video and images in ISO base media file format

• ISO/IEC 23090-14:2023/Amd 1:2023 Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 1: Support for immersive media codecs in scene description

• ISO/IEC 23090-18:2024 Information technology — Coded representation of immersive media — Part 18: Carriage of geometry-based point cloud compression data

• ISO/IEC 23090-6:2021/Amd 1:2024 Information technology — Coded representation of immersive media — Part 6: Immersive media metrics — Amendment 1: Immersive media metrics for V3C Data and OMAF

7.7.3 WG 4

• ISO/IEC 15938-17:2024 Information technology — Multimedia content description interface — Part 17: Compression of neural networks for multimedia content description and analysis

• ISO/IEC 23090-23:2023 Information technology — Coded representation of immersive media — Part 23: Conformance and reference software for MPEG immersive video

• ISO/IEC 23094-2:2021/Amd 1:2024 Information technology – General video coding — Part 2: Low complexity enhancement video coding — Amendment 1: Additional levels

7.7.4 WG 5

• ISO/IEC TR 23002-9:2024 Information technology — MPEG video technologies — Part 9: Film grain synthesis technology for video applications

• ISO/IEC 23008-2:2023 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding

• ISO/IEC 23090-3:2024 Information technology — Coded representation of immersive media — Part 3: Versatile video coding

• ISO/IEC 23090-15:2024 Information technology — Coded representation of immersive media — Part 15: Conformance testing for versatile video coding

7.7.5 WG 6

• ISO/IEC 23003-4:2020/Amd 2:2023 Information technology — MPEG audio technologies — Part 4: Dynamic range control — Amendment 2: Loudness leveling

• ISO/IEC 23008-6:2021/Amd 1:2024 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 6: 3D audio reference software — Amendment 1: Corrections for closest loudspeaker playout and increased software resilience

7.7.6 WG 7

• ISO/IEC 23090-20:2023 Information technology — Coded representation of immersive media — Part 20: Conformance testing for visual volumetric video-based coding (V3C) with video-based point cloud compression (V-PCC)

• ISO/IEC 23090-21:2024 Information technology — Coded representation of immersive media — Part 21: Reference software for Geometry-based Point Cloud Compression (G-PCC)

7.7.7 WG 8

• ISO/IEC 23092-2:2024 Information technology — Genomic information representation — Part 2: Coding of genomic information

• ISO/IEC 23092-6:2023 Information technology — Genomic information representation — Part 6: Coding of genomic annotations

7.8 SC 29 Publications

Since September 2023, the following deliverables have been published from SC 29:

7.8.1 WG 1

- ISO/IEC DIS 6048-1 Information technology — JPEG AI learning-based image coding system — Part 1: Core coding system

- ISO/IEC CD 6048-2 Information technology — JPEG AI learning-based image coding system — Part 2: Profiling

- ISO/IEC CD 6048-3 Information technology — JPEG AI learning-based image coding system — Part 3: Reference software

- ISO/IEC CD 6048-4 Information technology — JPEG AI learning-based image coding system — Part 4: Conformance

- ISO/IEC CD 6048-5 Information technology — JPEG AI learning-based image coding system — Part 5: File Format

- ISO/IEC DIS 18181-3 Information technology — JPEG XL image coding system — Part 3: Conformance testing

- ISO/IEC FDIS 15444-1 Information technology — JPEG 2000 image coding system — Part 1: Core coding system

- ISO/IEC 19566-5:2023/DAmd 1 Information technologies — JPEG systems — Part 5: JPEG universal metadata box format (JUMBF) — Amendment 1: JUMBF box compression and standalone JUMBF files

- ISO/IEC 19566-6:2019/DAmd 2 Information technologies — JPEG systems — Part 6: JPEG 360 — Amendment 2: Revision to the equirectangular projection constraints

- ISO/IEC 19566-7:2022/DAmd 1 Information technologies — JPEG systems — Part 7: JPEG linked media format (JLINK) — Amendment 1: Revision to the JLINK XMP expressions

- ISO/IEC 19566-8:2023/DAmd 1 Information technologies — JPEG systems — Part 8: JPEG Snack — Amendment 1: Revision of JPEG Snack content boxes

- ISO/IEC PRF 19566-10 Information technology — JPEG Systems — Part 10: Reference software

- ISO/IEC PRF 19566-10/DAmd 1 Information technology — JPEG Systems — Part 10: Reference software — Amendment 1: Additional reference software implementations

- ISO/IEC DIS 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 DIS 21617-1 Information technology — JPEG Trust — Part 1: Core Foundation

- ISO/IEC AWI 21617-2 Information technology — JPEG Trust — Part 2: Trust profiles catalogue

• ISO/IEC AWI 21617-3 Information technology — JPEG Trust — Part 3: Media asset watermarking

• ISO/IEC FDIS 21794-5 Information technology — Plenoptic image coding system (JPEG Pleno) — Part 5: Holography

• ISO/IEC DIS 21794-6 Information technology — Plenoptic image coding system (JPEG Pleno) — Part 6: Learning-based point cloud coding

• ISO/IEC AWI 29170-3 Information technology — Advanced image coding and evaluation — Part 3: Subjective quality assessment of high-fidelity images

7.8.2 WG 2

• ISO/IEC CD TR 23090-1 Information technology — Coded representation of immersive media — Part 1: Architectures for immersive media

• ISO/IEC CD TR 23090-27 Information technology — Coded representation of immersive media — Part 27: Media and architectures for render-based systems and applications

• ISO/IEC CD TR 23888-1 Information technology — Artificial intelligence for multimedia — Part 1: Vision and scenarios

7.8.3 WG 3

• ISO/IEC 13818-1:2023/DAmd 1 Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems — Amendment 1: Codec parameter clarifications and other improvements

• ISO/IEC DIS 14496-1 Information technology — Coding of audio-visual objects — Part 1: Systems

• ISO/IEC DIS 14496-12.2 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format

• ISO/IEC DIS 14496-12.2/DAmd 1 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format — Amendment 1: Support for T.35, original sample duration and other improvements

• ISO/IEC DIS 14496-12.2/AWI Amd 2 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format — Amendment 2: Tools for enhanced CMAF and DASH integration

• 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/DAmd 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 neural-network post-filter supplemental enhancement information and other improvements

• ISO/IEC CD 14496-22 Information technology — Coding of audio-visual objects — Part 22: Open Font Format

• ISO/IEC DIS 14496-32 Information technology — Coding of audio-visual objects — Part 32: File format reference software and conformance

• ISO/IEC DIS 14496-34 Information technology — Coding of audio-visual objects — Part 34: Syntactic description language

• ISO/IEC DIS 21000-3 Information technology — Multimedia framework (MPEG-21) — Part 3: Digital Item Identification

• ISO/IEC 23000-19:2024/AWI Amd 2 Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media — Amendment 2: Additional structural CMAF brand profile

• ISO/IEC FDIS 23000-22 Information technology — Multimedia application format (MPEG-A) — Part 22: Multi-image application format (MIAF)

• ISO/IEC FDIS 23000-22/AWI Amd 1 Information technology — Multimedia application format (MPEG-A) — Part 22: Multi-image application format (MIAF) — Amendment 1: Implementation based technologies for MIAF

• ISO/IEC AWI 23000-23 Information technology — Multimedia application format (MPEG-A) — Part 23: Decentralized media rights application format

• ISO/IEC AWI 23000-24 Information technology — Multimedia application format (MPEG-A) — Part 24: Messaging media application format (MeMAF)

• ISO/IEC 23001-10:2020/CD Amd 2 Information technology — MPEG systems technologies — Part 10: Carriage of timed metadata metrics of media in ISO base media file format — Amendment 2: Support for Attenuation Maps

• ISO/IEC 23001-11:2023/DAmd 2 Information technology — MPEG systems technologies — Part 11: Energy-efficient media consumption (green metadata) — Amendment 2: Energy-efficient media consumption for new display power reduction metadata

• ISO/IEC 23001-17:2024/DAmd 1 Information technology — MPEG systems technologies — Part 17: Carriage of uncompressed video and images in ISO base media file format — Amendment 1: High precision timing tagging

• ISO/IEC 23001-17:2024/DAmd 2 Information technology — MPEG systems technologies — Part 17: Carriage of uncompressed video and images in ISO base media file format — Amendment 2: Generic compression for samples and items in ISOBMFF

• ISO/IEC AWI 23001-19 Information technology — MPEG systems technologies — Part 19: Carriage of green metadata

• ISO/IEC 23008-1:2023/DAmd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT) — Amendment 1: Signalling of adaptive FEC scheme

• ISO/IEC DIS 23008-12 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format

• ISO/IEC DIS 23008-12/CD Amd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 1: Support for tone map derivation and other technologies

• ISO/IEC DIS 23008-12/AWI Amd 2 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 2: Low-overhead image file format

• ISO/IEC DIS 23009-1 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats

• ISO/IEC WD TR 23009-7 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 7: Delivery of CMAF contents with DASH

• ISO/IEC DIS 23009-8 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 8: Session-based DASH operations

• ISO/IEC DIS 23009-9.2 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 9: Redundant encoding and packaging for segmented live media (REaP)

• ISO/IEC 23090-2:2023/CD Amd 1 Information technology — Coded representation of immersive media — Part 2: Omnidirectional media format — Amendment 1: Server-side dynamic adaptation

• ISO/IEC 23090-6:2021/DAmd 2 Information technology — Coded representation of immersive media — Part 6: Immersive media metrics — Amendment 2: Additional latencies and other improvements

• ISO/IEC 23090-7:2022/DAmd 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 AWI 23090-10 Information technology — Coded representation of immersive media — Part 10: Carriage of visual volumetric video-based coding data

• ISO/IEC 23090-10:2022/DAmd 2 Information technology — Coded representation of immersive media — Part 10: Carriage of visual volumetric video-based coding data — Amendment 2: Clarification on brands and other improvements

• ISO/IEC CD TR 23090-11.2 Information technology — Coded representation of immersive media — Part 11: Network-based media processing implementation guidelines

• ISO/IEC AWI 23090-13 Information technology — Coded representation of immersive media — Part 13: Video decoding interface for immersive media

• ISO/IEC DIS 23090-14 Information technology — Coded representation of immersive media — Part 14: Scene description

• ISO/IEC DIS 23090-14/AWI Amd 1 Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 1: Support of MPEG-I audio, scene understanding and other extensions

• ISO/IEC 23090-18:2024/Amd 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 23090-18:2024/CD Amd 2 Information technology — Coded representation of immersive media — Part 18: Carriage of geometry-based point cloud compression data — Amendment 2: Point reliability indication and other improvements

• ISO/IEC PRF 23090-24 Information technology — Coded representation of immersive media — Part 24: Conformance and reference software for scene description

• ISO/IEC PRF 23090-24/AWI Amd 1 Information technology — Coded representation of immersive media — Part 24: Conformance and reference software for scene description — Amendment 1: Conformance and reference software for scene description on haptics, augmented reality, avatars, interactivity and lighting

• ISO/IEC FDIS 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 DIS 23090-26 Information technology — Coded representation of immersive media — Part 26: Conformance and reference software for carriage of geometry-based point cloud compression data

• ISO/IEC DIS 23090-32 Information technology — Coded representation of immersive media — Part 32: Carriage of haptics data

• ISO/IEC DIS 23090-33 Information technology — Coded representation of immersive media — Part 33: Conformance and reference software for haptics coding

7.8.4 WG 4

• ISO/IEC DIS 15938-18 Information technology — Multimedia content description interface — Part 18: Conformance and reference software for compression of neural networks
• ISO/IEC DIS 23090-12 Information technology — Coded representation of immersive media — Part 12: MPEG immersive video

• ISO/IEC 23094-3:2022/Amd 1 Information technology — General video coding — Part 3: Conformance and reference software for low complexity enhancement video coding — Amendment 1: Updated conformance data and reference software

• ISO/IEC 23094-4:2022/DAmd 1 Information technology — General video coding — Part 4: Conformance and reference software for essential video coding — Amendment 1: Green metadata supplemental enhancement information

• ISO/IEC AWI 23888-2 Information technology — Artificial intelligence for multimedia — Part 2: Video coding for machines

7.8.5 WG 5

• 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/CD Amd 1 Information technology — MPEG video technologies — Part 7: Versatile supplemental enhancement information messages for coded video bitstreams — Amendment 1: Additional SEI

• ISO/IEC 23008-2:2023/DAmd 1 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding — Amendment 1: New profiles, colour descriptors, and SEI messages

• ISO/IEC DIS 23008-2 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding

• ISO/IEC 23090-3:2024/CD Amd 1 Information technology — Coded representation of immersive media — Part 3: Versatile video coding — Amendment 1: Additions and corrections

• ISO/IEC CD 23090-16 Information technology — Coded representation of immersive media — Part 16: Reference software for versatile video coding

• ISO/IEC DIS 23091-2 Information technology — Coding-independent code points — Part 2: Video

• ISO/IEC CD TR 23888-3 Information technology — Artificial intelligence for multimedia — Part 3: Optimization of encoders and receiving systems for machine analysis of coded video content

7.8.6 WG 6

• ISO/IEC DTR 14496-24 Information technology — Coding of audio-visual objects — Part 24: Audio and systems interaction

• ISO/IEC FDIS 14496-26 Information technology — Coding of audio-visual objects — Part 26: Audio conformance

• ISO/IEC DIS 23003-4 Information technology — MPEG audio technologies — Part 4: Dynamic range control

• ISO/IEC DIS 23008-3 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio

• ISO/IEC DIS 23008-6 Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 6: 3D audio reference software

• ISO/IEC DIS 23090-4 Information technology — Coded representation of immersive media — Part 4: MPEG-I immersive audio

• ISO/IEC CD 23090-34 Information technology — Coded representation of immersive media — Part 34: Immersive audio reference software

7.8.7 WG 7

• 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 23090-19 Information technology — Coded representation of immersive media — Part 19: Reference software for V-PCC

• ISO/IEC 23090-22 Information technology — Coded representation of immersive media — Part 22: Conformance for G-PCC

• ISO/IEC DIS 23090-28 Information technology — Coded representation of immersive media — Part 28: Interchangeable scene-based media representations

• ISO/IEC CD 23090-29 Information technology — Coded representation of immersive media — Part 29: Video-based dynamic mesh coding (V-DMC)

• ISO/IEC CD 23090-30 Information technology — Coded representation of immersive media — Part 30: Low latency, low complexity LiDAR coding

• ISO/IEC FDIS 23090-31 Information technology — Coded representation of immersive media — Part 31: Haptics coding

• ISO/IEC AWI 23090-35 Information technology — Coded representation of immersive media — Part 35: Conformance and reference software for Low latency, low complexity LiDAR coding

• ISO/IEC AWI 23090-36 Information technology — Coded representation of immersive media — Part 36: Conformance and Reference Software for V-DMC

• ISO/IEC AWI 23090-37 Information technology — Coded representation of immersive media — Part 37: Conformance and reference software for carriage of haptics data

• ISO/IEC DIS 23093-1 Information technology — Internet of media things — Part 1: Architecture

• ISO/IEC DIS 23093-2 Information technology — Internet of media things — Part 2: Discovery and communication API

• ISO/IEC DIS 23093-3 Information technology — Internet of media things — Part 3: Media data formats and APIs

• ISO/IEC AWI 23093-4 Information technology — Internet of media things — Part 4: Reference software and conformance

• ISO/IEC DIS 23093-5 Information technology — Internet of media things — Part 5: IoMT autonomous collaboration

• ISO/IEC CD 23093-6 Information technology — Internet of media things — Part 6: IoMT Media data formats and API for distributed AI processing

7.8.8 WG 8

• ISO/IEC FDIS 23092-1 Information technology — Genomic information representation — Part 1: Transport and storage of genomic information

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

• ISO/IEC 23092-5:2020/CD Amd 1 Information technology — Genomic information representation — Part 5: Conformance — Amendment 1: Version 2 and Part 6 support