

ISO/IEC JTC 1
Information technology
Secretariat: ANSI (United States)

Document type: Business Plan

Title: Business Plan for JTC 1/SC 29, Coding of Audio, Picture, Multimedia and Hypermedia Information for the Period covered: September 2015 - August 2016

Status: This document is circulated for review and consideration at the November 2016 JTC 1 meeting in Norway.

Date of document: 2016-10-03

Source: SC 29 Chair

Expected action: ACT

Action due date: 2016-11-07

Email of secretary: lrajchel@ansi.org

Committee URL: <http://isotc.iso.org/livelink/livelink/open/jtc1>

JTC 1/SC 29 Business Report

Coding of Audio, Picture, Multimedia and Hypermedia Information

Period covered: September 2015 - August 2016

1.0 Executive summary

SC 29 has been developing and delivering the international standards which are basis of digital media services and applications. Those standards are designed to represent, package, record, preserve and convey digital media information. The standards support functionalities such as coding, composition, transport, filing and archiving of media and their control, interface, middleware for general and/or specific applications.

In the period SC 29 has made significant progress in coding of still pictures with extended bit-depth, lossless compression and alpha channel supporting backward compatibility with JPEG standard (JPEG-XT), coding of natural video with enhancement to screen content (MPEG-H HEVC-SCC), immersive audio (MPEG-H 3D Audio), 3D graphics objects (Web3DCoding), Application Formats (Augmented Reality, Multimedia Preservation, Publish/Subscribe, Media Linking), MPEG-V and other standards. In the marketplace HEVC has been implemented in smart-phones and devices for Ultra High Definition Television (UHDTV) such as 4K/8K. Adaptive streaming (MPEG-DASH) has been deployed in emerging VoD services. The ISO Base Media File Format (ISO/BMFF) celebrates its 25 years of vigorous life in development and active use.

In addition to the projects above, WG 1 has been studying on coding of advanced image representation such as light field, point cloud and holographic imaging possible as further extensions to the JPEG standard. This work is expected to bring new framework for novel markets of digital photography. WG 1 is also working on backward-compatible extension to the JPEG standard.

WG 11 has been working on coding of natural video (MPEG-H HEVC) and audio (MPEG-H 3D Audio) and 3D graphics objects (3D graphics coding for browsers), fonts (Open Font Format), file formats (ISO Base Media File Format), image description (Compact Description for Visual Search) and user description (User Description), application formats and media transport for streaming (DASH and MMT). WG 11 is currently engaged in video description (Compact Descriptors for Video Analysis), new application format (Common Media, Omnidirectional Media, Virtual Reality), media combination (Media Orchestration), 3D Media Formats, Future Video Coding (in collaboration with ITU-T SG 16 VCEG), Point Cloud Compression, Genomic Information compression and storage (in collaboration with TC 276/WG 5), Internet of Media Things and Media Wearables.

2.0 Chairman's Remarks

2.1 Market Requirements, Innovation

For WG 1, JPEG standard has been widely used in digital photography and other types of still image-based applications. Everyday JPEG-coded images are produced, sent, shared and seen in e-mails, photo-hosting websites and major social networking services. JPEG XT offers backward compatibility with traditional JPEG standard and extends it to support High Dynamic Range representation, which has been awaited by the digital camera industry and users. Other extensions to JPEG within this framework (e.g. privacy protection functionality) are studied.

JPEG 2000 has been used for digital photography, photo IDs, satellite images, medical images and Document databases. JPEG 2000 has been deployed for the compression of digital cinema and supports the cinema industry by developing appropriate amendments to meet their requirements.

For WG 11, MPEG-2 and Advanced Video Coding (AVC) have been widely used for digital broadcasting systems, visual communication equipment, IPTV servers and terminals, video discs, surveillance systems, recorders, mobile devices and so on. HEVC has been used in major smart-phones and content services for those devices. UHDTV services employing HEVC has started. There are still needs for further compression especially for mobile streaming environment.

MP3 (MPEG-1 audio layer3) and Advanced Audio Coding (AAC) have been widely used in the market of digital music distribution. These standards have been used for broadcasting, network audio services and various audio-capable devices. Unified Speech and Audio Coding (USAC) is used for generic applications such as unidirectional, conversational, communication, broadcasting, and playback types of applications. 3D audio is awaited for immersive audio presentation such as digital cinema and next generation broadcasting.

Compact Descriptors for Visual Search (CDVS) is a tool for search, detection and retrieval employing mobile devices, over large-scale databases and resources on the web. CDVS can be used in Augmented Reality (AR) type and Internet-of-Things (IoT) type applications. Applications of CDVS in more generic applications such as robotics, automotive, geo-location etc. could also be foreseen.

MPEG Media Transport (MMT) offers advanced solutions for media transport besides the ubiquitous MPEG-2 Transport Stream specification. It enables for example hybrid delivery services over heterogeneous network environment. MMT has been adopted for next generation broadcasting by major industry fora.

Dynamic adaptive streaming over HTTP (DASH) provides the standardized way of adaptive streaming in emerging Internet-based video delivery services known as Over The Top (OTT). Further standardization is needed for large scale of delivery of adaptable media presentation to a wide range of devices. DASH Industry Forum has been established to advance use of DASH by industry.

For both WGs, media data is expected to be more machine or system-friendly as rich sensory data in the aspect of Cyber Physical Systems or IoT-type applications. Standardized mechanism to support such systems and applications is required.

2.2 Accomplishments

2.3 Resources

Each WG has sufficient participation of experts so that we proceed the works on JPEG, JBIG-2, JPEG 2000, JPSearch, AIC, JPEG XR, JPEG XT, MPEG-2, MPEG-4, MPEG-7, MPEG-21, MPEG-A to MPEG-E, MPEG-H, MPEG-M, MPEG-V, MPEG-U and MPEG-DASH projects in addition to other standards being investigated.

2.4 Competition and Cooperation (including consortia)

SC 29 maintains many liaisons with other organizations to meet the requirements and expectations of the standards users from the other communities. SC 29 has been continuing productive collaboration with ITU-T. The Joint Photographic Experts Group (JPEG) and the Joint Collaborative Team on Video Coding (JCT-VC) are the current joint activities we share with ITU-T/SG 16. The Joint Video Exploration Team (JVET) is the latest joint activity with ITU-T aiming at the study on future video coding. The Joint Collaborative Team on 3D Video (JCT-3V) has completed its work and is to be discontinued. SC 29 works jointly with SC 24 in the Joint Ad hoc Group (JAhG) targeting Mixed and Augmented Reality Reference Model. See SC 29 Web site (<http://www.itscj.ipsj.or.jp/sc29/29w2l.htm>). A new collaboration is expected with TC 276 on genome compression.

While SC 29 and its WGs are open to welcome new members and willing to have opportunities to collaborate with new liaison partners, it is well recognized that this field of work requires rapid progress and has potential to bring about new competitors. SC 29 and its WGs should always communicate with the stakeholders in the industry and try to recruit appropriate members in order to keep its competitiveness.

3.0 Discussion of SC program of work

3.1 SC Opportunities and Strategies

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

The recipients or consumers of digital media information are not necessary to be human but machines or systems in the era of Internet of Things. Thus, we have a lot of opportunities.

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

SC 29 continues the practice of making their standards containing conformance testing bitstreams and reference software accessible as the freely available standards from ISO/IEC. SC 29 is pleased to hear the requirements from the industry, so that appropriate WG under SC 29 would study and work in case of observing good opportunities to deliver the standards that meet the requirements. The WGs may hold public seminars to introduce their activities and to align future standardization with users and industry needs.

3.2 SC and WG risks

- Management of documents, media data and software: The WGs depend on having good electronic repositories and systems capable to handle the high volume of traffic by members and the maintenance of these is important for the efficient work of the WGs. The volume of such data and traffic should be huge in WG 11. ISO e-committee is expected to be enhanced in its performance and functionalities so that the requirements of SC/WGs should be always met.

- Risk related to patents: In order to meet the requirements of industry and to compete with other standards/specifications developing organizations, SC 29's standards should incorporate new technologies and ideas, which may be associated with patents. The number of such patents could be many corresponding to the advanced features of the standards. The patents may be owned by the proponents, other parties that have not participated in the development process or even someone else we cannot know. The owners may have their respective policies. SC 29 and WGs continue to encourage their members to submit patent declarations, however it seems getting less practical to collect the declarations perfectly.

3.3 WG 1 Coding of Still Pictures

3.3.1 WG 1 accomplishments

JPEG XT Parts 6, 7 have been published. These standards provide backwards-compatible extensions to the popular JPEG image coding system and simultaneously support improved sample precisions and high dynamic range photography. Part 6 offers the extended bit-depth to JPEG images from 9 to 16 bits, while Part 7 offers even floating-point precision. Part 8 lossless coding and Part 9 alpha channels have been approved through their FDIS ballots.

3.3.2 WG 1 deliverables

JPEG XS is a new activity of WG 1. The targets of this standard may be mostly uncompressed video link at this moment, with increasing resolutions and frame rates. It aims at a low-latency lightweight image coding system while assuring visual quality and keeping power and bandwidth within a reasonable budget. WG 1 issued a Call for Proposals and received seven responses. WG 1 will review the results of the evaluation of the proposals in October 2016. Selection or creation of technologies for the upcoming standard is expected.

JPEG PLENO is another new activity that targets a standard framework for capture, representation and exchange of plenoptic image modalities. It has the long road map beyond this period, however visible progress is expected in the period.

3.4 WG 11 Coding of Moving Pictures and Audio

3.4.1 WG 11 accomplishments

HEVC Screen Content Coding (SCC), an extension to HEVC has been completed. Screen content is video containing a significant proportion of rendered graphics, text, or animation rather than camera-captured video scenes. The new SCC extension greatly improves the compression of screen content. Example applications include wireless displays, remote computer desktop access, and real-time screen sharing for videoconferencing.

MPEG-H 3D Audio Amendment 3 has been progressed to Final Draft Amendment status. It adds technology that increases coding efficiency and also adds features suitable for broadcast industry. Key compression technologies focus on intermediate bit rates of 128kb/s to 256 kb/s for immersive audio signals.

Web3DCoding has been added as Final Draft Amendment 3 to ISO/IEC 14496-11:2011. Web3DCoding extends the usage of the existing MPEG 3D mesh codecs to Web browsers. 3D objects are compressed by this codec and encapsulated in a 3D scene described using the GL Transmission Format.

Augmented Reality Application Format (ARAF) 2nd Edition has reached the stage of Final Draft International Standard. The ARAF enables augmentation of the real world with synthetic media objects by combining multiple existing WG 11 standards within a single specific application format addressing certain industry needs. In particular, ARAF comprises three components referred to as scene, sensor/actuator, and media.

Multimedia Preservation Application Format (MPAF) has reached the stage of Final Draft International Standard. This standard provides standardized description information to enable users to plan, execute, and evaluate preservation operations in order to achieve the objectives of digital preservation.

Publish/Subscribe Application Format (PSAF) has reached the stage of Final Draft International Standard. PSAF enables new forms of communication mediated by a Match Service Provider which provides matching services between subscriptions and publications.

Media Linking Application Format (MLAF) has reached the stage of Final Draft International Standard. MLAF enables a standard communication format between a main screen and a companion screen.

MPEG-V 3rd Edition has reached the stage of Final Draft International Standard. MPEG-V specifies the architecture and associated representations to enable interaction between digital content and virtual worlds with the physical one, as well as information exchange between virtual worlds. In this 3rd Edition, MPEG-V includes technology for environmental and camera-related sensors, and 4D- theater effects.

WG 11 and ITU-T SG 16's VCEG have agreed to join together to explore new technology possibilities for video coding that lie beyond the capabilities of the HEVC standard and its current extensions. The new partnership is known as the Joint Video Exploration Team (JVET), and the team is working to explore both incremental and fundamentally different video coding technology that shows promise to potentially become the next generation in video coding standardization.

WG 11 has issued a Joint Call for Proposals on genomic information compression and storage in conjunction with ISO/TC 276/WG 5. The genome includes the complete set of hereditary of living beings information and comprises of the DNA sequences that may contain hundreds billion base pairs. Genomic information compression and storage technologies are expected to enable the practical universal deployment of genomics in the daily clinic practice. This work will be carried out in collaboration with TC 276/WG 5.

3.4.2 WG 11 deliverables

WG 11 will continue its study on future video coding together with ITU-T in the JVET. Besides much needed further compression capabilities of standard rectangular video, the Joint Exploration Model (JEM) will cover 360° Video. When sufficient progress will be achieved, a new Call for Evidence (or Proposals) will be issued in 2017 and, if positive results will be achieved, initial JVET exploration will be converted into a formal video compression standard project.

WG 11 will work on 3D Audio for Virtual Reality and Augmented Reality applications.

WG 11 will continue its current work on Compact Descriptors for Video Analysis (CDVA) to develop a format that can accelerate objects searching for in video data bases, for example broadcasters' archives or videos available on the Internet. Visual content matching will be studied including matching of views of large and small objects and scenes, that is robust to partial occlusions as well as changes in vantage point, camera parameters, and lighting conditions.

WG 11 will continue its work on Common Media Application Format (CMAF) to define a standard for encoding and decoding of segmented data, which are used in environments that support adaptive streaming using HTTP(S). The standard should provide a format optimized for large scale of single encrypted, adaptable multimedia presentation.

WG 11 will continue its work on Omnidirectional Media Application Format (OMAF) to produce a 1st version of Virtual Reality format in response to urgent market need to avoid market fragmentation caused by the proprietary contents formats introduced by several companies. OMAF will support projection mapping, ROI based quality enhancements etc..

WG 11 will contribute to the vast IoT domain by working on data formats for Internet of Media Things, i.e. the particular class of Things with the capability to capture, process, transmit and store media streams. WG 11 will also specify data formats for Media Wearables, a class of wearable devices able to manipulate various types of media streams. Calls for Proposal on these issues are planned for October 2016.

WG 11 will engage in the standardization of Point Clouds Compression format. Point clouds represent a good compromise between ease of acquisition, size of the data, rapidity of rendering and support for user interaction, features that make PCC suitable for families of applications in the field of augmented and virtual reality.

Progress in the Genomic Information and Storage project is expected because some initial study has shown that integration of different tools available from the literature show better compression over the best state-of-the-art technology. Responses to a Call for Proposal are expected for October 2016.