

Question 6/21 – Visual, audio and signal coding

(Continuation of Question 6/21)*

1 Motivation

The goal of this Question is to produce Recommendations for visual, speech, audio and signal coding methods appropriate for conversational (e.g. videoconferencing and video telephony) and non-conversational (e.g. multimedia streaming, broadcast TV, IPTV, file download, media storage/playback, remote screen display, digital cinema, or virtual & augmented reality) audiovisual services and other services. The Question is to focus primarily on the coding of visual signals, including the compression of:

- video sequences;
- still images;
- graphics;
- stereoscopic, multi-view, depth maps, and free-viewpoint visual information;
- light fields, point clouds, and volumetric imagery;
- computer displays;
- medical imaging;
- 360 degree/panoramic/spherical-view video sequences;
- video and images for virtual and augmented reality.

This Question will primarily focus on the maintenance and extension of existing video and still-image coding Recommendations and the development of new Recommendations using advanced techniques to significantly improve the trade-offs between bit rate, quality, delay, and algorithm complexity. The Question will also be responsible for maintenance and the development of new Recommendations for the coding of speech, audio, biomedical and other signals and network-based signal processing. Video, still-image, speech, audio, biomedical and other signal coding standards will be developed with sufficient flexibility to accommodate a diverse number of transport types (Internet, LAN, 5G and other mobile networks, ITU-T H.222.0, etc.).

This Question will apply digital signing for verifying the integrity of multimedia content, enabling users to confirm its authenticity. The work includes integrating authentication technology into coded multimedia streams, including video, speech, audio, and other compressed signal types. Robust methods for synchronizing authenticity information across coded multimedia streams will be developed to ensure consistency and accuracy. The Question will also investigate the information that will be included in the coded multimedia stream to provide authentication.

2 Study items

Study items to be considered include, but are not limited to:

- new coding methods in order to achieve the following objectives:
 - improvements in compression efficiency;
 - robust operation in error/loss-prone environments (e.g. non-guaranteed-bandwidth packet networks or mobile wireless communication);
 - reduction of real-time delay, complexity, and of channel acquisition time and random access latency;
- organization of the compressed data format to support packetization and streaming;

* Update of Q6/21 "Visual, audio and signal coding" (WTSA-24) as prepared by ITU-T SG21 (Geneva, 13-24 January 2025) and endorsed by TSAG (Geneva, 26-30 May 2025).

- development of supplemental enhancement information to accompany source data for enabling enhanced functionality in application environments;
- study and specification of data for annotation, indexing, and searching;
- techniques to permit networks or terminals to adjust bit rates efficiently;
- techniques for object coding and multi-view operation;
- techniques to permit terminals to rapidly adjust the region-of-interest and/or field of view of video stream playback;
- techniques for efficient coding of 360-degree/panoramic/spherical-view video sequences, including those formed by stitching video sequences from multiple cameras with projection/rendering warping;
- techniques for efficient coding of video, images, audio, point clouds, and other signals for virtual and augmented reality, navigation, medical, and other applications;
- techniques for efficient compressed-digital to compressed-digital processing (including transcoding);
- artificial intelligence technology for encoding and decoding of video, images, audio, biomedical and other signals and the processing and analysis of coded data;
- the impact of colorimetry, video and image quality assessment, and quality control requirements on video and image codec development;
- computer graphics compression;
- security aspects that directly affect video, speech, audio and signal coding;
- coordination of video, still-image, speech, audio and signal coding matters not addressed in other coding Questions with other ITU study groups and other bodies;
- harmonization of video, still-image, speech, audio, biomedical and other signal coding activities with other standard development organizations (SDOs);
- enhancements to existing multimedia systems Recommendations including the addition of advanced audio and visual coding (e.g. ITU-T H.26x and G.72x extensions and beyond).
- use of digital signing techniques for verification of the integrity of coded multimedia content, enabling users to confirm the authenticity.
- integration possibilities of the authentication technology into coded multimedia streams containing video, speech, audio and other signals.
- synchronization methods of authenticity information among the different coded multimedia streams.
- investigation of the information needed in the authentication signal in the coded multimedia stream to enable authenticity verification.

3 Tasks

Tasks include, but are not limited to:

- development of extensions, additional profiles, and maintenance updates for ITU-T H.266 (VVC);
- work towards development of a future video coding Recommendation with compression capability substantially beyond that of ITU-T H.266;
- address needs for signal type identification for use with video and image coding Recommendations, including extensions and maintenance for ITU-T H.273;
- conformance and reference software development and maintenance for ITU-T H.264 (AVC), ITU-T H.265 (HEVC), and H.266, including ITU-T H.264.1, H.264.2, H.265.1, H.265.2, and conformance testing and reference software for H.266 (H.266.1 and H.266.2);

- development of guidelines and informative reports for effective use of video and still-image compression coding technology;
- in liaison with other ITU-T standardization groups or SDOs, recommend what video and still-image coding standards should be used in services/applications, networks, devices and specified in related ITU-T Recommendations;
- development of supplemental enhancement information to accompany video, still-image, speech, audio, and signal data, including data for image/video annotation, indexing, and searching, including maintenance and extension of ITU-T H.271 and H.274 (VSED);
- development of new image coding (T.8xx-sub-series) specifications;
- maintain the video, still-image, speech, and audio coding information in the ITU-T media coding database;
- maintenance of existing H-series video coding Recommendations and supplements, including ITU-T H.120, H.261, H.262 | ISO/IEC 13818-2, H.263, H.264 | ISO/IEC 14496-10, H.264.1, H.264.2, H.265 | ISO/IEC 23008-2, H.265.1, H.265.2, H.266 | ISO/IEC 23090-3, H.266.1, H.266.2, H.271, H.273, H.274 | ISO/IEC 23002-7, H-series Supplements 15, 18, and 19, and Technical Paper ITU-T HSTP-VID-WPOM;
- maintain and extend existing Recommendations and Supplements regarding still image coding, including ITU-T T.44, T.80, T.81, T.82, T.83, T.84, T.85, T.86, T.87, T.88, T.89, T.800, T.801, T.802, T.803, T.804, T.805, T.807, T.808, T.809, T.810, T.812, T.813, T.814, T.815, T.831, T.832, T.833, T.834, T.835, T.851, T.870, T.871, T.872, T.873 and T-series Supplement 2;
- maintenance of existing G-series regarding speech and audio coding and signal processing Recommendations including ITU-T G.711, G.711.0, G.711.1, G.718, G.719, G.720.1, G.722, G.722.1, G.722.2, G.723.1, G.726, G.727, G.728, G.729 and G.729.1;
- maintenance of related Recommendations to signal processing network equipment and functions: ITU T G.160, G.161, G.161.1, G.164, G.165, G.168, G.169, Q50-series, Q.115-series, G.799.1, G.799.2, G.799.3, G.776.1, G.776.4, G.763, G.764, G.765, G.766, G.767, G.768, G.769/Y.1242 and I.733;
- development of new coding Recommendations for speech, audio, biomedical and other signals.
- develop and specify methods for using digital signing techniques for verifying the integrity of coded multimedia content to enable users to confirm its authenticity.
- integrate authentication technology into coded multimedia streams, including video, speech, audio, and other signals, ensuring seamless operation within these formats.
- design and implement methods for synchronizing authenticity information across different coded multimedia streams to maintain consistent and reliable verification.
- define and document the necessary information required for inclusion in the authentication data to enable authenticity verification of coded multimedia streams.

An up-to-date status of work under this Question is found in the SG21 work programme (https://itu.int/ITU-T/workprog/wp_search.aspx?sp=18&q=6/21).

4 Relationships

Recommendations:

- ITU-T H.300 sub-series systems Recommendations
- ITU-T H.241, H.245 and H.248-series

Questions:

- 0/21, 2/21, 5/21, 7/21, 8/21, 11/21, 13/21, 21/21

Study groups

- ITU-T SGs 11, 12, 13
- ITU-R SG6

Other bodies:

- ISO/IEC JTC1/SC29 WGs 1-8 (JPEG and MPEG) on video, image, speech, and audio coding
- IETF, DVB, ATSC, ARIB, 3GPP, DICOM, EBU, SCTE, SMPTE, MC-IF, MEF, VESA, W3C, CTA, IEC TC100