

**Joint Video Experts Team (JVET)
of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC
29/WG 11**

Document: JVET-AO0284
MPEG: m75846

41st Meeting, by teleconference, 14-23 January 2026

Title:	A Bitstream Extension Mechanism for New Technologies Based on Video Coding
Status:	Input document to JVET
Purpose:	Proposal
Author(s) or Contact(s):	Y. Liao yiyi.liao@zju.edu.cn (Zhejiang Univ.) Lu Yu yul@zju.edu.cn submitter (Zhejiang Univ.) X. Jin jin.xin@sz.tsinghua.edu.cn (Tsinghua SIGS) Lei Zhao zhaoliliyjy@chinamobile.com (China Mobile)
Source:	Zhejiang Univ., Tsinghua SIGS, China Mobile

HEADER EXCERPT FROM SOURCE DOCUMENT

Abstract

This contribution proposes an extension method based on existing video coding standards, e.g. VVC/HEVC, to support emerging technologies such as Lenslet Video Coding (LVC) and Gaussian Splatting Coding (GSC). The core idea is to utilize existing extension mechanisms in the syntax structures of VVC/HEVC to embed LVC or GSC-related bitstreams as extended sub-bitstreams directly into standard VVC/HEVC bitstreams to form a unified carriage mechanism. The extended sub-bitstream shall start with a specific leading bit pattern, which is fully compliant with the specified syntax structure extension of VVC/HEVC and may avoid any potential conflict with any future extension of VVC and HEVC. Meanwhile, the syntax structures for LVC or GSC-related bitstreams and their decoding processes do not need to be additionally defined in the VVC/HEVC standards, but can be specified in the respective LVC or GSC standards.

Introduction

Conventionally, video-codec extensions such as Video-based Point Cloud Compression (V-PCC), Multi-view Image and Video (MIV), and Video Coding for Machines (VCM) have implemented extension logic by wrapping the standard video bitstream inside their respective codec-level containers or system-level frameworks. While effective, this "wrapper-based" approach requires significant extra design at the system level to manage, synchronize, and demultiplex the various sub-components (e.g., geometry, attributes, or occupancy maps) alongside the video data. Such architectural requirements can create barriers to rapid deployment in existing video ecosystems that are optimized for single-stream NAL unit delivery.

Built from JVET-AO0284.docx

MEETING NOTES

[JVET-AO0284](#) A Bitstream Extension Mechanism for New Technologies Based on Video Coding [Y. Liao, L. Yu (Zhejiang Univ.), X. Jin (Tsinghua SIGS), L. Zhao (China Mobile)] [late]

This contribution proposes an extension method based on existing video coding standards, e.g. VVC/HEVC, to support emerging technologies such as Lenslet Video Coding (LVC) and Gaussian Splatting Coding (GSC). The core idea is to utilize existing extension mechanisms in the syntax structures of VVC/HEVC to embed LVC or GSC-related data as extension data in standard VVC/HEVC bitstreams to form a unified carriage mechanism. The extension syntax is proposed to start with a specific leading bit pattern which would fit into the specified syntax structure for extension of VVC/HEVC and is intended to avoid any potential conflict with other future extensions of VVC and HEVC. The contribution asserts that the syntax structures for LVC or GSC-related bitstreams and their decoding processes do not need to be additionally defined in the VVC/HEVC standards, and can instead be specified in the respective LVC or GSC standards.

This was presented in a joint meeting on Mon. 19 Jan. at 1540. See the notes in section 7.3.1.

Plenary meetings, joint meetings, BoG reports, and liaison communications

General

Plenary on Monday 19 Jan.

The following topics were discussed in JVET plenary Monday 19 Jan. 1300–1500:

- Status of documents – 2 cross-checks (JVET-AO0261 and JVET-AO0263) still missing – owner are requested to withdraw in case they don't plan to finalize.
- Liaison communication (section 7.5) – 4 liaison inputs, responses to be reviewed on Wednesday, as the JVET sessions on Thursday start by the same time as the AG 3 session for reviewing liaison letters.
- Scheduling for the remaining week was discussed (further detail on scheduling is recorded in section 2.12)
- Joint meetings (different zoom link, see calendar)
 - Joint meeting with VCEG WG 4 on LVC (Mon. 1530-1600)
 - Joint meeting with VCEG WG 4/7 on GSC (Mon. 1600-1730)
 - Joint meeting with VCEG WG 2 AG 5 on NGVC (Thu. 1430-1600)
- Review status from main track
 - EE1 and EE2 review finalized, some remaining documents of non-EE categories and software aspects still to be reviewed.
 - Review of complexity aspects (AHG7/19) – a BoG was established
 - Review of CfP preparation (AHG17) – a BoG was established

- Review of ULL/ER (AHG18) – follow-up discussion about potential inclusion in CfP needs to be conducted (see notes under JMET-AO0052).
- Review status from HLS track and discussion on potential open issues
 - VSEI V4 FDIS – not yet submitted, some editorial changes as result of discussions at current meeting, and will also be integrated to H.274 before pre-publishing.
 - Aspects related to AVC mostly finished
 - Open issue: Text description SEI was requested to be included in AVC, but no software exists yet. Interest was also expressed to include in HEVC, but that has already been submitted for publication. It was agreed in plenary to defer this for future decision after availability of software (and potentially future versions of AVC and HEVC).
 - A request was made to start a WD on localization mapping SEI, but so far no software exists yet. It was agreed in the plenary that a minimum requirement is implementation in software (at minimum writing and parsing), and in cases where concern is raised about usefulness, a convincing showcase needs to be demonstrated.
 - Discussion about a potential timeline for VSEIv5 was deferred for a later meeting.
 - Roughly half of contributions has been reviewed (not yet in the category of new messages), various revisits, likely not contentious.
- Output document planning (section 10)
- Review of documents at plenary level: sections 4.1, 4.3/4.4, 4.5, 4.12, 4.14

Plenary on Wednesday 21 Jan.

The following topics were discussed in JMET plenary Wednesday 21 Jan. 2100–2340:

- Chairs of joint AHG: W. Husak, J. Jung. They should get in contact with the corresponding chairs from WG 4 and 7, to negotiate mandates. Proposal on mandates to be discussed in HLS track on Thursday 1300. JMET will install a “mirror” AHG which may reflect mandates of interest to its work, and may have a subset of the mandates of the joint AHG.
- Status and open issues from review
 - Main track: No specific issues – EE review and planning finalized, draft CfP development is progressing as expected.
 - HLS track: Extensibility of SEI messages in AVC: It was generally agreed that new SEI messages in AVC could have an extension mechanism, but SEI messages that are already existing should not be extended. There are two SEI messages that are currently extensible in the context of VVC and HEVC (NNPFA and FGC), but not extensible in AVC. For those a possible solution could be to define the intended extensions by new SEI messages which would be extensible in the future, but coexist with the previous ones having different codepoints. To allow more study of this issue, it appears appropriate to defer the FDIS until April or July.
- Draft liaison responses to JPEG and SMPTE were reviewed.
- The latest version of the draft CfP was reviewed during 2240–2340.

Plenary on Friday 23 Jan.

The following summary (including list of adoptions) was provided by the HLS track for the plenary on Friday 23 Jan.:

- Request to start ticketing systems for VSEI and TuC to handle minor editorial changes and report bugs, to improve meeting efficiency. K. Sühring will install this.
- From the subsequent adoptions, errata related to HEVC, VVC and VSEI should be included in JVET-AO1004. For the time being, also errata related to AVC which concern errata in the interface to VSEI shall be included in JVET-AO1004, as otherwise the interface would no longer be compatible as long as the counterpart in VSEI is not modified.
- Adoptions Summary from the HLS track (approved in plenary)

VSEI v4:

- - Tiny bug fixes to attempt to include in ITU-T publication process
 - Errata

AVC:

- - Extension mechanism for new SEI messages
 - Errata for interface text

VVC:

- - Errata for interface text

HEVC:

- - Errata for interface text

TuC:

- - New features, editorial improvements, additions of constraints, clarifications, and bug fixes of existing SEI messages in TuC
 - Extensions to existing SEI messages in VSEI v4; extend modality info SEI to support multispectral images
 - No new SEI messages added to TuC

1 VSEI

- - Tiny bug fixes to attempt to include in ITU-T publication process
 - JVET-AO0081 AHG9: Proposed corrections for miscellaneous aspects on VSEI V4
 - JVET-AO0085 AHG9: Editorial improvements for VSEI v4 and TuC
 - Errata
 - JVET-AO0081 AHG9: Proposed corrections for miscellaneous aspects on VSEI V4
 - JVET-AO0280 Editorial study of the extension mechanism for new AVC SEI messages proposed in JVET-AO0212

- JVET-AO0101 AHG9: Text description information SEI message in processing chain

2 VVC

- ○ Errata
 - JVET-AO0101 AHG9: Text description information SEI message in processing chain

3 HEVC

- ○ Errata
 - JVET-AO0081 AHG9: Proposed corrections for miscellaneous aspects on VSEI V4

4 AVC

- ○ JVET-AO0212 Extension mechanism for new AVC SEI messages

5 TuC

- ○ See list under JVET-AO2032
- No new messages

In the context of discussion mandates of AHG7, it was suggested that it might be desirable to conduct investigation with non-CTC sequences, potentially under CfE/CfP conditions (rate matching not necessary).

MPEG information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 19 Jan. 0500–0800, Wednesday 21 Jan. 0500–0600, and Friday 23 Jan. 2100–2300.

The status and plans for the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

Joint meetings

Joint session 1530-1620 Monday 19 Jan. on lenslet video coding: MPEG WG 4 / Video, MPEG WG 5 / JVET, and VCEG (ITU-T Q6/21)

The session description is based on notes by G. J. Sullivan.

This joint session was chaired by Lu Yu (MPEG WG 4 Convenor), Jens-Rainer Ohm (JVET Chair), Gary J. Sullivan (SC 29 Chair & Q6/21 Rapporteur), Jörn Ostermann (MPEG AG 2 Convenor), and Yan Ye (Q6/21 Associate Rapporteur).

WG 4 wants to create a new standard for lenslet video coding (LVC).

Additional coding tools are considered possible to be specified for LVC, but are planned to be outside of the core coding loop.

Thus, the new standard would be a post-processor (from the perspective of the “inner decoder”), producing transformed output pictures from each inner decoder’s output picture.

Some HLS would be added/inserted in the bitstream.

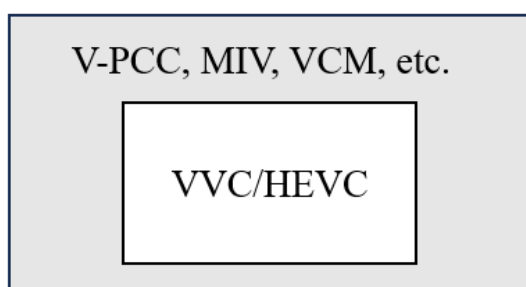
There was discussion about whether it is necessary to touch parameter sets, or if this be supported by SEI or new NAL unit types.

Compatibility with existing decoders was suggested to be a concern.

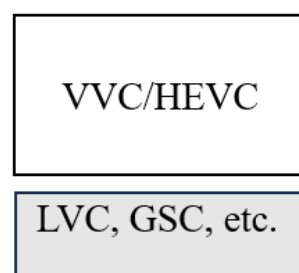
If defined as a new profile of HEVC or VVC (in the ordinary way), that could cause existing decoders to reject the bitstream.

A bitstream extension mechanism is proposed in JVET-AO0284. See the summary of that document provided elsewhere in this report.

As contrasted with the approach taken for MIV, V-PCC, and VCM, in which an HEVC or VVC bitstream would be encapsulated into some other bitstream that would not conform to HEVC or VVC, it is proposed to put extension data within an HEVC or VVC bitstream. These alternative approaches are illustrated in the figure below, where the left side depicts the approach taken for V-PCC, MIV, and VCM, and the right side shows the approach proposed in JVET-AO0284 with embedding of extension data within an HEVC or VVC bitstream.



(a) wrapping approach
(used in V-PCC, MIV and VCM)



(b) embedded extension data approach
(proposed in JVET-AO0284)

The proposal would identify a short bit pattern as an extension type code for parameter sets.

It was commented that a similar extension mechanism had previously been proposed.

There was discussion about whether the base video standard should identify what the particular pattern means.

It was asked whether the parameter sets really need to be changed. Several participants said that SEI might be a better place to specify a post processing syntax.

It was also suggested that we could just use one reserved bit as the extension mechanism if a parameter set extension is needed.

It was asked what would be the expected behaviour of an existing VVC or HEVC decoder that encounters the extended bitstream. It was commented that the scheme would be similar to frame packing – with no impact on low-level decoding process.

This would likely *not* be a new profile from the HEVC or VVC perspective.

Using this scheme could involve defining an extension of every relevant video coding standard (e.g., AVC, HEVC, and VVC).

It was asked how much extension data is involved.

It was commented that parameter sets are ordinarily for syntax relating to the inner decoding process, not post-processing.

It was suggested that post-processing could be normative even if this is done with an SEI message (e.g., like Green metadata, which would not necessarily need to be directly specified in the VSEI standard).

WG 4 suggested not to mention this extension in the relevant core video coding standards. Others thought that at least identifying the extension code and payload size would be needed, saying there is potential for so-called “collisions” in the future to occur.

The proposed extension patterns are currently specified in the core video coding standards as reserved and not to be used.

It was suggested that the interaction with layered coding and with schemes such as rectangle extractions and compositing would need to be considered.

It was suggested that integrating a specified functionality, possibly with a large amount of extra data, should be considered, together with schemes such as film grain synthesis.

Further study seemed needed. An SEI approach such as with Green metadata should be further studied.

SEI with normatively specified post-processing was suggested to be appropriate for this and other similar schemes.

Joint session 1620-1740 Monday 19 Jan. on Gaussian splat coding: MPEG WG 4 / Video, MPEG WG 5 / JVET, MPEG WG 7 3D Graphics, and VCEG (ITU-T Q6/21)

The session description is based on notes by G. J. Sullivan.

This session was chaired by Lu Yu (MPEG WG 4 Convenor), Gary Sullivan (SC 29 Chair and Q6/21 Rapporteur), Jörn Ostermann (MPEG AG 2 Convenor), Yan Ye (Q6/21 Associate Rapporteur), and Marius Preda (MPEG WG 7 Convenor).

The initial discussion included:

- The incoming liaison statement from SG21.
- A summary of status in MPEG WGs 4 & 7.
- ~7 technical documents for SEI that had been submitted.
- JVET-AO0152 was suggested for focus.

Liaison statement content from ITU-T SG 21:

“This liaison statement expresses to ISO/IEC JTC 1/SC 29 and its working groups an interest in collaboration on 3D Gaussian Splat coding technologies that take advantage of software and hardware implementations of existing and future video coding standards jointly developed by ITU-T and ISO/IEC

ITU-T Study Group 21 (SG21) wishes to inform ISO/IEC JTC 1/SC 29 and its working groups of its recent discussions and developments regarding 3D Gaussian Splat coding technologies.

Significant industry interest has been voiced during our meetings in support of standardizing technologies for 3D Gaussian Splat coding, particularly taking advantage of software and hardware implementations of existing and future video coding standards jointly developed by ITU-T and ISO/IEC.

According to proponents, such industry interest can be effectively addressed using standardized supplemental enhancement information (SEI) messages, and several related contributions have been submitted to recent meetings of the ITU-T/ISO/IEC Joint Video Experts Team (JVET). This approach has been discussed within SG21, and no objections to this approach have been raised.

We suggest that an appropriate standard for such SEI messages is the Video Supplemental Enhancement Information (VSEI) standard (Rec. ITU-T H.274 | ISO/IEC 23002-7), which provides a unified framework for SEI technologies applicable to the modern video coding standards jointly developed by ITU-T and ISO/IEC. We further note that VSEI is developed and maintained within JVET.

In light of the shared interests and potential for alignment, SG21 proposes a joint discussion between SG21 and SC 29 experts during the upcoming scheduled teleconference meetings of Q6/21 and MPEG in January 2026, to explore possible coordination and exchange of technical perspectives on this topic.

We appreciate the continued collaboration between our groups and look forward to further engagement.”

This liaison statement suggests VSEI as the approach.

It was commented that SG21 might not be aware of the status in SC 29, but it was commented that the status in SC 29 was generally known to the SG21 participants.

The status in WG 4 and 7 was reviewed, with three projects and active explorations that were noted:

- An amendment of V-PCC, with WG 7 planning to issue a CDAM at the current meeting
- An amendment of G-PCC, with WG 7 planning to issue a CDAM at the next meeting (in April 2026)
- Additional exploration activity, targeting development of a CfP, to include dynamic splatting as well as static content representation

Guide to abbreviations:

- PCC = point cloud compression
- V = video-based (e.g. using VVC or HEVC)
- G = geometry-based.

Gaussian splats can be interpreted as using attributes associated with points in a point cloud, where the centre of each 3D Gaussian is the corresponding point in the point cloud.

V3C is the framework used for V-PCC, MIV (MPEG immersive video) and V-DMC (dynamic mesh coding).

G-PCC has a recently added profile with extended transform precision, modified quantization, and other enhancements.

V-PCC uses four video frames (or single frames with regions) that carry the following information:

- Geometry
- Texture
- Occupancy

Atlas data syntax, describing patches, is specified.

The principal focus of the V-PCC planned near-term amendment is static Gaussian splats, coded using video I frames / 2D still pictures.

There is also the study of dynamic (i.e., moving) Gaussian splats, planned for a future CfP.

Bit depth support was discussed; even 8-bit video approaches are considered.

JVET-AO0152 was discussed, summarized as:

This contribution argues that it is within JVET's mandate to work on SEI messages for VSEI to interpret samples reconstructed from coded video pictures according to H.26[4/5/6] / AVC, HEVC, VVC and future video codec designs as Gaussian Splat parameters. Technical, business, and venue-related considerations are included. For this meeting, it is proposed to allow work, including proposal presentation and TuC-level decision making, in the JVET HLS breakout.

Use of a video approach was suggested by the proponent to leverage high-volume implementations of video standards.

Using VSEI was suggested by the proponent, due to its applicability to multiple video coding standards.

Basic architectural approaches as described for the lenslet handling could apply in principle, and there could be different groups where the work could be done.

It was commented that having good ways to compare approaches (e.g. common test conditions) and avoiding duplications are desirable. It was pointed out that multiple approaches are already being pursued.

It was commented that MPEG WG 2 (MPEG Requirements) is also relevant and studying related matters.

Discarding of SEI messages has been practiced. It was commented that the “no frame packing” flag can help. Conformance specification is also a matter of concern. A conformance specification for how to use the data may be needed.

We have already agreed that, in principle, there can be normative conformance requirements that involve an SEI message (or similar kinds of data, such as reserved ignored data or other NAL units).

The mapping between the 3D representation and the 2D representation used for video/image coding was suggested to be critical. It was suggested that different encapsulation types could be specified in a harmonized way. A basic scheme could be the same aside from how the data is encapsulated.

Having joint investigation of the topic by MPEG WG 7 with JVET (as well as WG 4) was suggested, starting at the current meeting, including consideration of the 6 other noted contributions (which are not all technically distinct). Daily joint sessions of WGs 4 & 7, of 2 hours per day [1300 UTC Tuesday] had been planned already.

The extension/encapsulation method for how to carry the syntax elements that are not put into 2D picture data was suggested to be lower immediate priority than the effectiveness of the basic scheme.

Dynamic splats were suggested to be very important.

Another participant emphasized a desire to develop something relatively quickly that can be implemented with relatively low resource requirements.

It was agreed to proceed with the joint meeting session as noted and as described in the next section.

Joint session 1300-1510 Tuesday 20 Jan. on review of Gaussian splatting technical inputs: MPEG WG 4 / Video, MPEG WG 5 / JVET, and MPEG WG 7 / 3D Graphics Coding

This joint meeting with WG 4 and WG 7 was held during 1300–1510 UTC on Tuesday 20 Jan. 2026 (chaired by JRO, M. Preda and L. Yu). Notes captured during this meeting can be found in section 6.3.1.

As a follow-up action, a Joint AHG was established on the topic (see section 9)

Joint session 1430-1600 Thursday 22 Jan. on next generation video standardization Call for Proposals: MPEG WG 2 / Requirements, MPEG WG 5 / JVET, MPEG AG 5 Visual Quality Assessment, and VCEG (ITU-T Q6/21)

This joint session was chaired by Jens-Rainer Ohm (JVET Chair and WG 5 Convenor), Gary Sullivan (VCEG Rapporteur and SC 29 Chair), Yan Ye (VCEG Associate Rapporteur), Mathias Wien (AG 5 Convenor), and Jörn Ostermann (AG 2 Convenor) and Mary-Luc Champel, on behalf of Igor Curcio (WG 2 Convenor).

This section is based on notes taken by GJS.

About 300 people attended this session. The following issues were discussed:

- The basic goal of the joint session was to assess whether the CfP draft was mature enough to declare and issue it as an official “Draft CfP” (not planned as final), including whether it was.
 - Sufficiently detailed on what will be done and when
 - Sufficiently aligned with requirements
- The draft was presented by M. Wien and was somewhat refined in the discussion. This included review of
 - timelines
 - categories of test content
 - anchors
- Aspects to be disclosed to proponents in advance and additional testing to be conducted later.
- Possible use of expert viewing for subjective evaluation was suggested to be clarified.
- Ultra-low latency and error/loss resilience and concealment – additional functionalities can be demonstrated in response to the CfP and will be considered in the collaborative phase.
- Further refinement of the requirements document can be conducted later, and can hypothetically be ongoing activities.
- Estimation of the expected amount of the testing fee.
- Potential use of a Docker container. It was commented that this might be unnecessary. (The current Ubuntu guidance is already only a “should”.)
- Requesting encoder executables and what we might do with them.
- Planned subjective evaluation methodology.
- The potential requirement for complete submissions for all test sequences in each test case, and what might occur if a submitter wishes to submit partial results.

Further refinements can be made until the final CfP is produced.

It was noted that further study will be conducted at an AHG meeting in February in Aachen.

In answer to the primary question of whether the CfP draft is mature enough to declare and issue it as a “Draft CfP” (not planned as final), the answer was Yes; this was approved.

An editing period for the draft was agreed – until January 30.

BoGs (2)

CITATION

Y. Liao, L. Yu, X. Jin, L. Zhao, "A Bitstream Extension Mechanism for New Technologies Based on Video Coding", JVET-AO0284, 41st Meeting: by teleconference, CH, January 2026.