

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION**



TITLE: Final Call for Proposals on JPEG Pleno Point Cloud Coding

SOURCE: WG1

PROJECT: ISO/IEC 21794 (JPEG Pleno)

STATUS: Approved

REQUESTED ACTION: For distribution

DISTRIBUTION: Public

Contact:

ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

Table of Contents

| | |
|---|----|
| Summary | 3 |
| Introduction | 4 |
| Scope | 4 |
| Use Cases and Requirements | 5 |
| Evaluation Conditions and Processes | 5 |
| Timeline | 6 |
| Proposal Composition and Requirements | 7 |
| <i>Proposal Elements</i> | 7 |
| <i>Proposal Registration and Delivery</i> | 8 |
| <i>IPR Conditions (ISO/IEC Directives)</i> | 9 |
| <i>Contribution to Standardization</i> | 9 |
| Royalty-free Goal | 9 |
| Participation | 10 |
| Contacts | 10 |
| References | 10 |
| Annex A – Further Details on Proposal Composition and Requirements | 12 |
| A.1. Proposal Description | 12 |
| A.2. Encoding and Decoding Software and Usage Information | 12 |
| A.3. Coded Bitstreams | 13 |
| A.4. Decoded Point Clouds | 13 |
| A.5. Computational Complexity Metrics Results | 13 |
| A.6. Optional Proposal Elements | 13 |
| A.7. Verification Model Source Code | 14 |

Final Call for Proposals on JPEG Pleno Point Cloud Coding

Summary

This document contains the Final Call for Proposals (CfP) on JPEG Pleno Point Cloud Coding issued in the context of the JPEG Pleno standardization project, which aims at developing next generation visual information coding standards that move beyond coding of 2D planar content by taking advantage of plenoptic representations.

This call addresses learning-based coding technologies for static point cloud content and associated attributes with emphasis on both human visualization and decompressed/reconstructed domain 3D processing and computer vision with competitive compression efficiency compared to point cloud coding standards in common use, with the goal of supporting a royalty-free baseline.

Additionally, contributions are encouraged in the form of:

- Use cases and requirements not yet identified in the JPEG Pleno Point Cloud Use Cases and Requirements document [2];
- Representative datasets for potential applications including those not currently identified by JPEG. Datasets should be supplied with conditions allowing usage for standardization and organization of special sessions and grand challenges in scientific events;
- Subjective evaluation methodologies and test-bed implementations that can be used to assess the various requirements identified (or new requirements if not already identified);
- Objective evaluation methodologies and test-bed implementations that can be used to assess the various requirements identified (or new requirements if not already identified).

Final Call for Proposals on JPEG Pleno Point Cloud Coding

1. Introduction

JPEG Pleno will integrate various modalities of plenoptic content under a single framework in a seamless manner. Efficient and powerful point cloud representation is a key feature of this vision [1]. In this document a *point cloud* refers to data representing positions of points in space, expressed in a given three-dimensional coordinate system, the so-called geometry. This geometrical data can be accompanied with per-point attributes of varying nature (e.g. color or reflectance). Such datasets are usually acquired with a 3D scanner, LIDAR or created using 3D design software and can subsequently be used to represent and render 3D surfaces. Combined with other types of data (like light field data), point clouds open a wide range of new opportunities, notably for immersive browsing and virtual reality applications.

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. Recently, learning-based point cloud coding solutions have shown great promise to achieve competitive compression efficiency compared to available conventional point cloud coding solutions at equivalent subjective quality.

Building on a history of successful and widely adopted coding standards, JPEG is well positioned to develop a standard for a learning-based point cloud coding.

This Call for Proposals solicits learning-based point cloud coding solutions for the purposes of development of relevant standards in this field.

Proponents are invited to first register their interest in the Call for Proposals by 23:59h UTC, 31st March 2022, followed by the submission of proposal encoder and decoder implementations by 31st May 2022 and then the submission of a full proposal by 23:59h UTC, 17th June 2022. Proposals deemed suitable to proceed to the following collaborative standardization process will be notified during the 96th JPEG Meeting in July 2022.

2. Scope

The scope of the JPEG Pleno Point Cloud activity is the creation of a learning-based coding standard for point clouds and associated attributes, offering a single-stream, compact compressed domain representation, supporting advanced flexible data access functionalities. This standard targets both interactive human visualization, with competitive compression efficiency compared to state of the art point cloud coding solutions in common use, and effective performance for 3D processing and machine-related computer vision tasks, and has the goal of supporting a royalty-free baseline.

This standard is envisioned to provide a number of unique benefits, including a single efficient point cloud representation for both humans and machines. The intent is to provide humans with the ability to visualize

and interact with the point cloud geometry and attributes while providing machines the ability to perform 3D processing and computer vision tasks in the decompressed/reconstructed domain, notably by enforcing error constraints, and in the compressed domain (latents after entropy decoding), notably by enabling lower complexity and higher accuracy through the use of compressed domain features extracted from the original instead of the lossy decoded point cloud.

To support the scope above, this activity will advance through a series of stages which shall develop as:

- Stage 1: A learning-based coding standard addressing human visualization and decompressed/reconstructed domain 3D processing and computer vision tasks;
- Stage 2: A learning-based coding standard additionally supporting compressed domain 3D processing such as visual enhancement and super-resolution and;
- Stage 3: A learning-based coding standard additionally supporting compressed domain computer vision tasks such as classification, recognition and segmentation.

This Call for Proposals is in regard to Stage 1: ***A learning-based coding standard addressing human visualization and decompressed/reconstructed domain 3D processing and computer vision tasks***. Proposals must demonstrate competitive compression efficiency compared to existing solutions.

Later Calls for Proposals will focus on Stages 2 and 3. While this Call for Proposals addresses and will assess technologies offering competitive compression efficiency targeting human visualization and decompressed/reconstructed domain machine consumption, proponents are welcome to bring any additional evidence regarding the functionalities in Stages 2 and 3, i.e. 3D processing and computer vision tasks in the compressed domain.

3. Use Cases and Requirements

Point cloud data supports a wide range of applications including computer-aided manufacturing, entertainment, virtual and augmented reality display, cultural heritage preservation, autonomous navigation and remote sensing and geographical information systems. A detailed list of use cases and requirements for this activity identified by JPEG is given in the JPEG Pleno Point Cloud Use Cases and Requirements document [2].

4. Evaluation Conditions and Processes

The JPEG Pleno Point Cloud Common Training and Test Conditions (CTTC) document [3] defines the training dataset (the test dataset will be provided after encoder and decoder implementations are provided by proponents), benchmarking codecs, coding conditions (especially target bitrates) and a set of reliable objective quality metrics and subjective assessment procedures. The JPEG Pleno Point Cloud Coding Common Training and Test Conditions will allow the exhaustive evaluation of multiple aspects of the proposed learning-based image codecs to fully understand their strengths and weaknesses, notably regarding already available point cloud coding technology.

JPEG plans to select technologies to be included in the JPEG Pleno Point Cloud Coding standard based on satisfying the requirements as well as compliance to the JPEG Pleno Point Cloud Coding Common Training

and Test Conditions, such as target bitrates. Several evaluation dimensions will be used for the assessment of the submissions and the decision-making process:

- Subjective quality evaluation for point clouds decoded with the proposed submission technology under the JPEG Pleno Point Cloud Coding Common Training and Test Conditions. The subjective evaluation results will be the primary element for the decision-making process;
- Objective quality evaluation for point clouds decoded with the proposed submission technology using the quality metrics and conditions defined in the JPEG Pleno Point Cloud Coding Common Training and Test Conditions;
- Complexity evaluation of both encoding and decoding processes according to the metrics defined in the JPEG Pleno Point Cloud Coding Common Training and Test Conditions.

5. Timeline

The following schedule is planned for the Call for Proposals on JPEG Pleno Point Cloud Coding (all times are given in Coordinated Universal Time (UTC)):

| | |
|--|--|
| 94th JPEG Meeting, 17th-21st January 2022 | <p>Release of Final Call for Proposals on JPEG Pleno Point Cloud Coding (WG1N100097).</p> <p>Release of the final version of the JPEG Pleno Point Cloud Use Cases and Requirements document (WG1N100096).</p> <p>Release of the final version of the JPEG Pleno Point Cloud Coding Common Training and Test Conditions (CTTC) document (WG1N100112).</p> |
| 23:59h UTC, 31st March 2022 | <p>Deadline for registration for participation in this Call for Proposals. Registrations should be sent by email to the persons listed in Section 9. Following registration, members of the JPEG Committee will advise proponents on how to submit elements of the proposal.</p> <p>Release of the training and validation parts of the datasets.</p> |
| 95th JPEG Meeting, 25th - 29th April 2022 | <p>Status evaluation of the Call for Proposals, review of registrations, organization of evaluation procedures and practical arrangements.</p> |
| 23:59h UTC, 31st May 2022 | <p>Deadline for submission of proposal encoder and decoder implementations. No (re)training is allowed after this date.</p> |
| 23:59h UTC, 3rd June 2022 | <p>Release of the test datasets for proponents to code.</p> |
| 23:59h UTC, 17th June 2022 | <p>Deadline for submission of encoded and reconstructed test set point clouds from proposals as well as proposal</p> |

presentation to be presented at the 96th JPEG Meeting, the proposal description and all other elements of the proposal as described in Section 6.1 and Annex A. JPEG begins subjective evaluation of proposals.

| | |
|--|--|
| 96th JPEG Meeting, 25th – 29th July 2022 | Presentation of proposals, together with review of subjective evaluation results from experiments performed by JPEG. |
| | Decision on future actions regarding point cloud coding standardization in JPEG. |
| October 2022 | Release of Working Draft (WD) on JPEG Pleno Point Cloud Coding. |
| July 2023 | Release of Committee Draft (CD) on JPEG Pleno Point Cloud Coding. |
| January 2024 | Release of Draft International Standard (DIS) on JPEG Pleno Point Cloud Coding. |
| July 2024 | Publication of the International Standard on JPEG Pleno Point Cloud Coding. |

6. Proposal Composition and Requirements

6.1. Proposal Elements

Proponents are asked to submit a detailed technical description of the entire point cloud codec, as well as encoder and decoder implementations in software, and the decoded point clouds. Regarding training material, proponents must use the defined JPEG Pleno Point Cloud training dataset (if a subset of the training dataset is used for training then this should be clearly described).

Once the point cloud test material is provided to the proponents as described in Section A.2, proponents are expected to code the provided test material, and provide the decoded point clouds along with their codec, targeting the bitrates specified as mandatory in JPEG Pleno Point Cloud Coding Common Training and Test Conditions [3]. In all cases, participants are required to submit enough material to validate the performance of their submission for the standard decoding task according to the procedure outlined below, notably the proponents are **required** to provide:

- A detailed technical description of the encoding and decoding algorithms in a document following a template which will be sent to the proponents after registration (a template for a JPEG meeting numbered input document);
- Details on the training process used to develop the proposal including training datasets, loss functions, hyper-parameter values and any other information required to reproduce the results given by the proposal.

Expect that JPEG members may reproduce your results by performing the training procedure, obtaining a model and running the encoder and decoder. **Only proposals trained exclusively with the datasets specified in the JPEG Pleno Point Cloud Common Training and Test Conditions [3], will be subject to subjective testing. Proposals trained using datasets outside of those specified in the JPEG Pleno Point Cloud Common Training and Test Conditions [3] will only be considered if the additional datasets used are supplied to JPEG and have usage arrangements that allow the free use of the datasets by JPEG for standardization purposes. Proposals using additional datasets that meet the above conditions will be considered by JPEG as far as possible based on objective measures and information supplied by the proponents;**

- Both an encoder and decoder implementation, including information on hardware and software requirements to run both implementations. Both the encoder and decoder implementations should be in a form allowing stand-alone inference/testing on a standard computer (both CPU and GPU) in a reasonable amount of time, preferably in source code form with clear instructions on how to build, run the software and configure it for CPU and GPU. Contributors are also expected to provide to JPEG sufficient rights to allow usage of the provided software for the purpose of evaluation;
- Easy to follow but complete instructions on how to use the implementation(s) provided, including parameter values to be used and configuration files, if called for by the implementations;
- Bitstreams representing each of the test point clouds supplied by JPEG to proponents, in the compressed format associated with the proposed solution;
- Full set of decoded point clouds obtained from the bitstreams provided, corresponding to each test point cloud;
- Computational complexity information for both the encoder and decoder implementations using the measures listed in the JPEG Pleno Point Cloud Coding Common Training and Test Conditions document [3].

Detailed submission requirements for the proposed solutions are listed in Annex A.

6.2. Proposal Registration and Delivery

Resources Available to Proponents

Proponents are directed to the following publicly available resources relevant to this Call for Proposals:

- JPEG Pleno Point Cloud Use Cases and Requirements document [2];
- JPEG Pleno Point Cloud Coding Common Training and Test Conditions document [3].
- Point cloud datasets for training submissions as described in the JPEG Pleno Point Cloud Coding Common Training and Test Conditions [3];

Following supply of an encoder and decoder implementation, proponents will be given access to test point clouds in the PLY file format (*.PLY). This test content will only be provided after the proponents provide the encoder and decoder implementations as described in Section 6.1.

Proposal Registration

Individuals or teams who desire to answer to this Call for Proposals propose solutions should inform JPEG of their intention by sending an email to the persons listed in Section 9 by the deadline indicated in Section 5. Proponents will then be given instructions regarding how to submit the required materials and where to download the training content.

Final Proposal Delivery

Once informed about how to deliver a proposal, proponents must send all materials listed in Annex A to JPEG, via the persons listed in Section 9, no later than the submission deadline as per the timeline detailed in Section 5. Point cloud test content will be provided to proponents only after the encoder and decoder implementations are provided.

Additionally, proponents shall give a detailed presentation of the submission during the 96th JPEG meeting, July 2022, where one of the proponents should attend the meeting to present the proposal and answer questions and requests for clarifications. Proponents should be aware that acceptance of the proposed technology for inclusion in the collaborative standardization process will require JPEG meeting participation for a few meetings.

Prior to the 96th JPEG Meeting, July 2022, JPEG will perform subjective evaluation of proposals and the computation of objective performance metrics according to the JPEG Pleno Point Cloud Coding Common Training and Test Conditions [3]. Hence it is mandatory that proponents provide the information detailed in Annex A by the submission deadline detailed in Section 5.

6.3. IPR Conditions (ISO/IEC Directives)

Proponents are advised that this call is being made in the framework and subject to the common patent policy of ITU-T/ITU-R/ISO/IEC and other established policies of these standardization organizations. The persons named below as contacts can assist potential submitters in identifying the relevant policy information.

6.4. Contribution to Standardization

Proponents are informed that based on the submitted proposals, a standard specification will be created. If they submit a proposal and (part of) the proposed technology is accepted for inclusion in the standard, they will hence have to attend subsequent JPEG meetings and contribute to the creation of the different standard documents. Within this process, evolution and changes are possible as several technologies may be combined to obtain a better performing solution.

7. Royalty-free Goal

The royalty-free patent licensing commitments made by contributors to previous standards, e.g. JPEG 2000 Part 1, have arguably been instrumental to their success. JPEG expects that similar commitments would be helpful for the adoption of a JPEG Pleno Point Cloud coding standard.

8. Participation

The Ad Hoc Group on JPEG Pleno Point Cloud Coding was established at the 88th JPEG meeting to develop standards in relation to point cloud coding. All interested parties are encouraged to register to the email reflector of the AhG (E-mail reflector: jpeg-pointcloud@jpeglists.org).

To subscribe to the E-mail reflector, please visit <http://jpeg-pointcloud-list.jpeg.org>; in case of problems, please contact lists@jpeg.org.

9. Contacts

The following contacts are available for clarifications regarding this Call for Proposals:

Touradj Ebrahimi (JPEG Convener)

Email: Touradj.Ebrahimi@epfl.ch

Fernando Pereira (JPEG Requirements Subgroup Chair)

Email: fp@lx.it.pt

Peter Schelkens (JPEG Plenoptic Coding & Quality Subgroup Chair)

Email: peter.schelkens@vub.be

Stuart Perry (Ad Hoc Group on JPEG Pleno Point Cloud Chair)

Email: Stuart.Perry@uts.edu.au

Luis Cruz (Ad Hoc Group on JPEG Pleno Point Cloud Co-Chair)

Email: lcruz@deec.uc.pt

References

- [1] P. Astola, L. A. da Silva Cruz, E. A. B. da Silva, T. Ebrahimi, P. G. Freitas, A. Gilles, K.-J. Oh, C. Pagliari, F. Pereira, C. Perra, S. Perry, A. M. G. Pinheiro, P. Schelkens, I. Seidel, I. Tabus, “JPEG Pleno: Standardizing a Coding Framework and Tools for Plenoptic Imaging Modalities”, ITU Journal: ICT Discoveries, Special Issue on The Future of Video and Immersive Media, vol. 3, no. 1, June 2020.

- [2] ISO/IEC JTC1/SC29/WG1 N100096, “JPEG Pleno Point Cloud – Use Cases and Requirements v1.6”, Online, January 2022.

- [3] ISO/IEC JTC1/SC29/WG1 N100112, “JPEG Pleno Point Cloud Coding Common Training and Test Conditions v1.1”, Online, January 2022.

Annex A – Further Details on Proposal Composition and Requirements

A.1. Proposal Description

Each proposal shall include a document with a detailed description of the proposed solution, including a high-level description of the proposal with block diagrams of the encoder and decoder. This information should contain sufficient details regarding the framework, training, software, usage rights and any other details needed to allow for JPEG to reproduce the results of the proposal on both the training and test datasets. In particular:

- Key features of the proposal, including the target quality range and covered bitrates;
- High-level description of the proposal, including the encoder/decoder architecture;
- Detailed analysis and explanation on how the JPEG Pleno Point Cloud Coding requirements listed in [2] are met by the proposed solution (not all requirements have to be addressed, although it is desirable that as many as possible are addressed, notably core requirements);
- Detailed architectures for every module, training procedure including loss function used, strategies used to handle key parts, such as the non-differentiable nature of quantization and bitrate allocation;
- Model size if applicable, which corresponds to the number of weights (and precision of each weight) in the encoder and decoder;
- Running time (encoder and decoder) for an example CPU and GPU platform. Include all the details of the platform (CPU model, clock rate and memory, GPU model and brand) but also the deep learning framework (e.g. Tensorflow or Pytorch). The recommended platform for the GPU is an NVIDIA 2080 Ti, but you may use other platforms.

This description shall be in Word document and PDF format. Providing insufficient details might jeopardize the adoption of the proposed technology.

A.2. Encoding and Decoding Software and Usage Information

Proponents need to provide software to run the encoder and decoder implementations in the final fully-trained forms, in addition to all information required to run the training process and run the encoder and decoder implementations. This includes a standalone executable package/docker file with all the libraries and tools needed to run the encoder and decoder with the submitted code-streams and preferably the decoder in source code form and sufficient description how to build and run the software. Proponents shall provide the command-line parameters intended to be used for encoding/decoding, and shall provide scripts to run their executable in the objective evaluation framework detailed below. All the information to run the decoder shall be provided, such as command line parameters and configuration files. If binaries are used, they should ideally correspond to statically linked Linux executables with all required libraries and system dependencies.

Once the proponents provide this software, JPEG will provide a set of point clouds for testing purposes. **Proponents are prohibited from re-training their solution following the provision of the test set point clouds. The codec training procedure must be described as faithfully as possible.**

Any auxiliary packages or software on which the submitted implementation(s) depend has to be clearly

identified (including version numbers) and, if not publicly available, also supplied.

Proponents can choose to use executable compression or similar tools to prevent reverse engineering or disassembly of the submitted executables at the time of submission, but as mentioned in Section A.7, proponents will need to provide source code in the event that all or part of their proposal is chosen to be part of a standard.

Proponents of codecs that support scalability, must provide clear instructions on how to decode bitstreams at various levels of scalability and on how to extract specific random access sets of points for solutions supporting random access and scalability.

A.3. Coded Bitstreams

Proponents need to submit coded bitstreams for each of the test content at each of the target bit rates described in the JPEG Pleno Point Cloud Coding Common Training and Test Conditions [3]. Test content will only be provided after encoding and decoding software and usage information as per Section A.2 has been provided.

A.4. Decoded Point Clouds

Proponents are required to provide decoded test point clouds corresponding to each bitstream supplied. In the case of random-access and scalability compatible bitstreams, the proponents should also include decoded point clouds for each scalability level and random access point/configuration deemed necessary to demonstrate the performance of the solution proposed, in terms of those functionalities.

A.5. Computational Complexity Metrics Results

The proponents need to submit tables and plots of each of the computational complexity metrics defined in JPEG Pleno Point Cloud Coding Common Training and Test Conditions [3], computed on each of the test point clouds at each of the test rates defined in [3].

A.6. Optional Proposal Elements

Additionally, this Call for Proposals encourages contributions on several **optional** elements, notably:

- Use cases and requirements not yet identified in the JPEG Pleno Point Cloud Use Cases and Requirements document [2];
- Representative datasets for potential applications including those not currently identified by JPEG. Datasets should be supplied with conditions allowing usage for standardization and organization of special sessions and grand challenges in scientific events;
- Subjective evaluation methodologies and test-bed implementations that can be used to assess the various requirements identified (or new requirements if not already identified);
- Objective evaluation methodologies and test-bed implementations that can be used to assess the various requirements identified (or new requirements if not already identified).

A.7. Verification Model Source Code

If parts of their technology are selected in the evaluation process, proponents agree to release source code (encoder and decoder) to serve as (part of) a Verification Model (VM), written in a high-level language, such as Python. The source code provided may include some project files needed to support compilation. Source code shall be documented and understandable. All libraries used by the source code shall be either public or provided in source code form with ISO/IEC and ITU-T compliant terms. The source code to be provided should run on a variety of operating systems (at least on Linux) and hardware, namely CPU processors and GPU boards.