ISO/IEC JTC1/SC29/WG1 **N92015**
92th Meeting – Online – 7-13 July 2021

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION

ISO/IEC JTC 1/SC 29/WG1
(ITU-T SG16)

Coding of Still Pictures

JBIG	**JPEG**
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

TITLE: JPEG Pleno Point Cloud – Use Cases and Requirements v1.4

SOURCE: Stuart Perry (Editor)

PROJECT: JPEG Pleno

STATUS: Approved

REQUESTED ACTION:

DISTRIBUTION: Public

Contact:
ISO/IEC JTC 1/SC 29/WG1 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

Editorial Comments

This is a living document that goes through iterations. Proposals for revisions of the text can be delivered to the editor Stuart Perry, by downloading this document, editing it using track changes and sending it to Stuart.Perry@uts.edu.au.

If you have interest in JPEG Pleno Point Cloud, please subscribe to the email reflector, via the following link: http://listregistration.jpeg.org/

ISO/IEC JTC1/SC29/WG1 N92015
92th Meeting – Online – 7-13 July 2021

JPEG Pleno Point Cloud - Use Cases and Requirements v1.4

July 13th, 2021

1 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 significant compression efficiency over state of the art point cloud coding solutions in common use at equivalent subjective quality, and effective performance for 3D processing and 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 visualise and interact with the point cloud geometry and attributes while providing machines the ability to perform 3D processing and computer vision tasks in the compressed domain 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:

Stage 1: A standard addressing learning-based coding with emphasis on human visualization;

Stage 2: A standard addressing learning-based coding supporting both human visualization and compressed domain 3D processing such as de-noising and super-resolution; and

Stage 3: A standard addressing learning-based coding supporting both human visualization and compressed domain 3D processing with the addition of support for compressed domain computer vision tasks such as classification, recognition and segmentation.

The first and second stages of this activity will not involve all of the requirements in this document. At each stage subsequent to stage 1, the set of requirements within this document applicable to that stage will grow, including requirements of earlier stages, but adding additional requirements. The eventual goal is to reach a solution for stage 3 that encompasses all of the requirements listed in this document.

2 Use cases

This section collects use cases where point cloud representations may play an important role. The use cases in this section refer primarily to static point clouds, however some use cases may also relate to dynamic point clouds. This report is not intended to be restricted to only static point clouds. This section provides short descriptions of each use case and the general features of each use case that inform requirements. Specific key requirements for each use case are listed in Appendix A.

2.1 Virtual, augmented and mixed reality

2.1.1 Computer graphics and gaming

Point clouds are instrumental to support the display of 3D content in virtual, augmented and mixed reality (VR/AR/MR) environments and 2D or 3D computer graphics and gaming scenarios.

The content may be generated artificially by computer-aided design (CAD) design software or created through the scanning of physical objects with 3D scanning equipment. The resolution and number of points of the content will be highly dependent on the type of object and industry. Geometrically, point clouds created using CAD software are likely to be arranged on regular grids and patterns, while data collected using 3D scanning equipment is likely to be arranged in an irregular geometric pattern unless re-meshed in some fashion. Attributes of points in this data is likely to include colour, gloss, bump maps or texture maps as well as bidirectional reflectance distribution function (BRDF) or other material appearance information. Point clouds may need to be rendered at frame rates of 5 to 30 fps and streaming capabilities may be also of importance as will the need to view the data at different scales and from random viewpoints. Figure 1**Error! Reference source not found.** provides an example workflow for this use case, illustrating the requirement to render the point cloud from changing viewpoints, scales and resolutions in real time in the course of the gaming or VR experience.

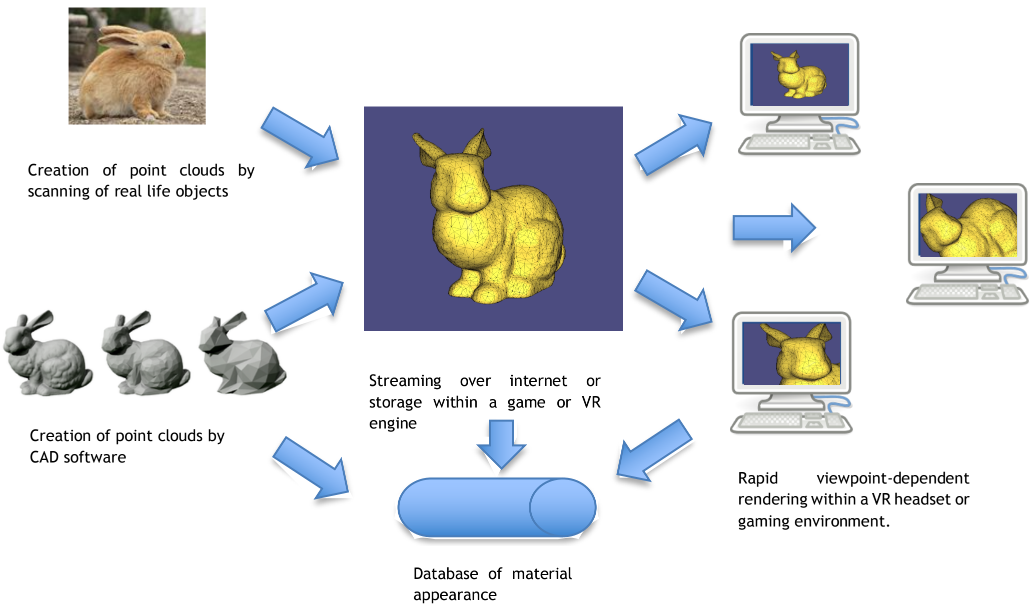

Figure 1: An example of use case 2.1.1. A 3D model is created by either the scanning of real objects or using Computer-Aided Design (CAD) software. This point cloud is streamed over the internet, a local network or stored within a game or VR engine to be displayed to the user. Depending on the viewpoint and scale of the object in the virtual environment, selective view selectivity or scalable bit stream capabilities may be needed for low latency display. In the case of computer generated content, a database of material appearance may be referenced at various points in the use case workflow. Image of live rabbit by Anna

Frodesiak [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)] and central image of rabbit mesh courtesy of G1malitm [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)]

2.1.2 Telepresence

This use case involves point clouds inside VR, AR or MR systems as a representation format for human shapes enabling interpersonal immersive interactions. Typical applications may feature bidirectional streaming of data between two endpoints (one-to-one), broadcasting from one to multiple users (one-to-many) or multilateral simultaneous interactions between groups of users (many-to-many). This use case requires support for low latency and real-time performance and may require 3D processing to merge multiple point clouds, correct artefacts or colour deficiencies in the compressed domain without time-consuming human intervention.

2.2 3D content creation

2.2.1 Design, manufacturing and 3D printing

This use case involves artificially created point clouds to support the manufacture of 3D content for 3D printing or traditional manufacture.

The resolution and number of points in the point clouds will be highly dependent on type of object and industry; however, geometrically, data is likely to be arranged on regular grids and patterns associated with CAD software. Attributes of points in this data are likely to include colour as well as other material appearance information, potentially linked to external databases containing a set number of materials. This use case requires accurate representation of point clouds so lossless representation will be important for geometry and attributes.
Figure 2 shows an example workflow for this use case. In this example, an object is created and then transmitted over an insecure channel for later 3D printing. It is essential that a means for ensuring the privacy and security of the content are available to protect the intellectual property of the creator.

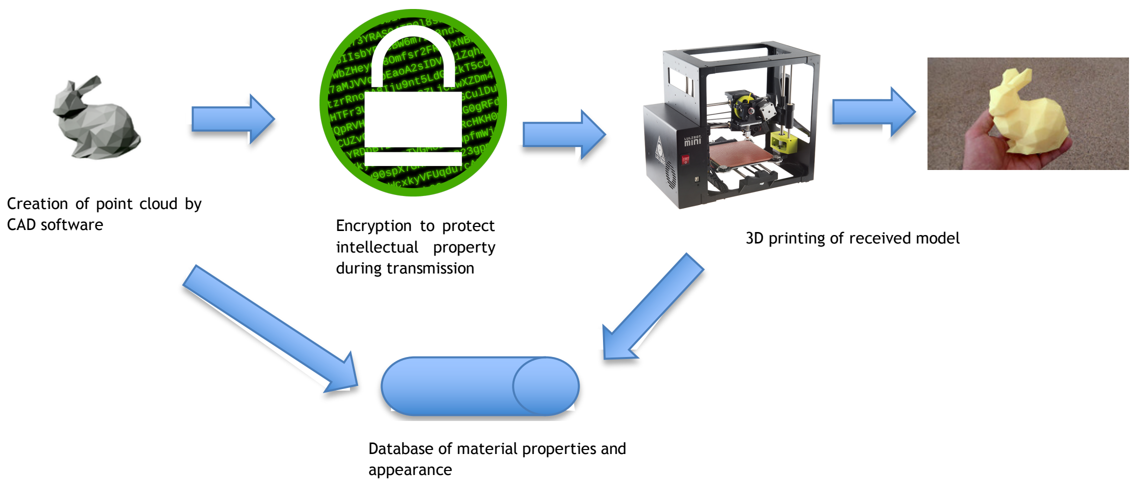

Figure 2: An example of use case 2.2.1. CAD software is used to create a 3D point cloud intended for printing. The point cloud requires security mechanisms to ensure protection of the creator's intellectual property during transmission. Image of 3D printed rabbit courtesy of Creative Tools [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/)]. Image of 3D printer courtesy of SparkFun Electronics

2.2.2 Motion capture

Motion capture with or without markers often results in point clouds where points correspond to each marker (either real or virtual).

Point clouds derived from motion capture data are relevant for motion analysis and synthesis for various applications such as sports analysis, video game production, choreography, and movie production. Motion capture systems link artificially generated content with scans of real life content and to allow this to occur without human intervention, support for 3D processing in the compressed domain is needed to perform operations such as filling gaps and holes as well as computer vision in the compressed domain to detect specific objects.

2.3 Medical applications

2.3.1 3D medical imaging

This use case refers to point clouds created by the 3D scanning of internal or external anatomy of human and animal subjects for the purpose of research, medical record keeping, diagnosis and the preparation of treatment plans.

The point cloud data may be tagged with important additional information such as tissue parameters or type, but traditional colour attributes may not play a strong role in some applications of this use case. Lossless or near lossless coding will be important for confidence in diagnosis. To support low latency visualisation by medical professionals, view selectivity, region of interest and resolution granularity/scalable bit stream capabilities may be of importance. In addition, there is a requirement to provide means to guarantee the privacy and security needs associated with content.

2.3.2 Prosthesis and body parts design and manufacture

This use case involves artificially created point clouds to support the manufacture of 3D prosthetics for medical purposes.

The resolution and number of points in the point clouds will be highly dependent on the type of object. Attributes of points in this data is likely to include colour as well as other material appearance information, potentially linked to external databases containing a set number of materials. Highly accurate colour and material appearance representations may be needed to match the prosthetic against skin. Lossless or near lossless coding will be important for 3D manufacture, and there is a requirement to provide means to guarantee the privacy and security needs associated with content.

2.4 Construction and manufacturing

2.4.1 Analysis of 3D structure for defect detection

This use case concerns point clouds resulting from the 3D scanning of construction material and industrial parts with the goal of detecting defects, damage, shape tolerance deviations or cracks. Lossless or very low loss coding will be important to preserve and detect small defects and coding that preserves sharp edges and fine detail is essential. This use case may often require the storage of multiple scans to handle object internal structure and occlusions. In addition, 3D processing and computer vision operations in the compressed domain are needed to automate defect detection in applications where human intervention in costly or unavailable.

2.4.2 3D scanning to support project management

Point clouds are used to support the display of 3D content for visualisation of architecture and infrastructure in context of a building site or surrounding area.

An example would be to visualise planned construction elements against already completed construction elements or visualise a planned construction in the context of existing surroundings. The content may be generated artificially by CAD design software or created through the scanning of physical objects with 3D scanning equipment or may be a merger of both sources of data. The resolution and number of points of the content will be highly dependent on the type of object and industry. Geometrically, point clouds created using CAD software are likely to be arranged on regular grids and patterns, while data collected using 3D scanning equipment is likely to be arranged in an irregular geometric pattern unless geometrically re-positioned in some fashion. Attributes of points in this data is likely to include colour, gloss, bump maps or texture maps as well as BRDF or other material appearance information. There may be applications that require 3D processing and computer vision for the matching of scanned data from real objects with the computer generated elements for verification of physical elements against a plan without human intervention.

2.4.3 Repair and analysis of confined or dangerous places

This use case refers to 3D scanning of cavities, confined, or dangerous places (for example inside pipes, grain silos or infrastructure) with the goal of detecting defects, damage, shape distortions or cracks.

For some diagnostic purposes, colour information may be very important. In addition, capture may be under ambient or uncontrolled lighting, so high-dynamic range colour may be required. For the automated detection of defects and damage, there will be a requirement for 3D processing and computer vision in the compressed domain.

2.4.4 Point clouds for urban planning and urban analysis

Point clouds are used to record 3D shape and details of objects greater than 20 metres in extent, especially in large urban areas. The data is intended for analysis and rendering, and capture is likely to be under ambient or uncontrolled lighting. This data is likely to consist of very large point clouds and random access and scalability of viewpoints will be needed for effective visualisation of this data. In addition, there may be a need to automatically detect and classify key elements such as buildings or roads from the data, which in turn requires 3D processing and computer vision in the compressed domain.

2.4.5 Crack and damage detection for roads, bridges and other infrastructure

Point clouds are used to record 3D shape and details on large items of infrastructure. The goal is to spot defects, damage or cracks, and capture is likely to be under natural lighting or circumstances of limited control and consist of a number of separate scans. Accurate colour may be required to spot various types of damage. For this use case, the storage of multiple point clouds will be needed as will the ability to visualise the data at various scales and from random viewpoints without decoding the entire point cloud. For automated defect and damage analysis, there will be a need to perform 3D processing and computer vision in the compressed domain.

2.5 Consumer and retail

2.5.1 Small objects: jewellery, decorations, shoes, etc.

Point clouds are used to record 3D shape and details of small objects, e.g. for on-line shopping. Point clouds for this use case are likely to be complicated in structure, high resolution with complex material appearance. Data from sources such as this will often be rendered on websites and mobile devices for retail. Lossless or near lossless coding will be a requirement of this use case. To protect intellectual property there is a requirement to provide means to guarantee privacy and security. Having a human visualise and process each scanned point cloud of a new product will not be feasible for many commercial applications, so automated systems that perform point cloud reconstruction gap-filling and noise removal in the compressed domain will be needed.

2.5.2 Mid-size objects: (cars, motorcycles, furniture, etc.) and large objects (houses, apartments, etc.)

Point clouds are used to record 3D shape and details of mid-size objects (up to around 10 metres in extent) and large objects (beyond 10 metres in extent), e.g. for online shopping, 3D content creation, semi-professional level manufacture, 3D printing, real estate, advertising and VR experiences.

This use case may often require multiple scans to handle object internal structure and occlusions. Point clouds for this use case are likely to be complicated in structure, and have high resolution with complex material appearance. In cases when objects need to be processed automatically, for example to support a large scale commercial sale scenario, 3D processing in the compressed domain may be required. To protect intellectual property there is a requirement to provide means to guarantee the privacy and security needs associated with content.

2.6 Cultural heritage

2.6.1 Jewellery, pottery, bones, fossils and other small to large-size artefacts

Point clouds are used to record 3D shape and details of small objects such as jewellery or large objects such as buildings. In this use case the point clouds are expected to be at high resolution. Point clouds for this use case are likely to be complicated in structure, with high geometric resolution and complex material appearance. Point clouds for cultural heritage often have a dual purpose of enabling detailed analysis and being visualised by the public on websites or mobile devices. Lossless or near lossless coding will be crucial, while capturing an entire object may often require multiple scans to handle object internal structure and occlusions, with potentially the need to keep point clouds from individual scans distinct in the coded format for record keeping or future reconstruction purposes. For processing large collections for visualisation automatically, there may be a need for 3D processing in the compressed domain in some applications, and means for ensuring the privacy and security of the point clouds will be needed.

2.6.2 Paintings

Point clouds are used to record 3D shape and details of paintings and murals. The goal of this use case is potentially the display of the artwork, and analysis of the artwork to recognise the artist, understand how the artwork was created, or detect forgeries. Multiple scans stored losslessly or near losslessly may be required to capture the entire object and point clouds for this use case are likely to be complicated in structure, with high geometrical resolution and complex material appearance. Highly accurate colour and material appearance is likely to be needed for this use case. There is a possibility that point clouds for this use case may be represented as a height map above a surface, reflecting the flat geometry of paintings and murals. For processing large collections for display automatically, there may be a need for 3D processing in the compressed domain in some applications, and means for ensuring the privacy and security of the point clouds will be needed.

2.7 Remote Sensing and Geographical Information Systems

2.7.1 Wide area scanning

Point clouds are used to record 3D shape and details of objects greater than 20 metres in extent. Lighting during capture is likely to be uncontrolled and multiple scans may be required to reconstruct the scene. Detailed colour capture is sometimes, but not always, a feature of this use case, and some use cases may have the need to capture detailed material appearance attributes such as transparency and gloss. This type of data often involves large point clouds collected frequently and human intervention to correct defects and merge smaller point clouds in larger wholes is expensive and time consuming. Automated 3D processing in the compressed domain will be a requirement of this use case. To support applications where automatic extraction of features such as buildings or roads is needed, computer vision in the compressed domain is also a requirement. Figure 3 shows an example of a point cloud typical of this use case. The data has a high resolution but does not encode the material appearance of the scene imaged. In this particular example, colour information has not been recorded.

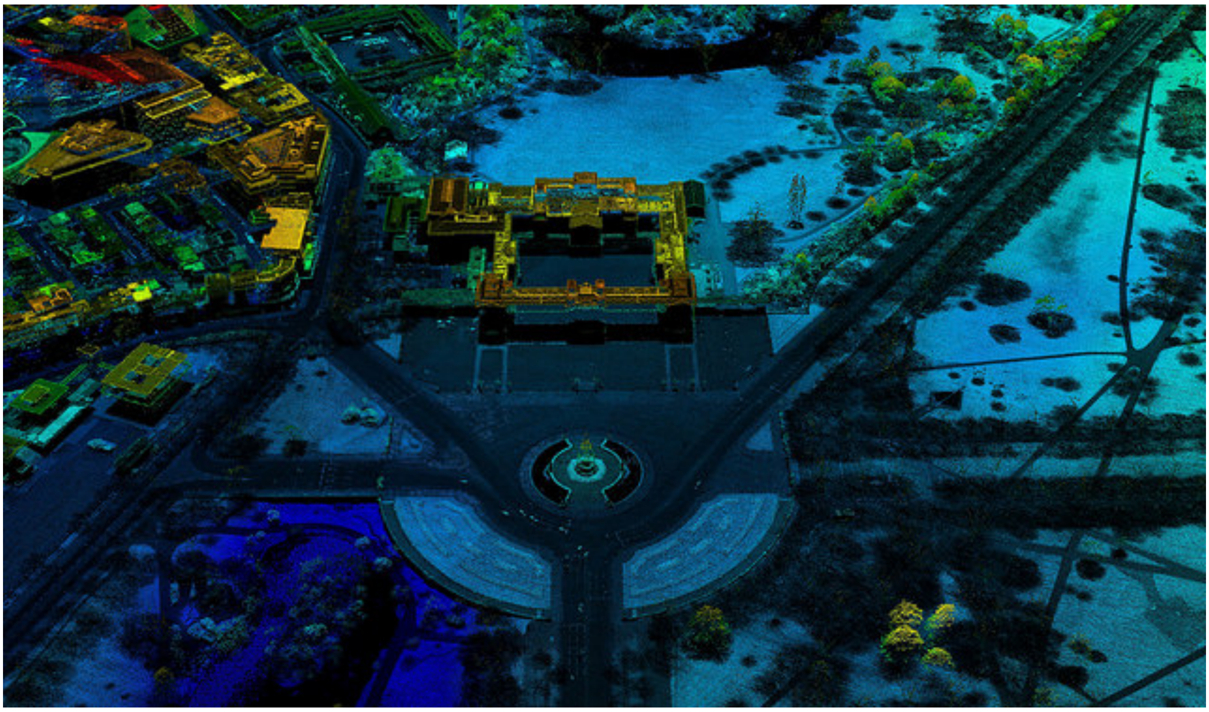

Figure 3: An example of a point cloud that would fall into use case 2.7.1. In keeping with this use case, the data is of high resolution, but without accurate material appearance information or colour. The colour in this figure represents height and is not related to the colour of the imaged scene. The image is a LIDAR scan of Buckingham Palace, UK and is courtesy of Environmental Agency (https://www.flickr.com/photos/environment-agency/27489358013) [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/)]

2.8 Automatic navigation systems

2.8.1 Autonomous vehicles

Point clouds are used to record 3D shape and details of the environment surrounding an autonomous vehicle. The goal of this use case is to analyse the point cloud to spot objects of interest such as pedestrians and obstructions without human intervention, hence 3D processing and computer vision in the compressed domain is needed. Resolution of the scanned area may vary with finer resolution in particular directions of interest, such as the direction of the motion of the vehicle. Capture is likely to be under uncontrolled lighting. Point clouds created for this use case may not have accurate colour and material appearance in some applications. Analysis requires accurate representation of point clouds so lossless or very low loss coding will be important. Coding that preserves sharp edges and fine detail is essential. Requirements for real-time processing and low latency are of importance to this use case.

2.8.2 Robotics

Point clouds are used to detect and model the 3D environment surrounding robotic devices. The extracted data can be noisy and incomplete, often being produced by multi-sensory systems. This use case can also include real-time surface estimation and object detection and classification to allow automatic navigation and exploration, and hence support for 3D processing and computer vision in the compressed domain will be essential to this use case.

2.9 Surveillance

2.9.1 Point clouds for search and rescue in confined spaces

Point clouds are used to record 3D shape and details of objects (in particular cavities up to around 1-10 metres in extent). The goal of this use case is to spot objects of interest such as trapped survivors. This may be performed by a human observer or automatically, so support for 3D processing and computer vision operations to be performed in the compressed domain is needed. Point cloud capture is likely to be under semi-controlled lighting, but in potentially chaotic environments and point clouds from this use case are not likely to require accurate colour or material appearance information, although colour may be important to aid object identification. Lossless or near-lossless coding of point clouds may be needed for some applications.

3 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 point cloud coding standard.

4 Requirements

This section presents the overall set of requirements that have been extracted from the above described use cases. Requirements are split between “core requirements” which are essential for the standard and “complementary requirements” which are desirable but not mandatory and will be decided depending on their cost. Depending on their type, the requirements will be classified as coding and systems requirements.

Requirements listed as “Systems requirements” and “Coding requirements for human visualisation” are common to all the stages of this activity as mentioned in Section 1. Requirements listed as “Additional requirements for 3D processing” apply to solutions to Stage 2 of the activity in addition to the requirements of Stage 1. Similarly, Requirements listed as “Additional requirements for computer vision tasks” apply to solutions to Stage 3 of the activity in addition to the requirements of Stage 1 and Stage 2.

When considering geometry and attribute scalability, it is important to be aware that there is a relationship between the scalability of attributes and the scalability of geometry. Although it is possible to have use cases that require attribute scalability without corresponding geometry scalability, it is difficult to visualise use-cases where geometry scalability is needed without corresponding attribute scalability.

In addition, in the description below:

- An object is set of points within the point cloud representing a semantically meaningful entity, e.g. a chair or the floor within a room setting. Object based scalability is expected to be addressed at the systems level.
- A region is understood to be a set of points within the point cloud selectable at the decoding stage. For example, a particular region might be the points associated with

the face in a point cloud representing a human being. This set of points may be contained within an object. This type of scalability is expected to be addressed at the coding level.

4.1 Core requirements

This section lists the set of core requirements derived from the full set of use cases.

4.1.1 Systems requirements

1. ***JPEG Systems Compatibility:*** The systems elements of the standard shall comply with the relevant JPEG Systems specification.

2. ***Multiple Point Clouds:*** The standard shall support the storage of multiple independent point clouds in a single stream or file. The standard shall allow selective access to an independent point cloud in the case that multiple point clouds are stored in the same stream or file.

3. ***Large Point Clouds***: The standard shall support the coding of point of clouds of any size, eventually using appropriate segmentation. The standard shall support the means for the appropriate decoding and recomposition of point clouds coded as multiple segments.

4. ***Metadata:*** The standard shall provide appropriate description tools for efficient search, retrieval, filtering and calibration of content. Strategies and technical solutions shall fit within the JPSearch and JPEG Systems frameworks.

5. ***Privacy and Security:*** The standard shall provide means to guarantee the privacy and security needs associated with content from the various imaging modalities. Strategies and technical solutions shall fit within the JPEG Privacy & Security and JPEG Systems frameworks.

4.1.2 Coding requirements for human visualisation

6. ***Data Types:*** The standard shall be able to code both the geometry data and the (global and local) attributes of a point cloud, where these data types are defined as follows:
    a. ***Geometry:*** Positional information of the points in the point cloud taken as integers or floats.
    b. ***Local/Point Attributes:*** Attributes on a point by point basis that may be integer or float such as colour, reflectance, links to external databases of materials, temperature or any indexes to lists. This includes attributes whose value changes with the angle of view to the associated point.
    c. ***Global Attributes:*** Global point cloud attributes that may be integer or float such as texture maps, bump maps and normal maps.

7. ***Coding Accuracy:*** The standard shall support both lossless and lossy coding of the geometry and attributes of the point cloud.

8. *Compression Efficiency:* The standard shall compress both geometry and attributes in an efficient manner, e.g. as measured in terms of bits per point.

9. *Low Complexity:* The standard shall allow for implementations of feasible encoding and decoding complexity, notably within reasonable expectations of time for the use cases listed above.

10. *Parallelisation:* The standard shall support efficient parallelisation of the encoding and decoding processes.

11. *Geometry Scalability*: The standard shall allow for decoding the point cloud geometry using only a part of the full bitstream wherein the compressed bitstream shall be structured with more than one layer to decode. This shall provide:
    a. A point cloud with increasing number of points as additional layers are decoded.
    b. A point cloud with increasing quality, for a specified bit depth, as additional layers are decoded.
    c. A point cloud with increasing precision, in terms of increasing bit depth, as additional layers are decoded.

    If increased scalability comes at the cost of a reduced compression efficiency, then it should be possible to tune/adjust the scalability granularity, notably to a minimum of scalability to reduce the compression efficiency penalty.

12. *Attributes Scalability*: The standard shall allow for decoding point cloud attributes using only a part of the full bitstream wherein the compressed bitstream shall be structured with more than one layer to decode. This shall provide point cloud attributes of increasing quality as additional layers are decoded.
    If increased scalability comes at the cost of a reduced compression efficiency, then it should be possible to tune/adjust the scalability granularity, notably to a minimum of scalability to reduce the compression efficiency penalty.

13. *Random Access:* The standard shall allow the selective decoding of a portion of the point cloud independently of the rest of the point cloud without decoding the entire bitstream. For example:
    a. Selectively decoding a portion of the point cloud corresponding to a volume in 3D space.
    b. Selectively decoding a portion of the point cloud visible from a given viewpoint.
14. *Tuneable Quality:* The standard shall allow tuning the quality of the decoded point cloud. Quality shall be measured by subjective and objective means as set forth by SC29/WG1.

4.1.3 Additional requirements for 3D processing

15. ***Denoising***: The standard shall allow for machines to perform denoising in the compressed domain. Denoising is defined as the removal of points in the point cloud that are outliers, and are inconsistent with the general shape and structure of the original point cloud.

16. ***Reconstruction***: The standard shall allow for machines to perform reconstruction in the compressed domain. Reconstruction of point clouds is defined as the alignment and merging of smaller point clouds representing partial reconstructions of a scene or object into a larger point cloud of the same scene or object.

4.1.4 Additional requirements for computer vision tasks

17. ***Object Detection***: The standard shall allow for machines to detect in the compressed domain an object represented by portion a point cloud indicate their location and identify to which semantic class they belong. Detection is defined as determining that a portion of the point cloud represents an instance of a single class of object of interest, such as "pedestrian".

18. ***Object Classification***: The standard shall allow for machines to classify an object in the compressed domain represented by portion a point cloud into one of several object classes of relevance to a particular use case. An example of relevant object classes may be "pedestrian", "road", "cyclist" or any other set of classes that may be applied to portions of the point cloud.

4.2 Complementary requirements

This section lists the set of complementary requirements derived from the full set of use cases.

4.2.1 Systems requirements

19. ***Metadata:*** The proposals for metadata must meet two concurrent requirements: i) support the JPEG Pleno Point Cloud use cases, and ii) be sufficiently flexible framework that is usable for other JPEG Pleno projects (e.g., Light Field and Holography), i.e., be a metadata framework. In addition, the proposals should be extensible to allow for emerging use cases for JPEG Pleno.

4.2.2 Coding requirements for human visualisation

20. ***Region-based Geometry Scalability***: The standard should allow for decoding the point cloud geometry using only a part of the full bitstream wherein the compressed bitstream should be structured with more than one layer to decode. This should provide a point cloud where an increasing number of regions are decoded as additional layers are decoded.

21. ***Hybrid Region-Based Geometry Scalability***: The standard should allow for decoding the point cloud geometry using only a part of the full bitstream wherein the compressed bitstream should be structured with more than one layer to decode. This should provide:
    a. A point cloud with increasing number of points as additional layers are decoded.
    b. A point cloud with increasing quality, for a specified bit depth, as additional layers are decoded.
    c. A point cloud with increasing precision, in terms of increasing bit depth, as additional layers are decoded.

    wherein the above provisions are controllable on a region by region basis.

22. ***Region-based Attributes Scalability***: The standard should allow for decoding point cloud attributes using only a part of the full bitstream wherein the compressed bitstream should be structured with more than one layer to decode. This should provide point cloud attributes of increasing quality as additional layers are decoded controllable on a region by region basis.
23. **Metrological Accuracy**: The standard should allow tuning the metrological accuracy of the decoded point cloud. Metrological accuracy refers to the mean and maximum deviations of point positions of decoded point clouds relative to the original point cloud. Metrological accuracy shall be measured by objective means as set forth by SC29/WG1.

4.2.3 Additional requirements for 3D processing

24. ***Super Resolution***: The standard should allow for machines to perform super resolution in the compressed domain. Super resolution is defined as increasing the number of points in the point cloud beyond that of the original point cloud with the following properties:
    d. The added points are consistent with the shape and structure of the original point cloud.
    e. The number of added points are adjustable within reasonable constraints.

25. ***Colour Correction***: The standard should allow for machines to perform colour correction in the compressed domain. Colour correction is defined as the application of a global transform to the values of the attributes.

26. ***Filling of Gaps and Holes***: The standard should allow for machines to perform the filling of gaps and holes in the point cloud in the compressed domain. Filling of gaps and holes is defined as increasing the number of points in the point cloud beyond that of the original point cloud such that the added points extend the surface, or close holes in the surface of the point cloud in a manner consistent with the shape and structure of the original point cloud.

4.2.1 Additional requirements for computer vision tasks

27. ***Scene retrieval***: The standard should allow for machines to identify and retrieve similar point clouds from a database based on topology and attribute features where

either the query or retrieved point cloud (or both) are in the compressed domain. A similar point cloud in this context is defined as a retrieved point cloud having a portion representing all or part of the query point cloud.

28. ***Scene classification***: The standard should allow for machines to classify a point cloud in the compressed domain into one of several scene classes of relevance to a particular use case. An example of relevant scene classes may be “outdoor”, “beach”, “building” or any other set of classes that may be applied to the entire point cloud.

29. ***Object Recognition***: The standard should allow for machines to recognise an object in the compressed domain represented by portion a point cloud. In this context, recognition is defined as determining that a portion of the point cloud represents an instance of an individual object of interest, such as a particular person or motor vehicle.

30. ***Semantic segmentation***: The standard should allow for machines to segment each portion of a point in the compressed domain into one of several classes of relevance to a particular use case. An example of relevant classes may be “road”, “building”, “park” or any other set of classes that may be applied to portions of the point cloud.

31. ***Event Detection***: The standard should allow for machines to detect in the compressed domain that a point cloud represents an event that is occurring. Event detection is defined as determining that the point cloud represents an instance of a single class of event of interest, such as “pedestrian stepping onto the road”, or “vehicle turning”.

32. ***Action Recognition***: The standard should allow for machines to recognise in the compressed domain that a portion of the point cloud represents one of several classes of action relevant to a particular use case. An example of relevant action classes may be “pedestrian walking”, “cyclist turning”, “vehicle stopping” or any other set of classes that may be applied to portions of the point cloud.

5 References

[1] JPEG Pleno – Scope, use cases and requirements Ver.1.6, Doc. ISO/IEC JTC 1/SC 29/WG1 N73030, Chengdu, China, Oct. 2016.

6 Appendix

Table 1: Use cases versus key core requirements. Key requirements listed here is intended to be representative of requirements of particular interest to the use case, however other requirements may be needed for use case. For example, the 1. JPEG Systems Compatibility would be considered common to all use cases.

Use case	Key Requirements
2.1.1 Computer Graphics and Gaming	9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality
2.1.2 Telepresence	9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction
2.2.1 Design, manufacturing and 3D printing	5. Privacy and Security 7. Coding Accuracy
2.2.2 Motion capture	7. Coding Accuracy 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection
2.3.1 3D medical imaging	2. Multiple Point Clouds 5. Privacy and Security 7. Coding Accuracy 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.3.2 Prosthesis and body parts design and manufacture	5. Privacy and Security 7. Coding Accuracy
2.4.1 Analysis of 3D structure for defect detection	7. Coding Accuracy 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.4.2 3D scanning to support project management	9. Low Complexity 11. Geometry Scalability

	12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.4.3 Repair and analysis of confined or dangerous places	7. Coding Accuracy 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection
2.4.4 Point clouds for urban planning and urban analysis	2. Multiple Point Clouds 9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Recognition
2.4.5 Crack and damage detection for roads, bridges and other infrastructure	7. Coding Accuracy 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.5.1 Small objects: jewellery, decorations, shoes, etc.	5. Privacy and Security 7. Coding Accuracy 9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction
2.5.2 Mid-size objects: (cars, motorcycles, furniture, etc.) and large objects (houses, apartments, etc.) 2.6.1 Jewellery, pottery, bones, fossils and other small artefacts 2.6.2 Paintings	2. Multiple Point Clouds 5. Privacy and Security 7. Coding Accuracy 9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction
2.7.1 Wide area scanning	2. Multiple Point Clouds

	5. Privacy and Security 7. Coding Accuracy 9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.8.1 Autonomous vehicles 2.8.2 Robotics	2. Multiple Point Clouds 5. Privacy and Security 7. Coding Accuracy 9. Low Complexity 11. Geometry Scalability 12. Attributes Scalability 13. Random Access 14. Tuneable Quality 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification
2.9.1 Point clouds for search and rescue in confined spaces	7. Coding Accuracy 15. 3D Processing: Denoising 16. 3D Processing: Reconstruction 17. Computer Vision: Object Detection 18. Computer Vision: Object Classification