**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2017/N16763**

**April 2017, Hobart, AU**

**Title: Call for Proposals for Point Cloud Compression V2**

**Source: MPEG 3DG and Requirements**

**Status: Approved**

# Abstract

This document is a Call for Proposals (CfP) for 3D point cloud compression technology, targeting an efficient representation of static objects and scenes, as well as dynamic objects and real-time acquisition environments.

# Introduction

Advanced 3D representations of the world are enabling more immersive forms of interaction and communication, and also allow machines to understand, interpret and navigate our world. 3D point clouds have emerged as an enabling representation of such information. A number of use cases associated with point cloud data have been identified [1], and corresponding requirements for point cloud representation and compression have been developed [2].

A point cloud is a set of points in a 3D space, each with associated attributes, e.g. color, material properties, etc. Point clouds can be used to reconstruct an object or a scene as a composition of such points. They can be captured using multiple cameras and depth sensors in various setups, and may be made up of thousands up to billions of points in order to realistically represent reconstructed scenes.

Compression technologies are needed to reduce the amount of data required to represent a point cloud. As such, technologies are needed for lossy compression of point clouds for use in real-time communications and six Degrees of Freedom (6 DoF) virtual reality. In addition, technology is sought for lossless point cloud compression in the context of dynamic mapping for autonomous driving and cultural heritage applications, etc. The standard to be developed will address compression of geometry and attributes such as colors and reflectance, scalable/progressive coding, coding of sequences of point clouds captured over time, and random access to subsets of the point cloud. The acquisition of point clouds is outside of the scope of this standard.

Recently, the investigation of new coding tools [3,5,6] for static and dynamic 3D point clouds have shown evidence that improved coding efficiency with respect to existing solutions are possible.

Companies and organizations are invited to submit proposals in response to this Call for Proposals.

Point cloud compression technologies will be evaluated based upon objective metrics. Results of these tests will be made public, taking into account that no direct identification of any of the contributors will be made (unless it is specifically requested or authorized by a contributor to be explicitly identified). Prior to having evaluated the results of the tests, no commitment to any course of action regarding the point cloud compression technology can be made. In addition, subjective evaluation of proposals will be performed.

Descriptions of proposals shall be registered as input documents to the proposal evaluation meeting in October 2017 (see timeline in next section). Proponents are strongly recommended to attend that meeting to present their proposals. For those organizations and individuals that are not accredited members of the MPEG working group, further information about logistical steps to attend the meeting can be obtained from the contact persons listed in Chapter 8.

# Timeline

**Timeline of the calls, deadlines and evaluation of the responses:**

|  |  |  |
| --- | --- | --- |
| **Action** | **Date** | **Remarks** |
| Call for proposals v2 | 2017.04.21 | Inclusion of more detailed test conditions and information on test procedure. |
| Declaration of intention of answering to the CfP. Payment for subjective evaluation has to be done at this time. | 2017.07.14 | An email should be sent to the contact addresses listed in chapter 8. An individual FTP account and a proponent number for subjective testing will be provided to each proponent by 2017.07.17. as confirmation. |
| Submission of testing material for subjective evaluation | 2017.10.09 | Testing material to be uploaded on the FTP site. Windows decoder and the compressed bistreams must be provided. |
| Submission of the completed objective evaluation spreadsheets | 2017.10.09 | To be uploaded on the individual FTP site |
| Compilation of submitted data in a unique spreadsheet and submission as an input contribution | 2017.10.10 | Action performed by 3DG Chair |
| Subjective evaluation with naive viewers | 2017.10.11 - 2017.10.21 | Laboratory involved: GBTech |
| Submission deadline for proposals documentation | 2017.10.18 | To be uploaded as input to the WG11 web site |
| Evaluation of responses | 2017.10.22–2017.10.27 | Action performed during the MPEG meeting week. Proponents are strongly advised to present their proposals in person. |
| Establishment of the first test model | 2017.10.27 |  |

Table 1 Timeline

**Preliminary Development Plan:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Month** | **Day** | **MPEG meeting** | **City** | **Country** | **Stage** |
| 2017 | 01 | 20 | 117 | Geneva | CH | Approval of CfP |
| 04 | 21 | 118 | Hobart | AU | Approval of CfP V2 |
| 10 | 27 | 120 | Macao | CN | Approval of Test Model |
| 2018 | 01 | 26 | 121 | Gwangju | KR | Approval of WD 1.0 |
| 04 | 20 | 122 | San Diego | US | Approval of WD 2.0 |
| 07 | 20 | 123 | Ljubljana | SI | Approval of CD |
| 10 | ?? | 124 | ?? | ?? | Approval of DIS |
| 2019 | 04 | 25 | 126 | ?? | ?? | Approval of FDIS |

Table 2: Preliminary Development Plan

# Definitions

The following terms are defined and used in the specification of the test conditions:

* **Point Cloud Frame**: A static point cloud representation at a given time instant or, for dynamically acquired point clouds, during a time range.
* **Geometry:** The locations in 3D space of the points in the point cloud, i.e. the (x,y,z) coordinates of the points.
* **Attribute:** A value or set of values, other than geometry, associated with a point, e.g., (R, G, B) color values, I for reflectance.
* **Lossy Geometry:** The decoded compressed geometry is not necessarily numerically identical to the uncompressed geometry. The number of points in the output cloud can differ from the number of points in the input cloud.
* **Lossless Geometry:** The decoded compressed geometry is numerically identical to the uncompressed geometry, in terms of (x,y,z) values. The number of points in the output cloud is identical to the number of points in the input cloud.
* **Lossy Attribute:** The decoded compressed attribute values are not necessarily numerically identical to the uncompressed attribute values.
* **Lossless Attribute:** The decoded compressed attribute values are numerically identical to the uncompressed attribute values.
* **Spatial Random Access:** It is possible to decode the point cloud corresponding to a pre-defined region from the compressed point cloud.
* **Progressive/Scalable:** It is possible to first decode a coarse point cloud and then refine it with additional data from the compressed bit stream.
* **Random Access:** Spatial Random Access and/or Temporal Random Access.
* **Intra-only:** Each point cloud frame can be decoded independently of any other encoded point cloud frames.
* **Temporal-random-access:** The ability to access any frame within a limited number of point cloud frames.
* **Low-delay:** There is no reordering of point cloud frames between decoding and display.
* **Operating Point:** Performance point of a particular codec for a particular data at a target bit rate or compression ratio.
* **Point Cloud:** A point cloud is defined as set of (x,y,z) coordinates, where x,y,z have finite precision and dynamic range. Each (x,y,z) can have multiple attributes associated to it (a1 ,a2, a3 …), where the attributes may correspond to color, reflectance or other properties of the object/scene that would be associated with a point. Typically, each point in a cloud has the same number of attributes attached to it.
* **1 Mbit/s:** 1,000,000 bits per second.

# Test Materials, Categories, Conditions and Anchors

This CfP addresses three categories of point cloud datasets:

* Category 1: Static Objects and Scenes
* Category 2: Dynamic Objects
* Category 3: Dynamic Acquisition

## Test Material Datasets

Below is a list of the 3D point cloud test material datasets to be used, organized based on the test category and test class. The test class is an indicator of how complex a point cloud is to encode, where A is the lowest and C the highest complexity. All test material datasets are available in the MPEG Content repository accessible under the following URL:

<http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/datasets/>

Test material datasets with normals are accessible under the following URL:

<http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/normals/>

Note: Downloaded test material datasets should be verified with MD5 checksums. Each zip file of a test material dataset contains an MD5 file (with the corresponding md5 sums for each file in the archive).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Test Category** | **Test Class** | **Test material dataset filename** | **# Frms** | **# Pts** | **Geometry**  **Precision** | **Attributes** | **Sequence**  **Number** |
| (1) Static Objects and Scenes | A | Egyptian\_mask | 1 | 272,689 | 32 bit | R,G,B | 1 |
| Statue\_Klimt | 1 | 499,886 | 32 bit | R,G,B | 2 |
| Arco Valentino Dense | 1 | 1,530,552 | 32 bit | R,G,B | 3 |
| Façade9 | 1 | 1,602,990 | 32 bit | R,G,B | 4 |
| Frog67 | 1 | 3,630,907 | 32 bit | R,G,B | 5 |
| Façade15 | 1 | 8,929,532 | 32 bit | R,G,B | 6 |
| Façade64 | 1 | 19,714,629 | 32 bit | R,G,B | 7 |
| Queen\_frame\_0200 | 1 | 1,000,993 | 10 bit | R,G,B | 8 |
| Loot\_vox10\_1200 | 1 | 805,285 | 10 bit | R,G,B | 9 |
| Redandblack\_vox10\_1050 | 1 | 787,237 | 10 bit | R,G,B | 10 |
| Soldier\_vox10\_0690 | 1 | 108,9091 | 10 bit | R,G,B | 11 |
| B | Shiva35 | 1 | 1,010,591 | 32 bit | R,G,B | 12 |
| House57 | 1 | 5,001,077 | 32 bit | R,G,B | 13 |
| Palazzo Carignano Dense | 1 | 4,203,962 | 32 bit | R,G,B | 14 |
| Head39 | 1 | 14,025,710 | 32 bit | R,G,B | 15 |
| Longdress\_vox10\_1300 | 1 | 857,966 | 10 bit | R,G,B | 16 |
| C | Landscape14 | 1 | 72,145,549 | 32 bit | R,G,B | 17 |
| Stanford Area2 | 1 | 54,989,822 | 32 bit | R,G,B | 18 |
| Stanford Area4 | 1 | 47,485,046 | 32 bit | R,G,B | 19 |
| (2) Dynamic Objects | A | Queen | 250 | ~1000,000 | 10 bit | R,G,B | 20 |
| 8i VFB – Loot | 300 | ~782,000 | 10 bit | R,G,B | 21 |
| 8i VFB – Red\_and\_Black | 300 | ~700,000 | 10 bit | R,G,B | 22 |
| 8i VFB – Soldier | 300 | ~1,500,000 | 10 bit | R,G,B | 23 |
| B | 8i VFB – Long\_dress | 300 | ~800,000 | 10 bit | R,G,B | 24 |
| (3) Dynamic Acquisition | Frame-based | Ford Campus Vision 1 | 1500 | ~100,000 / fr | 32 bit | I | 25 |
| Ford Campus Vision 2 | 1500 | ~100,000 / fr | 32 bit | I | 26 |
| Ford Campus Vision 3 | 1500 | ~100,000 / fr | 32 bit | I | 27 |
| Fused | Mitsubishi – citytunnel | 1 | 21,163,706 | 32 bit | R,G,B  I | 28 |
| Mitsubishi – overpass | 1 | 5,326,157 | 32 bit | R,G,B  I | 29 |
| Mitsubishi – tollbooth | 1 | 7,148,520 | 32 bit | R,G,B  I | 30 |

Table 3 Test material datasets

Notes:

1. The associated normals for each dataset are available to compute distortion metric D2 (point-to-plane), as described in Annex B.
2. The order of the points as they are stored in the file is not necessary to be maintained in the decoded versions.
3. The Geometry Precision in Table 3 indicates the known geometry precision of the test material. The data type in the PLY file describes the format stored in the file.

## Test Conditions and Parameters

### Test Conditions including target bit rates

The point cloud test material will be tested under the following conditions:

* Lossless Geometry & No Attributes: This test condition applies to Categories 1 and 3.
* Lossless Geometry & Lossy Attributes: This test condition applies to Categories 1 and 3.
* Lossy Geometry & Lossy Attributes: This test condition applies to Categories 1, 2 and 3.

#### Test Conditions for Lossless Geometry & No Attributes

Compression rates for all test material datasets from Category 1 and 3 datasets shall be provided.

#### Test Conditions for Lossless Geometry & Lossy Attributes

Submissions under this test condition shall be capable of associating the decoded attributes with the given geometry, so that the decoded attributes can be displayed as a point cloud and distortion metrics can be correctly computed.

Category 1

The following static test data from Category 1 will be evaluated under this condition:

* Longdress\_vox10\_1300
* Soldier\_vox10\_0690
* Queen\_frame\_0200
* Shiva35
* Façade15

Given a lossless geometry as provided by the input data, the target bit rates for the color attributes of Category 1 in terms of bits per point, are:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test Dataset | R1 | R2 | R3 | R4 |
| Longdress | 1.6 | 2 | 2.8 | 5.12 |
| Soldier | 1.12 | 1.36 | 1.92 | 3.68 |
| Queen | 0.8 | 1.04 | 1.44 | 2.88 |
| Shiva | 2.24 | 2.72 | 3.6 | 6 |
| Façade | 1.36 | 1.76 | 2.48 | 4.4 |

Table 4 Target rates in bits per color per point

The data submitted to this test condition will only undergo objective evaluation; however, some informal subjective viewing may be conducted on the results.

Category 3

The following static test data from Category 3 will be evaluated under this condition:

* Mitsubishi - citytunnel
* Mitsubishi - tollbooth

There are 2 test sub-conditions:

* Given a lossless geometry as provided by the input data, the target bit rates for the color attributes of Category 3 in terms of bits per input point (bpp) are: 8, 4, 2, 1, 0.5, 0.25.
* Given a lossless geometry as provided by the input data, the target bit rates for the reflectance attributes of Category 3 in terms of bits per input point (bpp) are: 8, 4, 2, 1, 0.5, 0.25.

The data submitted to these test conditions will only undergo objective evaluation, however some informal subjective viewing may be conducted on the results.

#### Test Conditions for Lossy Geometry & Lossy Attributes

Category 1

The target bit rates under this condition for Category 1 in terms of bits per input point (bpp) are given by the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | R1 | R2 | R3 | R4 |
| A | 0.4 | 1.25 | 3.5 | 6.75 |
| B | 0.6 | 2 | 4.5 | 8 |
| C | 0.07 | 0.25 | 1.0 | 1.6 |

Table 5 Target rates in bits per point (geometry + color), for each point cloud class

For data that will undergo subjective evaluation (see Table 10 in Annex D), proposals shall not exceed the target rates for each class.

For data that will only undergo objective evaluation (all Category 1 test material datasets listed in Table 9 in Annex B), proposals shall submit bitstreams near the target rate points (+/- 10%) for each class. Additional rate points beyond the designated rate, especially lower rates, are encouraged. Bitstreams should be submitted so that cross checks are possible.

Category 2

The target bit rates under this condition for Category 2 in terms of Mbit/s are indicated in the table below. These rates are inclusive of geometry and attribute (color) data. The data submitted to this test condition will undergo both objective and subjective evaluation, with the sequences selected for subjective evaluation indicated in Annex D. For the subjective evaluation, it is permitted to use temporal prediction with a maximum intra period of 32 frames. For the objective evaluation, both all-intra and random access (with a maximum period of 32 frames) bit rates and distortion values should be provided.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test Dataset |  |  |  |  |  |  |  |
| Longdress | 3 | 6\* | 10 |  | 16\* | 27\* |  |
| Redandblack |  | 4\* | 7 | 11\* |  | 34\* | 72\* |
| Soldier | 3 | 5\* | 10 |  | 17\* |  | 57\* |
| Queen | 3 | 4\* | 12.5\* |  | 28\* |  | 72\* |
| Loot | 3 | 4\* | 11\* |  | 29\* |  | 57\* |

Table 6 Target bitrates for Category 2 in Mbit/s

Note: bit rates marked with a \* are covered by anchors. When anchors are not available, the subjective evaluation will be performed only between technologies provided by proponents.

Note: the rate numbers shall be associated from left to right. For example, with Redandblack, R1=4Mbit/s and R5=72Mbit/s.

For data that will only undergo objective evaluation (all Category 2 test material datasets listed in Table 9 in Annex B), proposals shall submit bitstreams near the target rate points (+/- 10%). Additional rate points beyond the designated rate, especially lower rates, are encouraged.

Category 3

Two test sub-conditions are considered for Category 3, with target rates provided in terms of bits per input point (bpp):

1. Geometry only: The target rates are 24, 12, 8, 4, 2, 1, 0.5 bpp.
2. Geometry & Attributes: The target rates are 32, 16, 12, 6, 3, 1.5, 0.75 bpp. These rates are inclusive of geometry and all associated attribute data. It is mandatory for submissions to this condition for Category 3 to encode both reflectance and color attributes, if present in the input point cloud.

The data submitted to this test condition (all Category 3 test material datasets listed in Table 9 in Annex B) will only undergo objective evaluation; however, some informal subjective viewing may be conducted on the results.

### Summary of all Test Conditions

The following table summarizes all test conditions and assigns codes that shall be used for naming the submitted bitstreams:

|  |  |
| --- | --- |
| Test Sub-Condition | Code for bitstream naming convention |
| Category 1: lossless geometry – no attributes | 0 |
| Category 1: lossless geometry – lossy color | 1 |
| Category 1: lossy geometry – lossy color | 2 |
| Category 2: lossy geometry – lossy color / All Intra | 3 |
| Category 2: lossy geometry – lossy color / Randon Access | 4 |
| Category 3: lossless geometry – no attributes | 5 |
| Category 3: lossless geometry – lossy color | 6 |
| Category 3: lossless geometry – lossy reflectance | 7 |
| Category 3: lossy geometry | 8 |
| Category 3: lossy geometry – lossy color & reflectance | 9 |

Table 7 Test Sub-Conditions

### Restrictions for all test categories and conditions

Point cloud compression technologies shall obey the following additional constraints:

1. No post-processing shall be applied after decoding, i.e., before the computation of any metric or assessment of visual quality. Supplemental results that demonstrate benefits of pre- or post-processing may also be provided. While this information can be considered during the evaluation process, the main focus of the evaluation will be on results that do not apply any pre- or post-processing. If any pre-processing is done, the comparison shall be done with the original data.
2. Low-level programming optimizations, such as assembly code/intrinsics and external compression libraries, are discouraged. If any such optimization is implemented, then the rationale for, and extent of, the optimization shall be described.
3. Optimization of encoding parameters using non-automatic means is discouraged. In case that optimization is done, then it needs to be described in detail.
4. The coding test set shall not be used as the training set for training large entropy coding tables, VQ codebooks, etc. If any processes in the encoder or decoder are designed using training data, then the coding test set shall not be used as part of the training set.
5. For all bitstreams for which subjective evaluation is conducted, submitted bitstreams must have bit rates less than or equal to the corresponding target bit rate.
6. For all testing categories, the intra period shall be less than or equal to 32 frames.

## Anchors

Anchors were generated by using the experimental software tagged “hobart”. The software is based on the Point Cloud Library version 1.8 extended with features for lossy color coding, bit rate parameter settings and inter-predictive coding. The codec is described in [3] and the software is available on MPEG SVN:

<http://wg11.sc29.org/svn/repos/MPEG-04/Part16-Animation_Framework_eXtension_(AFX)/trunk/3Dgraphics/3DG-PCC/tags/hobart>

The compressed anchor bitstreams for different settings and MD5 checksums are available at:

<http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/anchors>

The decoded anchor point clouds can be found under the following URL for each of the categories and files:

<http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/decoded>

Each zip file of a test material dataset contains an MD5 file (with the corresponding md5 sums for each file in the archive).

For non-MPEG members, the software and the anchors will be available upon request to the contact persons. Annex C contains information on how to regenerate the anchors.

# Evaluation Procedure

The evaluation procedure will be conducted to select a baseline set of technologies. Subsequent core experiments will be used to obtain the best technical solution to fulfill the requirements.

## Objective Evaluation

Proposals will be evaluated using a set of objective metrics, including PSNR of a point-to-point error (D1) and a point-to-plane error (D2) for geometry, as well as PSNR of color and reflectance attributes. The distortion metrics compare the original data with the decoded data and provide numerical values. For Category 1 and Category 3, the rate shall be reported as bpp (bits per input point). For Category 2 and the Ford data set of Category 3, the rate shall be reported as Mbit/s. The evaluation will be made based on the rate-distortion (RD) performance, and RD curves shall be plotted using PSNR as the quality measure for all categories.

### Geometric Distortions

For D1 and D2, both MSE and PSNR shall be reported. The calculation of PSNR is specified in Annex B, with the peak value determined for each sequence depending on the case:

* For cases where the geometry is specified with high precision (e.g., 32 or 64 bits) and there is greater emphasis on local geometric errors, the peak value of the PSNR calculation corresponds to an intrinsic resolution that is computed as the maximum of the nearest neighbor distances amongst all the points in the reference point cloud.
* For cases where the geometry is specified with lower precision (e.g., 10 bits) and there is greater emphasis on the shape within the bounding box, the peak value of the PSNR calculation corresponds to the maximum possible range of the x, y, z point location values (e.g., 1023 in the case of 10-bit values).

In the case of dynamic content, the reported distortion measures shall be averaged over all coded frames (i.e., averaged over I frames, averaged over P frames, averaged over all frames). In addition, proponents shall report the PSNR per frame.

### Attributes Distortion

The color distortion is measured in YUV space with 3 separate MSE distortions, which are reported as PSNRs for each channel: Y, U, and V. The reflectance distortion is also computed using MSE, but only for a single component, and the corresponding PSNR values are reported. The PSNRs for each component will be individually used to compare methods.

### Software & Usage

The metric software for performing the objective evaluation can be found in the MPEG SVN :

<http://wg11.sc29.org/svn/repos/MPEG-04/Part16-Animation_Framework_eXtension_(AFX)/trunk/3Dgraphics/3DG-PCC/trunk/PccAppQualityMetric>

Additional details about the metrics to be reported and the usage of the metric software is briefly described in Annex B .

### Reporting of Results

See Section 6.2 for information regarding evaluation spreadsheets.

## Subjective Evaluation

Subjective testing will only be conducted for selected test material datasets of Category 1 and Category 2.

The test material datasets will be rendered using the point cloud renderer selected by MPEG. During preparation of the subjective evaluation sessions, three test material datasets for Category 1 and three test material datasets for Category 2 have been selected from a larger set of test material datasets. The rendering view-point/camera path will follow a pre-defined path selected by the MPEG Test chair for each test data set, which is not known to the proponents. This prevents any possibility of optimizing a codec for these tests. The output of this camera path will be stored as high quality video sequences of a length as close as possible to 10s. These video sequences will be viewed and evaluated by naïve viewers using method DCR Rec. ITU-T P.910 (<https://www.itu.int/rec/T-REC-P.913/en>). More detailed information can be found in Annex D.

**Software & Usage**

The renderer software that will be used during the subjective evaluation can be found in the MPEG SVN :

<http://wg11.sc29.org/svn/repos/MPEG-04/Part16-Animation_Framework_eXtension_(AFX)/trunk/3Dgraphics/3DG-PCC/trunk/PccAppRenderer>

The usage of the renderer software is described in [4].

### Reporting of results

The data extracted from the scoring sheets will be collected in an Excel spreadsheet by the MPEG Test chair. Mean Opinion Score (MOS) and Confidence Interval (CI) data will be computed. Also, a subjects’ post-screening will be applied to verify the level of reliability of their answers (data from subjects with a correlation index lower than 0.75 will be excluded from the MOS and CI values computation). Graphs reporting, for example, MOS on the y-axis and bits per input point on the x-axis, will be drafted.

# Submission Requirements

Proposals should include submissions to at least one of the three test Categories and to at least one of the three test conditions defined in this document. In the following sections, details on the coded test material and documentation that form a complete proposal are provided.

## Coded test material

The following material must be made available by proponents of technologies:

1. Bitstreams for all datasets in target test Categories, which follow the associated test conditions and satisfy the rate constraints specified in this document. Proponents shall use the following naming scheme:

PnnSmmCxRy.bin (.bin is used for compressed datasets)

nn = Proponent number; to the Anchor is typically assigned the code P00

mm = Sequence number from Table 3

x = a code identifying the test sub-condition which is defined in Table 7

y = the rate number, in our case typically from 1 to 4

1. Binary decoder executable (Windows 64 bits), configuration files (if required for decoder usage), and usage documentation allowing for decoding of the bitstreams.

## Evaluation spreadsheets

Complete submissions shall include results of the objective tests that shall be reported by using the Excel sheet that has been obtained as part of this CfP.

## Documentation

Complete submissions shall include the following elements:

1. An information form must be submitted within each proposal. This form can be found in Annex A of this document.
2. A technical description for full conceptual understanding and generation of equivalent performance results by experts. This description should include all data processing paths and individual data processing components used to generate the bitstreams. It does not need to include complete bitstream format or implementation details, although as much detail as possible is desired.
3. The technical description shall state how the proposed technology behaves in terms of random access to any frame within the sequence. For example, a description of the GOP (Group of Pictures), such as structure and the maximum number of frames that must be decoded to access any frame, could be given.
4. The technical description shall specify the expected encoding and decoding delay characteristics of the technology, including structural delay, e.g., due to the amount of frame reordering and buffering and the degree by which the delay can be minimized by parallel processing.
5. The technical description shall contain information suitable to assess the complexity of the implementation of the technology, including the following:

* A complete complexity analysis for encoder and decoder is not required to be provided by contributors, but compression/decompression time should be reported per test material dataset (a relative time with respect to ALL anchors from the category/test condition(s) they target for encode/decode process should be provided). Time for input/output operations to storage systems are excluded in the calculations.
* Expected memory usage of encoder and decoder.
* Complexity of encoder and decoder, in terms of number of operations, dependencies that may affect throughput, etc.
* Degree of capability for parallel processing.
* Degree to which bitstreams can be considered progressive.

1. The technical description should include an analysis of how the rendering setting (especially the point size) affects the perceived quality of the compressed data. For example, proponents could provide images/videos of reconstructed point clouds with different rendering parameters and analyze the codec behavior with respect to them.

## Source code

* Proponents are encouraged (but not required) to allow other committee participants to have access, on a temporary or permanent basis, to their encoded bit streams and binary executables or source code.
* Proponents are encouraged to submit a statement about the programming language in which the software is written, e.g. C/C++, and the platform(s) on which the binaries were compiled.

Proponents are advised that, upon acceptance for further evaluation, it will be required that certain parts of any proposed technology be made available in source code format to participants in the core experiments process and for potential inclusion in the prospective standard as reference software. When a particular technology is a candidate for further evaluation, commitment to provide such software is a condition of participation. The software shall produce identical results to those submitted to the test. Additionally, submission of improvements (bug fixes, etc.) is strongly encouraged.

# IPR

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

# Contacts

Contact persons:

Dr. Marius Preda

MPEG 3D Graphics Chair

Institut Mines-Télécom

37-39 rue Dareau 75014 Paris, France

Email: [marius.preda@it-sudparis.eu](mailto:marius.preda@it-sudparis.eu)

Prof. Dr.-Ing. Jörn Ostermann

MPEG Requirements Chair

Leibniz Universität Hannover, Institut für Informationsverarbeitung

Appelstr. 9A 30167 Hannover, Germany

Email: [ostermann@tnt.uni-hannover.de](mailto:ostermann@tnt.uni-hannover.de)

# References

1. Use Cases for Point Cloud Compression, ISO/IEC JTC1/SC29 WG11 Doc. N16331, Geneva, CH, June 2016
2. Requirements for Point Cloud Compression, ISO/IEC JTC1/SC29 WG11 Doc. N16330, Geneva, CH, June 2016
3. Rufael Mekuria, Kees Blom and Pablo Cesar. "Design, Implementation and Evaluation of a Point Cloud Codec for Tele-Immersive Video." *IEEE Transactions on Circuits and Systems for Video Technology* April 2017. ISO/IEC JTC1/SC29 WG11 doc m38136 MPEG, San Diego, USA, 2016
4. Documentation for Point Cloud Renderer, ISO/IEC JTC1/SC29 WG11 Doc. N16902, Hobart, AU, April 2017
5. Philip A. Chou, Ricardo L. de Queiroz “Transform Coder for Point Cloud Attributes” ISO/IEC JTC1/SC29 WG11 Doc. m38674, Geneva, CH, 2016
6. Philip A. Chou, Ricardo L. de Queiroz “Rate-Distortion Optimized Coder for Dynamic Voxelized Point Clouds” ISO/IEC JTC1/SC29 WG11 Doc. m38675, Geneva, CH, 2016

# Annex A: Information Form

1. Title of the proposal
2. Organization (name of proposing company, name and email of contact person)
3. What does your proposal apply to? More than one application area can be selected. In case of “other”, please describe which application area you envision.

|  |  |
| --- | --- |
| (a) Tele-immersive | (b) Interactive parallax |
| (c) Free viewpoint sports | (d) GIS |
| (e) Cultural Heritage | (f) Automotive |
| (g) Other |  |

1. What is the main functionality of your proposal in relation to the application areas?
2. Which Categories and which test conditions are covered by your proposal?
3. Do you plan to attend the MPEG#120 meeting and explain your proposal, show how it meets the PCC requirements, and answer questions about it?
4. Proponents shall provide information on how the MPEG-PCC requirements [2] are fulfilled, in the form of a table. This table should list the requirement, the fulfillment (yes/no) and give a short rationale why the requirement is fulfilled.

|  |  |  |
| --- | --- | --- |
| Requirements on PCC | Fulfillment (yes/no) | Rationale for fulfillment |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Table 8 Fulfillment of PCC requirements

# Annex B: Object Evaluation Metrics & Usage of Metric Software

## B.1 Geometric Distortions

Let and denote the original and the compressed point cloud, respectively. Consider evaluating the compression errors, denoted as in point cloud relative to reference point cloud. The steps to compute both point-to-point error (D1) and point-to-plane error (D2) for geometric errors are summarized in the following and illustrated in the below figure.

For each point in point cloud, i.e., the black point in the figure, identify a corresponding point in point cloud, i.e. the red point in the figure. Nearest neighbor is used to locate the corresponding point. In particular, a KD-tree search is used to perform the nearest neighbor search in order to reduce the computation complexity.

### B.1.1 Computing D1

Determine an error vector by connecting the identified point in reference point cloud  to point in point cloud . The length of the error vector is the point-to-point error, i.e.,

(A-1)

Based on the point-to-point distances for all points , the point-to-point error (D1) for the whole point cloud, withas the number of points in point cloud, is defined as:

(A-2)

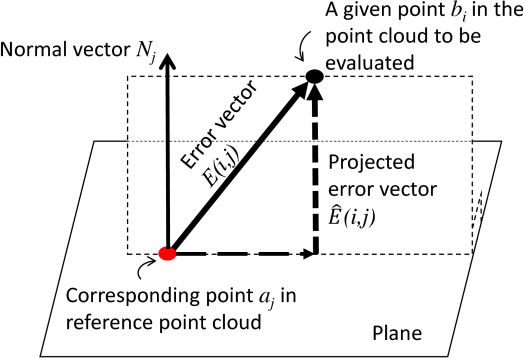
### B.1.2 Computing D2

Project the error vector along the normal direction and get a new error vector. In this way, the point-to-plane error is computed as,

(A-3)

The point-to-plane error (D2) for the whole point cloud is then defined as,

(A-4)



**Illustration of point-to-point distance (D1) and point-to-plane distance (D2)**

### B.1.3 PSNR Calculation

The geometric PSNR value is computed as:

(A-5)

where is the peak constant value defined for each reference point cloud as specified in Table 9 , and is the mean squared point-to-point (D1) or point-to-plane (D2) error. For dynamic content, the peak value is unchanged over the frames of a sequence.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test Category** | **Test Class** | **Test material dataset filename** | **Precision** | **Peak Value (p)** |
| (1) Static Objects and Scenes | A | Egyptian\_mask | 32 bit | 0.142735 |
| Statue\_Klimt | 32 bit | 0.30491 |
| Arco Valentino Dense | 32 bit | 0.393073 |
| Façade9 | 32 bit | 0.0337024 |
| Frog67 | 32 bit | 0.177377 |
| Façade15 | 32 bit | 0.161909 |
| Façade64 | 32 bit | 0.0413915 |
| Queen\_frame\_0200 | 10 bit | 1023 |
| Loot\_vox10\_1200 | 10 bit | 1023 |
| Redandblack\_vox10\_1550 | 10 bit | 1023 |
| Soldier\_vox10\_0690 | 10 bit | 1023 |
| B | Shiva35 | 32 bit | 0.432619 |
| House57 | 32 bit | 0.386349 |
| Palazzo Carignano Dense | 32 bit | 2.39912 |
| Head39 | 32 bit | 0.35906 |
| Longdress\_vox10\_1300 | 10 bit | 1023 |
| C | Landscape14 | 32 bit | 0.82408 |
| Stanford Area2 | 32 bit | 0.0701 |
| Stanford Area4 | 32 bit | 0.195681 |
| (2) Dynamic Objects | A | Queen | 10 bit | 1023 |
| 8i VFB - Loot | 10 bit | 1023 |
| 8i VFB - Red\_and\_Black | 10 bit | 1023 |
| 8i VFB - Soldier | 10 bit | 1023 |
| B | 8i VFB - Long\_dress | 10 bit | 1023 |
| (3) Dynamic Acquisition | Frame-based | Ford Campus Vision 1 | 32 bit | 96.5736 |
| Ford Campus Vision 2 | 32 bit | 83.9194 |
| Ford Campus Vision 3 | 32 bit | 60.2512 |
| Fused | Mitsubishi - citytunnel | 32 bit | 16.9967 |
| Mitsubishi - overpass | 32 bit | 8.50033 |
| Mitsubishi - tollbooth | 32 bit | 27.8963 |

Table 9 Peak values used for PSNR calculations for each reference point cloud

Note that the metric software dynamically determines the intrinsic resolution and uses it as normalizer if it is not specified in the command line (option -r).

## B.2 Attribute Distortions

The attribute PSNR value is computed as:

For color attributes, the MSE for each of the three color components is calculated. A conversion from RGB space to YUV space is conducted using ITU-R BT.709, since YUV space correlates better with human perception. A symmetric computation of the distortion is utilized, in the same way as is done for geometric distortions. The maximum distortion between the two passes is selected as the final distortion. Since the color attributes for all test data have a bit depth of 8 bits per point, the peak value for PSNR calculation is 255.

For reflectance attributes, the MSE for a single component is calculated. Since the reflectance attribute for all test data has a bit depth of 16 bits per point, the peak value for PSNR calculation is 65535.

## B.3 Metric Software Usage

Example command lines to use the evaluation metric tool are provided below.

./pc\_error -a pointcloud1.ply

This will load the point cloud and report nearest neighbor distances – the intrinsic resolution to be used for PSNR computation.

./pc\_error -a pointcloudOrg.ply -b pointcloudDec.ply

This will produce D1.

./pc\_error -a pointcloudOrg.ply -b pointcloudDec.ply -n normalOrg.ply

This will produce both D1 and D2. The normal is provided in normalOrg.ply for the original point cloud. It could be the same as pointcloudOrg.ply.

./pc\_error -a pointcloudOrg.ply -b pointcloudDec.ply -n normalOrg.ply -c

This will produce D1, D2, as well as the color distortion metric.

./pc\_error -a pointcloudOrg.ply -b pointcloudDec.ply -n normalOrg.ply -c -r intrinsicFloatNumber

This will produce D1, D2 as well as color distortion metric, using an imported intrinsic resolution for PSNR computation rather than an internally determined intrinsic resolution.

# Annex C: Anchor tool usage

The anchor can be compiled on Linux and/or Windows from the experimental software. The tool needs a configuration file called parameter\_config.txt in the same directory. The configuration files used for the anchors and a pre-compiled 64-bit Windows 7/8/10 version is available in the MPEG SVN at (no support on the pre-compiled binary will be given):

<http://wg11.sc29.org/svn/repos/MPEG-04/Part16-Animation_Framework_eXtension_(AFX)/trunk/3Dgraphics/3DG-PCC/trunk/anchor_config>

The gop coder tool codes an intra frame and a predictive frame in one bitstream. The frame coder tool codes a single intra coded frame in a single bitstream.

Usage of gop coder encoding a gop:

pcl\_mpeg\_pcc\_gop\_coder\_release.exe -e Iframe.ply pFrame.ply GopIP.pcc

Alternatively the parameter\_config file can be given as an argument when name is different:

pcl\_mpeg\_pcc\_gop\_coder\_release.exe -c parameter\_config.txt -d Iframe.ply pFrame.ply GopIP.pcc

Usage of gop coder decoding a gop:

pcl\_mpeg\_pcc\_gop\_coder\_release.exe -d Iframe.ply pFrame.ply GopIP.pcc

Usage of frame coder encoding a frame

pcl\_mpeg\_pcc\_frame\_coder\_release.exe -e intraframe.ply intraframe.pcc

Alternatively the parameter\_config file can be given as an argument when name is different:

pcl\_mpeg\_pcc\_frame\_coder\_release.exe -c parameter\_config.txt -d Iframe.ply pFrame.ply GopIP.pcc

Usage of frame coder decoding a frame

pcl\_mpeg\_pcc\_frame\_coder\_release.exe -d intraframe.ply intraframe.pcc

# Annex D: Details on subjective testing

Table 10 indicates the test datasets that have been selected for subjective evaluation, including point sizes to be used for rendering each operating point for the anchor codec, as well as the command lines used by the point cloud renderer for creating the anchors.

|  |  |  |  |
| --- | --- | --- | --- |
| Test Data Set | Operating Point | Point Size\* | Renderer Command line |
| Egyptian\_mask  (Class A) | R1 | 16.9 | PccAppRenderer.exe -f <full\_path\_to\_filename> -s ‑c ‑x Egyptian\_mask\_camera\_path ‑o –width=1920 –height=1080 –size=<size> --type=0 |
| R2 | 8.33 |
| R3 | 5.2 |
| R4 | 4.62 |
| Façade9  (Class A) | R1 | 7.4 | PccAppRenderer.exe -f <full\_path\_to\_filename> -s ‑c ‑x Facade9\_camera\_path ‑o –width=1920 –height=1080 –size=<size> --type=0 |
| R2 | 3.65 |
| R3 | 2.03 |
| R4 | 1 |
| Longdress\_vox10\_1300  (Class B) | R1 | 7.7 | PccAppRenderer.exe -f <full\_path\_to\_filename> -s ‑c ‑x Longdress1300\_camera\_path ‑o –width=1920 –height=1080 -p –size=<size> --type=0 |
| R2 | 3.65 |
| R3 | 2.03 |
| R4 | 1.8 |
| 8i VFB – Long\_dress  (Class B) | R1 | 7.7 | PccAppRenderer.exe -d <full\_path\_to\_directory> ‑n 300 ‑s ‑c ‑x Long\_dress\_camera\_path ‑o –width=1920 –height=1080 -p -–size=<size> --type=0 –fps=<fps> |
| R2 | 3.65 |
| R3 | 2.03 |
| R4 | 1.8 |
| 8i VFB – Red\_and\_Black  (Class A) | R1 | 7.7 | PccAppRenderer.exe -d <full\_path\_to\_directory> -n 300 ‑s ‑c ‑x Red\_and\_Black\_camera\_path ‑o –width=1920 –height=1080 -p -–size=<size> --type=0 –fps=<fps> |
| R2 | 3.65 |
| R3 | 2.03 |
| R4 | 1.8 |
| 8i VFB – Soldier  (Class A) | R1 | 7.7 | PccAppRenderer.exe -d <full\_path\_to\_directory> -n 300 ‑s ‑c ‑x Soldier\_camera\_path ‑o –width=1920 –height=1080 -p -–size=<size> --type=0 –fps=<fps> |
| R2 | 3.65 |
| R3 | 2.03 |
| R4 | 1.8 |

Table 10 Anchors creation

\* proponents are expected to provide the point size per operation point to be used when rendering the decoded bitstreams and producing the videos to be subjectively evaluated. Otherwise, the point size as indicated in the table above will be used. The table template to be used should be based on Table 10.

## D.1 Generation of video sequences

The set of video sequences to be subjectively assessed will be generated according to what is specified in Table 10. The video sequences will be generated with the following video parameters:

* Video resolution: progressive uncompressed full-range HD format (possibly 1920x1080; as alternative 1280x720).
* Frame rate: For category 2 point clouds, the frame rate will be aligned with the frame rate in the test data set. For category 1 point clouds, the frame rate is fixed to 25 fps.
* Color space: ITU-R BT.709
* Sub-sampling: 4:4:4 RGB planar

## D.2 Delivery of submissions

The encoded bitstreams shall be submitted to the FTP site provided to each proponent. Proponents shall provide: decoder executable (including everything that is needed for using the decoder, such as decoding batch files), compressed bitstreams, and corresponding MD5 checksums. Decoded point clouds have to be provided for verification needs. Due to the large file size of decoded point clouds, only their MD5 checksums are required to be provided to the FTP site in a text file by the response date. Decoded point clouds have to be provided at the Macao meeting in October 2017.

## D.3 Laboratory Setup

The test will be done using a high-performance video server and a professional dedicated video player. Ambience will be quiet and screened for any visual and audible disturbance.

A high-end quality TV set (2017 or 2018 model) will be used as display. All local TV set features (e.g. image improvement signal processing) will be disabled.

Three subjects will be seated at 2.5 H (H = height the active part of the screen) under a visual angle lower than 60°.

Walls, ceiling and floor will be made of dark grey (or black) non-reflecting material.

Ambient light will be lower than 20 nits with no light directly pointed to the display.

A 30 nits light will be put between the back wall and the display.

## D.4 Scoring Sheets

Scoring will be done on paper scoring sheets. An example of a scoring sheet is provided as an attachment to this CfP.

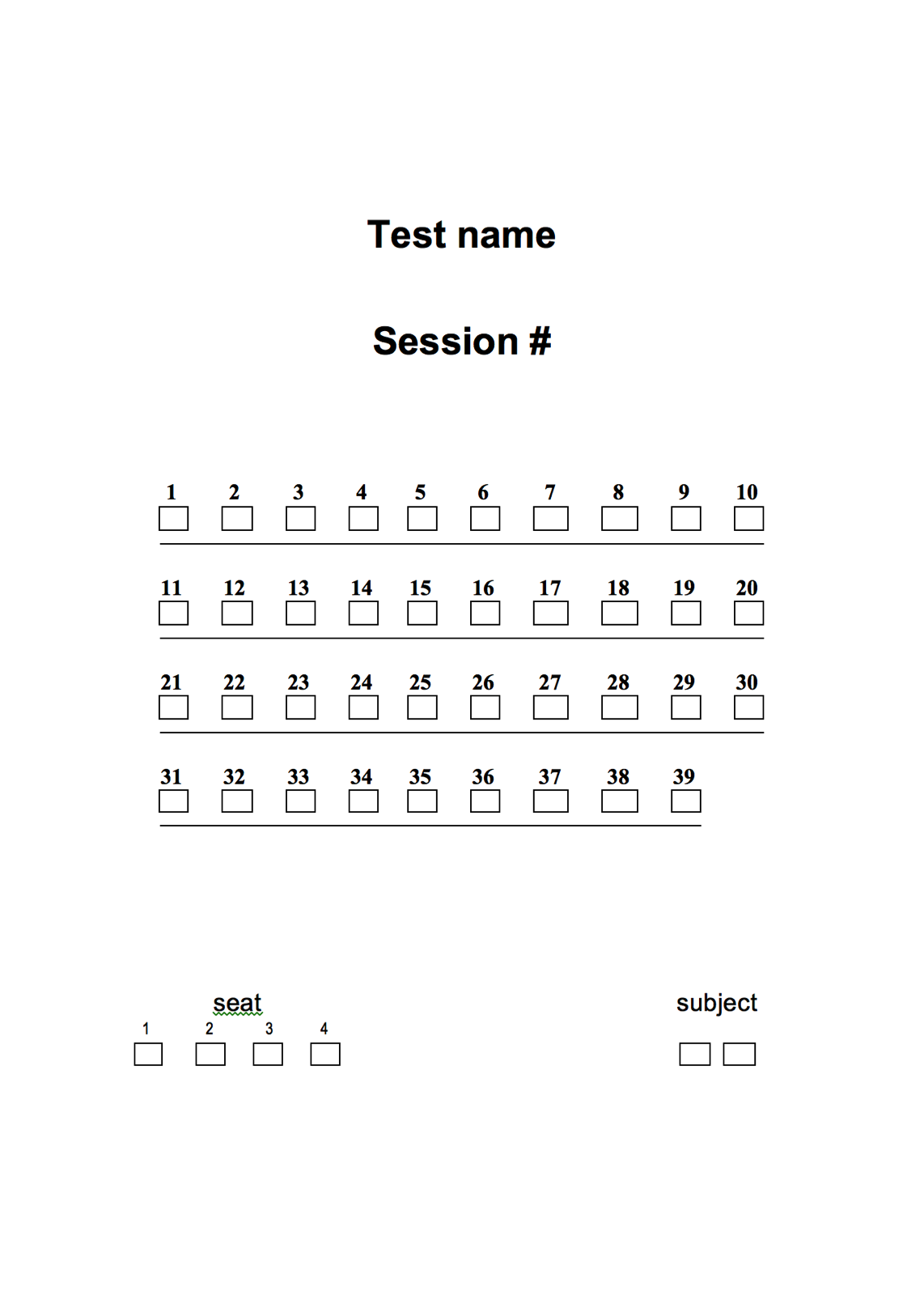


Figure 1 Scoring Sheet

Subjects will be asked to write a number (in the range from 0 to 10) in the box associated to the number of the “Vote N” message they see on the display.

The scoring sheet will also allow to identify the subject in a seedless way (no name, only a code), and the position of the viewer in front of the screen (left, center, right).

## D.5 Training of viewing subjects

Before participating in a test experiment, all subjects will participate in a training activity during which a detailed explanation of the test scope, the test method, the voting procedure, and the kind of assessment they are expected to do, will be provided.

The training activity will include a short training session conducted after the training explanation, to let the subjects practice with the scoring procedure and to allow them to familiarize themselves with the video sequences; some explanations will be provided about the kind of impairments to look for in the videos that they will see.

The video sequences included in the training session will be selected from those used for the test, trying to cover, as much as possible, the whole range of quality of the test experiment.

## D.6 Overall test effort and subjects’ involvement

The number of test sessions will cover the evaluation of all the received operating points.

Namely a DCR (i.e. the test method selected for tests with naïve subjects) test session consists of 38 BTCs (Basic Test Cell) each allowing assessing an operating point. 35 BTCs will allow to assess 35 operating points, and three dummy BTCs will be introduced at the beginning of each test session to allow a smooth creation of quality opinion by the viewing subjects.

The subjects will be grouped by three and will work for one day only.

During a working day, two groups of subjects will participate in the test running a total of four test sessions; while one group is running a test session, the other group will be resting in a quiet area.

The test sessions are designed for a total length of around 17 minutes each. Therefore, during a testing day, the subjects will be active for a total of around 4 hours (including training).

## D.7 Subjective evaluation with naïve viewers

### D.7.1 Test location

The test will be conducted in Rome at the GBTech laboratories.

### D.7.2 Related costs

The minimum testing fee when responding to the CfP is set to EUR 1000. If a proponent submits more than 35 operating points, then an additional fee of EUR 1000 will be applied for each set of 35 (or less) operating points.

The fee is used to compensate the efforts of the recruited test subjects.

### D.7.3 Selection of test subjects

Twenty four young university students (all aged between 18 and 30 years) will be selected as the test subjects. Each subject will be carefully screened for visual acuity (Snellen Chart) and color blindness (Ishihara tables). Viewing subjects will be carefully trained on both the test protocol and the kind of impairments they have to take into account and detect.