**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

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

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2016/n16332**

**February 2016, Geneva, CH**

|  |  |
| --- | --- |
| **Status** | Draft |
| **Title** | Evaluation criteria for PCC (Point Cloud Compression) |
| **Authors** | Rufael Mekuria, Zhu Li, Christian Tulvan , Phil Chou |

# Aim and scope of the document

This document presents an objective quality evaluation metric used to quantify the performance of point cloud compression.

# Introduction

While so far the most common way of representing the visual component of the world has been to take the output of a camera, compress it for transmission and storage using one of the MPEG video coding standards and eventually decode it and present it on 2D displays, there are now more and more devices that capture and present 3D representations of the world.

A point cloud is a set of points in a 3D space each with associated data relative to the value of the two angles (phi and theta) used in the acquisition, e.g. color, material properties and/or other attributes. Point clouds can be used to reconstruct an object or a scene as a composition of such points. Point clouds 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 represent realistically reconstructed scenes.

As compression technologies are needed to reduce the amount of data required to represent a point cloud, MPEG is planning to develop a Point Cloud Compression standard targeting lossy compression for use in real-time communications, lossless compression for GIS, CAD and cultural heritage applications, with attributes of efficient geometry and attributes compression, 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.

# Quality Evaluation Metric for Point Cloud Compression

The objective evaluation will be based on a full reference quality metric. The primary metric used is heavily based on standard practices for quality evaluation of polygonal mesh models, a closely related data type [1].

The point cloud consists of a set of points represented by (x,y,z) and various attributes of which color components (y,u,v) are of critical importance.

For a detailed description of the point cloud representation as defined in MPEG we refer to [4]

First, let us define the point *v,* it has as a mandatory position in a 3D space *(x,y,z)* and an optional colour attribute *c*  that has components *r,g,b* or *y,u,v* and optional other attributes possibly representing normal or texture mappings.

$point v=\left(\left((x,y,z\right),\left[c\right],\left[a\_{0}..a\_{A}\right]\right): x,y,z\in R ,\left[c \in \left(r,g,b\right) \right|r,g,b \in N], [a\_{i} \in \left[0,1\right]])$ (def. 1)

The point cloud is then simply a set of K points without a strict ordering:

$Original Point Cloud V\_{or}=\{\left(v\_{i}\right):i=0…..K-1\}$ (def. 2)

The point cloud is a set of (x,y,z), and individual attributes can be supported multiple attributes. The original cloud will act as the reference for determining the quality of a second degraded cloud Vdeg. Vdeg consists of N points, where not necessarily N = K. It is a version with a lower quality possibly resulting from lossy encoding and decoding Vor.

$Degraded Point Cloud V\_{deg}=\{\left(v\_{i}\right):i=0…..N-1\}$ (def. 3)

The quality metric $Q\_{point\\_cloud}$ is computed from $V\_{or}$ and $V\_{deg}$ and used for assessment as shown in .



Figure 1 Schematic view of full reference quality estimation of point cloud codec

In Table 1 we outline the metrics used for the assessment of the quality of a point cloud. The geometric distortion metrics are similar as the standard ones for meshes based on haussdorf (Linf) and root mean square (L2), instead of distance to surface we take the distance to the most nearby point in the cloud (see defs 4,5,6,7). We defined PSNR as the peak signal of the geometry over the symmetric rms distortion (def 8.). For colors we defined a similar metric, we compare the color of the original cloud to the most nearby colour in the degraded cloud and compute PSNR per YUV component in the YUV color space (def. 10). The key advantage of this metric is that it corresponds to PSNR in Video Coding. The quality metric is supported in the 3DG PCC software [3].

Table 1 Assessment criteria for assessment of the point cloud quality of Vdeg, $Q\_{point\\_cloud}$

|  |  |
| --- | --- |
| d\_symmetric\_rms | Symmetric rms distance between the point clouds (def. 5.) |
| d\_symmetric\_haussdorf | Symmetric haussdorf distance between the clouds (def. 7.) |
| psnr\_geom | Peak signal to noise ratio geometry (vertex positions) (def. 8.) |
| psnr\_y | Peak signal to noise ratio geometry (colors Y) (def. 10) |
| psnr\_u | Peak signal to noise ratio geometry (colors U) (as def. 10 rep. y for u) |
| psnr\_v | Peak signal to noise ratio geometry (colors V) (as def. 10 rep. y for v) |

$d\_{rms}\left(V\_{or},V\_{deg}\right)=\sqrt{\frac{1}{K}\sum\_{vo\in Vor}^{}\left⟦vo-vd\\_nearest\\_neighbour\right⟧^{2}}$ (def.4)

$d\_{symmetric\\_rms}\left(V\_{or},V\_{deg}\right)=max⁡(d\_{rms}\left(V\_{or},V\_{deg}\right),d\_{rms}\left(V\_{deg},V\_{or})\right)$ (def.5)

$d\_{haussdorf}\left(V\_{or},V\_{deg}\right)= max\_{v\_{o} \in V\_{or, }}( ||v\_{o}-v\_{d\\_nearest\\_neighbour}||\_{2} , v\_{d} is the point in Vdeg closest to v\_{o} (L2)$) (def.6)

$d\_{symmetric\\_haussdorf}\left(V\_{or},V\_{deg}\right)=max⁡(d\_{rms}\left(V\_{or},V\_{deg}\right),d\_{rms}\left(V\_{deg},V\_{or}\right)$ (def. 7)

$BBwidth=max⁡(\left(xmax-xmin\right),\left(ymax-ymin\right),(zmax-zmin)$ (def.8)

$psnr\_{geom}=10log\_{10}(\left|BBwidth\right||\_{2}^{2}/ (d\_{symmetric rms}\left(V\right))^{2})$ (def. 9)

$d\_{y}\left(V\_{or},V\_{deg}\right)=\sqrt{\frac{1}{K}\sum\_{vo\in Vor}^{}\left⟦y(vo)-y(v\_{dnearest\_{neighbour}})\right⟧^{2}}$ (def.10)

$psnr\_{y}=10log\_{10}(\left|255\right||\_{}^{2}/ (d\_{symmetric\_{y}}\left(V\right))^{2})$ (def. 11)

Additional metrics that define the performance of a codec are outlined in Table 2. Geometry and color byte size are mandatory (if colors are available). Encoding and decoding times are optional as they can only be an indicator of complexity. The number of input and output points should also be reported.

Table 2 Assessment criteria for assessment of the performance of point cloud compression ratio

|  |  |
| --- | --- |
| Compressed size | Complete compressed mesh size |
| In point count | K, the number of vertices in Vor |
| Out point count | N, number of vertices in Vdeg |
| Bytes\_geometry\_layer | Number of bytes for encoding the vertex positions |
| Bytes\_color\_layer (opt) | Number of bytes for encoding the colour attributes |
| Bytes\_att\_layer (opt) | Number of bytes for encoding the other attributes |
| Encoder time (opt) | Encoder time in ms on commodity hardware (optional) |
| Decoder time (opt) | Decoder time in ms on commodity hardware (optional) |

All these metrics are supported in the mpeg svn reference software quality evaluation metric.

# PreProcessing

As points can be located anywhere on the real line, it is recommended that any input cloud should be pre-processed to make x $\in $ [0,1],y $\in $ [0,1],z $\in $ [0,1], via normalization to the bounding box of the cloud (min\_x, min\_y, min\_z – min\_x, min\_y,min\_z). By doing this geometric distortions will be consistent.

For sequences of point clouds, an expanded bounding box (enlarging the bounding box by a factor f) can be used to keep consistent bounding box and normalization across frames instead of adapting the bounding box each frame. In addition outlier filtering is a pre-processing operation that might be used on some noisy datasets. This is implemented in the MPEG PCC software [3]. Alternatively fixed bounding boxes over finite length sequences can be defined.

# Conclusion

This document presents the primary objective quality assessment metric for assessment of point cloud compression in MPEG PCC. Other objective and subjective metrics are still open for consideration for additional evaluation. This is recommended as objective evaluation of point cloud compression might be application dependent.

# References

[1] Bulbul, A.; Capin, T.; Lavoue, G.; Preda, M., "Assessing Visual Quality of 3-D Polygonal Models," *Signal Processing Magazine, IEEE* , vol.28, no.6, pp.80,90, Nov. 2011 doi: 10.1109/MSP.2011.942466
[2] Rusu, R.B.; Cousins, S., "3D is here: Point Cloud Library (PCL)," *Robotics and Automation (ICRA), 2011 IEEE International Conference on* , vol., no., pp.1,4, 9-13 May 2011 doi: 10.1109/ICRA.2011.5980567

[3] http://wg11.sc29.org/content//MPEG-04/Part16 Animation\_Framework\_eXtension\_(AFX)/3DGraphics/pointCloud

[4] Requirements for Point Cloud Compression ISO/JCT1 SC29 WG11 n16330 Geneva, CH, June 2016