**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

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

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2015/** **N16331**

**June 2016, Geneva, CH**

|  |  |
| --- | --- |
| **Status** | Draft |
| **Title** | Use Cases for Point Cloud Compression (PCC) |
| **Authors** | Christian Tulvan, Rufael Mekuria, Zhu Li, Sebastien Laserre |

# 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.

# Use Cases for Point Cloud Compression

This document presents the use cases targeted for the MPEG standard on point cloud compression. Each of the use cases are used to illustrate the requirements defined in [1].

## Real-Time 3D immersive telepresence

Advances in 3D capture and reconstruction enable real-time generation of highly realistic 3D representations for 3D tele-presence. 3D Point clouds are an efficient representation as they can be seamlessly integrated and rendered in 3D virtual worlds enabling a convergence between real and virtual realities. As point cloud capture and reconstruction from single or multiple calibrated cameras are simpler compared to 3D mesh reconstruction, it makes the representation particularly suitable for real-time applications. In this case, point clouds are reconstructed, compressed transmitted and rendered in real-time as in video conferencing systems, enabling conversational style communication.

3D Recon

Struction

Software

Reconstructed 3D Human

3D Point Cloud Representation

3D Source Encoding

IP

Network

Multi Depth Camera

Capture

Or other 3D Capture

Real-Time 3D Rendering

Composition

In Virtual World

Packetization

&

Transmission

3D Source Decoding

N RGB + Depth Images

Or other sensor data



Reception

&

Synchronization



Figure 1: Transmission pipeline for conferencing with 3D geometry

Figure 1 shows a typical tele-immersive media pipeline. Multiple calibrated color plus depth cameras are used to acquire streams that are fed in to a reconstruction module (sensed data converter [2]). The reconstruction module produces 3D point clouds based on a reconstruction algorithm. Reconstructed point clouds are then compressed and transmitted and at the receiver side decoded and rendered compositely in a virtual world or 3D space. Some examples of these systems in the industry, 8i [3], Microsoft holo-portation [4], and in ongoing industry oriented research projects [5].

**Key requirements for this application are:**

1. *Lossy compression* with bit-rate control is needed
2. *low complexity and/or support for real-time* encoding/decoding is needed
3. *error resilience* to transmission errors is needed
4. *color attributes* coding is needed for realistic rendering
5. *material/appearance related attributes* coding is needed to code additional attributes to support the rendering
6. *view dependence:* for streaming optimization view dependence can be used to optimize the transmission process

**Typical point clouds in this use case have the following characteristics:**

1. Between 100,000 and 10,000,000 points to represent a reconstructed human
2. Color attributes with 8-10 bits per color component
3. Normals and or material properties to support the rendering using a shader

## Content VR viewing with Interactive Parallax

New VR generation now provides to end-users equipped with Head Mounted Display (HMD) an immersive experience for viewing specific contents like large field of view movies. One major improvement for improving the visual comfort relates to the viewing of content with interactive parallax, where the rendering viewport is updated for each new position of the end-user head. This use case is very much related to the previous one, except that

* It does not require a real time preparation of the source, but is rather typically based on a long and sophisticated production workflow mixing shooting and computer graphics
* The point cloud only needs to be rendered from a range of viewport position, typically corresponding to a bounding box around the average position of the end-user head. This corresponds to the view dependence mentioned here below.
* The point cloud may be a part only of the total content transmitted to the rendering sites.

Packetization

& Transmission

Content production workflow

Real time 3D rendering with HMD

and composition in virtual world

3D Source (non real time) Encoding

IP network



Reception & synchronization

3D Source (real time) decoding

Figure 2: Point Cloud for interactive parallax broadcasting

**Key requirements for this application are:**

1. *Lossy compression* with bit-rate control is needed
2. *error resilience* to transmission errors is needed
3. *color attributes* coding is needed for realistic rendering
4. *material/appearance related attributes* coding is needed to code additional attributes to support the rendering
5. *view dependence:* the view dependence can be used to optimize the encoding process

**Typical point clouds in this use case have the following characteristics:**

1. Between 100,000 and 10,000,000 points to represent closeby objects in the scene
2. Color attributes with 8-10 bits per color component
3. Global parameters defining the spatial constraints of the rendering viewport

## 3D Free viewpoint Sport Replays Broadcasting

Point cloud capture of sports events like basketball, baseball for free viewpoint playback and interaction on mobile devices and TV. This requires a compression and file format standard such that the point cloud data can be streamed or stored in interoperable manner. Desired features are good compression efficiency, progressive download for different device capabilities. The industry has already moved in this direction, examples are Replay Technology [14] (see Figure 3). The full scene can also contain textures and/or video. The coding of these data is outside the scope of PCC and can be performed using alternative MPEG standards. The scope of PCC will only be to compress the point cloud data. Composition and synchronization can be achieved using MPEG Systems standards.

|  |  |  |
| --- | --- | --- |
|  |  |  |

*Figure 3* Free viewpoint sport capture as a point cloud

**Key requirements for this application are:**

1. *Low Delay* Encoding and Decoding
2. *Color* attributes coding is needed, preferable 8-12 bits per component

**Typical point clouds in this use case have the following characteristics:**

1. 100,000 – 100,000,000 points
2. Color attributes of 8-12 bits per color component
3. Can contain multiple clusters/groups of points (different players)

## Geographic Information Systems

Geographic information is often represented as point clouds (resulting from airborne SAR, Lidar) and are often too massive to be handled efficiently in current geo-ICT infrastructures [5]. Point Cloud compression can improve point cloud data management systems (PCDMS) [5]. Typically, point clouds representing geographic information are captured using Lidar, SAR or other techniques. These data are often stored in servers providing a service such as remote rendering, or querying based on specific geographic information [5]. Compression can reduce the storage requirements at the server, and the traffic exchanged between the client and server. An example of a commercial compression tool for GIS PCC is LASZip [7] (lossless compression for geographic information). Examples of applications that could benefit from a standardized solution are Oracle spatial and graphical databases [8], ArcGIS [9], PointScene [10].

**Key requirements for this application are:**

1. *region selectivity* is important to query subsets of the point cloud representing a certain geographic area
2. *lossless* is important to use the best representation available for interpreting geographic information. Height, water level, forestry etc.
3. *progressive* is important to enable efficient download.

**Typical point clouds in this use case have the following characteristics:**

1. Billions of points (e.g. height map Netherlands, 640 Billion points [11])
2. Points often have a planar structure
3. Points have additional attributes related to geographic properties.

## Cultural heritage

Point cloud scans be used to visualize and archive cultural heritage objects and collections. Examples are the Shipping Gallery in the London Museum [12]. Another example is the project Culture 3D Cloud [13] shown in Figure 4, where pictures taken by smart phones of cultural heritage objects are transformed into 3D point clouds that can be viewed from any angle. Point cloud compression and streaming can make these objects available to the wider public.

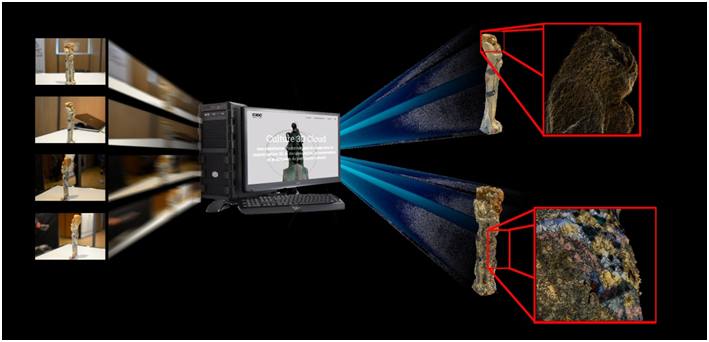


Figure 4 Cultural heritage in the 3D Cloud project [13]

**Key requirements for this application are:**

1. *Progressive coding* to enable increasing quality.
2. *Color* attributes coding is needed, preferable 8-12 bits per component
3. *Generic* attributes coding such as for material properties.
4. *Lossless* is important to enable the best representation when possible

**Typical point clouds in this use case have the following characteristics:**

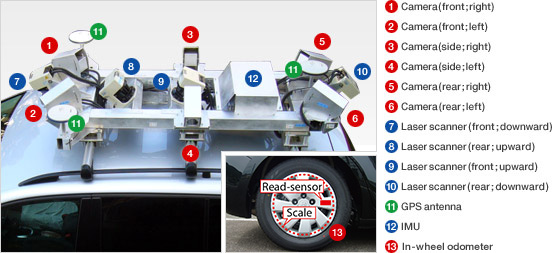
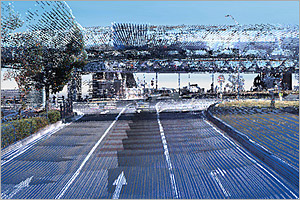
1. 1 Million upto billions of points (e.g. [13])
2. Color attributes of 8-12 bits per color component
3. Can contain multiple clusters/groups of points

## Autonomous Navigation Based on Large-Scale 3D Dynamic Maps

Large-scale 3D maps of indoor and outdoor environments can be created using devices that provide localization combined with depth and color measurements of the surrounding environment. Localization could be achieved with GPS, inertial measurement units (IMU), cameras, or combinations of these and other devices, while the depth measurements could be achieved with time-of-flight, radar or laser scanning systems. Example mapping systems are already commercially available and come in various forms such as the high-end mobile mapping system [16] shown in Figure 5(a). There are also lightweight mobile platforms under development based on sensors mounted on drones or in mobile phones/tablets.

Irrespective of the specific platform that is used to acquire the measurements, it is possible to generate a 3D map by combining the depth measurements, such as the high-density laser-scanned point cloud in Figure 5(b), with camera images. This combination of point cloud data with camera images to generate a 3D map is illustrated in Figure 5(c). These maps can further be combined with road markings such as lane information and road signs to create maps to enable autonomous navigation of vehicles around a city as shown in Figure 5(d).

Multiple map layers will be stored and exchanged across the network, including static maps that do not change very frequently and dynamic maps that include real-time information about dynamic objects in the scene such as vehicles or pedestrians.

(a) (b)

(c) (d)

Figure 5: Illustrations of large-scale 3D mapping equipment and data. (a) Mobile mapping system for data acquisition, (b) sample of a high-density laser-scanned point cloud, (c) map creation where images are projected by superimposing laser point cloud data onto camera images, and (d) sample of map with roadside features added.

**Key requirements for this application are:**

1. *High precision* is needed to support localization needs
2. *Low complexity and/or support for real-time* encoding/decoding is needed
3. *Low delay* is needed for real-time communication of dynamic parts of the map
4. *Region selectivity* is important to maintain and access the map data
5. *Color attributes* coding is needed for realistic rendering and visualization
6. *Additional attributes coding* for reflectance and other scene properties

**Typical point clouds in this use case have the following characteristics:**

1. Millions to billions of 3D points with up to 1cm precision
2. Color attributes with 8-12 bits per color component
3. Normals and/or reflectance properties as additional attributes

# References

[1] n16330 Requirements for Point Cloud Compression ISO JCT1 SC29 WG11 Geneva June 2016

[2] m37050 3D Visual Capture Inventory ISO JCT1 SC29 WG11 Geneva October 2015

[3] “Making virtual Reality Human” Create Volumetric video of real people <http://8i.com/>

[4] “Microsoft Holoportation” <http://research.microsoft.com/en-us/projects/holoportation/>

[5] “Real and Virtual engagement in Realistic Virtual Environments” [www.reveriefp7.com](http://www.reveriefp7.com)

[6] P. van Oosterom, O. Martinez-Rubi, M. Ivanova, M. Horhammer, D. Geringer, S. Ravada, Theo Tijssen, Martin Kodde, Romulo Gonçalves, Massive point cloud data management: Design, implementation and execution of a point cloud benchmark, Computers & Graphics, Volume 49, June 2015, Pages 92-125, ISSN 0097-8493, http://dx.doi.org/10.1016/j.cag.2015.01.007

[7] Isenburg, Martin. "LASzip." Photogrammetric Engineering & Remote Sensing 79.2 (2013): 209-217.

[8] Oracle Spatial and Graph <https://www.oracle.com/database/spatial/index.html>

[9] ArcGis <https://www.arcgis.com/features/>

[10] PointScene https://pointscene.com/

[11]Actueel Hoogtebestand AHN2, Netherlands Height Map data <http://www.ahn.nl/pagina/open-data.html>

[12] London Shipping Galleries, Scanlab <http://scanlabprojects.co.uk/projects/sciencemuseum>

[13] Culture 3D Cloud <http://c3dc.fr/>

[14] Replay technologies <http://replay-technologies.com/>

[15] <https://en.wikipedia.org/wiki/Z-order_curve>

[16] Mitsubishi Electric Mobile Mapping Systems (MMS), <http://www.mitsubishielectric.com/bu/mms/>