|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  36th Meeting: Gothenburg, SE, 6–12 July 2019 | Document: JCTVC-AJ0028-v2 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Encoder-only Supplemental Motion Vector Estimation for Point cloud Coding content** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or*  *Contact(s):* | **V. Zakharchenko**  **J. Chen** 2330 Central expy, Santa Clara CA 95050, USA | Tel:  Email: | +1-408-828-6368  vzakharc@futurewei.com |
| *Source:* | Futurewei Technologies, Inc. | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This contribution proposes an encoder-only motion estimation based on external information for the point cloud content. The motion estimation is performed at the encoder side based on the supplementary information provided by the video-based point cloud compression engine. Source pictures generated by the video-based point cloud compression engine can be accompanied with an additional information that is generated during projection and atlas composition stage in the video-based point cloud compression solution. Samples in the selected picture are allocated based on 3d coordinates thus the 2d location can be inferred from the information that is present in coded point cloud bitstream.

The nature of the picture that contains projected point cloud atlas can’t guarantee temporal consistency in patch allocation. The method suggests introducing additional start point for the motion estimation process to solve this problem. As the result the precise motion vector shall be generated and encoded, providing good spatial candidate for motion vector candidate list construction for the subsequent coded units.

The method was reportedly tested under V-PCC TMC2 v6 RA, common test configurations. The average D1/D2/Y/U/V total BD-rates are reported to be −6.4%/−6.6%/−10.8%/−16.7%/−15.3% for the dynamic point cloud content [3-5].

It is proposed to adopt the proposed method into the HM software and keep the method disabled by default for HM CTC and enabled for V-PCC CTC only.

# Introduction

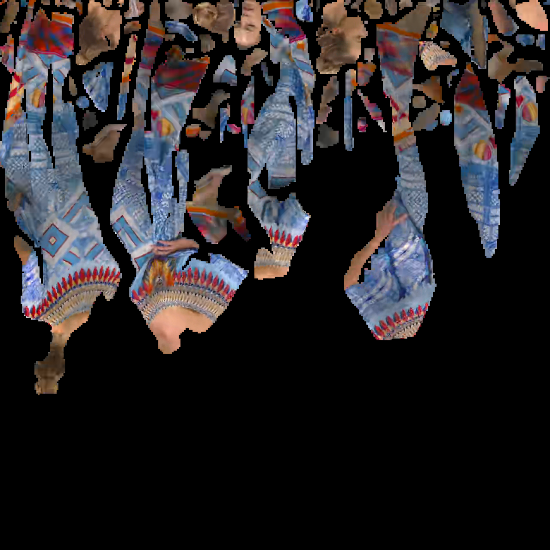
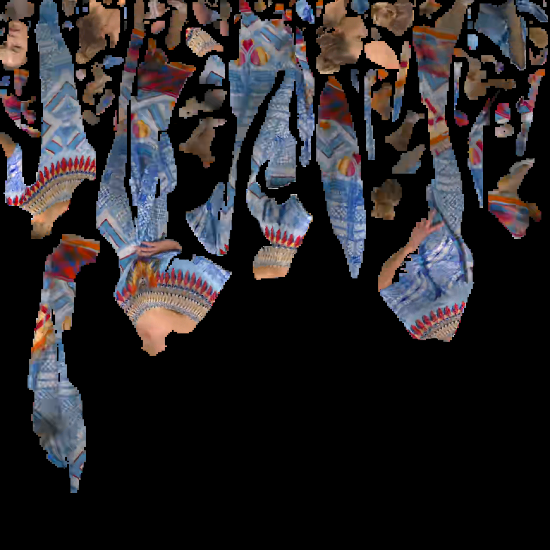
The HM encoder performs motion estimation based on existing motion vector candidates, however the HEVC coding standard supports any motion vector difference, that would satisfy the requirement of cost function minimization: rate\*lambda+distortion for a given block.

If the next picture differs significantly and motion is inconsistent the probability of an INTRA block to be generated increase, unless motion estimation process can find a good predictor. Rather than performing an exhaustive search for the texture or geometry atlas, the projection information, present in video-based point cloud compression (V-PCC) solution can be used. This is the motivation for adding the patch data frame information to motion estimation process in HM solution.

# Proposal

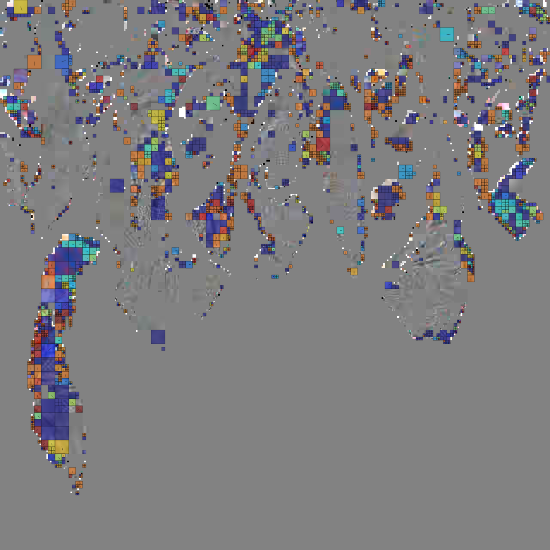
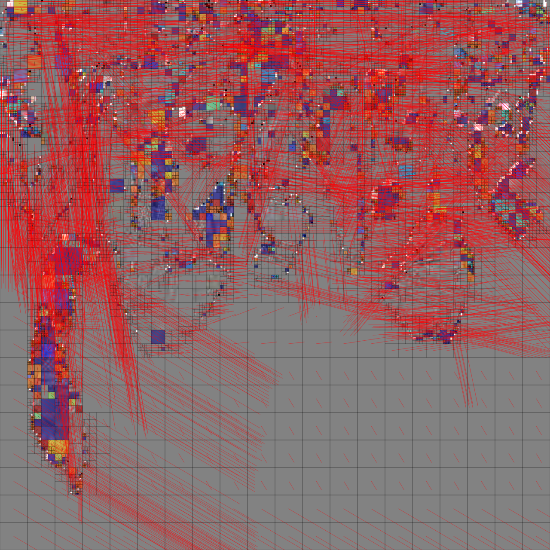
## 3D motion estimation improvement

The patch allocation on two-dimensional canvas is an n-p hard problem and can’t be directly solved under constraints as live streaming, conferencing systems, etc. A motion estimation process in hybrid video compression is limited with the search windows and maximum motion vector size does not allow the true motion representation, especially for patches that switch their position within canvas in subsequent pictures. The effect of patch allocation (figure 1) and corresponding motion vectors is demonstrated on figure 2 below.

a) b)

Figure 1 – Texture atlas for dynamic point cloud content subsequent frames (near layer only): a – pcc frame 0, video picture 0; b – pcc frame 1, video picture 2.

a) b)

Figure 2 – The results of picture 2 coding: a – intra blocks; b – motion vector field.

In the non-normative solution [6], instead of using the proposed 3D MV as a MV predictor candidate, we simply set the proposed 3D MV as a candidate of the start position of the motion estimation search range. After adding the derived MV as the candidate, the encoder will choose from the MV predictor, the zero MV, the proposed 3D MV, and the MV from the partition 2Nx2N using rate distortion optimization to determine the center of the search range. If the proposed 3D MV is with the smallest rate-distortion (R-D) cost and chosen as the start point, it will be more probable for the encoder to find the corresponding block in the reference frame thus can provide a significant performance improvement. The non-normative solution also has the benefit of no changes to the standard.

The method operates based on introduction of relation for 3d to 2d information conversion in V-PCC solution [2] the algorithm workflow is described in figure 3.

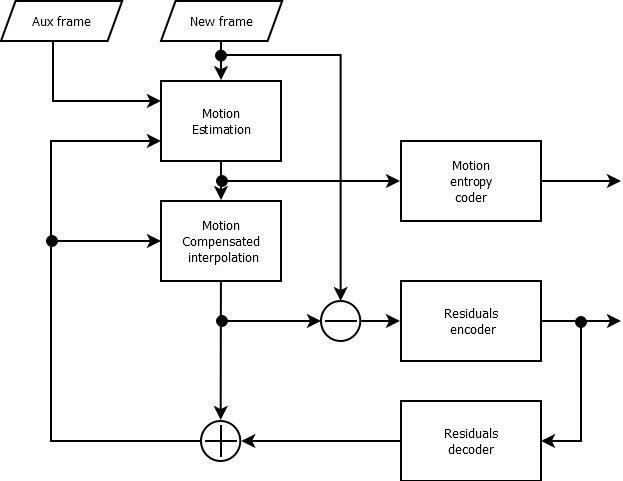
 ****

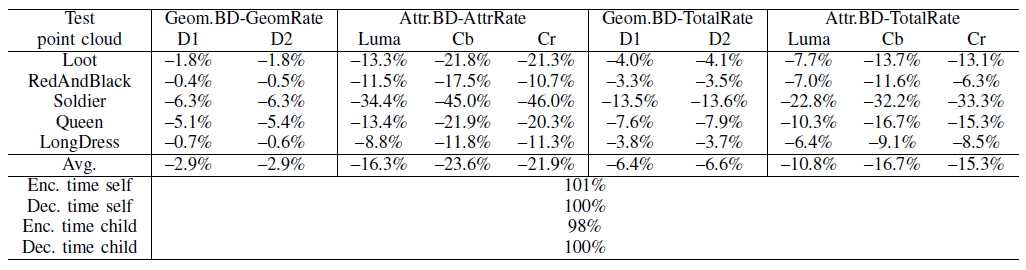
Figure 3 – 3D motion vector derivation process based in 3d information.

|  |
| --- |
| Algorithm. PCC coder additional data export |
| **while** frame *k* in GoF frames **do**  generate patches in frame *k*.  **for each** patch in frame *k*  generate projection map from patches  store patch2D data: U0, sizeU0, V0, sizeV0  store patch3D data: U1, V1, D1, axis  store occupancy Map data  **end for**  **end while** |

|  |
| --- |
| Algorithm. Find matched patches in reference frames |
| **while** frame *k* in GoP frames **do**  **for each** patch in patches list of reference frame *[refIdx]*  **for each** patch in patches list of frame *[k]*  estimate cost in RDO based on prediction  **if** (patchIdx[k] = patchIdx[refIdx])  do motion refinement  **end if**  **if** (dist < bestDist)  update pcMvFieldNeighbours  **end if**  **end for**  **end for**  **end while** |

# Results

Proposed method was implemented in and compared to HM-16.20. The method is applied for RA coding, and not to AI coding. The method has no impact on the HM-16.20 CTC results.



# Conclusion

In this document an external motion estimation method was introduced providing improved compression efficiency for dynamic point cloud content. We recommend adopting the proposed method into the HM software, and disable the method in the common test conditions by default.

# References

[1]. HM-16.20 reference software, <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20>

[2]. V-PCC-5.1 reference software, <http://mpegx.int-evry.fr/software/MPEG/PCC/TM/mpeg-pcc-tmc2/tree/release-v5.1>

[3]. Dynamic point cloud content submitted by 8i 8iVFB, <http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/datasets/Dynamic_Objects/People/8i/>

[4] Dynamic point cloud content submitted by Technicolor, <http://mpegfs.int-evry.fr/MPEG/PCC/DataSets/pointCloud/CfP/datasets/Dynamic_Objects/People/Technicolor/>

[5] Dynamic point cloud content submitted by Owlii, <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Part05-PointCloudCompression/dataSets/Owlii/VoxelizedPointCloud/>

[6] V. Zakharchenko, Li Li, [V-PCC] [EE2.1 Response] 3D motion estimation for video compression”, ISO/IEC JTC1/SC29 WG11 Doc. m45967, Marrakech, MA, January 2019

# Patent rights declaration(s)

**Futurewei Technologies may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**