Lossless Point Cloud Attribute Compression with Normal-based Intra Prediction

Qian Yin, Qingshan Ren, Lili Zhao, Wenyi Wang, and Jianwen Chen*

School of Information and Communication Engineering
University of Electronic Science and Technology of China
Chengdu, China
chenjianwen@uestc.edu.cn

Abstract—The sparse LiDAR point clouds become more and more popular in various applications, e.g., the autonomous driving. However, for this type of data, there exists much under-explored space in the corresponding compression framework proposed by MPEG, i.e., geometry-based point cloud compression (G-PCC). In G-PCC, only the distance-based similarity is considered in the intra prediction for the attribute compression. In this paper, we propose a normal-based intra prediction scheme, which provides a more efficient lossless attribute compression by introducing the normals of point clouds. The angle between normals is used to further explore accurate local similarity, which optimizes the selection of predictors. We implement our method into the G-PCC reference software. Experimental results over LiDAR acquired datasets demonstrate that our proposed method is able to deliver better compression performance than the G-PCC anchor, with 2.1% gains on average for lossless attribute coding.

Index Terms—3D point cloud, G-PCC, attribute compression, normal-based prediction

I. INTRODUCTION

With rapid development of 3D sensing and capturing technologies, point clouds, which have the capacity of representing spatial structures and surface properties of 3D objects or scenes, are often encountered in various fields, e.g., the autonomous driving, the heritage reconstruction and 3D immersive communication [1]. However, it is well-known that point clouds have the unorganized distribution in 3D space and consist of millions of points, which imposes the burden on the limited transmission bandwidth and storage space. Therefore, the compression of point clouds is indispensable but challenging. Therefore, it is necessary to explore more effective point cloud compression (PCC) schemes.

Due to the potential needs of point cloud related applications, the Moving Picture Experts Group (MPEG) establishes the MPEG-3DG subgroup to exploit universal PCC frameworks. Two standardized test models are proposed: the video-based PCC (V-PCC) [2] and the geometry-based PCC (G-PCC) [3]. For V-PCC, 3D point clouds are projected into the 2D domain, and then coded by using the existing video codec (e.g., High Efficiency Video Coding, HEVC [4]), which is more suitable for the dense point cloud compression. The G-PCC, by contrast, performs better for the sparse point cloud by compressing point clouds in the original 3D space. It is noted that point clouds typically consist of geometry information (i.e., 3D coordinates) and attribute information (e.g., colors, reflectances and normals). In the existing PCC schemes, the geometry and attribute information are coded separately. In this work, we focus on the attribute lossless coding in G-PCC for point clouds acquired by LiDAR sensors. Some examples can be seen in Fig. 1.

Among the existing methods for the point cloud attribute coding, one popular strategy is the Graph Fourier Transform (GFT) (e.g., [5], [6], [7]). In [5], Zhang et al. proposed an attribute compression method based on the graph transform. This scheme delivers a more effective PCC by using GFT instead of the traditional DCT, but generates many isolated sub-graphs when it comes to the sparse point clouds. To address this issue, Robert et al. used the k-nearest neighbours (KNNs) method in [6] and Shao et al. introduced Laplacian sparsity in [7] to optimize the graph transform respectively. However, due to much higher computational complexity introduced by eigenvalue decompositions, it is difficult for these graph-based methods to achieve real-time PCC.

Besides the GFT-based methods, Queiroz et al. proposed a region-adaptive hierarchical transform (RAHT) scheme by
The proposed prediction framework for the lossless attribute compression, as shown in Fig. 3, consists of the level of detail (LOD) generation and the enhanced Predicting Transform scheme. The LOD structure divides the whole point cloud into a set of refinement levels. Then, the re-organized point cloud is processed by the enhanced Predicting Transform scheme, which incorporates the original Predicting Transform scheme in G-PCC shown by the dashed box (left) and our proposed normal-based prediction scheme shown by the dashed box.
A. The Integration of the Original Distance-based Prediction and the Proposed Normal-based Prediction

In G-PCC, the point cloud is first sorted according to their associated Morton codes in an ascending order, and then divided by the LOD generation process. In our proposed scheme, we preprocess the re-ordered point clouds by calculating the normals before the LOD process, which is prepared for the subsequent normal-based prediction. To be specific, for each point $P_i$, we use the k-d tree to find $N$ nearest neighbors (i.e., $N = 15$), which defines the local plane of $P_i$. Then, the normal of the local plane is calculated by using the eigenvalue decomposition, which serves as the approximate normal of the point $P_i$. The point clouds along with calculated normals are next re-organized by the LOD structure process.

Fig. 3 presents the enhanced Predicting Transform framework, where the original Predicting Transform scheme is as a part shown by the dashed box (left). It is known that there are mainly three stages in the Predicting Transform scheme: the distance-based predictor generation, the calculation of the attribute value range and the selection of predictor modes. At the stage of the distance-based predictor generation, the distances from the current point $P_i$ to previously encoded points are computed and then $k$ ($k = 3$) nearest neighbour points of $P_i$ can be selected.

Based on 3 selected points, the distance-based predictor is generated, which is then utilized for the selection of predictor modes. Specifically, the differences between each pair of three nearest neighbors are computed in turn, and then the maximum difference can be available, denoted as $\text{max}\_\text{diff}$. By comparing $\text{max}\_\text{diff}$ with a pre-defined threshold, different predictor modes are selected accordingly. In G-PCC, there are four predictor modes, denoted as Mode 0, Mode 1, Mode 2 and Mode 3 respectively. Mode 0 represents an interpolation-based prediction with the Inverse Distance Weighted (IDW) method by using 3 nearest-neighbors. For Mode 1, Mode 2 and Mode 3, 1st, 2nd and 3rd nearest neighbors are directly used to predict the current point respectively.

From Fig. 3, it can be observed that when the $\text{max}\_\text{diff}$ of the neighbor’s reflectance exceeds the pre-defined threshold, the current point $P_i$ is adaptively predicted from three predictor candidates (Mode 1, Mode 2 and Mode 3) by using...
the rate-distortion optimization (RDO) procedure. Otherwise, Mode 0 will be chosen. In this case, instead of using Mode 0, our proposed normal-based predictor is applied, which aims to optimize the selection of predictor candidates. The specific procedure is shown in Fig. 3 denoted with the dashed box (right). The main novelty is that, in addition to the predictor in the original G-PCC using the similarity in distances between coordinates, our proposed method further explores the local similarity among points by introducing the angle between normals.

B. The Generation, Enhancement and Selection of the Normal-based Predictor

The normal-based prediction method is shown in Fig. 3 with the dashed box (right), which mainly consists of three stages: the normal-based predictor generation, the predictor enhancement with an angle selection and the predictor modes decision. To be specific, the normal-based predictor is generated on top of the previous distance-based predictor, by introducing the normals of 3 nearest-neighbors.

The normal-based predictor obtained above is then improved by an angle selection. The calculation of the angle between normals is shown in Fig. 4. Let \( V_i \) and \( V_j \) be the normals of the points \( P_i \) and \( P_j \) respectively. Then, the angle \( \theta \) between normals \( V_i \) and \( V_j \) is calculated by

\[
\theta = \arccos\left(\frac{V_i \cdot V_j}{\|V_i\| \cdot \|V_j\|}\right). 
\]

At the stage of the predictor enhancement, considering the similarity of the attributes decreases with the increase of the distance between point clouds, we only use the 1st nearest neighbouring point, which is to enhance the normal-based predictor with an angle selection. Specifically, we calculate the angle between normals of 1st nearest neighbor and the current point based on Equation (1), denoted as the normal_angle.

From Fig. 3, it can be observed that when the normal_angle is greater than 90\(^\circ\), Mode 0 is selected by using the weighted average of neighbors’ attributes for the prediction. If not, Mode 1 is selected by using the 1st neighbor to predict the current point. Note that Mode 0 and Mode 1 follow the same paradigm adopted in G-PCC.

Finally, the prediction residuals are processed by the quantizer and arithmetic encoder. It is worth to mention that no extra flags or parameters are required to be written into the final bitstream, because the decoder is implemented by repeating the operation of the normal-based prediction scheme.

IV. EXPERIMENTAL RESULTS

To evaluate the effectiveness of our proposed normal-based intra prediction scheme, extensive simulations have been conducted on the test dataset provided by MPEG. We incorporate our method into the MPEG G-PCC reference software, i.e., TMC13v10 [8], and compare its compression performance [4] of attribute compression with the original TMC13.

In our experiments, the test dataset consists of seven dynamically acquired point clouds (denoted as Category 3-frame). Among them, three ford sequences, ford_01_q1mm, ford_02_q1mm, and ford_03_q1mm can be available by [15], while the other four qnxadas sequences, qnxadas-junction-approach, qnxadas-junction-exit, qnxadas-motorway-join, and qnxadas-navigating-bends, can be obtained in [16]. More details of the Category3-frame dataset are listed in Table II. All the experiments are conducted under CW condition of the Common Test Conditions (CTC) [10], where CW represents lossless geometry and lossless attribute.

Table II shows the comparison results between proposed method and the G-PCC anchor on the Category 3-frame under the CW condition. Since this work aims for lossless compression with no distortion of data, the commonly-used evaluation metric, measuring the rate in terms of bits for attributes, bits per input point (denoted as \( \text{bpip} \)), is used for evaluation. Specifically, we compute the \( \text{bpip} \) of the proposed method and that of the TMC13, and then the bit saving ratio (\( \Delta R \)) can be obtained for the final performance evaluation, which is defined as

\[
\Delta R = \frac{\text{bpip}_{\text{ours}} - \text{bpip}_{\text{TMC13}}}{\text{bpip}_{\text{TMC13}}} \times 100\%
\]

From equation (2), we can see that the proposed method outperforms the anchor (TMC13) when the bit saving rate \( \Delta R \) is negative. Otherwise, it means that the anchor shows better performance than the proposed method.

From Table II, it can be obviously observed that the bit savings can be achieved with 2.1% on average, especially for sequence qnxadas-junction-exit, up to 5.6%. From the results, we can see that our proposed method is effective on the LiDAR datasets for attribute lossless compression in G-PCC. The rationale behind the gains is that by introducing our normal-based prediction scheme, the attribute similarity can be explored more accurately, which leads to better selection of the predictor mode.
TABLE II
THE COMPARISON RESULTS BETWEEN OUR METHOD AND THE G-PCC ANCHOR ON CATEGORY 3-FRAME UNDER THE CONDITION CW.

| Sequences                              | Frame Number | Input Points | Geometry Precision (bits) | Peak Value | ∆R (%) |
|----------------------------------------|--------------|--------------|---------------------------|------------|--------|
| ford_01_q1mm                           | 1500         | 123940658    | 18                        | 30000      | -1.3   |
| ford_02_q1mm                           | 1500         | 125751705    | 18                        | 30000      | -1.0   |
| ford_03_q1mm                           | 1500         | 126093865    | 18                        | 30000      | -0.9   |
| qnxadas-junction-approach              | 74           | 2233793      | 18                        | 30000      | -2.8   |
| qnxadas-junction-exit                  | 74           | 2016190      | 18                        | 30000      | -5.6   |
| qnxadas-motorway-join                  | 500          | 14430189     | 18                        | 30000      | -4.5   |
| qnxadas-navigating-bends               | 300          | 8167066      | 18                        | 30000      | -4.0   |
| Cat3-frame average                     |              |              |                           |            | -2.1   |

V. CONCLUSION

In this paper, to improve the efficiency of the lossless attribute compression in G-PCC, a normal-based intra prediction scheme is proposed, which further exploits the geometrical correlations among neighbors in point clouds. Based on the original distance-based Predicting Transform scheme, the normals of each point, as an additional descriptor, are introduced to optimize original predictors. By computing the angle between normals, a better predictor mode can be selected. Experimental results have demonstrated that our method is able to consistently deliver a better performance than the G-PCC test model.

REFERENCES

[1] C. Tulvan, R. Mekuria, and Z. Li, “Use cases for point cloud compression (PCC), document N16331,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Geneva, Jun. 2016.
[2] “V-PCC codec description, document N19332,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Alpbach, Apr. 2020.
[3] “G-PCC codec description, document N19331,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Alpbach, Apr. 2020.
[4] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” IEEE Trans. Circuits Syst. Video Techn., vol. 22, no. 12, pp. 1649–1668, Sep. 2012.
[5] C. Zhang, D. Florêncio, and C. Loops, “Point cloud attribute compression with graph transform,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2014, pp. 2066–2070.
[6] R. A. Cohen, D. Tian, and A. Vetro, “Attribute compression for sparse point clouds using graph transforms,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 1374–1378.
[7] Y. Shao, Z. Zhang, Z. Li, K. Fan, and G. Li, “Attribute compression of 3D point clouds using laplacian sparsity optimized graph transform,” in Proc. IEEE Int. Conf. Visual Commun. Image Process. (VCIP), Dec. 2017, pp. 10–13.
[8] R. L. de Queiroz and P. A. Chou, “Compression of 3D point clouds using a region-adaptive hierarchical transform,” IEEE Trans. Image Process., vol. 25, no. 8, pp. 3947–3956, Aug. 2016.
[9] “Lifting scheme for lossy attribute encoding in TMC1, document m42640,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, San Diego, Apr. 2018.
[10] “Common test conditions for point cloud compression, document N19324,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Alpbach, Apr. 2020.
[11] “CE13.20 report on neighbor’s weight modification on Lifting and Predicting Scheme, document m50773,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Geneva, Oct. 2019.
[12] “PCC Adaptive predictor selection for attributes coding in TMC13 related to CE13.3, document m43665,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Ljubljana, Jul. 2018.
[13] “[G-PCC][New Proposal] on improvement for adaptive reflectance predictor selection, document m50765,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Geneva, Oct. 2019.
[14] “G-PCC performance evaluation and anchor results, document N19326,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Alpbach, Apr. 2020.
[15] G. Pandey, James R. Mcbride, and Ryan M. Eastice, “Ford Campus vision and lidar data set,” The International Journal of Robotics Research, vol. 30, no. 13, pp. 1543–1552, 2011.
[16] “PCC Cat3 test sequences from BlackBerry—QNX, document m23647,” ISO/IEC JTC 1/SC 29/WE 11 MPEG, Ljubljana, July. 2018.