WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 2, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n2.03
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Artificial Intelligence in Point Cloud Compression

Bowen Lv *

School of Computer Science and Technology, Beijing Jiaotong University, Beijing, 100044, China
*Corresponding Author: 22722063@bjtu.edu.cn

ABSTRACT

Point cloud is a data representation of a three-dimensional object or scene. It is created by collecting a large number of spatial coordinate points, which are used to discretely represent the surface information of a three-dimensional object. Compared to traditional two-dimensional images, point clouds are particularly suited to the presentation of three-dimensional models and geospatial information. They have had a significant impact on the development of smart cities, autonomous driving applications, and augmented reality technology. However, due to the irregularity, disorganisation, diversity and other shortcomings of the point cloud, the traditional point cloud compression faces significant challenges. This paper will review the traditional coding methods and focus on the current cutting-edge artificial intelligence technology to bring the point cloud compression advances. It will also consider the future development direction.

KEYWORDS

3D point cloud; Artificial intelligence; Compression algorithm; 3D convolution

1. INTRODUCTION

Point Cloud is one of the newest data to represent 3D objects or realistic scenes, which utilises a huge collection of 3D space. It uses a large collection of spatial coordinates to discretely represent the surface information of three-dimensional objects. Compared with traditional two-dimensional images, point cloud can directly and flexibly represent three-dimensional information, and has a wide range of application scenarios in the fields of robotics, immersive media, free-viewpoint broadcasting, geographic mapping, and automated driving, etc. It is a kind of data that will be used in the future. It has a wide range of application scenarios in the fields of immersive media, free viewpoint broadcasting, geographic mapping, automatic driving, etc. It is a kind of data object with great application prospects in the future, and is worthy of deep research. The characteristics of high dimensional data and the demand for high precision and large scenes have resulted in a huge amount of data in point cloud, which also brings pressure and pressure on its storage and transmission. storage and transmission brings pressure and challenges, which is a major bottleneck restricting the development of related applications. Therefore, point cloud data compression of point clouds has become an urgent research topic. Compared with traditional data such as images, videos, etc., which have a long history of research and mature compression methods, the compression of point clouds has become an urgent research topic.

However, the characteristics of high dimensional data and the demand for high precision and large scenes cause the huge amount of data in point cloud, which also brings pressure and challenge to its storage and transmission, which is a major bottleneck restricting the development of related applications. Therefore, data compression of point cloud has become an urgent research topic. Compared with traditional data such as images and videos, which have a long history of research and mature compression methods, point cloud compression is still in the development stage because of its irregular distribution and sparse characteristics, which make compression difficult.

In the evolution of point cloud compression techniques, we observe two major trends developing in parallel. Firstly, there is the clever use and adaptive improvement of traditional data compression techniques, especially image compression algorithms. This approach is represented by the Video-based Point Cloud Compression (V-PCC) technique introduced by MPEG (Motion Picture Experts Group), which provides an efficient means of compression for point cloud data by drawing on proven image compression strategies. Secondly, with vtinuous advancement of in-depth research on the characteristics of point cloud data, a series of new coding techniques for its uniqueness have emerged. These techniques are usually based on structures such as octrees, grids, graphs, etc., and aim to achieve more efficient data compression by capturing the intrinsic geometric and topological relationships of the point cloud data. MPEG's Geometry-Based Compression of Point Clouds (G-PCC) is an outstanding representative in this direction. Whether based on adaptive improvements of traditional techniques or new coding techniques for point cloud data characteristics, these different schemes have demonstrated excellent compression performance in their respective application scenarios, providing strong support for the storage, transmission and processing of point cloud data [1-3].

Artificial intelligence technology has made significant advances in the field of traditional image and video coding [4]. These advances have resulted in a reduction in the bit rate and an improvement in reconstruction quality. Additionally, they have enabled the integration of intelligent analysis and processing functions, which foreshadow the future direction of video coding. Similarly, in the field of point cloud compression, technology based on artificial intelligence is receiving increasing attention and in-depth research in the industry. The challenge of optimising the compression performance of point cloud data while maintaining its accuracy and integrity has become a significant research topic, given the potential of intelligent compression techniques to address this issue.

The objective of this paper is to provide a comprehensive overview of the challenges associated with point cloud compression, to examine in depth the traditional point cloud compression algorithms, and to analyse the current research progress and future development trend of artificial intelligence in the field of point cloud compression [5]. In recent years, the rapid development of artificial intelligence technology has led to significant advantages in the application of AI in point cloud compression. In the case of dense point clouds, the use of voxel-based strategies combined with three-dimensional convolutional neural networks (3D CNNs) has surpassed traditional compression techniques in terms of geometric compression. This innovation not only improves the compression efficiency, but also dramatically improves the quality of the reconstructed point cloud. In the case of sparse point cloud geometry compression, the direct use of point set data in combination with advanced network structures such as PointNet has also shown impressive potential. This approach better captures the intrinsic structure and features of the point cloud data, leading to more efficient compression. Furthermore, for LiDAR point cloud compression, researchers have ingeniously borrowed ideas from intelligent image coding and achieved remarkable results [6]. These innovative applications not only enhance the compression performance, but also facilitate real-time transmission and processing of point cloud data. In addition to the aforementioned applications, AI methods have also been extensively employed in the optimisation of entropy coding, post-processing and other modules within the traditional point cloud compression framework. These enhancements not only enhance the compression efficiency, but also improve the accuracy and visual effect of the reconstructed point cloud. Looking ahead, the application of AI in the field of point cloud compression will become more profound and extensive.

2. CURRENT CHALLENGES

Despite the development of a variety of applications based on point clouds in a number of scenarios, and the emergence of point cloud compression techniques that can be applied over short distances, the inherent characteristics of point clouds present significant challenges.

2.1. Sparsity

Due to the limitations of LiDAR sampling frequency, it is challenging to achieve uniform sampling in all directions when sampling objects. Furthermore, the distance, occlusion, reflections, and other factors contribute to the inhomogeneity of points sampled by LiDAR in practical applications. This results in high point cloud density in some areas, while there may be almost no point cloud in others. This presents significant challenges to 3D target detection.

2.2. Disorder

The generation of point cloud data typically occurs through the capture of raw depth information utilising sophisticated devices such as RGB-D sensors or LiDAR scanners. These devices are capable of measuring the distance from an object's surface to the sensor and combining it with other sensor data (e.g., RGB images) to construct 3D point-space coordinates. Following a series of data processing steps, including data fusion and noise filtering, the raw data is transformed into a high-quality point cloud product. These point cloud data are subsequently employed in a multitude of downstream tasks, including 3D reconstruction, object recognition, and scene understanding. The point cloud generated in these processes is a collection of points, lacking any inherent order between the points. Instead, it is simply a set of points in space, representing the original ordered arrangement of the original data as a collection of unordered points. The compression of the point cloud has introduced challenges in this regard.

2.3. Variegation

The diversity of the point cloud is reflected in two main aspects. Firstly, the point cloud includes more than just geometric coordinates, which are the most basic information. Secondly, it includes laser reflection intensity and colour information, which reflect the material and other characteristics of the object. This makes the point cloud data richer and more intuitive. Moreover, point cloud data can be utilised in a multitude of applications, including automated driving and robot navigation. This data can be employed for environment sensing and obstacle detection, thereby providing accurate 3D spatial information for automated driving systems. Additionally, point cloud data can be employed in augmented reality and virtual reality, whereby it can be used to build real 3D environments and integrate them with virtual objects, thus creating immersive and interactive experiences. Furthermore, point cloud data can be employed in the field of building design and management. The use of point cloud data in the design and planning of new buildings and the digital management and maintenance of existing buildings is becoming increasingly prevalent. The diverse range of applications for point cloud data, however, introduces a significant degree of variability into the data itself. This, in turn, necessitates the use of a multitude of compression methods to ensure that a single solution can adequately address the diverse demands of these applications.

3. TRADITIONAL ALGORITHM

3.1. Octree-based Point Cloud Compression

The core starting point of the technique is the recursive division of the 3D space into 8 sub-cubes, each of which is called a voxel unit (Voxel). When a voxel cell is occupied by a point in the point cloud data, the voxel cell is further subdivided into smaller 8 subvoxels. This process represents the spatial division information recursively by constructing an octree structure [7].

In an octree, each node represents a voxel unit, and if the density of point cloud data within that node exceeds a certain threshold, the node is subdivided into eight sub-nodes [8], each representing an eighth of the original node space. In this way, the octree is able to adaptively divide the space according to the distribution of the point cloud data, resulting in an efficient data representation.

Importantly, there is correlation between parent-child nodes and neighbouring nodes in the octree. This correlation can be used to make efficient predictions of voxel occupancy information. Through prediction, the coding algorithm is able to determine more accurately which voxels are occupied and which are vacant, leading to more efficient arithmetic coding. This prediction and arithmetic coding based approach can further reduce the data storage and transmission requirements and improve the compression efficiency of point cloud data. The representative algorithms are MPEG's G-PCC and V-PCC methods.

3.2. KD-tree Based Point Cloud Compression

The KD-tree based point cloud compression method represents an efficient solution for the rapid indexing and searching of point cloud data utilising the KD-tree data structure. The KD-tree enables the identification of similar points or regions, as well as the selection of representative points for compression, thereby reducing data redundancy. This method not only improves the compression efficiency, but also ensures the data quality. The flexibility of the KD tree allows the compression strategy to be adjusted according to the specific needs, which is applicable to various point cloud processing scenarios. Consequently, it can be considered a practical point cloud compression technique.

3.3. Mapping-based Point Cloud Compression

Another innovative approach to point cloud compression is to map the 3D point cloud into a 2D image, which can then be efficiently compressed using a mature video image encoder. The research of this approach focuses on the development of a lossless mapping and reconstruction process from 3D to 2D, as well as on ensuring that the mapped image exhibits good spatio-temporal correlation for video coding techniques. Currently, a leading mapping technique is to project the point cloud slice by slice onto a cube surface and convert it to a depth map for compression. This approach has been adopted by the MPEG V-PCC standard and is especially suitable for dense dynamic point clouds. For sparse LiDAR point clouds, the original depth image can also be compressed directly by mapping. This mapping strategy greatly improves the compression efficiency of point cloud data [9].

3.4. Point Cloud Compression Based on Attribute Coding

Once the geometric coding of the point cloud data has been completed, various transform and predictive coding methods can be employed for the further compression of the attribute information. Among the transform coding techniques, Region Adaptive Hierarchical Transform (RAHT) is an effective method, which is based on the octree model and performs a weighted wavelet transform on the attribute information according to the number of nodes to achieve efficient compression. Another transform coding method is the Graph Fourier Transform (GFT), which constructs the graph structure on the point cloud and converts the attribute information from the spatial domain to the frequency domain for decomposition. This is achieved by decomposing the graph Laplace matrix and constructing the transform using the eigenvectors, which improves the compression efficiency [10].

In contrast, predictive coding methods focus on compression using correlations within the point cloud. The point cloud is typically divided into distinct sections or layers, such as by grouping points according to their distance from one another. This enables the coding of each layer individually, with the attributes between layers then predicted to reduce the amount of redundant information.

In addition to the aforementioned techniques that encode attributes directly in 3D space, mapping-based coding schemes map the point cloud to a 2D image and then compress the attributes directly using a sophisticated image-video encoder. This approach makes attribute coding more efficient by simplifying the data structure and capitalising on the maturity and efficiency of image and video coding techniques.

4. ARTIFICIAL INTELLIGENCE IN POINT CLOUD COMPRESSION

4.1. Artificial Intelligence Applied to Image Compression

In the field of image and video compression, intelligent techniques have been extensively investigated and yielded significant outcomes [2,3,6]. The integration of neural network technology has not only facilitated the advancement of traditional coding frameworks, such as feature transformation, intra- and inter-frame prediction, entropy coding, and pre- and post-processing, but also enabled the development of an end-to-end full neural network framework. This technological innovation not only optimises compression efficiency, but also optimises for different evaluation metrics. Furthermore, image and video compression techniques can also be organically combined with computer vision tasks, such as target detection and recognition, to promote each other and show broader application prospects.

4.2. CNN-based Point Cloud Compression

The CNN-based point cloud compression technique employs a convolutional neural network (CNN) to achieve efficient compression of 3D point cloud data. The technique enables the CNN to learn the feature extraction and representation of point clouds, resulting in a high compression ratio and low distortion rate. Compared to traditional compression methods [11], CNN-based point cloud compression significantly reduces the bandwidth and storage space required for storage and transmission while maintaining data integrity and detail information. Furthermore, the technology is capable of GPU parallel processing, which enhances the compression efficiency and is well-suited to scenarios requiring high real-time performance, such as autonomous driving and virtual reality.

4.3. Point-based Point Cloud Geometry Compression Algorithm

Voxel-based compression algorithms are constrained by fixed accuracy, particularly when processing point cloud data with uneven and sparse density distributions. In contrast, neural networks that directly handle point sets [12], such as PointNet, offer novel solutions for geometric compression of point cloud data. These networks accept point cloud coordinate values as direct inputs, effectively resolving the issue of unordered arrangement of point clouds by pooling isosymmetric functions. In the decoding stage, the coordinate values of the point cloud are generated directly through network structures such as fully connected layers and optimised using point-to-point distance as a loss function. This point-based method is free from the limitations of voxelisation and can flexibly handle point cloud data with various distribution densities. Furthermore, it directly optimises for the objective evaluation metric of point-to-point distance, showing significant advantages and great potential. In the case of sparse small point cloud datasets (e.g., ShapeNet), the scheme has already demonstrated performance beyond that of the octree coding algorithm. Furthermore, the scheme is capable of being extended to the processing of large point cloud data through the strategy of chunked compression.

Nevertheless, despite the satisfactory performance of the scheme at low bit rates, there is still room for improvement in the decoder part of the point cloud generation network at present. Specifically, there are still challenges in high-precision decoding and reconstruction. Future research should therefore focus on optimising the decoder network structure and improving the decoding accuracy in order to better meet the needs of practical applications.

4.4. Artificial Intelligence Applied to LiDAR Point Cloud Compression

The extensive utilisation of LiDAR sensors in the domains of autonomous driving, robot navigation and 3D modelling has led to the emergence of compression techniques for sparse point cloud data collected by LiDAR as a subject of significant interest within the research community. Given that LiDAR point clouds are typically characterised by a high degree of sparsity in 3D space, traditional compression methods are unable to achieve efficient and lossless results.

The current research landscape can be broadly categorised into two main areas of focus. On the one hand, researchers have attempted to achieve efficient compression by compressing LiDAR data in its original 2D format, such as distance images, and using deep learning models such as recurrent neural networks (RNNs). This method can improve the compression efficiency to a certain extent by reducing the dimensionality of the data and combining the powerful learning ability of deep learning models [13].

On the other hand, the improved method based on octree encoder has also received wide attention. Researchers have optimised the entropy model through deep learning methods, which in turn improves the performance of entropy coding. This method makes full use of the spatial structure information of point cloud data to achieve more efficient compression. Furthermore, these improvements are not only applicable to static point cloud data, but also able to be extended to the compression of dynamic point cloud sequences. This provides the possibility of real-time applications such as autonomous driving.

These novel compression techniques not only enhance the compression efficiency of point cloud data, but also provide a superior quality database for subsequent point cloud processing and analysis. As research progresses, it is anticipated that more innovative methods will emerge in the future to further advance the development of LiDAR point cloud compression technology.

5. A RETROSPECTIVE AND PROSPECTIVE ANALYSIS

In general, after years of researchers studying various aspects of point cloud compression, the traditional approach has reached a point of maturity. Similarly, in the field of artificial intelligence, point cloud compression has also made significant progress. However, the research in this area is still in its infancy, with numerous avenues for future exploration. For instance, there is a need to develop point cloud attribute compression, sparse point cloud compression, large-scale point cloud compression, dynamic point cloud compression, and point cloud compression optimised for human eyes or machine perception. These needs are close to our lives, and new research is urgently needed to satisfy them. At the same time, it is possible to put literally as much research as possible into real production and life. The resolution of these issues necessitates the investigation of novel concepts and techniques, including the integration of existing deep learning-based tools.

For instance, the extension of existing deep learning-based geometric compression schemes to the compression of attributes is a promising avenue for further research. Similarly, the exploration of neural network structures with more expressive capabilities, the introduction of artificial intelligence approaches to projection-based compression, and the incorporation of quality evaluation methods more in line with human eye or machine perception as a loss function to be applied in the optimisation of end-to-end learning are all areas that warrant further investigation. Finally, the application of point cloud compression with other point cloud processing tasks such as scene segmentation and target detection represents a promising avenue for future research. It is our contention that with further research into these matters, the future prospects of AI application in point cloud compression will become increasingly promising.

REFERENCES

[1] MPEG 3DGC. (2020). G-PCC codec description v9. In International Organization for Standardization Meeting Document ISO/IEC JTC1/SC29/WG7 MPEG/N0011.

[2] MPEG 3DGC. (2020). V-PCC codec description v12. In International Organization for Standardization Meeting Document ISO/IEC JTC1/SC29/WG7 MPEG/N0012.

[3] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, 31(1).

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA. pp. 770-778.

[5] Guarda, A. F., Rodrigues, N. M., & Pereira, F. (2019). Point cloud coding: Adopting a deep learning-based approach. In 2019 Picture Coding Symposium (PCS). pp. 1-5.

[6] Wang, J., Ding, D., Li, Z., & Ma, Z. (2021). Multiscale point cloud geometry compression. In 2021 Data Compression Conference (DCC). pp. 73-82.

[7] Fu, C., Li, G., Song, R., Gao, W., & Liu, S. (2022) Octattention: Octree-based large-scale contexts model for point cloud compression." In Proceedings of the AAAl. 36(1).

[8] Chen, Z., Qian, Z., Wang, S., & Chen, Q. (2022). Point cloud compression with sibling context and surface priors. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), Computer Vision – ECCV 2022 . ECCV 2022. Lecture Notes in Computer Science, vol. 13698.

[9] He, Y., Ren, X., Tang, D., Zhang, Y., Xue, X., & Fu, Y. (2022). Density-preserving deep point cloud compression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2333-2342 .

[10] Zhang, J., Liu, G., Ding, D., & Ma, Z. (2022). Transformer and upsampling-based point cloud compression. In Proceedings of the 1st International Workshop on Advances in Point Cloud Compression, Processing, and Analysis. pp. 33-39.

[11] X. Li, O. A. Hanna, C. Fragouli, S. Diggavi, G. Verma and J. Bhattacharyya. (2023) "Feature Compression for Multimodal Multi-Object Tracking," MILCOM 2023 - 2023 IEEE Military Communications Conference (MILCOM). Boston, MA, USA. pp. 139-143.

[12] Z. Yu, L. Zhang, S. Li, Y. Zhang and X. Ning. (2021) "2D-3DMatchingNet: Multimodal Point Completion with 2D Geometry Matching," 2021 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS). Macau, China. pp. 94-99.

[13] Charles, R. Q., Su, H., Mo, K., & Guibas, L. J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. pp. 77-85.