International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 18th Dec 2024. Link

https://ijiemr.org/downloads.php?vol=Volume-13&issue= Issue12

DOI:10.48047/IJIEMR/V13/ISSUE 12/31

Title: " ADVANCED 3D MOTION PREDICTION FOR VIDEO BASED DEEP DYNAMIC POINT CLOUD COMPRESSION"

Volume 13, ISSUE 12, Pages: 248- 253

Paper Authors
Mr. A.Shiva Prasad, Achagatla Sri Sai Shri Valli,
Nidamanuri Jeshnavi Lakshmi, Pettemu Meghana

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

ADVANCED 3D MOTION PREDICTION FOR VIDEO BASED DEEP DYNAMIC POINT CLOUD COMPRESSION

Mr. A.Shiva Prasad[1], Achagatla Sri Sai Shri Valli[2] , Nidamanuri Jeshnavi Lakshmi[3], Pettemu Meghana[4]

[1,2,3,4]Department of ECE, CMR Institute of Technology, Medchal, Hyderabad, Telangana, India.
21r01a04a1@cmritonline.ac.in

Abstract - The non-uniformly distributed nature of the 3D dynamic point cloud (DPC) brings significant challenges to its high-efficient inter-frame compression. This paper proposes a novel 3D sparse convolution-based Deep Dynamic Point Cloud Compression(D-DPCC) network to compensate and compress the DPC geometry with 3D motion estimation and motion compensation in the feature space. In the proposed D-DPCC network, we design a Multiscale Motion Fusion (MMF) module to accurately estimate the 3D optical flow between the feature representations of adjacent point cloud frames. Specifically, we utilize a 3D sparse convolution based encoder to obtain the latent representation for motion estimation in the feature space and introduce the proposed MMF module for fused 3Dmotion embedding. Besides, for motion compensation, we propose a 3D Adaptively Weighted Interpolation (3DAWI) algorithm with a penalty coefficient to adaptively decrease the impact of distant neighbours. To our knowledge, this paper is the first work proposing an end-to-end deep dynamic point cloud compression framework. The experimental result shows that the proposed D-DPCC frame work.

Keywords: Auxiliary information, Feature Extraction, Inter Prediction, Residual Compression, point cloud Reconstruction.

I. Introduction

In recent years, the dynamic point cloud (DPC) has become a promising data format for representing sequences of 3D objects and scenes, with broad applications in AR/VR, autonomous driving, and robotic sensing . However, compared with pixelized 2D image/video, the nonuniform distribution of DPC makes the exploration of temporal correlation extremely difficult, bringing significant challenges to its inter-frame compression. This paper focus on the dynamic point cloud geometry compression with 3D motion estimation.The existing dynamic point cloud compression (DPCC) methods can be concluded as 2D video based and 3D-model based methods. The 3D-model-based methods rely on motion estimation and motion compensation on 3D volumetric models like octree [8]. However, the above methods are rule based, with hand-crafted feature extraction modules and assumption-based matching rules, resulting in unsatisfactory coding efficiency [9]. To this end, we propose a Deep Dynamic Point Cloud Compression (D-DPCC) framework, which optimizes the motion estimation, motion compensation, motion compression, and residual compression module in an end-to-end manner [10]. Our contributions are summarized as follows:

1. We first propose an end-to-end Deep Dynamic Point Cloud Compression framework (D-DPCC)

for the joint optimization of motion estimation, motion compensation, motion compression, and residual compression.

2. We propose a novel Multi-scale Motion Fusion (MMF)module for point cloud inter-frame prediction, which extracts and fuses the motion flow information at different scales for accurate motion estimation. For motion compensation, we propose a novel 3D Adaptively Weighted Interpolation (3DAWI) algorithm which utilizes the neighbor information and adaptively decreases the impact of distant neighbors to produce a point-wise prediction of the current frame's.

II. Literature review

Dynamic Point Cloud Compression. Existing DPCC works can be mainly summarized as 2D-video-based and 3D-model-based methods [1]. The 2D-video-based methods perform 3D-to-2D projection to utilize the 2D motion estimation (ME) and motion compensation (MC) algorithms. However, the 3D-model-based methods fail to fully exploit the temporal correlations, leading to inferior results than 2D-video-based methods [2]. Learnt Static Point Cloud Compression. Inspired by the implementation of deep learning techniques in image/video compression, recently learnt static point cloud compression (SPCC) networks have emerged [4]. However, these works are limited to small-scale point cloud datasets with only thousands of points. Thanks to the development of 3D sparse convolution propose an end-to-end sparse convolution-based multi-scale SPCC network, which reports state-of-the-art performance on large-scale data sets [3].

Point Cloud Scene Flow Estimation. Recently, 3D scene flow estimation has received much attention, which inspires us to compress DPC geometry by 3D motion estimation. FlowNet3D is the representative work of 3D scene flow estimation, which integrates a set conv layer based on Point Net++ for feature extraction, followed by a flow embedding layer for point matching [5]. However, the above methods are trained on manually labelled scene flow data [5]. Learnt Video Compression. There are massive attempts to apply deep learning techniques to video compression and build an end-to-end framework [6]. Among them, DVC is the first video compression deep model that jointly optimizes all the components for video compression. For more accurate motion estimation, further proposes FVC by performing inter-frame prediction in the feature space, reporting superior performance over DVC [7].

III. System Model

1.Video Acquisition Module: Video acquisition is the initial step in the pipeline for advanced 3D motion prediction and compression of dynamic point clouds. This step involves capturing raw data from the physical environment. It involves using specialized 3D capture devices, ensuring synchronization of data from multiple sensors, and preprocessing the raw data to enhance its quality. Proper acquisition and preprocessing of video data are essential for accurate feature extraction, motion estimation, and compression, ultimately leading to high-quality reconstruction and efficient transmission of 3D dynamic scenes.

2.Feature Extraction: The feature extraction module consists of two serially connected Down sample Blocks for the hierarchical reduction of spatial redundancies, which encodes the current frame x_t and the previously reconstructed frame $\hat{x}_{t-1}$ as latent representation y_t and $\hat{y}_{t-1}$.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Inspired by Wang et al, we adopt the sparse CNN-based Down sample Block for low-complexity point cloud down sampling. The Down sample Block consists of a stride-two sparse convolution layer for point cloud.

3.Inter Prediction: The inter prediction module takes the latent representation of both the current frame and the previously reconstructed frame, i.e., yt and yˆt−1 as input, analysing the temporal correlation and producing the feature prediction of yt, i.e., y¯t. The existing 3D scene flow estimation networks estimate a motion flow between two frames p1 and p2to minimize the distortion between D and the ground truth D∗ . However, without end-to-end optimization, the predicted motion flow is unsatisfactory for motion compensation and motion compression.

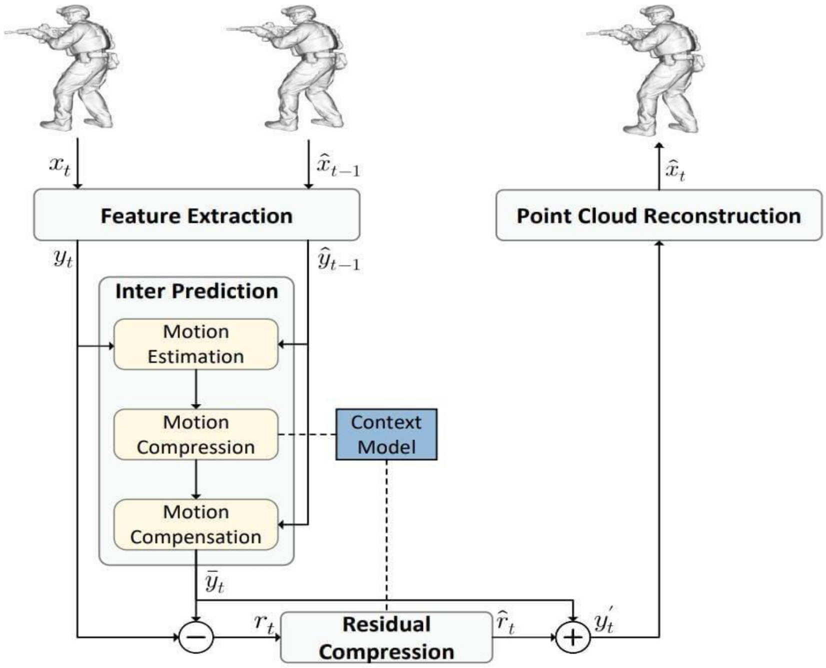

Figure 1: The overall architecture of D-DPCC. xt and xˆt−1 are the current frame and the previously reconstructed frame. yt and yˆt−1are the associated latent representations in feature space. y¯t is the prediction of yt. rt and rˆt are the feature residual and reconstructed residual. y′t is the reconstruction value of yt in decoder.

i.Motion Estimation: This step determines the movement of points between consecutive frames. Accurate motion estimation is crucial for effective compression, as it allows for the prediction of future frames based on past data.

Block Matching-Search Window: Define a search area in the previous frame to find the best match for each block. Cross Correlation (NCC) to evaluate matches.

ii.Motion Compression: It Compresses the motion information derived from motion estimation to reduce its size while retaining accuracy.

Fixed Quantization: Apply a fixed quantization step size to reduce precision.

Adaptive Quantization: Adjust the quantization step size based on the importance variability of the motion data.

iii. Motion Compensation: It applies the motion vectors to the predicted frame to align it with the current frame, reducing prediction errors and improving the quality of the predicted data.
Apply Motion Vectors: Point Displacement and Interpolation.

4.Residual Compression: It Compresses the residual data, which represents the difference between the predicted frame and the actual frame. To efficiently encode the residual data, which captures details not accurately predicted, minimizing the data required to represent the entire point cloud sequence.

5.Point Cloud Compression: It reconstructs the original point cloud data from the compressed information during the decompression phase with minimal loss of quality from the compressed bitstream, ensuring high fidelity to the original data.

IV. Result

We compare the proposed D-DPCC with the current state-of-the-art dynamic point cloud geometry compression framework: V-PCC Test Model v13, with the quantization parameter (QP) setting as 18, 15, 12, 10, 8, respectively. We also compare with Wang's framework which is state-of-the-art on static object point cloud geometry compression. For the fairness of comparison, we retrain Wang's framework using our training data and strategy, and the network parameters of each module are set the same as Wang's except for the proposed inter prediction module. When using the proposed D-DPCC for interframe coding, the first frame of the sequence is encoded using Wang's network with the same λ.

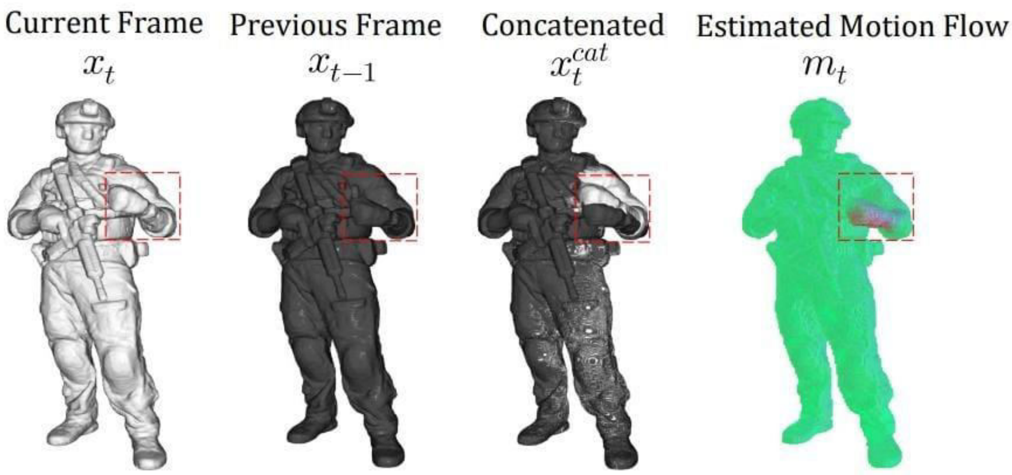

Figure 2: Visualization of the estimated motion flow. The leftmost column is the current frames; the left middle column is the previous frames; the middle right column is the concatenations of the current frames and the previous frames. The rightmost column visualizes the predicted motion flow, with green indicating stillness and purple indicating movement.

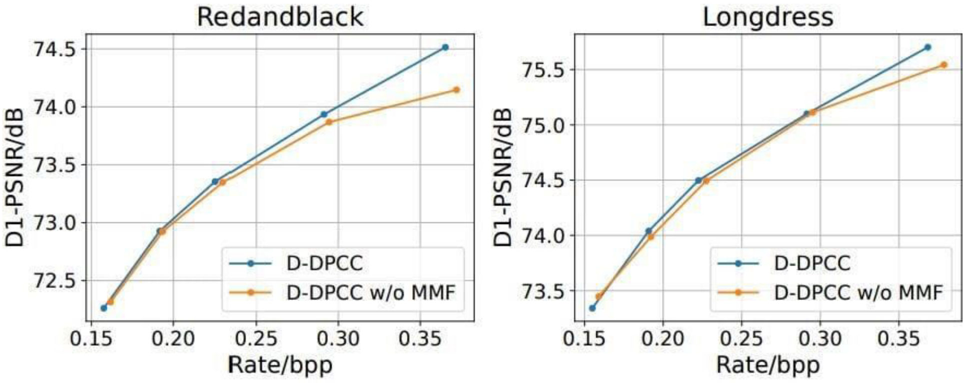

Figure 3: Ablation study on Red and black and Long dress for Multi-scale Motion Fusion (MMF) module.

V. Conclusion:

This paper proposes a Deep Dynamic Point Cloud Compression (D-DPCC) framework for the compression of dynamic point cloud geometry sequence. We introduce an inter prediction module to reduce the temporal redundancies. We also propose a Multi-scale Motion Fusion (MMF) module to extract the multi-scale motion flow. For motion compensation, a 3D Adaptively Weighted Interpolation (3DAWI) algorithm is introduced. The proposed D-DPCC achieves 76.66% BD-Rate gain against state-of-the-art V-PCC Test Model v13. information to derive a better motion vector predictor to deal with the patch inconsistency problem for both the geometry and attribute videos under the MPEG-I video-based point cloud compression (V-PCC) framework. In essence, we use the geometry or auxiliary information to derive the geometry of the current prediction unit to find the corresponding block in the reference frame. We provide both normative and non-normative solutions to adapt the proposed algorithm to different use cases. The experimental results show that the proposed solutions can provide both significant objective quality and subjective quality improvements. The experimental results obviously demonstrate the effectiveness of the proposed algorithm. In the future, we will further dive into finding better video compression algorithms for the V-PCC framework.

References

1. K. Radhakrishna, D. Satyaraj, H. Kantari, V. Srividhya, R. Tharun and S. Srinivasan, "Neural Touch for Enhanced Wearable Haptics with Recurrent Neural Network and

IoT-Enabled Tactile Experiences," *2024 3rd International Conference for Innovation in Technology (INOCON)*, Bangalore, India, 2024, pp. 1-6,

2. Karne, R. K., & Sreeja, T. K. (2023, November). Cluster based vanet communication for reliable data transmission. In *AIP Conference Proceedings* (Vol. 2587, No. 1). AIP Publishing.
3. Karne, R., & Sreeja, T. K. (2023). Clustering algorithms and comparisons in vehicular ad hoc networks. *Mesopotamian Journal of Computer Science*, *2023*, 115-123.
4. Karne, R. K., & Sreeja, T. K. (2023). PMLC-Predictions of Mobility and Transmission in a Lane-Based Cluster VANET Validated on Machine Learning. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*, 477-483.
5. Mohandas, R., Sivapriya, N., Rao, A. S., Radhakrishna, K., & Sahaai, M. B. (2023, February). Development of machine learning framework for the protection of IoT devices. In *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1394-1398). IEEE.
6. Kumar, A. A., & Karne, R. K. (2022). IIoT-IDS network using inception CNN model. *Journal of Trends in Computer Science and Smart Technology*, *4*(3), 126-138.
7. Karne, R., & Sreeja, T. K. (2022). Routing protocols in vehicular adhoc networks (VANETs). *International Journal of Early Childhood*, *14*(03), 2022.
8. Karne, R. K., & Sreeja, T. K. (2022). A Novel Approach for Dynamic Stable Clustering in VANET Using Deep Learning (LSTM) Model. *IJEER*, *10*(4), 1092-1098.
9. RadhaKrishna Karne, D. T. (2021). COINV-Chances and Obstacles Interpretation to Carry new approaches in the VANET Communications. *Design Engineering*, 10346-10361.
10. RadhaKrishna Karne, D. T. (2021). Review on vanet architecture and applications. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(4), 1745-1749.