Compression Methods of NeRF Model Implicitly Representing 3D Video Using NNC

Dong-Ha Kim, Jun Young Jeong, Gwangsoon Lee, and Jae-Gon Kim

***Abstract*— Implicit Neural Representation (INR), which parametrizes various types of signals into a single function through Neural Networks, has been actively researched. A Neural Radiance Field (NeRF) model utilizes the INR approach to represent a 3D space. MPEG established the ad-hoc called INVR (Implicit Neural Visual Representation) and is currently exploring the potential standardization of 3D video compression using NeRF-based techniques. In this paper, we proposed two compression methods of NeRF model using Neural Network Compression (NNC) codec.**

I. Introduction

Recently, the Implicit Neural Representation (INR) method, which parametrizes various types of signals into a single function through neural networks, has been actively researched. A Neural Radiance Field (NeRF) model [1] utilizes the INR approach to represent a 3D space. Multiple views captured from various positions are used in the training stage, and the trained model can synthesize views from any positions within the 3D space. Recognizing that the representation of new time-series data could serve as a novel compression approach [2], the Moving Picture Expert Groups (MPEG) established the ad-hoc called Implicit Neural Visual Representation (INVR) and is currently exploring the potential standardization of 3D video compression using NeRF-based techniques.

In this paper, we proposed two compression methods of NeRF model using Neural Network Compression (NNC) [3].

II. NeRF Model Compression

The NeRF model trained on a given multi-view video can synthesize and render arbitrary viewpoints. The NeRF model consists of four parameter parts, among which the parameters for the coarse and fine networks are essential for inference. Therefore, the compression is applied to only two essential parameter parts. In the first NeRF model compression method involves using the same Quantization Parameter (QP) to compress the two essential parameters with NNC [4]. NNC includes only quantization and entropy coding, excluding preprocessing steps such as pruning.

The coarse and fine networks utilize the same multi-layer perceptron (MLP) architecture and exhibit similar data sizes. However, between the two parameters, the fine network has a more crucial role in inference compared to the coarse network. Therefore, by allocating the given bitrate differently based on the significance of each network, it is possible to compress the NeRF model with enhanced rendering performance while maintaining the target bitrate. In other words, compressing the fine network with a smaller QP and compressing the coarse network with a higher QP in NNC can maintain better rendering performance at the same compression bitrate. That is, the second method compress the NeRF model with network-adaptive bit allocation.

NeRF is trained using the 'mirror' sequence from the Test Model MPEG Immersive Video (TMIV) dataset [5]. The model parameters are then compressed using NNC through the two methods. Performance comparison is based on Peak Signal-to-Noise Ratio (PSNR) versus Bits Per Pixel (BPP), calculated by dividing compressed the bitstream data size by the number of total rendered pixels. TMIV Common Test Conditions (CTCs) [5] uses 5 predefined QPs for compression and measures PSNR quality at each bit rate. Figure 1 shows the experimental results. Applying different QPs to compress the two parameters using NNC get higher PSNR compared to compressing with the same QP. And at some QP points, better performance is confirmed than TMIV.

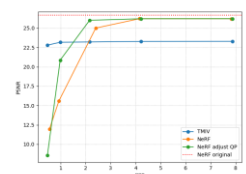

Figure 1. Comparison of NeRF compression performance.

Acknowledgment

This work was supported by IITP grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory).

References

[1] Mildenhall, Ben, et al., "Nerf: Representing scenes as neural radiance fields for view synthesis," in Proc. European Conf. Computer Vision (ECCV 2020), Aug. 2020.

[2] G. Bang, Y. Liao, and P. Hellier, "Exploration experiments on implicit neural visual representation," ISO/IEC JTC1/SC29/WG04, N0302, Jan. 2023.

[3] D. Kim, G. Han, J. Kim, J. Jeong, and G. Lee, "[INVR] EE 2.3-related: Results of Static NeRF Model Compression," ISO/IEC JTC1/SC29/WG04, M63089, Apr. 2023.

[4] MPEG Liaison and communication, "White paper on neural network coding," ISO/IEC JTC1/SC29/AG3, N0057, Jan. 2022.

[5] J. Jung and B. Kroon, "Common Test Conditions for MPEG Immersive Video," ISO/IEC JTC1/SC29/WG4, N0232, Jul. 2022.

Dong-Ha Kim is with the Electronics and Information Engineering Department, Korea Aerospace University, Goyang, Korea (e-mail: donghakim@kau.kr)

Jun young Jeong and Gwangsoon Lee are with the Immersive Media Research Laboratory, Electronics and Telecommunications Research Institute, Daejeon, Korea (e-mail: {jyj0120, gslee}@etri.re.kr)

Jae-Gon Kim is a professor with the Electronics and Information Engineering Department, Korea Aerospace University, Goyang, Korea (corresponding author to provide phone: +82-2-300-0414, e-mail: jgkim@kau.ac.kr)