MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

IMPLICIT NEURAL REPRESENTATION FOR LOW-OVERHEAD GRAPH-BASED HOLOGRAPHIC-TYPE COMMUNICATIONS

Fujihashi, Takuya; Kato, Sorachi; Koike-Akino, Toshiaki

TR2024-022 March 19, 2024

Abstract

Point cloud delivery over wireless and mobile channels will be a key technology for untethered users to realize extended reality via wire- less and mobile terminals. A key challenge of point cloud delivery is efficiently delivering the point cloud over unstable and band-limited channels. Graph Fourier Transform (GFT) is a potential solution to compress such non-uniformly and non-orderly distributed signals in a 3D space, whereas GFT-based solutions require large communication overhead to share the graph information with the receiver. This paper proposes a novel point cloud delivery scheme that introduces implicit neural representation (INR) to reduce the overhead. Specifically, the INR of the proposed scheme trains a mapping between the indices and the corresponding weight in the adjacency matrix and the proposed scheme sends the parameter set of the INR as the metadata. Evaluations demonstrate that the proposed scheme can improve the point cloud quality under the same amount of communication over- head because the proposed INR can predict most of the elements in the adjacency matrix using a small parameter set.

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2024

IMPLICIT NEURAL REPRESENTATION FOR LOW-OVERHEAD GRAPH-BASED HOLOGRAPHIC-TYPE COMMUNICATIONS

Takuya Fujihashi[1], *Sorachi Kato*[1], *Toshiaki Koike-Akino*[2]

[1]Graduate School of Information Science and Technology, Osaka University, Suita, Osaka, Japan
[2]Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA

ABSTRACT

Point cloud delivery over wireless and mobile channels will be a key technology for untethered users to realize extended reality via wireless and mobile terminals. A key challenge of point cloud delivery is efficiently delivering the point cloud over unstable and band-limited channels. Graph Fourier Transform (GFT) is a potential solution to compress such non-uniformly and non-orderly distributed signals in a 3D space, whereas GFT-based solutions require large communication overhead to share the graph information with the receiver. This paper proposes a novel point cloud delivery scheme that introduces implicit neural representation (INR) to reduce the overhead. Specifically, the INR of the proposed scheme trains a mapping between the indices and the corresponding weight in the adjacency matrix and the proposed scheme sends the parameter set of the INR as the metadata. Evaluations demonstrate that the proposed scheme can improve the point cloud quality under the same amount of communication overhead because the proposed INR can predict most of the elements in the adjacency matrix using a small parameter set.

1. INTRODUCTION

Point cloud [1] is a typical format to represent 3D scenes and objects using numerous 3D points distributed in the 3D space. Sending high-quality point clouds over wireless and mobile channels is a key technology to realize holographic-type communications. One key issue is efficiently delivering the point cloud over band-limited and unreliable channels. Since each point cloud consists of many 3D points for 3D scene representation, it requires large traffic. Such significant traffic causes low reconstruction quality of the point cloud in band-limited wireless channels. Graph signal processing (GSP) [2, 3] is one of the typical solutions to compress the point cloud. GSP-based solutions regard the point cloud signal as the graph signal and define the graph basis matrix using the adjacency matrix based on the edge between each 3D point. The derived graph basis matrix is used to carry out graph Fourier transform (GFT) to map the graph signals onto frequency representations. Integrating GFT, quantization, and entropy coding reduces the required traffic to deliver the point cloud.

However, the integration causes catastrophic wireless point cloud delivery errors due to unreliable wireless channels and unrecoverable compression errors. In wireless channels, channel quality between the sender and receiver is time-varying. Suppose the wireless channel quality drops a certain threshold. In that case, bit errors occur in the transmitted bitstream even with channel coding, and even a single bit error may fail the entropy decoding [4]. The decoding failure causes significant quality degradation, referred to as the cliff effect. On the other hand, the point cloud quality will be leveling-off even when the wireless channel quality improves during the delivery. This is because the quantization noise does not recover at the receiver, irrespective of wireless channel quality.

In addition, the GSP-based solutions may cause a significant communication overhead for signal decoding. The adjacency matrix for GFT is signal-dependent [5, 6], and the sender needs to send the adjacency matrix to carry out inverse GFT at the receiver. Specifically, the sender needs to send M^2 real values for the point cloud with M points since the adjacency matrix is $M \times M$ matrix.

This paper proposes a novel wireless point cloud delivery scheme to solve the existing graph-based point cloud delivery. The proposed scheme consists of GFT coefficient and metadata transmission methods to solve the above-mentioned issues. For the GFT coefficient transmission, the proposed scheme skips quantization and entropy coding for the GFT coefficients and directly maps the GFT coefficients onto the transmission symbols inspired by joint source-channel coding studies [4, 7–9]. By skipping the quantization and entropy coding, the proposed scheme reduces the cliff effect and leveling-off effect even with the fluctuation of wireless channel quality during point cloud delivery.

For metadata transmission, the proposed scheme utilizes an implicit neural representation (INR) to transmit the adjacency matrix with low communication overhead. INR [10–13] is a novel memory-efficient format for representing multidimensional signals. Our INR uses a small Multi-Layer Perceptron (MLP)-based Neural Network (NN) architecture to train the mapping between the indices of each element in the adjacency matrix and the corresponding edge weight. The weights of the trained INR are then transmitted to the receiver as the metadata for decoding the adjacency matrix at the receiver.

Evaluations show that the proposed INR-based metadata transmission significantly reduces the required communication overhead to decode the adjacency matrix and the corresponding graph basis matrix with a slight error on the receiver side. The decoded adjacency matrix will contribute to reconstructing a better-quality point cloud than the other graph-based delivery schemes under the same wireless channel quality.

Related Work and Our Contributions 3D points in each point cloud are non-uniformly distributed in 3D space to represent various 3D scenes and objects. Some studies [14–22] utilized GSP to compress and deliver the 3D coordinate and color attributes over band-limited and unreliable channels. They regarded the 3D points as a graph signal and defined the graph basis matrix or the graph-based NN to compress the graph signal. For example, HoloCast in [17] compacts the 3D coordinates and color components using GFT and maps the GFT coefficients onto the transmission symbols to prevent the above-mentioned cliff and leveling-off effects. A key issue to utilizing the graph basis matrix for compression is to re-

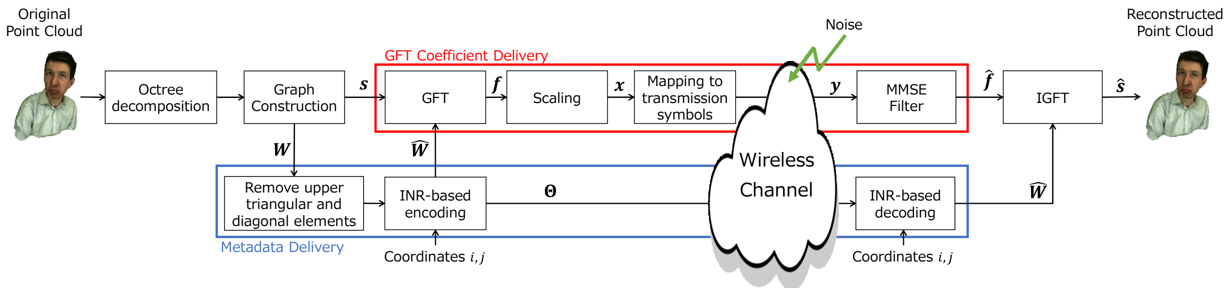

Fig. 1. End-to-end architecture of proposed graph-based point cloud delivery scheme.

construct the graph basis matrix at the receiver side. The typical solutions are 1) to define the graph basis matrix based on the 3D coordinates and transmit the original/compressed 3D coordinates as metadata [15, 16, 20, 21], or 2) to transmit the defined graph basis matrix as the metadata using the combination of Givens rotation and uniform/non-uniform quantization [8, 23]. However, both solutions still cause large communication overhead to send high-quality point clouds.

Our study aims to prevent the cliff and leveling-off effects for future holographic-type communication. Specifically, we define the adjacency matrix based on the bilateral Gaussian kernel for the 3D coordinate and color components and the graph basis matrix according to the random-walk graph shift operator to efficiently compress the 3D coordinate and color component signals. The main contribution of this paper is first to introduce the concepts of INR-based compression [24, 25] for overhead reduction in graph-based point cloud delivery. Unlike the above solutions, the perceptrons of our INR overfit the target adjacency matrix using a small NN architecture. However, the straightforward way that overfits the full-size adjacency matrix requires a large NN architecture and, thus large communication overhead. Since the adjacency matrix is symmetric and the diagonal elements are one, the proposed scheme considers the upper triangular and diagonal elements of the adjacency matrix as zero, and the INR aims to overfit the lower triangular elements. This modification allows the lower triangular elements to be regressed using an NN architecture, even with a single hidden layer.

2. PROPOSED SCHEME

2.1. Overview

Fig. 1 shows the overview of the proposed scheme. The proposed scheme consists of GFT coefficient and basis matrix transmission parts. The proposed scheme uses octree decomposition to decompose 3D points into multiple octree blocks. The proposed scheme constructs an undirected and weighted graph for each octree block. The following operations are performed in each octree block. Therefore, the block index b is omitted for simplicity. Based on the constructed graph, it then makes the adjacency matrix using a bilateral Gaussian kernel [5]. The adjacency matrix is then fed into the metadata transmission part. The metadata transmission part considers an MLP-based NN architecture based on sinusoidal representation networks (SIREN) [11] as an INR, and the neurons overfit the lower triangular elements of the adjacency matrix. After overfitting, each lower triangular element can be decoded by inputting the corresponding index into the INR, and the decoded elements are assigned to the upper triangular elements since the adjacency matrix is symmetric. The GFT coefficient transmission part uses the decoded adjacency matrix to derive the graph basis matrix to convert the attributes of the point cloud into frequency representations. Each frequency representation is unequally scaled for error protection to minimize distortion due to channel noise. Finally, two-by-two scaled coefficients are mapped to I (in-phase) and Q (quadrature-phase) components for coefficient transmission. On the other hand, the metadata transmission part sends the weights of the overfitted INR as the metadata.

The receiver side regards the received symbols as the received GFT coefficients and removes the channel noise using a filter that minimizes the error between the original and decoded GFT coefficients. At the same time, the adjacency matrix is decoded from the received weights of the INR, and the corresponding graph basis matrix is obtained from the decoded adjacency matrix. Finally, the receiver can reconstruct the 3D coordinates and color components from the denoised GFT coefficients by taking inverse GFT.

2.2. Graph Construction

We consider that each 3D point comprises the attributes of the 3D coordinates and color components, and the 3D points have already been organized into octree blocks where each octree block may contain N 3D points. The 3D points in each octree block can be regarded as a weighted and undirected graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{W})$, where $\boldsymbol{V}$ is the vertex set, $\boldsymbol{E}$ is the edge set, and $\boldsymbol{W} \in \mathbb{R}^{N\times N}$ is an adjacency matrix. Each vertex has the 3D coordinates $\boldsymbol{p} = [\mathrm{x}, \mathrm{y}, \mathrm{z}]^{\mathrm{T}} \in \mathbb{R}^{3\times N}$ and the corresponding color components $\boldsymbol{c} = [\mathrm{y}, \mathrm{u}, \mathrm{v}]^{\mathrm{T}} \in \mathbb{R}^{3\times N}$. In addition, an element of the adjacency matrix $w_{i,j}$ represents the edge weight between vertices i and j. We use the bilateral Gaussian kernel for the adjacency matrix [5] as follows:

$$\boldsymbol{W}_{i,j} = \exp\left(-\left(\frac{\|\boldsymbol{p}_i - \boldsymbol{p}_j\|_2^2}{\epsilon_p} + \frac{\|\boldsymbol{c}_i - \boldsymbol{c}_j\|_2^2}{\epsilon_c}\right)\right), \tag{1}$$

where ϵ_p and ϵ_c are the sample variance of 3D coordinates and color components, respectively.

2.3. GFT Coefficient Transmission

The 3D coordinates and color components are then transformed into frequency representations using GFT. The GFT is defined through eigenvalue decomposition or singular value decomposition of the random-walk graph shift operator $\boldsymbol{L} \in \mathbb{R}^{N\times N}$ [26] based on adjacency matrix $\boldsymbol{W}$ and degree matrix $\boldsymbol{D}$ as follows:

$$\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{W}, \tag{2}$$

where $\boldsymbol{I}$ is an identity matrix of proper dimension and $\boldsymbol{D}$ is the diagonal degree matrix which is represented as:

$$\boldsymbol{D} = \mathrm{diag}(D_1, \ldots, D_N),\ D_i = \sum_{n=1}^{N} W_{i,n}. \tag{3}$$

The graph basis matrix is then derived from the graph shift operator using the singular value decomposition as:

$$\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T. \tag{4}$$

Here, $\boldsymbol{\Sigma}$ is a diagonal matrix of the singular values arranged in ascending order, and $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ is a unitary matrix that comprises the right singular vectors, i.e., graph basis matrix. The GFT coefficients of each attribute $\boldsymbol{f} \in \mathbb{R}^{1\times N}$ are obtained by multiplying the graph basis matrix by the attribute signals $\boldsymbol{s} \in \mathbb{R}^{1\times N}$ as:

$$\boldsymbol{f} = \boldsymbol{s}\ \boldsymbol{V}^T. \tag{5}$$

The proposed scheme maps the GFT coefficients onto the transmission symbols without quantization and entropy coding. However, such GFT coefficient transmission without error protection will cause significant quality degradation due to channel noise.

The proposed scheme scales the GFT coefficients before mapping to realize error protection against channel noise. Here, the scaling factor for each GFT coefficient is unequal to minimize the mean-square error (MSE) between the original and received GFT coefficients. Let x_i denote the ith transmission symbol, which is the ith GFT coefficient f_i of an attribute scaled by a factor of g_i for noise protection as:

$$x_i = g_i \cdot f_i. \tag{6}$$

The optimal scale factor g_i is obtained by minimizing the MSE under the power constraint with an average power budget of P. The optimal solution [27] is expressed as:

$$g_i = \lambda_i^{-1/4}\sqrt{\frac{NP}{\sum_j^N \sqrt{\lambda_j}}}, \tag{7}$$

where λ_i is the power of the ith GFT coefficient.

The transmitted symbols are impaired via wireless links. Let y_i denote the ith received symbol and n_i denote an effective additive white Gaussian noise (AWGN) with a variance of σ^2. The received symbol y_i over wireless links can be modeled as follows:

$$y_i = x_i + n_i. \tag{8}$$

The GFT coefficients are then extracted from the I and Q components through a minimum MSE (MMSE) filter [27]:

$$\hat{f}_i = \frac{g_i\lambda_i}{g_i^2\lambda_i + \sigma^2} \cdot y_i, \tag{9}$$

where $\hat{f}_i$ is a receiver estimate of i-th transmitted GFT coefficient. The decoder finally reconstructs each attribute of the point cloud $\hat{\boldsymbol{s}} \in \mathbb{R}^{1\times N}$ by taking IGFT for the extracted GFT coefficients as:

$$\hat{\boldsymbol{s}} = \hat{\boldsymbol{f}}\,\boldsymbol{V}. \tag{10}$$

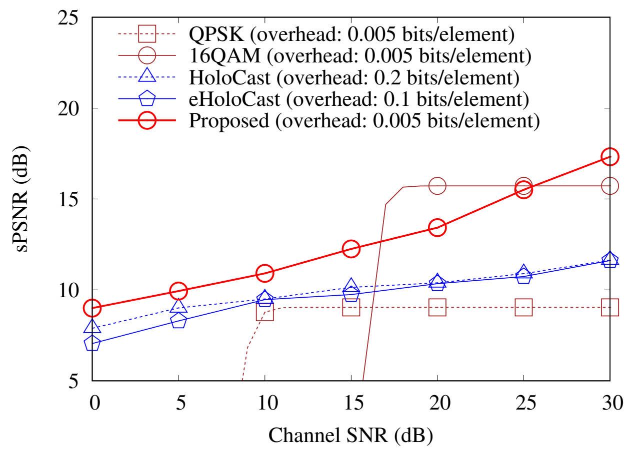

Fig. 2. Symmetric PSNR performance of the proposed and baseline schemes as a function of wireless channel SNRs under certain communication overhead.

2.4. INR-based Adjacency Matrix Compression

The proposed scheme designs the INR to represent the adjacency matrix $\boldsymbol{W}$ with few bits. The proposed INR Φ_Θ: $\mathbb{R}^2 \to \mathbb{R}^1$ is an MLP-based NN architecture with a set of parameters Θ. The input to the proposed INR is the coordinates of the adjacency matrix $\boldsymbol{\rho}_{\boldsymbol{W}_{i,j}} = [i, j]$. The proposed INR comprises L fully connected (FC) layers, each with 28 neurons. The activation functions for the FC layers are sine functions based on SIREN and the final FC layer outputs the edge weight of the coordinate $\Phi_\Theta(\boldsymbol{\rho}_{\boldsymbol{W}_{i,j}}) = [\hat{\boldsymbol{W}}_{i,j}]$ with an activation function of the identity function. For example, the amount of metadata for the parameter set Θ at $L = 1$ is approximately 4 Kbits.

The proposed INR aims to reconstruct each coordinate's edge weight as close to the original edge weight as possible. Here, the reconstruction quality of the INR may depend on the number of the non-zero elements of the target signal and loss function. The proposed scheme assigns zero to the upper triangular and diagonal elements of the adjacency matrix $\boldsymbol{W}$ because the adjacency matrix $\boldsymbol{W}$ is a symmetric matrix. From preliminary evaluations, this modification improves the reconstruction quality of the edge weight even with a small number of FC layers. For the loss function, the proposed INR uses the normalized MSE (NMSE) to optimize the parameter set of the INR Θ to minimize the error between the original edge weights $\boldsymbol{W}_{i,j}$ and the reconstructed ones $\Phi_\Theta(\boldsymbol{\rho}_{\boldsymbol{W}_{i,j}})$ over all coordinates.

3. EVALUATIONS

Performance Metric in 3D Point Cloud: We evaluate the 3D reconstruction quality of point cloud delivery in terms of the symmetric peak signal-to-noise ratio (sPSNR) [18] in color components $\boldsymbol{c}$.
Point Cloud Dataset: We use publicly available point cloud data, namely, *pencil_10_0* whose number of points M is 2,731.

3.1. Baseline Performance

We consider four baselines of the GSP-based joint source-channel coding for comparison. The first and second baselines are graph-based schemes with quantization and entropy coding [15, 16]. Specifically, the 3D coordinate attributes are compressed using

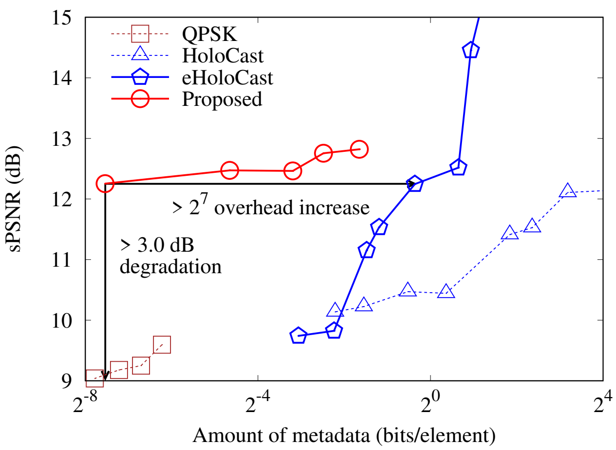

Fig. 3. Symmetric PSNR performance of the proposed and baseline schemes under the different communication overheads. Here, the wireless channel SNR between the sender and receiver is fixed to 15 dB.

the octree-based compression [28, 29], and the compressed 3D coordinates are used to construct the graph basis matrix to decorrelate the color components. The decorrelated color components are uniformly quantized and entropy-coded for compression and modulated using the digital modulation formats of Quadrature Phase-Shift Keying (QPSK) and 16 Quadrature Amplitude Modulation (16QAM) for transmission. Here, the compressed 3D coordinates are transmitted as the metadata. The third baseline is HoloCast [17], and the fourth is extended HoloCast (eHoloCast) [23]. HoloCast sends the elements of the graph basis matrix in each octree block as the metadata. eHoloCast takes Givens rotation and uniform quantization for the graph basis matrix to reduce the overhead. Both schemes first use the octree decomposition with the different values of N from 500 through 3000 to discuss the tradeoff between the communication overhead and reconstruction quality. The proposed scheme sets N to 1000; thus, each octree block contains up to 1000 3D points. For fair comparisons, we align the traffic of the proposed and baseline schemes by controlling the quantization factor and the level of octree decomposition.

Fig. 2 shows the reconstructed quality of the color components for the proposed and baseline schemes as a function of wireless channel SNRs. Here, the communication overhead is obtained by dividing the total number of bits for metadata by the number of elements in the adjacency matrix. We can see that the proposed scheme gradually improves the reconstruction quality by improving the wireless channel quality. In addition, it achieves the best reconstruction quality in low SNR regimes. The QPSK and 16QAM schemes achieve high reconstruction quality in high SNR regimes, whereas they suffer from the cliff and leveling-off effects and low reconstruction quality at low SNR regimes due to the limitation of compression efficiency.

Fig. 3 shows the reconstructed quality of the color components under the different amounts of communication overhead. Here, we consider the wireless channel SNR between the sender and the receiver to be fixed to 15 dB. We note that the 16QAM scheme does not decode the point cloud at the channel SNR due to large bit errors. The figure shows that the proposed scheme achieves more than 3.0 dB improvement compared with the QPSK scheme under the same amount of communication overhead. In addition, the proposed scheme realizes approximately 2^7 overhead reduction compared with the eHoloCast scheme under the same reconstruction quality. In future work, we will discuss the performance gain of the proposed scheme for point clouds with a larger number of 3D points.

4. CONCLUSION

This paper proposes a low-overhead graph-based point cloud delivery scheme for realizing holographic-type communication. The proposed scheme consists of the GFT coefficient and metadata transmission parts to prevent cliff and leveling-off effects due to unreliable wireless channels and reduce the communication overhead for decoding the graph basis matrix at the receiver. Specifically, the proposed scheme skips quantization and entropy coding from the sender operations and uses a dense modulation format to directly map the GFT coefficients onto the transmission symbols. In addition, it designs the INR-based adjacency matrix compression for overhead reduction. Evaluation results show that the proposed INR-based adjacency matrix compression can decode an accurate graph basis matrix with low overhead. The decoded graph basis matrix can realize high-quality and low-overhead graph-based point cloud delivery for future holographic-type communication.

5. REFERENCES

[1] Z. Liu, Q. Li, X. Chen, C. Wu, S. Ishihara, J. Li, and Y. Ji, “Point cloud video streaming: Challenges and solutions,” *IEEE Network*, vol. 35, no. 5, pp. 202–209, 2021.

[2] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[3] G. Cheung, E. Magli, Y. Tanaka, and M. K. Ng, “Graph spectral image processing,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 907–930, 2018.

[4] T. Fujihashi, T. Koike-Akino, and T. Watanabe, “Soft delivery: Survey on a new paradigm for wireless and mobile multimedia streaming,” *ACM Computing Survey*, vol. 56, no. 2, pp. 1–37, 2023.

[5] X. Liu, G. Cheung, X. Wu, and D. Zhao, “Random walk graph Laplacian based smoothness prior for soft decoding of JPEG images,” *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 509–524, 2017.

[6] W. Hu, X. Gao, G. Cheung, and Z. Guo, “Feature graph learning for 3d point cloud denoising,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2841–2856, 2020.

[7] C. He, Y. Hu, Y. Chen, X. Fan, H. Li, and B. Zeng, “MUcast: Linear uncoded multiuser video streaming with channel assignment and power allocation optimization,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 4, pp. 1136–1146, 2020.

[8] S. Ueno, T. Fujihashi, T. Koike-Akino, and T. Watanabe, “Point cloud soft multicast for untethered XR users,” *IEEE Transactions on Multimedia*, vol. 25, pp. 7185–7195, 2022.

[9] E. Bourtsoulatze, D. B. Kurka, and D. Gunduz, “Deep joint source-channel coding for wireless image transmission,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567–579, 2019.

[10] S. Saito, T. Simon, J. Saragih, and H. Joo, "Pifuhd: multi-level pixel-aligned implicit function for high-resolution 3d human digitization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 84–93.

[11] V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020, pp. 1–12.

[12] P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura, and W. Wang, "Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction," *Advances in Neural Information Processing Systems*, vol. 34, pp. 27 171–27 183, 2021.

[13] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[14] P. de Oliveira Rente, C. Brites, J. Ascenso, and F. Pereira, "Graph-based static 3D point clouds geometry coding," *IEEE Transactions on Multimedia*, vol. 21, no. 2, pp. 284–299, 2019.

[15] C. Zhang, D. Florêncio, and C. Loop, "Point cloud attribute compression with graph transform," in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 2066–2070.

[16] P. Gao, L. Zhang, L. Lei, and W. Xiang, "Point cloud compression based on joint optimization of graph transform and entropy coding for efficient data broadcasting," *IEEE Transactions on Broadcasting*, pp. 1–13, 2023.

[17] T. Fujihashi, T. Koike-Akino, T. Watanabe, and P. Orlik, "HoloCast: Graph signal processing for graceful point cloud delivery," in *IEEE International Conference on Communications*, 2019, pp. 1–7.

[18] T. Fujihashi, T. Koike-Akino, T. Watanabe, and P. V. Orlik, "HoloCast+: hybrid digital-analog transmission for graceful point cloud delivery with graph fourier transform," *IEEE Transactions on Multimedia*, vol. 24, pp. 2179–2191, 2021.

[19] T. Fujihashi, T. Koike-Akino, S. Chen, and T. Watanabe, "Wireless 3D point cloud delivery using deep graph neural networks," in *IEEE International Conference on Communications*, 2021, pp. 1–6.

[20] Y. Xu, W. Hu, S. Wang, X. Zhang, S. Wang, S. Ma, Z. Guo, and W. Gao, "Predictive generalized graph fourier transform for attribute compression of dynamic point clouds," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 5, pp. 1968–1982, 2020.

[21] R. A. Cohen, D. Tian, and A. Vetro, "Attribute compression for sparse point clouds using graph transforms," in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 1374–1378.

[22] E. Pavez, B. Girault, A. Ortega, and P. A. Chou, "Region adaptive graph fourier transform for 3d point clouds," in *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 2726–2730.

[23] T. Fujihashi, T. Koike-Akino, T. Watanabe, and P. Orlik, "Overhead reduction in graph-based point cloud delivery," in *IEEE International Conference on Communications*, 2020, pp. 1–7.

[24] E. Dupont, A. Golinski, M. Alizadeh, Y. W. Teh, and A. D. an an, "COIN: COmpression with implicit neural representations," in *Neural Compression: From Information Theory to Applications – Workshop ICLR 2021*, 2021.

[25] E. Dupont, H. Loya, M. Alizadeh, A. Goliński, Y. W. Teh, and A. Doucet, "COIN++: neural compression across modalities," *Transactions on Machine Learning Research*, vol. 2022, no. 11, 2022.

[26] R. Horaud, "A short tutorial on graph Laplacians, Laplacian embedding, and spectral clustering." [Online]. Available: http://csustan.csustan.edu/ tom/Lecture-Notes/Clustering/GraphLaplacian-tutorial.pdf

[27] S. Jakubczak and D. Katabi, "A cross-layer design for scalable mobile video," in *ACM Annual International Conference on Mobile Computing and Networking*, Las Vegas, NV, sep 2011, pp. 289–300.

[28] J. Kammerl, N. Blodow, R. B. Rusu, S. Gedikli, M. Beetz, and E. Steinbach, "Real-time compression of point cloud streams," in *IEEE International Conference on Robotics and Automation*, 2012, pp. 778–785.

[29] K. Muller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. H. Rhee, G. Tech, M. Winken, and T. Wiegand, "3D is here: Point cloud library (PCL)," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1–4.