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

Point Cloud Soft Multicast for Untethered XR Users

Soushi Ueno; Fujihashi, Takuya; Koike-Akino, Toshiaki; Watanabe, Takashi

TR2022-164 December 16, 2022

Abstract

3D point cloud data formats are used to express three-dimensional (3D) information using numerous points in a 3D space. A key challenge is the delivery of high-quality 3D point cloud for the users under a diverse channel quality and available bandwidth to share the same 3D space across multiple untethered extended reality (XR) users. The existing digital-based schemes suffer from two issues owing to the diversity: cliff and leveling-off effects. This paper proposes a novel soft multicasting scheme of point cloud data for untethered XR users. The key ideas of the proposed scheme are three-fold: 1) integration of graph signal processing and analog modulation to adaptively improve the 3D reconstruction quality according to the channel quality for all individual XR users, 2) integration of Givens rotation and non- uniform adaptive quantization to reduce metadata overhead for the graph Fourier transform, and 3) prioritized transmission of the metadata to realize adaptive quality improvement based on the bandwidth available for each XR user. This paper reveals that the proposed scheme prevents cliff and leveling- off effects even when the XR users experience different channel qualities. Furthermore, the proposed transmission exhibits better 3D reconstruction quality compared with the state-of-the-art graph-based delivery scheme in band-limited environments.

IEEE Transactions on Multimedia 2022

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Point Cloud Soft Multicast for Untethered XR Users

Soushi Ueno*, Takuya Fujihashi*, Toshiaki Koike-Akino†, Takashi Watanabe*
*Graduate School of Information and Science, Osaka University, Osaka, Japan
†Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA

***Abstract*—3D point cloud data formats are used to express three-dimensional (3D) information using numerous points in a 3D space. A key challenge is the delivery of high-quality 3D point cloud for the users under a diverse channel quality and available bandwidth to share the same 3D space across multiple untethered extended reality (XR) users. The existing digital-based schemes suffer from two issues owing to the diversity: cliff and leveling-off effects. This paper proposes a novel soft multicasting scheme of point cloud data for untethered XR users. The key ideas of the proposed scheme are three-fold: 1) integration of graph signal processing and analog modulation to adaptively improve the 3D reconstruction quality according to the channel quality for all individual XR users, 2) integration of Givens rotation and non-uniform adaptive quantization to reduce metadata overhead for the graph Fourier transform, and 3) prioritized transmission of the metadata to realize adaptive quality improvement based on the bandwidth available for each XR user. This paper reveals that the proposed scheme prevents cliff and leveling-off effects even when the XR users experience different channel qualities. Furthermore, the proposed transmission exhibits better 3D reconstruction quality compared with the state-of-the-art graph-based delivery scheme in band-limited environments.**

***Index Terms*—Point Cloud Delivery, Graph Fourier Transform, Givens Rotation, Non-Uniform Quantization, XR Devices**

I. Introduction

Volumetric contents are attractive media to reproduce large and detailed three-dimensional (3D) environments on extended reality (XR) and holographic displays. The reproduced 3D environments may aid in the digitization of the real world for numerous applications, such as e-learning, virtual tours, and online meetings and interactions. Point cloud [1] is among the typical data structures used for the volumetric contents. The point cloud uses numerous points for the 3D scene reproduction, and each 3D point is defined by 3D coordinates (X, Y, Z) and color attributes (R, G, B).

The typical applications of point cloud delivery are multicasting delay-sensitive and quality-sensitive contents for untethered XR displays such as live entertainment and collaborative work. In live entertainment, for instance, a server shares the same 3D point cloud to multiple XR users through wireless channels, allowing each XR user to observe the received point cloud from their preferred perspective. In this case, the key challenge is to send such numerous 3D points via band-limited and unreliable wireless channels while keeping the high 3D reconstruction quality on displays, as shown in Fig. 1. In general, the 3D points will be non-uniformly distributed in 3D space to reproduce various real-world settings. Such irregular structure of the signals causes a low efficiency in the compression because the conventional compression schemes for the conventional 2D images and videos assume uniformly sampled pixels. Some compression methods have

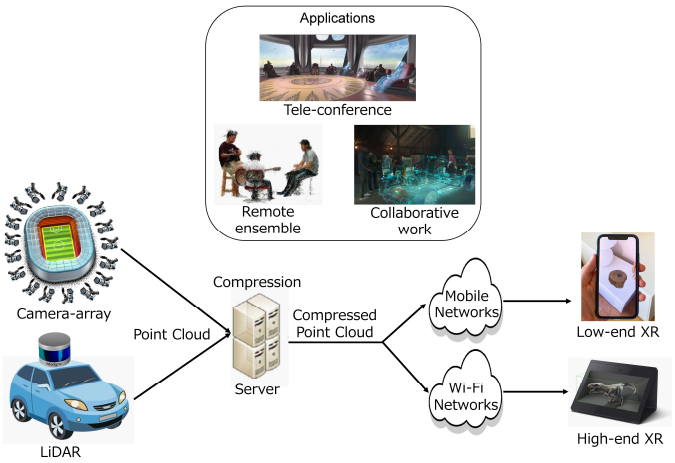


Fig. 1. Exemplification of end-to-end architecture in wireless point cloud delivery.

recently been proposed for point clouds; e.g., point cloud library (PCL) [2], [3], Draco [4], and point cloud compression (PCC) [5], [6] to efficiently compress the non-uniformly sampled signals. PCL uses octree-based coding and scalar quantization for the 3D coordinates and color components. Draco uses Kd-tree data structure and quantization for the compression of each attribute. However, such solutions have limited efficiency, especially in color components, because the conventional solutions do not fully exploit the correlations between the 3D points. Despite the fact that conventional transform techniques, such as the discrete cosine transform (DCT), can still be used for the 3D point cloud, they do not fully exploit the underlying irregular and non-ordered structure of the 3D points.

More recently, graph signal processing (GSP) [7] was introduced to exploit the correlations for 3D point cloud compression. GSP employing the graph Fourier transform (GFT), for example, was used for the color components in [8] and 3D coordinates in [9] of the graph signals. The GFT-based schemes obtain a unitary basis matrix from the 3D coordinates and color components for graph-signal decorrelation. The digital-based source coding, i.e., quantization and entropy coding, is then carried out for the GFT coefficients to remove the spatial redundancy. Another study in [10] designed the graph-domain prediction before decorrelation for further energy compaction.

Nevertheless, the existing studies on digital-based source coding resulted in low reconstruction quality, especially in multi-user environments; it is because the wireless channel quality at the end of each XR user may unpredictably change

over time owing to mobility, fading, and interference. Furthermore, the available bandwidth may vary depending on the user's device and connected wireless resource. There are three major issues in multicasting the 3D point cloud with GFT to multiple XR users via diverse wireless channels.

The first issue is the cliff effect [11]: significant quality degradation owing to wireless channel fading. Specifically, several bit errors will occur in the transmitted bits when the experienced channel quality, i.e., signal-to-noise ratio (SNR), drops below a certain threshold. The bit errors cause a catastrophic impact on the signal restoration; thus, some XR users having faded channel quality may be unable to reconstruct the 3D point clouds because the entropy decoding in the point cloud compression has an all-or-nothing feature.

The second issue is the leveling-off effect: the reconstruction quality will saturate even when the channel quality improves considerably. This is attributed to the fact that the XR user cannot restore the quantization errors at the sender's end. The quantization level is generally determined based on the user with the lowest channel quality, i.e., lowest available data rate, across multiple users. Therefore, other XR users will experience lower-quality 3D point clouds even when the wireless channel quality is considerably high.

The third issue is high communication overhead owing to the GFT basis matrix. The GFT-based solutions need to send the corresponding basis matrix for the GFT operation as metadata despite realizing better energy compaction. Some studies utilized Givens rotation [12], [13], which represents any basis matrices with fewer angle parameters. However, all the XR users require an identical bandwidth because all the angle parameters are required to be transmitted to each XR user for the basis matrix reconstruction. The quality of the 3D reconstruction would significantly degrade if some XR users failed to receive a fraction of angle parameters owing to bandwidth limitations.

Compressing and multicasting the 3D point cloud while considering the lowest channel quality and bandwidth across the XR users is the typical solution for preventing a catastrophic impact on signal restoration. However, it limits 3D reconstruction quality for XR users with better channel quality and sufficient bandwidth. This paper proposes a soft multicast scheme motivated by soft delivery, namely, HoloCast [14], to solve the above-mentioned issues of wireless point cloud delivery for multiple users. For this purpose, the proposed scheme utilizes GSP for energy compaction, unequal power allocation across GFT coefficients, near-analog modulation for the diversity of the channel quality, Givens rotation with adaptive and non-uniform quantization for the compression of the GFT basis matrix, and stacked transmission for the diversity of the bandwidth.

Based on the evaluation results, the proposed scheme gradually improves 3D reconstruction quality against the wireless channel quality improvement. The proposed non-uniform quantization achieves up to a 40% reduction in communication overhead of the GFT basis matrix compared with the conventional uniform quantization. Additionally, the proposed stacked transmission yields adaptive 3D reconstruction quality according to the available bandwidth for the overhead.

The proposed scheme is an extension of our previous studies [14]–[16]. Specifically, HoloCast in [14] integrated the energy compaction using GFT, unequal power allocation, and near-analog modulation to deliver the point cloud for multiple users. However, the integration needs a large metadata overhead for the GFT. Givens rotation and uniform quantization were adopted in [15] and uniform quantization was extended to non-uniform quantization [16] for overhead reduction. This paper proposes stacked transmission of the angle parameters to yield adaptive 3D reconstruction quality according to the bandwidth available for each XR user. Furthermore, the proposed scheme realizes adaptive quantization across the angle parameters to reconstruct high-quality 3D point cloud under the same amount of overhead.

The major contributions of this study are as follows:

- To the best of our knowledge, this is the first study to consider channel quality and bandwidth heterogeneity across multiple XR users for wireless point cloud delivery.
- The authors investigate the distribution of the angle parameters and conclude that the angle parameters follow the two-phase hyper-exponential distribution.
- Non-uniform quantization based on the two-phase hyper-exponential distribution with three parameters is adopted for the efficient compression of the angle parameters.
- The authors discuss the column order of the GFT basis matrix to deal with the bandwidth diversity across multiple XR users.
- Adaptive quantization is designed to discuss the effect of unequal bit allocation across the angle parameters.

II. Related Work

A. GSP-Based Point Cloud Delivery

Many 3D points are non-uniformly sampled in the 3D space to represent 3D scenes and objects in the point cloud. One of the key issues in the point cloud delivery is to find methods to compact the energy of each 3D point signal. Some studies have proposed GSP-based solutions for this purpose. For example, GFT is utilized for the energy compaction of 3D coordinates [9], color components [8], or both [14], [17]. In these studies, the decorrelated components are compressed and delivered in digital [8], [9], analog [14], and digital–analog hybrid [17] manners. In [18], region adaptive GFT (RA-GFT) has been proposed for a similar purpose on the point cloud delivery using multiresolution decorrelation based on the region of interest (ROI). A graph neural network (GNN) [19] architecture has been designed for point cloud delivery in recent years. Specifically, the GNNCast in [20] was designed using an auto-encoder architecture based on the graph convolutional neural network (GCNN) for energy compaction and then the latent variables were mapped into the transmission symbols in an analog manner.

B. Soft Delivery

Soft delivery schemes [11], [21]–[23] have been proposed to solve the cliff and leveling-off effects across wireless and mobile receivers. Several studies on energy compaction [24],

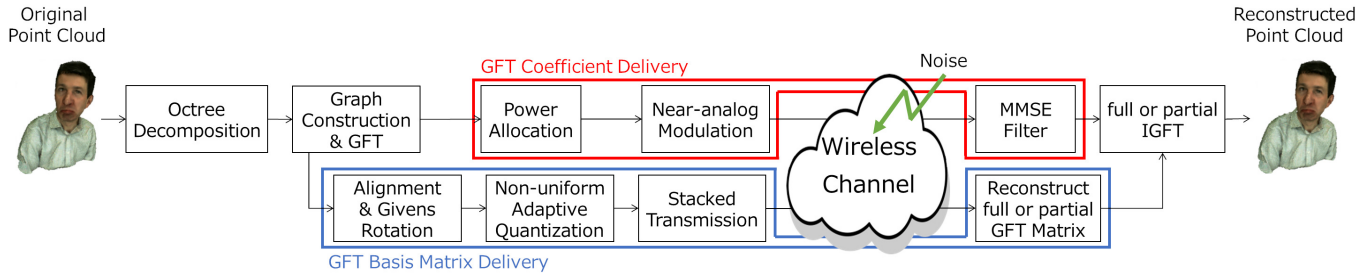


Fig. 2. Overview of the proposed end-to-end point cloud delivery scheme.

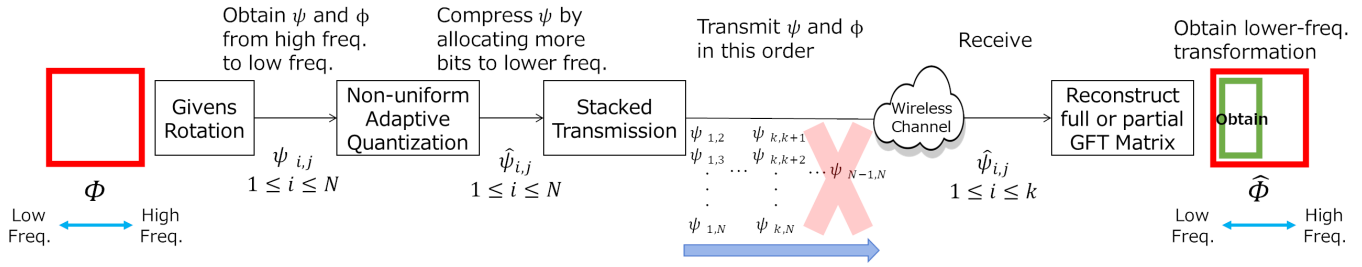


Fig. 3. Overview of GFT basis matrix transmission.

[25] and power allocation [26]–[30] have been conducted to enhance the 2D image and video reconstruction quality. For example, DCast in [25] employed coset coding to remove inter-frame redundancy, and the proposed scheme in [24] adopted motion-compensated temporal filtering (MCTF) in addition to 2D-DCT for energy compaction. Other studies in [29], [30] designed unequal transmission power assignment for each analog-modulated symbol to maximize the human-perceptual quality considering human vision systems [29] and the ROI and non-ROI parts in each image [30]. Furthermore, power allocation problems were studied in modern wireless systems, including multiple-input and multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) [26], unmanned aerial vehicle (UAV)-enabled [28], and mmWave lens MIMO channels [27]. Some soft-delivery studies proposed energy compaction methods for immersive image and video signals, such as, free-viewpoint video [31], [32], 360-degree video [33]–[35], and point cloud [14].

The authors previously proposed HoloCast [14] to solve the cliff and leveling-off effects for the point cloud delivery by integrating unequal power allocation and near-analog modulation with GFT to deal with the irregular structure of the 3D points. The GFT-based operation can realize better energy compaction; however, HoloCast will cause a large communication overhead to exchange the GFT basis matrix. Overhead reduction methods [15], [16] have been proposed as a preliminary step to decrease the communication overhead.

C. Contributions over Existing Studies

This study proposes a novel soft multicast scheme for point cloud delivery by extending our preliminary investigations [15], [16]. The proposed scheme adopts GFT, unequal power allocation, and near-analog modulation, motivated by HoloCast. In contrast to the HoloCast, the proposed scheme reduces communication overhead and deals with bandwidth heterogeneity across multiple XR users. Specifically, a significant overhead reduction can be achieved using the Givens rotation and non-uniform quantization. Furthermore, the proposed scheme introduces a new rate-adaptive scheme with a prioritized transmission of angle parameters in Givens rotation to reconstruct higher-quality point clouds according to the available bandwidth.

III. Overview of Proposed Soft Multicast Scheme

The objectives of the proposed scheme are: 1) to compact the energy of irregular and non-ordered 3D points, 2) to prevent cliff and leveling-off effects owing to wireless channel quality fluctuation in each XR user, 3) to reduce a communication overhead owing to the GFT basis matrix, and 4) to prevent quality degradation because of bandwidth differences between the XR users.

Fig. 2 exhibits the overview of the proposed scheme. The proposed scheme mainly consists of the GFT coefficient and basis matrix delivery parts. First, the proposed scheme takes octree decomposition to decompose the 3D points into multiple octree blocks as necessary. The proposed scheme constructs an undirected and weighted graph for each octree block and obtains the GFT basis matrix from the graph Laplacian matrix. The proposed scheme takes GFT across the attributes of the 3D points based on the GFT basis matrix for decorrelation. The GFT coefficients are then modulated to a dense constellation—near-analog modulation—for transmission; this is followed by the unequal power allocation across the GFT coefficients to protect unequal signals against channel noises for maximizing the reconstruction quality. The analog-

modulated GFT coefficients are then transmitted to each XR user via wireless channels.

Fig. 3 shows an overview of a part of the GFT basis matrix delivery. The proposed scheme sorts the columns of the GFT basis matrix in the ascending order of the corresponding graph frequency. The sorted basis matrix is transformed into the angle parameters using Givens rotation. Further, based on the XR user with the highest available bandwidth, the angle parameters are non-uniformly quantized into adequate bit depth, and the quantization intervals follow a two-phase hyper-exponential distribution. In this case, the proposed scheme can set unequal quantization intervals to the angle parameters to enhance the reconstruction quality. Specifically, more bits should be allocated to the angle parameters contributing to the low frequencies than those contributing to the high frequencies. Further, the quantized angle parameters are progressively broadcast to multiple XR users using the stacked transmission technique. Each XR user receives the full or partial angle parameters based on the available bandwidth.

Next, each XR user reconstructs the partial column vectors corresponding to the low-frequency GFT coefficients based on the received angle parameters. The column vectors are then used to reconstruct the attributes of the 3D coordinates and color components from the received analog-modulated GFT coefficients by taking full or partial inverse GFT (IGFT). Consequently, each XR user can play adequate quality of the 3D point cloud according to its wireless channel quality and bandwidth.

IV. GFT Coefficient Delivery

A. GFT-Based Energy Compaction

1) Graph Construction: The 3D point cloud comprises the 3D coordinates and color components. Both 3D coordinates and color components can be regarded as a weighted and undirected graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{W})$. Here, $\boldsymbol{V}$ is the vertex set and $\boldsymbol{E}$ is the edge set. $\boldsymbol{W}$ is an adjacency matrix representing the edge weights. Specifically, $W_{i,j}$ represents the edge weight between vertices i and j. The 3D coordinates $\boldsymbol{p} = [\mathrm{x}, \mathrm{y}, \mathrm{z}]^{\mathrm{T}} \in \mathbb{R}^{3\times N}$ and the corresponding color components $\boldsymbol{c} = [\mathrm{y}, \mathrm{u}, \mathrm{v}]^{\mathrm{T}} \in \mathbb{R}^{3\times N}$ can be regarded as graph signals that reside on the vertices. Here, N is the number of vertices in the octree block.

2) Graph Fourier Transform (GFT): Each edge weight $W_{i,j}$ of the graph signals can be obtained by Gaussian kernel as follows:

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

where ϵ_p is the sample variance of 3D coordinates. The graph signals are then transformed into frequency-domain coefficients using GFT. The GFT is defined through the graph Laplacian operator $\boldsymbol{L}$ based on adjacency matrix $\boldsymbol{W}$ and degree matrix $\boldsymbol{D}$ as follows: $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}$, where $\boldsymbol{D}$ is the diagonal degree matrix which is represented as:

$$\boldsymbol{D} = \mathrm{diag}(D_1, \ldots, D_N),\ D_i = \sum_{n=1}^{N} W_{i,n}, \tag{2}$$

In general, the graph Laplacian is a real symmetric matrix with a complete set of orthonormal eigenvectors with corresponding nonnegative eigenvalues. The eigenvectors and eigenvalues of the graph Laplacian operator can be obtained by the eigenvalue decomposition as $\boldsymbol{L} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{-1}$. Here, $\boldsymbol{\Lambda}$ is a diagonal matrix of the eigenvalues arranged in ascending order and $\boldsymbol{\Phi} \in \mathbb{R}^{N\times N}$ is a unitary matrix that comprises the eigenvectors, i.e., GFT basis matrix. The GFT coefficients of each attribute $\boldsymbol{f} \in \mathbb{R}_N$ are obtained by multiplying the GFT basis matrix by the attribute signals $\boldsymbol{s} \in \mathbb{R}_N$ as:

$$\boldsymbol{f} = \boldsymbol{s}\,\boldsymbol{\Phi}. \tag{3}$$

Here, each attribute can be reconstructed from GFT coefficients using the GFT basis matrix as:

$$\boldsymbol{s} = \boldsymbol{f}\,\boldsymbol{\Phi}^{-1}. \tag{4}$$

B. Near-Analog Modulation and Demodulation

The proposed scheme performs GFT for the attributes of each 3D point to obtain the corresponding GFT coefficients. Unequal transmission power is assigned to each GFT coefficient to minimize the mean-square error (MSE), and the power-assigned GFT coefficients are then mapped, two by two, to I (in-phase) and Q (quadrature-phase) components, i.e., near-analog modulation, for the coefficient transmissions.

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 $x_i = g_i \cdot f_i$. The optimal scale factor g_i is obtained by minimizing the MSE under the power constraint with a total power budget of P. The near-optimal solution [21] is expressed as:

$$g_i = \lambda_i^{-1/4} \sqrt{\frac{NP}{\sum_j^N \sqrt{\lambda_j}}}, \tag{5}$$

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. Here, the fading attenuation is also considered in the noise variance. The received symbol y_i over wireless links can be modeled as follows:

$$y_i = x_i + n_i. \tag{6}$$

The GFT coefficients are then extracted from the I and Q components through a minimum MSE (MMSE) filter [21]:

$$\hat{f}_i = \frac{g_i \lambda_i}{g_i^2 \lambda_i + \sigma^2} \cdot y_i. \tag{7}$$

The decoder finally reconstructs the corresponding graph signals $\hat{\boldsymbol{s}}$ by taking IGFT for the filtered GFT coefficients in each attribute $\hat{\boldsymbol{f}}$ as: $\hat{\boldsymbol{s}} = \hat{\boldsymbol{f}}\,\boldsymbol{\Phi}^{-1}$.

V. GFT Basis Matrix Delivery

A. Givens Rotation

The sender needs to send the GFT basis matrix to the receiver for signal reconstruction in the GSP-based point cloud delivery schemes. The overhead for this metadata transmission

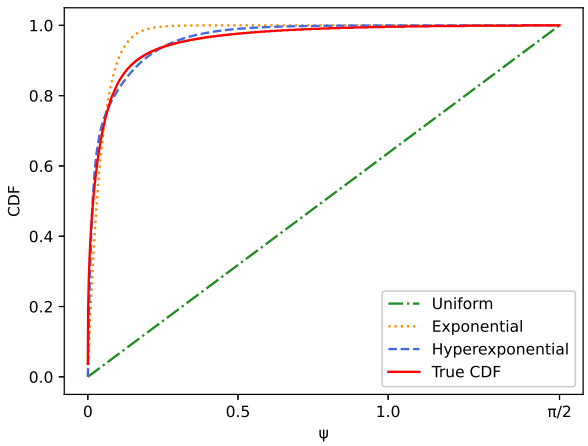


Fig. 4. CDF of the angle parameters in GFT basis matrix $\boldsymbol{\Phi}$ for the 3D point cloud of *pencil_10_0*, whose number of points is $N = 2{,}731$.

is generally non-negligible. Specifically, the sender transmits N^2 real elements to the receiver for N points. The proposed scheme employs the Givens rotation for the GFT basis matrix to reduce the amount of metadata. The GFT basis matrix is orthonormal; therefore, it can be decomposed into a product of $N(N-1)/2$ Givens rotations, represented by $N(N-1)/2$ angle parameters, and $N(N+1)/2$ binary values in total. Givens rotation for the sorted GFT basis matrix can be derived as follows:

$$\boldsymbol{\Phi} = \prod_{i=1}^{N} \Big[\boldsymbol{D}_i(\phi_{i,i}, \ldots, \phi_{i,N}) \prod_{j=i+1}^{N} \boldsymbol{G}_{i,j}(\psi_{i,j}) \Big] \boldsymbol{I}_N, \quad (8)$$

$$\boldsymbol{D}_i(\phi_{i,i}, \ldots, \phi_{i,N}) = \mathrm{diag}(\mathbf{1}_{i-1}, \mathrm{e}^{\jmath\phi_{i,i}}, \ldots, \mathrm{e}^{\jmath\phi_{i,N}}), \quad (9)$$

where $\mathbf{1}_{i-1}$ is an all-ones vector of size $i-1$, and $\boldsymbol{G}_{i,j}(\psi)$ is the Givens rotation matrix for rotating ith and jth components as follows:

$$\boldsymbol{G}_{i,j}(\psi) = \begin{bmatrix} \boldsymbol{I}_{i-1} & & & & \\ & c_\psi & & -s_\psi & \\ & & \boldsymbol{I}_{j-i-1} & & \\ & s_\psi & & c_\psi & \\ & & & & \boldsymbol{I}_{N-j} \end{bmatrix}, \quad (10)$$

where $\boldsymbol{I}_k$ is $k \times k$ identity matrix, $c_\psi = \cos(\psi)$ and $s_\psi = \sin(\psi)$. Hence, the GFT basis matrix can be represented with angle variables of $\phi_{i,j}$ and $\psi_{i,j}$. Here, the angle parameters $\phi_{i,j}$ are either 0 or π because the GFT basis matrix only comprises the real elements, and thus 1 bit is enough to be assigned for each $\phi_{i,j}$. Thus, it can drastically reduce the amount of metadata. For example, the 3×3 GFT basis matrix $\boldsymbol{\Phi}$ can be decomposed as follows:

$$\begin{aligned} \boldsymbol{\Phi} = \boldsymbol{D}_1(\phi_{1,1}, \phi_{1,2}, \phi_{1,3}) \boldsymbol{G}_{1,2}(\psi_{1,2}) \boldsymbol{G}_{1,3}(\psi_{1,3}) \\ \boldsymbol{D}_2(\phi_{2,2}, \phi_{2,3}) \boldsymbol{G}_{2,3}(\psi_{2,3}) \boldsymbol{D}_3(\phi_{3,3}) \boldsymbol{I}_3. \end{aligned} \quad (11)$$

Here, there are 3 angle parameters and 6 binary parameters in the Givens rotation decomposition, whereas the original GFT matrix has 9 real values.

B. Non-Uniform Quantization

The angle parameters ψ in the Givens rotation decomposition are usually quantized uniformly over a range between 0 and $\pi/2$. However, such a uniform quantization is inefficient because the probability distribution of ψ is typically not a uniform distribution for most 3D point cloud data. Fig. 4 shows the empirical cumulative distribution function (CDF) of the angle parameters ψ for a sample 3D point cloud data with $2{,}731$ number of points. In this case, most of the angle parameters are close to zero. It encourages us to employ a non-uniform quantization to minimize distortion at a limited quantization level.

Here, uniform, exponential (one-phase hyper-exponential), and the two-phase hyper-exponential CDFs are considered to discuss on the CDFs that agree well with the empirical CDF. The authors found that the two-phase hyper-exponential CDF agrees well with the empirical CDF compared with other CDFs. Specifically, the root mean square error (RMSE) between the empirical CDF and the uniform, exponential, and two-phase hyper-exponential CDFs are 0.53, 0.040, and 0.014, respectively, using the best fitting parameters. The two-phase hyper-exponential CDF is given as follows:

$$F_{\mathrm{h.exp}}(\psi) = 1 - \beta_1 e^{-\gamma_1 \psi} - \beta_2 e^{-\gamma_2 \psi}, \quad (12)$$

where β_l and γ_l are the fitting parameters. Here, the best fitting parameters can be found with a least-squares method under the condition of $\beta_1 + \beta_2 = 1$.

The quantization points Δ_n are determined according to a bit depth b for the quantization of ψ based on the fitted CDF as follows:

$$\Delta_n = \frac{1}{2}(x_{n-1} + x_n), \ n = 1, \ldots, 2^b, \quad (13)$$

$$x_n = \begin{cases} F^{-1}_{\mathrm{h.exp}}\left(\frac{n}{2^b}\right), & \text{if } n = 0, 1, \ldots, 2^b - 1, \\ \pi/2, & \text{if } n = 2^b, \end{cases} \quad (14)$$

where $F^{-1}_{\mathrm{h.exp}}(\cdot)$ is an inverse function of $F_{\mathrm{h.exp}}(\psi)$, which returns the value between 0 and $\pi/2$ from the input value between 0 and 1 based on the two-phase hyper-exponential CDF. Given the quantization points $\boldsymbol{\Delta} = \{\Delta_1, \Delta_2, \ldots, \Delta_{2^b}\}$, the proposed scheme selects the one in $\boldsymbol{\Delta}$ closest to the angle parameter $\psi_{i,j}$ as the quantized angle parameter:

$$\hat{\psi}_{i,j} = \Delta_n, \text{ if } x_{n-1} < \psi_{i,j} \leq x_n, \ n = 1, \ldots, 2^b. \quad (15)$$

C. Progressive Stacked Transmission

The quantized angle parameters are shared with all XR users at an adequate bit depth to minimize distortion because of quantization error. However, some users may not be able to receive all the GFT coefficients and the angle parameters depending on the available bandwidth of individual XR devices.

For such heterogenous multicasting scenarios, the authors propose a stacked transmission technique to enable XR users to progressively reconstruct partial GFT basis vectors corresponding to low-frequency GFT coefficients within the individual bandwidth.

Fig. 5 shows an overview of the proposed stacked transmission for multiple XR users. Let $\boldsymbol{T}_n$ be a set of angle parameters

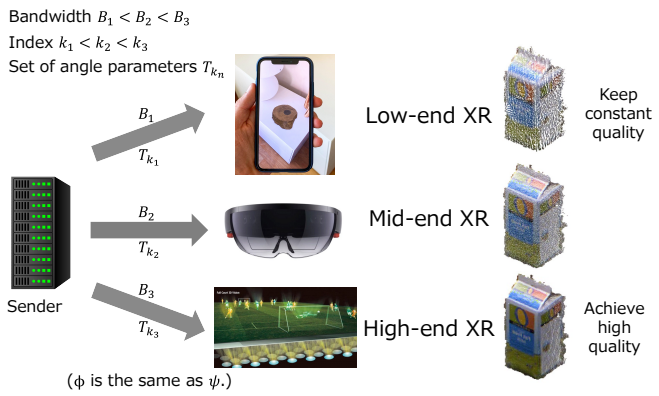


Fig. 5. Overview of stacked transmission. The sender progressively sends the angle parameters corresponding to low-frequency GFT coefficients.

for the nth GFT basis vector, i.e., $\boldsymbol{T}_n = \{\phi_{n,*}, \psi_{n,*}\}$, where $*$ represents an arbitrary index. The proposed scheme progressively broadcasts the angle parameters in priority order from $\boldsymbol{T}_1$ to $\boldsymbol{T}_N$ through wireless channels as the GFT basis vectors are sorted in the ascending order of the graph frequency. One of the XR users can reconstruct the first through nth column vectors using Givens rotations upon receipt of the subsequent set from $\boldsymbol{T}_1$ to $\boldsymbol{T}_n$. The XR user then takes the partial IGFT based on the received GFT coefficients as in Eq. (4). The GFT coefficients corresponding to the unreconstructed column vectors, i.e., high-frequency GFT coefficients, are regarded as 0 in the partial IGFT.

For an example of 3×3 GFT basis matrix $\boldsymbol{\Phi}$ in (11), the angle parameters can be divided to three sets as follows:

$$\begin{aligned}\boldsymbol{T}_1 &= \{\phi_{1,1}, \phi_{1,2}, \phi_{1,3}, \psi_{1,2}, \psi_{1,3}\},\\ \boldsymbol{T}_2 &= \{\phi_{2,2}, \phi_{2,3}, \psi_{2,3}\},\\ \boldsymbol{T}_3 &= \{\phi_{3,3}\}.\end{aligned}$$

Consider the case when there are three types of XR devices: 1) low-end XR devices can receive only $\boldsymbol{T}_1$; 2) mid-end XR devices receive $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$; 3) high-end XR devices can obtain all the sets. Each type of the XR user can reconstruct its own GFT basis matrix $\boldsymbol{\Phi}_1$ through $\boldsymbol{\Phi}_3$ using the received angle parameters, respectively, as follows:

$$\begin{aligned}\boldsymbol{\Phi}_1 &= \boldsymbol{D}_1(\phi_{1,1}, \phi_{1,2}, \phi_{1,3})\boldsymbol{G}_{1,2}(\psi_{1,2})\boldsymbol{G}_{1,3}(\psi_{1,3}),\\ \boldsymbol{\Phi}_2 &= \boldsymbol{\Phi}_1\boldsymbol{D}_2(\phi_{2,2}, \phi_{2,3})\boldsymbol{G}_{2,3}(\psi_{2,3}),\\ \boldsymbol{\Phi}_3 &= \boldsymbol{\Phi}_2\boldsymbol{D}_3(\phi_{3,3}) = \boldsymbol{\Phi}.\end{aligned}$$

The lower-end XR devices reconstruct 3D points using the partial IGFT, while the highest-end XR users may enjoy the highest quality of 3D point reconstruction with the full IGFT.

D. Adaptive Bit Allocation

The proposed non-uniform quantization can improve the reconstruction quality of the 3D points by utilizing the distribution of the angle parameters. The efficiency can be further improved using unequal bit allocation across angle parameters. Specifically, higher bit depths are assigned to the angle parameters corresponding to lower-frequency GFT coefficients; in general, this will impact the reconstruction quality more significantly.

To simplify, all the angle parameters ψ are divided into two sets for bit allocation: 1) higher-priority angle parameters corresponding to lower-frequency GFT coefficients and 2) lower-priority angle parameters corresponding to higher-frequency GFT coefficients. Here, bit depths of b_{low} and b_{high} are used for high-priority and low-priority angle parameters, respectively ($b_{\mathrm{low}} \geq b_{\mathrm{high}}$).

VI. Performance Evaluation

A. Simulation Settings

1) Performance Metric: This section evaluates the 3D reconstruction quality for point cloud delivery in terms of the symmetric MSE based on [36] in the attributes of 3D coordinates $\boldsymbol{p} = [\mathrm{x}, \mathrm{y}, \mathrm{z}]^{\mathrm{T}}$ and color components $\boldsymbol{c} = [\mathrm{y}, \mathrm{u}, \mathrm{v}]^{\mathrm{T}}$. The symmetric MSE of the 3D coordinates, $\mathsf{sMSE}_{\mathsf{xyz}}$, can be obtained as follows:

$$\mathsf{sMSE}_{\mathsf{xyz}} = \frac{1}{2}\big(d(\boldsymbol{p}_{\mathrm{org}} \to \boldsymbol{p}_{\mathrm{dec}}) + d(\boldsymbol{p}_{\mathrm{dec}} \to \boldsymbol{p}_{\mathrm{org}})\big), \quad (16)$$

where $\boldsymbol{p}_{\mathrm{org}}$ is the original 3D coordinates and $\boldsymbol{p}_{\mathrm{dec}}$ is the decoded 3D coordinates. Here, the asymmetric MSE metrics in the 3D coordinates are defined as follows:

$$d(\boldsymbol{p}_{\mathrm{org}} \to \boldsymbol{p}_{\mathrm{dec}}) = \frac{1}{N}\sum_{\boldsymbol{p}\in\boldsymbol{p}_{\mathrm{org}}}\left(\min_{\boldsymbol{p}'\in\boldsymbol{p}_{\mathrm{dec}}}\|\boldsymbol{p}-\boldsymbol{p}'\|_2^2\right), \quad (17)$$

$$d(\boldsymbol{p}_{\mathrm{dec}} \to \boldsymbol{p}_{\mathrm{org}}) = \frac{1}{N}\sum_{\boldsymbol{p}\in\boldsymbol{p}_{\mathrm{dec}}}\left(\min_{\boldsymbol{p}'\in\boldsymbol{p}_{\mathrm{org}}}\|\boldsymbol{p}-\boldsymbol{p}'\|_2^2\right). \quad (18)$$

The symmetric MSE of the color components, $\mathsf{sMSE}_{\mathsf{yuv}}$, is derived as follows:

$$\mathsf{sMSE}_{\mathsf{yuv}} = \frac{1}{2}\big(d(\boldsymbol{c}_{\mathrm{org}} \to \boldsymbol{c}_{\mathrm{dec}}) + d(\boldsymbol{c}_{\mathrm{dec}} \to \boldsymbol{c}_{\mathrm{org}})\big), \quad (19)$$

where $\boldsymbol{c}_{\mathrm{org}}$ and $\boldsymbol{c}_{\mathrm{dec}}$ are the original and decoded color components, respectively. Here, the asymmetric MSE metrics of the color component are defined as follows:

$$\begin{aligned}d(\boldsymbol{c}_{\mathrm{org}} \to \boldsymbol{c}_{\mathrm{dec}}) &= \frac{1}{N}\sum_{\boldsymbol{c}\in\boldsymbol{c}_{\mathrm{org}}}\big\|\boldsymbol{c}-\boldsymbol{c}_{\mathrm{dec}}(\boldsymbol{p}'_{\mathrm{dec}})\big\|_2^2,\\ \boldsymbol{p}'_{\mathrm{dec}} &= \arg\min_{\boldsymbol{p}'\in\boldsymbol{p}_{\mathrm{dec}}}\big\|\boldsymbol{p}_{\mathrm{org}}-\boldsymbol{p}'\big\|_2^2, \quad (20)\end{aligned}$$

$$\begin{aligned}d(\boldsymbol{c}_{\mathrm{dec}} \to \boldsymbol{c}_{\mathrm{org}}) &= \frac{1}{N}\sum_{\boldsymbol{c}\in\boldsymbol{c}_{\mathrm{dec}}}\big\|\boldsymbol{c}-\boldsymbol{c}_{\mathrm{org}}(\boldsymbol{p}'_{\mathrm{org}})\big\|_2^2,\\ \boldsymbol{p}'_{\mathrm{org}} &= \arg\min_{\boldsymbol{p}'\in\boldsymbol{p}_{\mathrm{org}}}\big\|\boldsymbol{p}_{\mathrm{dec}}-\boldsymbol{p}'\big\|_2^2, \quad (21)\end{aligned}$$

where $\boldsymbol{c}_{\mathrm{dec/org}}(\boldsymbol{p})$ represents the color components of the corresponding 3D coordinate $\boldsymbol{p}$.

2) Point Cloud Dataset: The following benchmark point clouds are used in evaluations: *pencil__10_0* with 2,731 points and *milk_color* with 13,704 points. The point cloud of *pencil__10_0* is used to clarify the baseline performance. The point cloud of *milk_color* is used for the discussion on the visual quality. Octree decomposition is used to decompose the 3D points into multiple octree blocks in this case. Here, the maximum number of 3D points in each block is set to 5,000.

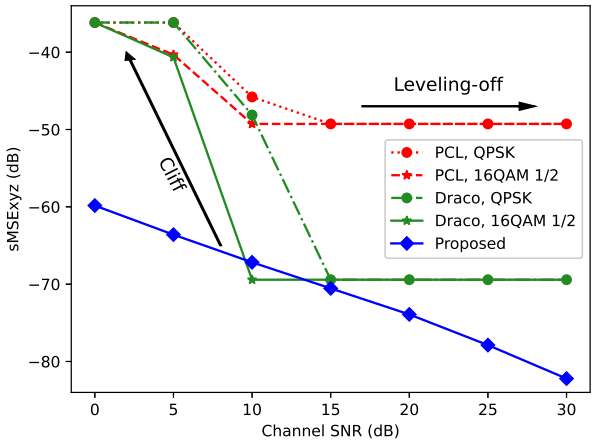


Fig. 6. Symmetric MSE of 3D coordinate attributes as a function of wireless channel SNRs.

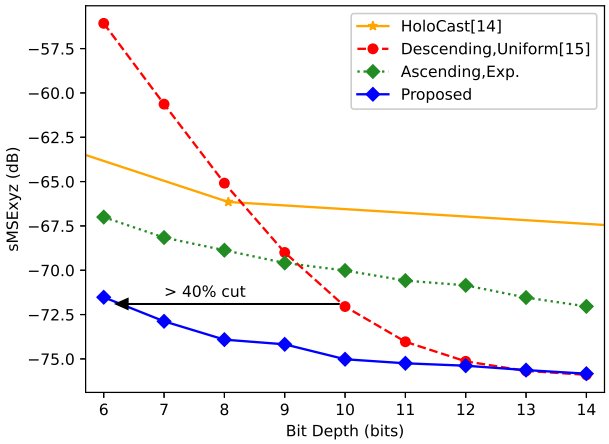


Fig. 8. Symmetric MSE of 3D coordinates as a function of bit depths at the wireless channel SNR of 20 dB.

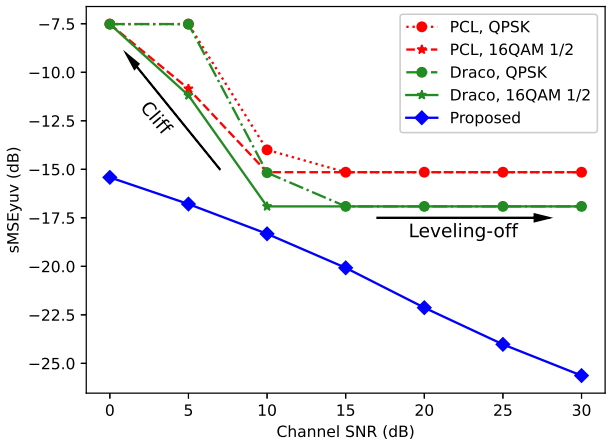


Fig. 7. Symmetric MSE of color attributes as a function of wireless channel SNRs.

3) Hyper-exponential Fitting: We set the fitting parameters of *pencil__10_0* in (12) to $\gamma_1 = 83.62$, $\gamma_2 = 7.31$, $\beta_1 = 0.62$.

B. Digital vs. Proposed Scheme

This section discusses the baseline quality of the proposed scheme. We consider the baseline digital-based solutions based on PCL 1.12.0 [2] and Draco 1.4.3 [4], and the proposed scheme, for performance comparison. Each point cloud is firstly compressed into the bitstream using the digital encoder of PCL and Draco in the digital-based solutions. The bitstream is then channel-coded and modulated for transmissions. A default profile of PCL, LOW_RES_OFFLINE_COMPRESSION_WITH_COLOR, is used for the compression. In Draco, the bit depth for the 3D coordinates is set to 7 bits, and other parameters are set to default values. The compressed bitstream is then channel coded by a 1/2-rate convolutional code with a constraint length of 10 and digitally modulated by Quadrature Phase-Shift Keying (QPSK) and 16-Quadrature Amplitude Modulation (QAM) formats, respectively.

The proposed scheme directly maps the GFT coefficients onto the transmission symbols using the near-analog modulation. The proposed scheme takes non-uniform quantization for angle parameters ψ with the bit depth of 8 bits in terms of the overhead reduction. We closely fit the number of transmission symbols for a fair comparison across the digital-based and proposed schemes. Here, the metadata transmissions of angle parameters are considered error-free.

Figs. 6 and 7 show the symmetric MSE performance of 3D coordinate and color attributes, respectively, as a function of wireless channel SNRs. The evaluation results show the following observations in common with the 3D coordinate and color attributes.

- The proposed scheme achieves the best 3D reconstruction quality under the same available bandwidth in low and high wireless SNR regimes.
- The digital-based schemes using PCL and Draco suffer from the cliff effect at low channel SNR regimes and the leveling-off effect at high channel SNR regimes.
- The proposed schemes gradually improve the 3D reconstruction quality with the improvement of wireless channel quality.

C. Overhead Reduction

This section evaluates the communication overhead associated with the GFT basis matrix in the conventional HoloCast [14] and three soft multicast schemes with different quantization methods. HoloCast adopts the different levels of $N = 1501, 1000, 500$ in the octree decomposition to obtain the overhead and corresponding reconstruction quality. Here, three schemes are considered to discuss the effects of the sorting of GFT basis matrix and quantization methods. The first scheme sorts the GFT basis matrix in descending order of the graph frequency and takes uniform quantization for

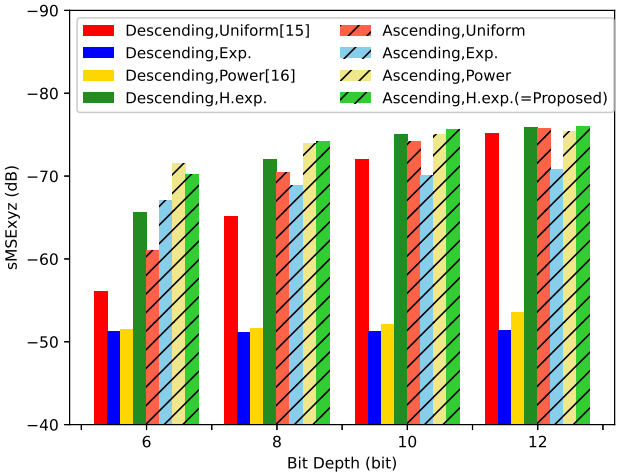


Fig. 9. Effect of the order of GFT basis matrix on the uniform and non-uniform quantization at wireless channel SNR of 20 dB.

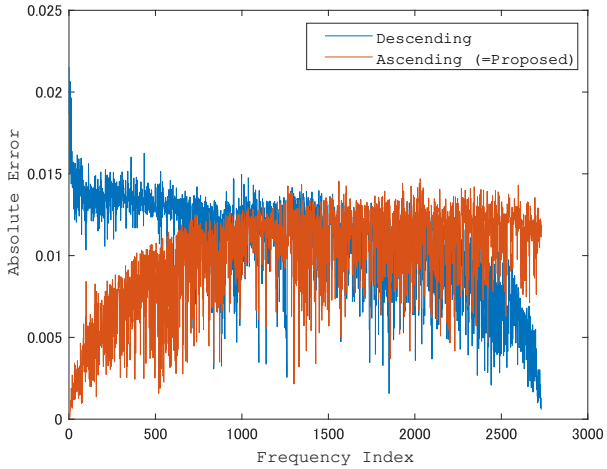


Fig. 10. Quantization error of the proposed scheme in each frequency index under the different orders of the GFT basis matrix. Here, the bit depth is set to 8 bits.

overhead reduction [15]. The second and third schemes sort the GFT basis matrix in ascending order of the graph frequency and take non-uniform quantization based on the exponential and two-phase hyper-exponential CDFs, respectively. Here, the third scheme is the proposed scheme.

Fig. 8 shows the reconstruction quality of the 3D coordinates as a function of bit depths for each angle parameter at a wireless channel SNR of 20 dB. It is observed that the proposed scheme achieves the best 3D reconstruction quality under the same communication overhead. In other words, the proposed scheme achieves the same 3D reconstruction quality as other methods with a lower bit depth. For example, the proposed scheme with the bit depth of 6 bits reconstructs the point cloud with the same quality as the scheme with descending order of the GFT basis matrix and uniform quantization with the bit depth of 10 bits. Additionally, it is discovered that, regardless of bit depths, two-phase hyper-exponential CDF-based quantization realizes lower communication overhead than exponential CDF-based quantization.

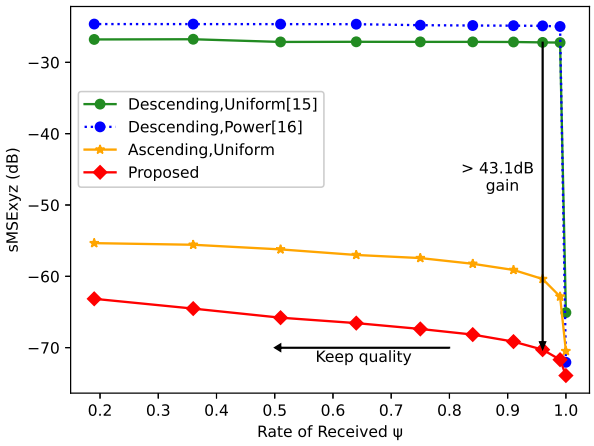


Fig. 11. Symmetric MSE of the 3D coordinates as a function of the rate of the received angle parameters to all the angle parameters. Here, the bit depth is 8 bits, and wireless channel SNR is 20 dB.

D. Effect of Column Order of GFT Basis Matrix

This section further discusses the effect of the column order of the GFT basis matrix on the 3D reconstruction quality and communication overhead. For the comparison, the proposed scheme adopts the different quantization methods and column orders: uniform CDF-based [15], exponential CDF-based, power CDF-based [16], and two-phase hyper-exponential CDF-based quantization after the GFT basis vectors are sorted in the ascending/descending order of the graph frequency.

Fig. 9 shows the symmetric MSE performance of the 3D coordinates for each method at a wireless channel SNR of 20 dB under a bit depth of 6, 8, 10, and 12 bits, respectively. Here, the parameter of the power CDF is set to 0.50. The ascending order of the GFT basis vectors performs well, irrespective of the bit depths and quantization methods. Especially, the two-phase hyper-exponential CDF-based quantization for the ascending order of the GFT basis vectors outperforms our preliminary method [16].

Fig. 10 shows the quantization error under the ascending and descending orders of the GFT basis vectors in each frequency index to discuss the effect of the column order of the GFT basis matrix in detail. Here, bit depth is 8 bits. The descending order causes a significant quantization error on column vectors corresponding to low-frequency GFT coefficients, resulting in a significant distortion in the reconstructed point cloud.

E. Effect of Bandwidth Heterogeneity

The aforementioned evaluations demonstrate that the proposed scheme performs well when each XR user experiences diverse channel quality under an identical bandwidth. This section evaluates the effect of bandwidth heterogeneity across the XR users on the wireless point cloud delivery performance.

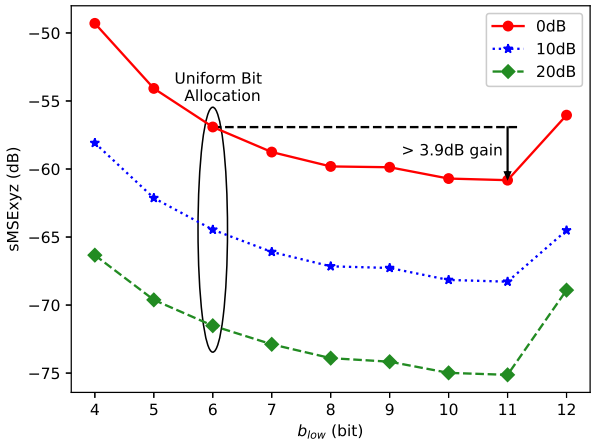


Fig. 12. Symmetric MSE of the 3D coordinates as a function of the bit depths for the angle parameters corresponding to lower-frequency GFT coefficients. The average bit depth is 6 bits.

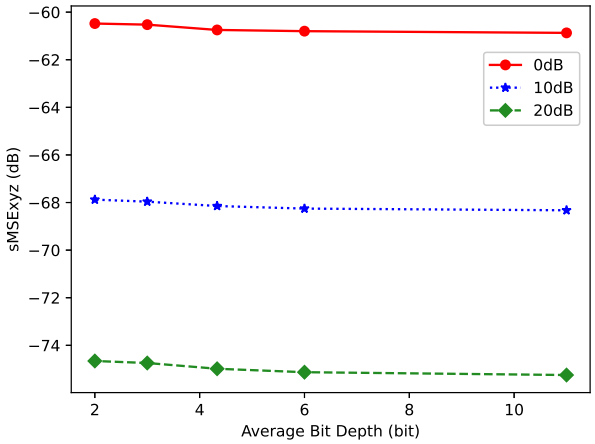


Fig. 13. Symmetric MSE of the 3D coordinates as a function of the average bit depth over the entire angle parameters. The bit depth for lower-frequency b_{low} is 11 bits whereas the bit depth for higher-frequency b_{high} is 1 bit.

Fig. 11 shows the 3D reconstruction quality as a function of the rate of the received angle parameters to all the angle parameters. Here, a bit depth is 8 bits, and a wireless channel SNR is 20 dB. The results show the following key observations:

- The proposed scheme can keep the best 3D reconstruction quality even though an XR user only receives 19% of the angle parameters.
- In our preliminary schemes [15], [16], the 3D reconstruction quality of an XR user is significantly low even when the XR user receives 99% of the angle parameters due to the loss of the low-frequency information.
- The uniform quantization can provide an adaptive 3D reconstruction quality according to the available bandwidth of the XR user. However, the reconstruction quality is lower than the proposed scheme.

F. Discussion on Bit Adaptation

This section discusses the effect of adaptive bit assignment for the angle parameters to further improve the 3D reconstruction quality without sacrificing an extra communication overhead. Fig. 12 shows the 3D reconstruction quality under the different bit depths for the angle parameters corresponding to the lower-frequency GFT coefficients and wireless channel SNRs. Here, half of the angle parameters correspond to the lower-frequency GFT coefficients, and an average bit depth is 6 bits over the entire angle parameters. In other words, the results of 6 represent uniform bit allocation for all the angle parameters.

The best 3D reconstruction quality is when 11 bits and 1 bit are assigned to the angle parameters corresponding to low-frequency and high-frequency GFT coefficient for each wireless SNR of 0 dB, 10 dB, and 20 dB, respectively. For example, the adaptive bit allocation achieves quality improvement by up to 3.9 dB compared with the uniform bit allocation at wireless SNR of 0 dB.

Fig. 13 shows the 3D reconstruction quality as a function of the average bit depths over the entire angle parameters. Here, 10 % through 100 % of angle parameters are considered the higher-priority angle parameters to adjust the average bit depth. The bit depths for higher-priority angle parameters b_{low} and lower-priority angle parameters b_{high} are set to 11 bits and 1 bit, respectively. The quality degradation against the uniform 11-bit allocation across all the angle parameters is only 0.5 dB irrespective of wireless SNRs, even though the average bit depth is 2 bits. In this case, the adaptive bit assignment achieves up to 81.8 % overhead reduction compared with the uniform bit allocation.

G. Visual Quality

Finally, this section discusses the visual quality of the proposed scheme to clarify further the effect of the non-uniform quantization and stacked transmission on the reconstructed 3D point clouds. We adopt 2D projected point cloud quality [17] to discuss the visual quality in detail. Specifically, we project original and decoded 3D point clouds on the 2D view plane from particular angles and then measure the performance of peak signal noise-to-ratio (PSNR) and structural similarity (SSIM) by using the original and decoded 2D images. The 2D projected metrics may be more relevant to discuss the perceptual distortion for practical holographic display systems compared with 3D symmetric MSE. PSNR is defined as follows:

$$\mathsf{PSNR} = 10\log_{10}\frac{(2^L-1)^2}{\epsilon_{\mathsf{MSE}}}, \tag{22}$$

where L is the number of bits used to encode pixel luminance (typically eight bits), and ε_{MSE} is the MSE between all pixels of the decoded and the original 2D images. SSIM can predict the perceived quality of the original and decoded point cloud images. Larger values of SSIM close to 1 indicate higher perceptual similarity between original and decoded images.

Figs. 14 (a)-(e) present the visual quality at a wireless channel SNR of 25 dB with a different number of received

(a) Original

(b) Descending, Uniform [15] (100% received) $\mathsf{sMSE}_{\mathsf{xyz}}$: -56.25 dB $\mathsf{sMSE}_{\mathsf{yuv}}$: -28.37 dB PSNR: 23.02 dB SSIM: 0.941

(c) Proposed (100% received) $\mathsf{sMSE}_{\mathsf{xyz}}$: -62.83 dB $\mathsf{sMSE}_{\mathsf{yuv}}$: -35.62 dB PSNR: 26.17 dB SSIM: 0.947

(d) Descending, Uniform [15] (19% received) $\mathsf{sMSE}_{\mathsf{xyz}}$: -10.33 dB $\mathsf{sMSE}_{\mathsf{yuv}}$: -9.04 dB

(e) Proposed (19% received) $\mathsf{sMSE}_{\mathsf{xyz}}$: -38.79 dB $\mathsf{sMSE}_{\mathsf{yuv}}$: -27.81 dB

Fig. 14. Snapshots at wireless channel SNR of 25 dB under the received ratio for the angle parameters when allocating angle parameter ψ to 8 bits.

angle parameters. Furthermore, Figs. 14 (b) and (c) show the average PSNR and SSIM performance across the different 2D projected angles. The original and decoded point clouds are projected onto a 2D plane at different angles horizontally rotated by 5-degree steps. Our preliminary scheme [15] and the proposed scheme are used for comparison. Both schemes use a bit depth b of 8 bits. It is observed that the proposed scheme may achieve better perceptual quality compared with the preliminary scheme, irrespective of the 2D projected angles. In particular, the proposed scheme can keep the structure of the point clouds when the low-end XR users receive a limited number of angle parameters, whereas the preliminary scheme fails to reconstruct the shape.

VII. Conclusion

This paper proposed a novel delivery scheme for untethered XR users to maintain better 3D reconstruction quality. The proposed scheme integrates GFT with analog modulation to deal with channel quality diversity. The integration ensures that the reconstruction quality of the 3D points is monotonically increased according to the quality of the wireless channel. The proposed scheme designs the prioritized transmissions of the quantized angle parameters obtained via Givens rotation and non-uniform and adaptive quantization to deal with bandwidth diversity. It yields better quality than existing digital-based and graph-based schemes, even when the channel quality fluctuates for each XR user. Furthermore, it achieved a significant reduction of communication overhead of up to 40% compared with the existing graph-based scheme.

There are several future directions in our study. The first direction is the extension for the point cloud video. The inter-frame redundancy between the point cloud video frames [37] may cause rate and power losses in wireless point cloud delivery. Typical solutions are predictive coding in the graph-domain [37]–[39] and adaptive graph signal processing [40] to remove the inter-frame redundancy in the graph-domain signals. The extension of the proposed scheme considering inter-frame redundancy reduction is left as future work.

The second direction is to realize prioritized operations based on the perspective of XR users to further enhance the visual quality. XR users may observe the point cloud from a certain perspective [41] in point cloud playback. We will extend the power allocation considering the user's perspective to enhance the visual quality as a future work as the unequal power allocation in the proposed scheme minimized the MSE between the original and reconstructed point clouds.

Acknowledgment

T. Fujihashi's work was partly supported by JSPS KAKENHI Grant Number JP22H03582.

References

[1] R. Mekuria and L. Bivolarsky, "Overview of the MPEG activity on point cloud compression," in *Data Compression Conference*, 2016, p. 620.

[2] 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.

[3] 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.

[4] "Draco 3d data compression." [Online]. Available: https://google.github.io/draco/

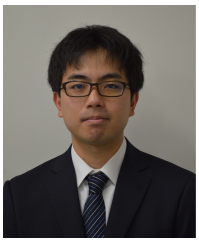
[5] D. B. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, "An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC)," *APSIPA Transactions on Signal and Information Processing*, vol. 9, pp. 1–17, 2020.

[6] F. Shen and W. Gao, "A rate control algorithm for video-based point cloud compression," in *International Conference on Visual Communications and Image Processing*, 2021, pp. 1–5.

[7] 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.

[8] 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.
[9] 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.
[10] S. Gu, J. Hou, H. Zeng, and H. Yuan, "3D point cloud attribute compression via graph prediction," *IEEE Signal Processing Letters*, vol. 27, pp. 176–180, 2020.
[11] T. Fujihashi, T. Koike-Akino, T. Watanabe, and P. V. Orlik, "High-quality soft video delivery with GMRF-based overhead reduction," *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 473–483, Feb. 2018.
[12] M. A. Sadrabadi, A. Khandani, and F. Lahouti, "Channel feedback quantization for high data rate MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3335–3338, 2006.
[13] J. C. Roh and B. D. Rao, "Efficient feedback methods for MIMO channels based on parameterization," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 282–292, 2007.
[14] 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.
[15] ——, "Overhead reduction in graph-based point cloud delivery," in *IEEE International Conference on Communications*, 2020, pp. 1–7.
[16] S. Ueno, T. Fujihashi, T. Koike-Akino, and T. Watanabe, "Overhead reduction for graph-based point cloud delivery using non-uniform quantization," in *IEEE International Conference on Consumer Electronics*, 2022, pp. 1–6.
[17] 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. PP, no. 99, pp. 1–13, 2021.
[18] E. Pavez, B. Girault, A. Ortega, and P. A. Chou, "Region adaptive graph Fourier transform for 3D point clouds," 2020. [Online]. Available: https://github.com/STAC-USC/RA-GFT.
[19] C. T. Duong, T. D. Hoang, H. H. Dang, Q. V. H. Nguyen, and K. Aberer, "On node features for graph neural networks," *arXiv e-prints*, pp. 1–6, Nov. 2019.
[20] T. Fujihashi, T. K. 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.
[21] 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.
[22] J. Wu, J. Wu, H. Cui, C. Luo, X. Sun, and F. Wu, "DAC-Mobi: Data-assisted communications of mobile images with cloud computing support," *IEEE Transactions on Multimedia*, vol. 18, no. 5, pp. 893–904, 2016.
[23] T. Fujihashi, T. Koike-Akino, and T. Watanabe, "Soft delivery: Survey on a new paradigm for wireless and mobile multimedia streaming," *arXiv preprint arXiv:2111.08189*, 2021.
[24] H. Cui, R. Xiong, C. Luo, Z. Song, and F. Wu, "Denoising and resource allocation in uncoded video transmission," *IEEE Journal on Selected Topics in Signal Processing*, vol. 9, no. 1, pp. 102–112, 2015.
[25] X. Fan, F. Wu, D. Zhao, and O. C. Au, "Distributed wireless visual communication with power distortion optimization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 6, pp. 1040–1053, 2013.
[26] X. L. Liu, W. Hu, C. Luo, Q. Pu, F. Wu, and Y. Zhang, "ParCast+: Parallel video unicast in MIMO-OFDM WLANs," *IEEE Transactions on Multimedia*, vol. 16, no. 7, pp. 2038–2051, 2014.
[27] Y. Gui, L. Hancheng, F. Wu, and C. W. Chen, "LensCast: Robust wireless video transmission over mmWave MIMO with lens antenna array," *IEEE Transactions on Multimedia*, vol. PP, no. 99, pp. 1–16, 2020.
[28] X.-w. Tang, X.-l. Huang, and F. Hu, "QoE-driven UAV-enabled pseudo-analog wireless video broadcast: A joint optimization of power and trajectory," *IEEE Transactions on Multimedia*, vol. 23, pp. 2398–2412, 2021.
[29] J. Shen, L. Yu, L. Li, and H. Li, "Foveation-based wireless soft image delivery," *IEEE Transactions on Multimedia*, vol. 20, no. 10, pp. 2788–2800, 2018.
[30] X.-W. Tang, X.-L. Huang, F. Hu, and Q. Shi, "Human-perception-oriented pseudo analog video transmissions with deep learning," *IEEE Transactions on Vehicular Technology*, pp. 1–14, may 2020.
[31] T. Fujihashi, T. Koike-Akino, T. Watanabe, and P. V. Orlik, "FreeCast: Graceful free-viewpoint video delivery," *IEEE Transactions on Multimedia*, vol. PP, no. 99, pp. 1–11, 2019.
[32] L. Luo, T. Yang, C. Zhu, Z. Jin, and S. Tang, "Joint texture/depth power allocation for 3-D video SoftCast," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 2973–2984, dec 2019.
[33] T. Fujihashi, M. Kobavashi, K. Endo, S. Saruwatari, S. Kobayashi, and T. Watanabe, "Graceful quality improvement in wireless 360-degree video delivery," in *IEEE Global Communications Conference*, 2018, pp. 1–7.
[34] J. Zhao, R. Xiong, and J. Xu, "OmniCast: Wireless pseudo-analog transmission for omnidirectional video," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 58–70, mar 2019.
[35] Y. Lu, T. Fujihashi, S. Saruwatari, and T. Watanabe, "360Cast: Foveation-based wireless soft delivery for 360-degree video," in *IEEE International Conference on Communications*, jun 2020, pp. 1–6.
[36] P. A. Chou, E. Pavez, R. L. de Queiroz, and A. Ortega, "Dynamic polygon clouds: Representation and compression for VR/AR," Microsoft Research Technical Report, Tech. Rep., 2017.
[37] A. L. Souto and R. L. de Queiroz, "On predictive RAHT for dynamic point cloud coding," in *IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 2701–2705.
[38] 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, 2021.
[39] D. Thanou, P. A. Chou, and P. Frossard, "Graph-based compression of dynamic 3D point cloud sequences," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1765–1778, 2016.
[40] P. Di Lorenzo, P. Banelli, E. Isufi, S. Barbarossa, and G. Leus, "Adaptive graph signal processing: Algorithms and optimal sampling strategies," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3584–3598, 2018.
[41] B. Han, Y. Liu, and F. Qian, "ViVo: visibility-aware mobile volumetric video streaming," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–13.

Soushi Ueno received his Bachelor of Engineering from Osaka University in 2020 and M.S. degree in information science and technology in 2022. He joined Daikin Industries, Ltd., Japan. His research interests include point cloud compression and delivery.



Takuya Fujihashi (M'16) received the B.E. degree in 2012 and the M.S. degree in 2013 from Shizuoka University, Japan. In 2016, he received Ph.D. degree from the Graduate School of Information Science and Technology, Osaka University, Japan. He was an assistant professor at the Graduate School of Science and Engineering, Ehime University between Jan. 2017 and Mar. 2019. He is currently an assistant professor at the Graduate School of Information Science and Technology, Osaka University, Japan since Apr. 2019. He was research fellow (PD) of Japan Society for the Promotion of Science in 2016. From 2014 to 2016, he was research fellow (DC1) of Japan Society for the Promotion of Science. From 2014 to 2015, he was an intern at Mitsubishi Electric Research Labs. (MERL) working with the Electronics and Communications group. He selected one of the Best Paper candidates in IEEE ICME (International Conference on Multimedia and Expo) 2012. He received Young Professional Award from IEEE Kansai Chapter in 2021. His research interests are in the area of video compression and communications, with a focus on multi-view video coding and streaming over high and low quality networks.

Toshiaki Koike-Akino (M'05–SM'11) received the B.S. degree in electrical and electronics engineering, M.S. and Ph.D. degrees in communications and computer engineering from Kyoto University, Kyoto, Japan, in 2002, 2003, and 2005, respectively. During 2006–2010 he was a Postdoctoral Researcher at Harvard University, and is currently a Distinguished Research Scientist at Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA. He received the YRP Encouragement Award 2005, the 21st TELECOM System Technology Award, the 2008 Ericsson Young Scientist Award, the IEEE GLOBECOM'08 Best Paper Award in Wireless Communications Symposium, the 24th TELECOM System Technology Encouragement Award, and the IEEE GLOBECOM'09 Best Paper Award in Wireless Communications Symposium. He is a Fellow of Optica.

Takashi Watanabe (S'83–M'87) is a Professor of Graduate School of Information Science and Technology, Osaka University, Japan since 2013. He received his B.E. M.E. and Ph.D. degrees from Osaka University, Japan, in 1982, 1984 and 1987, respectively. He joined Faculty of Engineering, Tokushima University in 1987 and moved to Faculty of Engineering, Shizuoka University in 1990. He was a visiting researcher at University of California, Irvine from 1995 through 1996. He has served on many program committees for networking conferences, IEEE, ACM, IPSJ, IEICE. His research interests include mobile networking, ad hoc sensor networks, IoT/M2M networks, intelligent transport systems, specially MAC and routing. He is a member of IEEE, IPSJ and IEICE.