HoloCast: Graph Signal Processing for Graceful Point Cloud Delivery

Fujihashi, T.; Koike-Akino, T.; Watanabe, T.; Orlik, P.V.

TR2019-026   June 05, 2019

Abstract

In conventional point cloud delivery, a sender uses octree-based digital video compression to stream threedimensional (3D) points and the corresponding color attributes over band-limited links, e.g., wireless channels, for 3D scene reconstructions. However, the digital-based delivery schemes have an issue called cliff effect, where the 3D reconstruction quality is a step function in terms of wireless channel quality. We propose a novel scheme of point cloud delivery, called HoloCast, to gracefully improve the reconstruction quality with the improvement of wireless channel quality. HoloCast regards the 3D points and color components as graph signals and directly transmits lineartransformed signals based on graph Fourier transform (GFT), without digital quantization and entropy coding operations. One of main contributions in HoloCast is that the use of GFT can deal with non-ordered and non-uniformly distributed multidimensional signals such as holographic data unlike conventional delivery schemes. Performance results with point cloud data show that HoloCast yields better 3D reconstruction quality compared to digital-based methods in noisy wireless environment.

IEEE International Conference on Communications (ICC)

HoloCast: Graph Signal Processing for Graceful Point Cloud Delivery

Takuya Fujihashi[†*], Toshiaki Koike-Akino[†], Takashi Watanabe[‡], and Philip V. Orlik[†]

[†]Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA

[*]Graduate School of Science and Engineering, Ehime University, Ehime, Japan

[‡]Graduate School of Information and Science, Osaka University, Osaka, Japan

Abstract—**In conventional point cloud delivery, a sender uses octree-based digital video compression to stream three-dimensional (3D) points and the corresponding color attributes over band-limited links, e.g., wireless channels, for 3D scene reconstructions. However, the digital-based delivery schemes have an issue called cliff effect, where the 3D reconstruction quality is a step function in terms of wireless channel quality. We propose a novel scheme of point cloud delivery, called HoloCast, to gracefully improve the reconstruction quality with the improvement of wireless channel quality. HoloCast regards the 3D points and color components as graph signals and directly transmits linear-transformed signals based on graph Fourier transform (GFT), without digital quantization and entropy coding operations. One of main contributions in HoloCast is that the use of GFT can deal with non-ordered and non-uniformly distributed multi-dimensional signals such as holographic data unlike conventional delivery schemes. Performance results with point cloud data show that HoloCast yields better 3D reconstruction quality compared to digital-based methods in noisy wireless environment.**

I. Introduction

Holographic displays [1], [2] have emerged as attractive interface techniques for reconstructing three dimensional (3D) scenes that provide full parallax and depth information for human eyes. 3D holographic display can be widely used for many applications: entertainment, remote device operation, medical imaging, and simulated training as shown in Fig. 1. Point cloud [3] is one of data structures to reconstruct 3D scenes/objects on the holographic display [4]. Point cloud is a set of 3D points, and each point is defined by 3D coordinates, i.e., (X, Y, Z) and color attributes, i.e., (R, G, B) or (Y, U, V).

In contrast to conventional two dimensional (2D) images, 3D points in point cloud data are not ordered and are non-uniformly distributed in space. One of major issues in point cloud delivery is how to compress and send such numerous and irregular structure of 3D points while maintaining high 3D reconstruction quality on displays. For example, when the number of 3D points is 800,000, the amount of traffic without any compression is approximately 38 Mbits [5]. Large traffic causes low reconstruction quality in point cloud delivery over limited data rate links, especially, wireless communications.

For point cloud compression over wireless links, conventional encoders, such as popular Point Cloud Library (PCL) [6], [7], use octree decomposition, prediction, quantization, and entropy coding. Specifically, a sender first decomposes point cloud into multiple 3D point sets [8] and takes quantization and entropy coding for each point set to
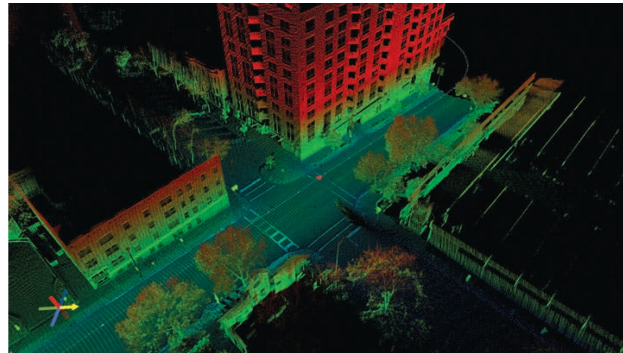
(a) light detection and ranging (LIDAR) scenario [9]

(b) AR/VR scenario [10]

Fig. 1. Examples of holographic applications.

generate the compressed bitstream for transmissions. Here, the compression rate of the bitstream is adaptively selected according to the wireless channel quality. After the compression, the transmission part uses a channel coding and digital modulation scheme to reliably transmit the compressed bitstream over wireless channels. High-quality transmissions of point clouds over wireless links can realize immersive video applications such as virtual reality and augmented reality on wireless devices as shown in Fig. 2.

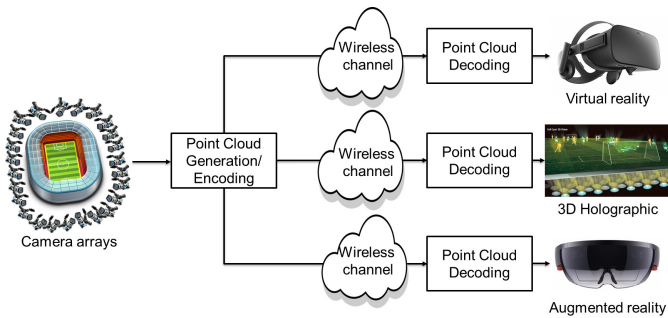However, the conventional schemes of point cloud delivery

Fig. 2. Wireless point cloud delivery for immersive video applications.

suffer from the following problems due to the wireless channel unreliability. First, the encoded bitstream is highly vulnerable for bit errors [11]. When the channel signal-to-noise ratio (SNR) falls under a certain threshold, possible few bit errors occurred in the bitstream during communications can cause a synchronization problem in point cloud decoding. As a result, the display does not reconstruct 3D scenes, and thus the reconstruction quality degrades significantly. This phenomenon is called cliff effect [12]. Second, the reconstruction quality does not improve even when the wireless channel quality is improved unless an adaptive rate control of source and channel coding is performed in real-time according to the rapid fading channels. This is called leveling effect. Finally, quantization is a lossy process and its distortion cannot be recovered at the receiver.

As mentioned above, conventional point cloud transmissions have two challenging issues over wireless links: 1) cliff effect and 2) leveling effect. To overcome these issues, we propose a new point cloud transmission scheme to reconstruct 3D scenes in high quality on holographic displays. The key idea of this scheme is skipping nonlinear operations, i.e., quantization and entropy coding, in point cloud coding. For high-quality delivery, this study considers 3D points as vertices in a graph $\mathcal{G}$, with edges between nearby vertices to deal with irregular structure motivated by [13], [14]. Each point p has attributes of 3D coordinates and color components, and those attributes are regarded as signals residing on the vertices of the graph. The proposed scheme takes graph Fourier transform (GFT) [15], [16] for each attribute in graph signals to compact the signal power, whose output is then scaled and directly mapped to transmission signals without relying on digital modulation schemes. The advantage of this modification lies in a fact that the point distortion due to communication noise is proportional to the magnitude of the noise, resulting into graceful reconstruction quality according to the wireless channel quality, without any cliff effect and leveling effect. We demonstrate that the proposed point cloud delivery scheme achieves graceful reconstruction quality with the improvement of wireless channel quality and better reconstruction performance compared to the conventional digital-based schemes. For example, HoloCast achieves 39.45 dB and 32.48 dB

improvement in the attributes of the 3D coordinates and color components, respectively, in terms of mean squared error (MSE) compared with the digital-based delivery schemes.

Related Works and Our Contributions: Soft video delivery schemes have been recently proposed for multi-dimensional ordered video signals in [17]–[20]. For example, SoftCast [17] was designed for 3D ordered video signals to realize graceful video delivery. They skip quantization and entropy coding, and uses 3D discrete-cosine transform (DCT) and analog modulation, which maps DCT coefficients directly to transmission signals, to ensure that the received video quality is proportional to wireless channel quality. FoveaCast [18] considers the foveation characteristic of human vision into soft video delivery of 2D ordered video signals to achieve higher visual perceptual quality. FreeCast [20] extended the soft video delivery towards 5D ordered multi-view video plus depth (MVD) signals. They use 5D-DCT for decorrelation and directly send the coefficients to realize graceful quality improvement with the improvement of wireless channel quality.

Our study realizes soft coding and decoding for point cloud delivery. Although existing schemes of soft video delivery deal with ordered and uniformly distributed video signals, point cloud delivery needs to handle non-ordered and non-uniformly distributed points in coding and decoding. To this end, the proposed scheme has the following major contributions:

- We regard point clouds as graph signals with the attributes of 3D coordinates and color components to deal with irregular structure of holographic data formats.
- We introduce GFT and analog modulation for graph signals to exploit correlations in graph-domain for performance improvement.
- We discuss an impact of graph Laplacian variants and adjacency hyperparameters on 3D scene reconstruction quality.
- We demonstrate that GFT-based HoloCast achieves graceful 3D reconstruction quality with a significant performance improvement over digital-based point cloud delivery.

II. HoloCast: Graceful Point Cloud Delivery

The objectives of our study are 1) to prevent cliff effect in 3D scene reconstruction and 2) to gracefully improve reconstruction quality with the improvement of wireless channel quality.

Fig. 3 shows the overview of proposed HoloCast. The encoder first performs GFT for 3D points and the corresponding colors, i.e., graph signals. The GFT coefficients are then scaled and analog-modulated according to the signal power information for wireless transmissions. Next, the encoder sends the analog modulated symbols to the receiver over a wireless channel, which is often impaired with additive white Gaussian noise (AWGN) and time-varying fading. At the receiver side, the decoder uses minimum mean-square error (MMSE) filter to obtain the transmitted GFT coefficients. The decoder finally takes inverse GFT to reconstruct 3D coordinates and color components for display.
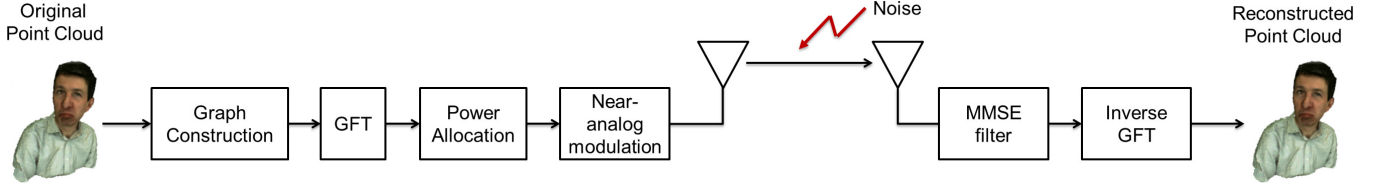
Fig. 3. Overview of HoloCast for graceful point cloud delivery.

A. Encoder

We first represent 3D points and color components using a weighted and undirected graph $\mathcal{G} = (\boldsymbol{V}, \mathcal{E}, \boldsymbol{W})$ where $\boldsymbol{V}$ and $\mathcal{E}$ are the vertex and edge sets of $\mathcal{G}$, respectively. $\boldsymbol{W}$ is an adjacency matrix having positive edge weights and the (i, j)th entry $\boldsymbol{W}_{i,j}$ represents the weight of an edge connecting vertices i and j. For the graph $\mathcal{G} = (\boldsymbol{V}, \mathcal{E}, \boldsymbol{W})$, we consider the attributes of the point cloud, i.e., the 3D coordinates $\boldsymbol{p} = [\mathsf{x}, \mathsf{y}, \mathsf{z}]^{\mathrm{T}} \in \mathbb{R}^{3 \times N}$ and the color components $\boldsymbol{c} = [y, u, v]^{\mathrm{T}} \in \mathbb{R}^{3 \times N}$ as signals that reside on the vertices in the graph (N is the number of vertices). From the attributes, each weight $\boldsymbol{W}_{i,j}$ can be calculated, e.g., by the Gaussian kernel as follows:

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

where $\boldsymbol{p}_i$ represents the 3D coordinates of point i and κ is a hyperparameter. In HoloCast, we use either the sample variance or the standard deviation of distances across all the points for the hyperparameter κ. A sender then transforms the graph signals into spectral representation using GFT. The GFT is defined through the graph Laplacian operator $\boldsymbol{L}$ using edge weight matrix $\boldsymbol{W}$ and degree matrix $\boldsymbol{D}$, where $\boldsymbol{D}$ is the diagonal degree matrix whose ith diagonal element is equal to the sum of the weights of all the edges incident to vertex i. Specifically, the diagonal matrix is represented as:

$$\boldsymbol{D}_{i,j} = \begin{cases} \sum_{n=1}^{N} \boldsymbol{W}_{i,n}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Based on the degree matrix, we can calculate some variants of the graph Laplacian matrix [21]:

$$\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}, \tag{3}$$

$$\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-1/2} \boldsymbol{W} \boldsymbol{D}^{-1/2}, \tag{4}$$

$$\boldsymbol{L} = \boldsymbol{D}^{-1} \boldsymbol{W}, \tag{5}$$

$$\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-1} \boldsymbol{W}, \tag{6}$$

where $\boldsymbol{I}$ denotes an identity matrix of proper dimension. We refer to each graph Laplacian matrix as regular, normalized, transition, and random-walk Laplacian, respectively. We will discuss the impact of those graph Laplacian matrices on the delivery quality in Section III-C. In general, the graph Laplacian is a real symmetric matrix that has a complete set of orthonormal eigenvectors with corresponding nonnegative eigenvalues. To obtain the eigenvectors and eigenvalues, the eigen decomposition of the Laplacian matrix is performed as:

$$\boldsymbol{L} = \boldsymbol{\Phi} \boldsymbol{\Delta} \boldsymbol{\Phi}^{-1}, \tag{7}$$

where $\boldsymbol{\Phi}$ is the eigenvectors matrix and $\boldsymbol{\Delta}$ is a diagonal matrix containing the eigenvalues.[1] The multiplicity of the smaller eigenvalue indicates the number of connected components of the graph. The GFT coefficients of each attribute of graph signals $\boldsymbol{f} \in R^N$ are obtained by multiplying the graph-based transform basis matrix by the corresponding attribute vector as follows:

$$\boldsymbol{s} = \boldsymbol{f} \boldsymbol{\Phi}, \tag{8}$$

where $\boldsymbol{s}$ is a vector of GFT coefficients corresponding to the graph signals of $\boldsymbol{f}$. After power allocation for each GFT coefficient, the GFT coefficients are mapped to I (in-phase) and Q (quadrature-phase) components for analog wireless transmissions.

Let x_i denote the ith analog-modulated symbol, which is the ith GFT coefficient s_i of all the attributes scaled by a factor of g_i for noise reduction as follows:

$$x_i = g_i \cdot s_i. \tag{9}$$

The optimal scale factor g_i is obtained by minimizing the MSE under the power constraint with a total power budget of P as follows:

$$\min_{\{g_i\}} \quad \mathsf{MSE} = \mathbb{E}\left[(s_i - \hat{s}_i)^2\right] = \sum_{i}^{N} \frac{\sigma^2 \lambda_i}{g_i^2 \lambda_i + \sigma^2}, \tag{10}$$

$$\text{s.t.} \quad \frac{1}{N} \sum_{i}^{N} g_i^2 \lambda_i = P, \tag{11}$$

where $\mathbb{E}[\cdot]$ denotes expectation, $\hat{s}_i$ is a receiver estimate of the transmitted GFT coefficient, λ_i is the power of the ith GFT coefficient, N is the number of GFT coefficients, and σ^2 is a receiver noise variance. As shown in [17], the near-optimal solution is expressed as

$$g_i = \lambda_i^{-1/4} \sqrt{\frac{NP}{\sum_j \sqrt{\lambda_j}}}. \tag{12}$$

[1] For non-diagonalizable graph Laplacian matrix, the singular value decomposition (SVD) is instead used to express as $\boldsymbol{L} = \boldsymbol{\Psi} \boldsymbol{\Delta} \boldsymbol{\Phi}^{-1}$ where $\boldsymbol{\Psi}$, $\boldsymbol{\Delta}$ and $\boldsymbol{\Phi}$ denote left singular vectors matrix, diagonal matrix containing singular values, and right singular vectors matrix, respectively. In this case, we use the right singular vectors of $\boldsymbol{\Phi}$ as the graph-based transform basis matrix $\boldsymbol{\Phi}$.

B. Decoder

Over the wireless links, the receiver obtains the received symbol, which is modeled as follows:

$$y_i = x_i + n_i, \tag{13}$$

where y_i is the ith received symbol and n_i is an effective AWGN with a variance of σ^2 (which is already normalized by wireless channel strength in the presence of fading attenuation). The GFT coefficients are extracted from I and Q components via an MMSE filter [17]:

$$\hat{s}_i = \frac{g_i \lambda_i}{g_i^2 \lambda_i + \sigma^2} \cdot y_i. \tag{14}$$

The decoder then reconstructs corresponding graph signals $\hat{\boldsymbol{f}}$, i.e., attributes of 3D coordinates and color components, by taking the inverse GFT for the filtered GFT coefficients in each attribute $\hat{s}$ as follows:

$$\hat{\boldsymbol{f}} = \hat{\boldsymbol{s}} \boldsymbol{\Phi}^{-1}. \tag{15}$$

C. Analog Compression for Limited Bandwidth

The previous designs assume that the sender has enough bandwidth to transmit all the coefficients in the spectral domain over the wireless medium. If the available bandwidth and/or time resources are restricted for wireless channel use, it has to selectively transmit the coefficients to fit the available bandwidth. For such cases, our scheme sorts the coefficients in descending order of the power and picks higher-power coefficients to fill the bandwidth. When the sender discards a coefficient, the receiver regards the discarded coefficient as zero. As a result, a sort of data compression can be accomplished even for analog-based video delivery. Even when some coefficients are discarded to reduce the amount of data, the receiver can still achieve a graceful video quality until reaching the distortion limit due to the compression.

III. PERFORMANCE EVALUATION

A. Simulation Settings

Performance Metric: We evaluate the reconstruction quality of point cloud delivery in terms of the symmetric MSE based on [22] in each attribute of 3D coordinates $\boldsymbol{p}$ and color components $\boldsymbol{c}$. The symmetric MSE of the 3D coordinates, $\mathsf{sMSE_{xyz}}$, can be obtained as follows:

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

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

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

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

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

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

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

$$d(\boldsymbol{c}_{\mathrm{org}} \to \boldsymbol{c}_{\mathrm{dec}}) = \frac{1}{N} \sum_{\boldsymbol{c} \in \boldsymbol{c}_{\mathrm{org}}} \left(\left\| \boldsymbol{c} - \boldsymbol{c}_{\mathrm{dec}}(\boldsymbol{p}'_{\mathrm{min}}) \right\|_2^2 \right),$$
$$\boldsymbol{p}'_{\mathrm{min}} = \arg \min_{\boldsymbol{p}' \in \boldsymbol{p}_{\mathrm{dec}}} \left\| \boldsymbol{p}_{\mathrm{org}} - \boldsymbol{p}' \right\|_2^2, \tag{20}$$

$$d(\boldsymbol{c}_{\mathrm{dec}} \to \boldsymbol{c}_{\mathrm{org}}) = \frac{1}{N} \sum_{\boldsymbol{c} \in \boldsymbol{c}_{\mathrm{dec}}} \left(\left\| \boldsymbol{c} - \boldsymbol{c}_{\mathrm{org}}(\boldsymbol{p}''_{\mathrm{min}}) \right\|_2^2 \right),$$
$$\boldsymbol{p}''_{\mathrm{min}} = \arg \min_{\boldsymbol{p}'' \in \boldsymbol{p}_{\mathrm{org}}} \left\| \boldsymbol{p}_{\mathrm{dec}} - \boldsymbol{p}'' \right\|_2^2, \tag{21}$$

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

Point Cloud Dataset: We use the reference point clouds, namely, *pencil__10_0*, *pencil__9_0*, and *office1*, whose number of points N is 2,731, 6,712, and 307,200, respectively. We first focus on *pencil__10_0* to compare between HoloCast and digital-based delivery. The performance at a different number of 3D points will then be evaluated with *pencil__9_0*. In addition, we demonstrate the visual quality for the point cloud data of *office1* in Section III-E.

Wireless Settings: The received symbols are impaired by an AWGN channel. For digital-based schemes, we use a rate-$1/2$ convolutional codes with a constraint length of 10. The digital modulation formats are either quadrature phase-shift keying (QPSK), 16-ary quadrature-amplitude modulation (16QAM), or 256-ary QAM (256QAM).

Digital Point Cloud Coder: We compare HoloCast with the conventional digital-based delivery, which is based on point cloud digital compression used in PCL [6]. We consider two default profiles for compression: LOW_RES_OFFLINE_COMPRESSION_WITH_COLOR and MED_RES_OFFLINE_COMPRESSION_WITH_COLOR. Note that this digital-based PCL point cloud delivery does not exploit GFT, while there exist recent work of GFT-based digital compression schemes to improve the efficiency, e.g., [13], [14]. Because the primary objective of our paper is a preliminary demonstration of a new soft delivery technique for point cloud data, we focus on widely used PCL-based digital delivery for benchmark performance comparisons in this paper. Nevertheless, we plan to compare our HoloCast with GFT-based digital compression methods in near future.

B. HoloCast vs. Digital-based Schemes

We first evaluate the quality of HoloCast and conventional digital-based schemes. Fig. 4 (a) shows the MSE of 3D coordinate attributes $\boldsymbol{p}$ in the digital-based scheme and HoloCast as a function of wireless channel SNRs. Here, HoloCast uses
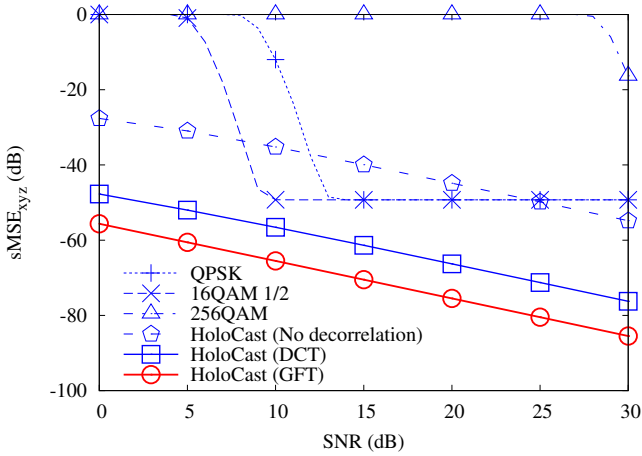
(a) 3D coordinates p
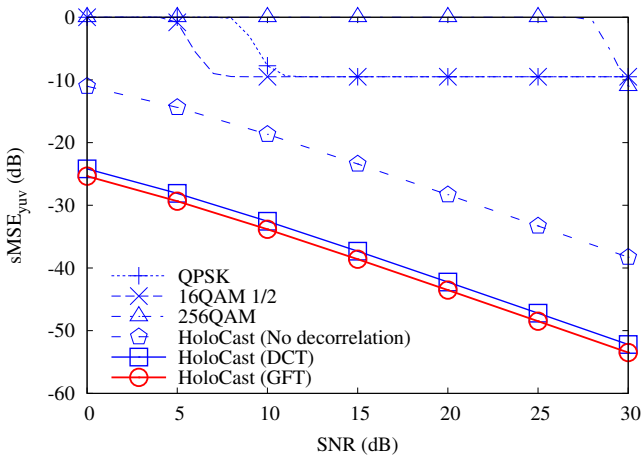
(b) Color components c

Fig. 4. MSE of 3D coordinates and color attributes in digital-based schemes and HoloCast for *pencil__10_0* ($N = 2731$).

the sample variance of point distances as the hyperparameter κ and the random-walk matrix for the graph Laplacian L. In addition, we consider additional two HoloCast schemes to demonstrate an impact of GFT on quality improvement: DCT-based decorrelation and no decorrelation. From evaluation results in Fig. 4 (a), we can see the following observations:

- HoloCast gracefully improves the reconstruction quality of 3D coordinate attributes with the improvement of wireless channel quality.
- Digital-based schemes suffer from cliff effect at low channel SNR regimes because bit errors cause synthesis errors of entropy decoding and leveling effect at high channel SNR regimes due to quantization errors.
- GFT-based HoloCast can achieve better MSE compared with DCT-based HoloCast and HoloCast w/o decorrelation. GFT can utilize correlations of non-ordered and non-uniformly distributed 3D points by treating the 3D point data as graph signals.

For example, HoloCast achieves 30.1 dB and 8.9 dB improve-

ment compared with HoloCast without decorrelation and DCT-based HoloCast on average across the channel SNRs between 0 dB and 30 dB, respectively.

Fig. 4 (b) also shows the MSE of color component attributes in the digital-based scheme and HoloCast as a function of wireless channel SNRs. Even in the attributes of the color components, HoloCast realizes graceful quality improvement with the improvement of wireless channel quality. In digital-based schemes, they have low reconstruction quality even in high channel SNR regimes. It suggests that GFT-based decorrelation has a great advantage to represent point clouds with higher reconstruction quality.

For further quality improvement in color components, we can consider the bilateral Gaussian kernel [23] in each weight $W_{i,j}$ to decorrelate color components more efficiently. Specifically, Eq. (1) will be modified as follows:

$$W_{i,j} = \exp\left(-\left(\frac{\|p_i - p_j\|_2^2}{\kappa_p} + \frac{\|c_i - c_j\|_2^2}{\kappa_c}\right)\right), \quad (22)$$

where κ_p and κ_c are hyperparameters for 3D coordinates and color components, respectively. Our evaluation in the quality of color components verified that the use of bilateral kernel in (22) instead of (1) can offer additional 6.64 dB gain on average across the channel SNRs between 0 dB and 30 dB.

### C. Impacts of Graph Laplacian Matrix and Adjacency Hyperparameters

In the previous section, we evaluated the performance of HoloCast using the random-walk graph Laplacian matrix L and variance-based hyperparameter κ. In HoloCast, different types of graph Laplacian matrix can be used to encode/decode graph signals. In addition, the weight matrix W under consideration in Eq. (1) highly depends on the value of Gaussian kernel hyperparameter κ. For the calculation of κ, the sample variance (var) or the standard deviation (std) of point distances is often used. In this section, we discuss the effects of graph Laplacian matrix and hyperparameter on the reconstruction quality in detail.

Figs. 5 (a) and (b) show the MSE of 3D coordinate and color component attributes in HoloCast, respectively, with different graph Laplacian matrix and hyperparameter as a function of wireless channel quality. The key results from these figures are summarized as follows:

- The random-walk Laplacian matrix with the hyperparameter of standard deviation achieves the best performance in 3D coordinate attributes while the regular Laplacian matrix with the hyperparameter of variance yields the best quality in the color component attributes.
- When we use the normalized and transition matrices as the graph Laplacian operator, the sender should use the standard deviation as the hyperparameter.
- Interestingly, HoloCast with the random-walk Laplacian matrix achieves better 3D coordinate reconstruction using the hyperparameter of the variance, while achieving high-quality color components reconstruction using the hyperparameter of the standard deviation.
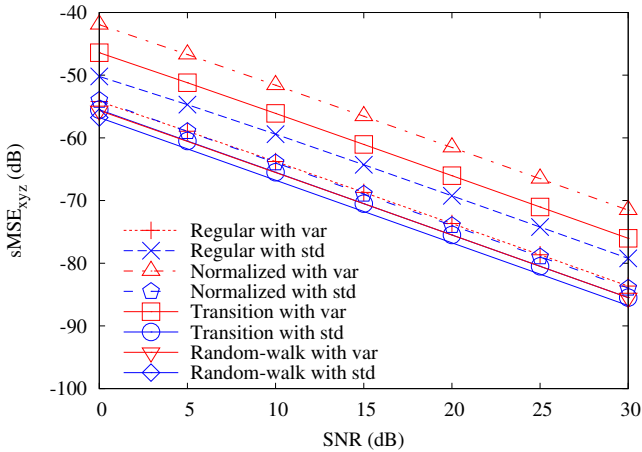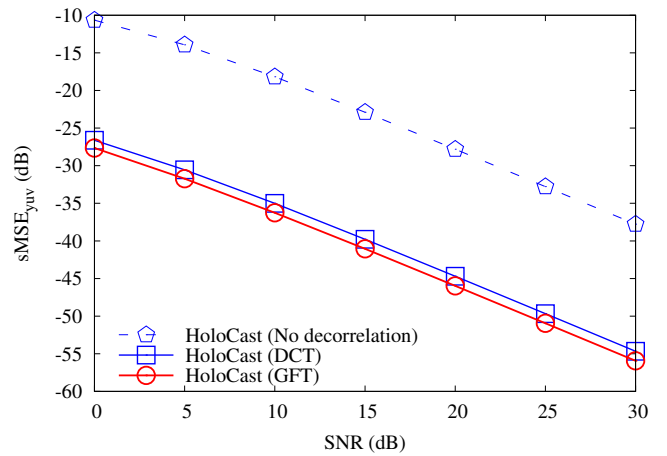
(a) 3D coordinates p
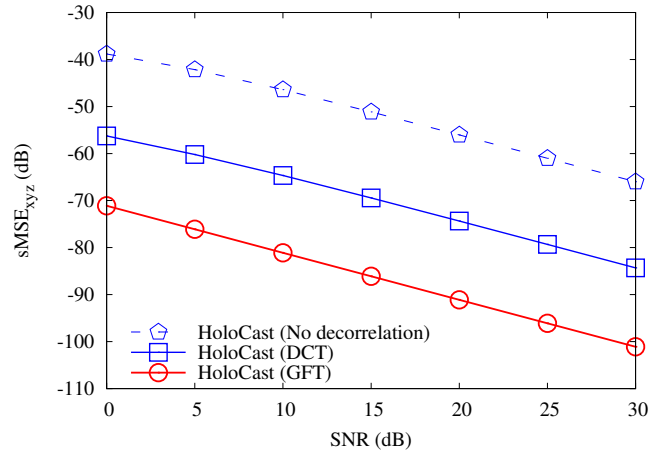
(a) 3D coordinates p
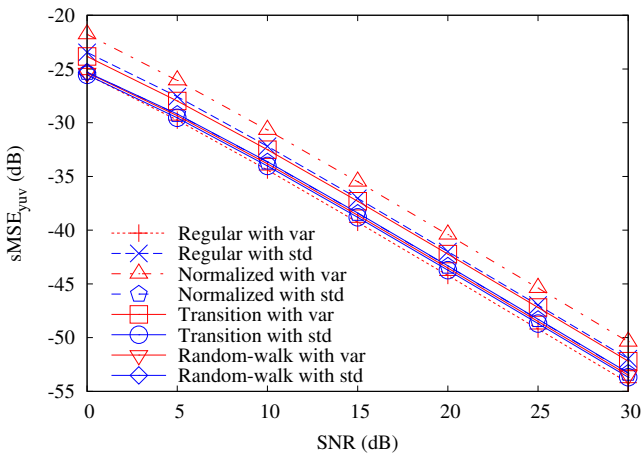
(b) Color components c

(b) Color components c

Fig. 5. MSE of 3D coordinates and color attribute in HoloCast with different graph Laplacian matrix and hyperparameters for *pencil__10_0* ($N = 2731$).

Fig. 6. MSE of 3D coordinates and color attribute in HoloCast with different decorrelation transform for *pencil__9_0* ($N = 6712$).

How to optimize weight matrix and Laplacian matrix is still an open problem. We leave rigorous analysis as future work.

D. Impacts of Different Point Clouds

Previous sections use relatively small number of 3D points $N = 2731$ to demonstrate the benefit of HoloCast. In this section, we consider a larger number of 3D points as the test point cloud to show the scalability of the proposed HoloCast. Figs. 6 (a) and (b) show the MSE of 3D coordinate and color component attributes in HoloCast, respectively, for the point cloud data of *pencil__9_0* ($N = 6712$). Compared with a small number of 3D points in Figs. 4 (a) and (b), HoloCast achieves better reconstruction quality in both 3D coordinates and color components. For example, HoloCast achieves 15.6 dB and 2.4 dB improvement on average across the channel SNRs between 0 dB and 30 dB, respectively.

E. Visual Quality

Finally, Fig. 7 compares the visual quality of HoloCast and digital-based schemes for the reference point cloud of

office1. We consider the digital-based scheme with QPSK modulation format at a channel SNR of 10 dB, where the compressed bitstream can be successfully transmitted to the receiver over wireless channels. Here, HoloCast uses DCT for the decorrelation of the point cloud. The MSE of color attributes achieved by the digital-based scheme is -12.93 dB, whereas -56.38 dB and -72.01 dB are achieved by HoloCast at wireless channel SNRs of 10 dB and 20 dB, respectively. From the snapshots, we can observe that the digital-based scheme provides lower-quality point cloud (color degradation at the door). In contrast, HoloCast gracefully improves the reconstruction quality according to available wireless channel quality. Specifically, HoloCast can reproduce a clean 3D scene with details at a higher channel SNR of 20 dB.

IV. CONCLUSION

In this paper, we proposed HoloCast to realize graceful point cloud delivery over wireless links/networks. In contrast to conventional 2D images, 3D point cloud data are not ordered

(a) Original

(b) Digital QPSK (SNR: 10dB)
sMSE$_{xyz}$: -50.75 dB
sMSE$_{yuv}$: -12.93 dB

(c) HoloCast (SNR: 10dB)
sMSE$_{xyz}$: -40.47 dB
sMSE$_{yuv}$: -56.38 dB

(d) HoloCast (SNR: 20dB)
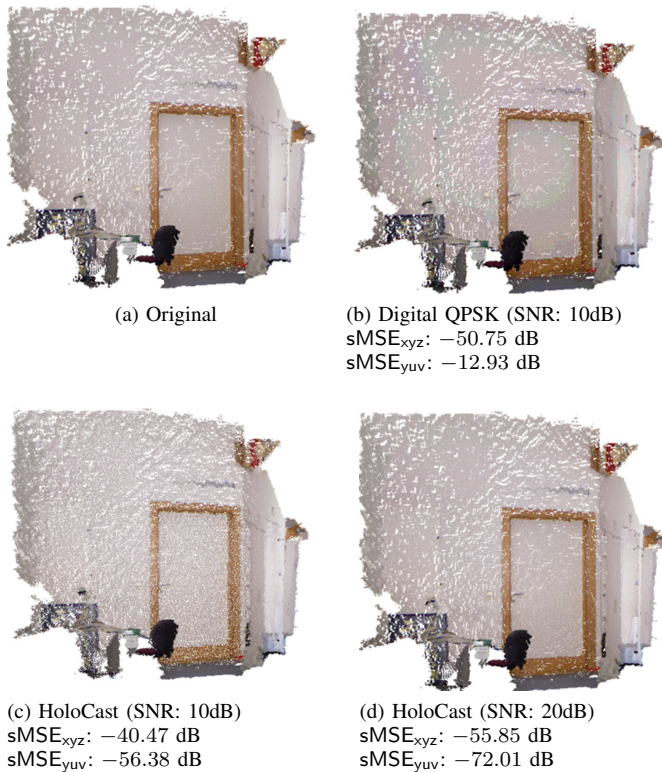sMSE$_{xyz}$: -55.85 dB
sMSE$_{yuv}$: -72.01 dB

Fig. 7. Snapshot of *office1* in digital-based and HoloCast schemes.

and are non-uniformly distributed in space. HoloCast regards the 3D points and color components as graph signals and directly transmits linear-transformed signals based on GFT. Evaluation results with several point cloud data showed that HoloCast yields better reconstruction quality even at low wireless channel SNR regimes. Feasibility study over practical experiments for various datasets with reduced amount of metadata will be conducted as future work.

ACKNOWLEDGMENT

REFERENCES

[1] P. A. Blanche, A. Bablumian, R. Voorakaranam, C. Christenson, W. Lin, T. Gu, D. Flores, P. Wang, W. Y. Hsieh, M. Kathaperumal, B. Rachwal, O. Siddiqui, J. Thomas, R. A. Norwood, M. Yamamoto, and N. Peyghambarian, "Holographic three-dimensional telepresence using large-area photorefractive polymer," *Nature*, vol. 468, no. 7320, pp. 80–83, 2010.

[2] H. Yu, K. Lee, J. Park, and Y. Park, "Ultrahigh-definition dynamic 3D holographic display by active control of volume speckle fields," *Nature Photonics*, vol. 11, no. 3, pp. 186–192, 2017.

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

[4] P. Su, W. Cao, J. Ma, B. Cheng, X. Liang, L. Cao, and G. Jin, "Fast computer-generated hologram generation method for three-dimensional point cloud model," *Journal of Display Technology*, vol. 12, no. 12, pp. 1688–1694, 2016.

[5] M. Preda, "Point cloud compression in MPEG," 2017.

[6] 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.

[7] 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.

[8] R. schnabel and R. Klein, "Octree-based point-cloud compression," in *Eurographics Symposium on Point-Based Graphics*, 2006, pp. 111–121.

[9] B. Schwarz, "Lidar: Mapping the world in 3D," *Nature Photonics*, vol. 4, no. 7, pp. 429–430, 2010.

[10] K. Rematas, I. Kemelmacher-Shlizerman, B. Curless, and S. Seitz, "Soccer on your tabletop," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1–10.

[11] S. Pudlewski, N. Cen, Z. Guan, and T. Melodia, "Video transmission over lossy wireless networks: A cross-layer perspective," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 1, pp. 6–21, 2015.

[12] 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.

[13] 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.

[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. PP, no. 99, pp. 1–16, 2018.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] D. He, C. Luo, F. Wu, and W. Zeng, "Swift: A hybrid digital-analog scheme for low-delay transmission of mobile stereo video," in *ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, Cancun, Mexico, nov 2015, pp. 327–336.

[20] 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.

[21] 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

[22] 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.

[23] 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.