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

Graph-Based Analog Joint Source Channel Coding for 3D Haptic Communication

Fujihashi, Takuya; Koike-Akino, Toshiaki; Corcodel, Radu

TR2024-070 June 11, 2024

Abstract

Haptic communication will be a key technology for extended reality (XR), robotics, and remote manipulation. The deformation magnitude of 3D deformable objects is a key attribute for realizing fine-grained remote manipulation. The typical solutions for sending the deformation magnitudes over wireless channels are to perform digital compression and transmission considering the channel quality, i.e., digital source- channel coding. However, the key problems of the solutions are 1) still large traffic and 2) catastrophic quality degradation due to channel quality fluctuation. This paper proposes a graph- based analog joint source-channel coding for 3D haptic com- munication. Specifically, Graph Fourier Transform (GFT)-based energy compression efficiently removes the redundancy across deformation magnitudes. In addition, the integration of unequal error protection and analog modulation prevents catastrophic degradation and gradually improves the reconstruction quality according to the instantaneous channel quality. Evaluation results using the deformation magnitude of various 3D objects show that the proposed scheme prevents quality degradation due to channel quality fluctuations and provides accurate deformation magnitude for remote users.

IEEE International Conference on Communications (ICC) 2024

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

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

Graph-Based Analog Joint Source Channel Coding for 3D Haptic Communication

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

***Abstract*—Haptic communication will be a key technology for extended reality (XR), robotics, and remote manipulation. The deformation magnitude of 3D deformable objects is a key attribute for realizing fine-grained remote manipulation. The typical solutions for sending the deformation magnitudes over wireless channels are to perform digital compression and transmission considering the channel quality, i.e., digital source-channel coding. However, the key problems of the solutions are 1) still large traffic and 2) catastrophic quality degradation due to channel quality fluctuation. This paper proposes a graph-based analog joint source-channel coding for 3D haptic communication. Specifically, Graph Fourier Transform (GFT)-based energy compression efficiently removes the redundancy across deformation magnitudes. In addition, the integration of unequal error protection and analog modulation prevents catastrophic degradation and gradually improves the reconstruction quality according to the instantaneous channel quality. Evaluation results using the deformation magnitude of various 3D objects show that the proposed scheme prevents quality degradation due to channel quality fluctuations and provides accurate deformation magnitude for remote users.**

I. Introduction

Haptic communication is an emerging technique to provide a sense of touch and object deformation to remote users and robots [1]. In particular, untethered transmission of object deformation is required for dexterous manipulation of 3D objects by remote users [2]. In addition, the reproduced deformation plays a key role in many applications as shown in Fig. 1. To reproduce the object deformation to remote users, a sender sends the magnitude of the external contact force for the 3D deformable object, i.e., deformation magnitude, at each instant. The deformation magnitude for the 3D deformable object is represented as a time series of 3D point clouds with the attribute of the magnitude [3]. Fig. 2 shows an example of the deformation magnitude for the 3D object. Here, the 3D object receives an external contract force in the x-axis, and the 3D points with block color in Fig. 2 (b) through (d) represent strong deformation magnitude at the 3D point.

The deformation magnitudes corresponding to 3D samples are distributed unevenly in 3D space to reproduce the various deformations at remote users. To efficiently provide the non-uniformly sampled signals, graph signal processing (GSP) based schemes [4]–[6] for non-uniform 3D samples have been proposed as shown in Fig. 3 (a). The existing GSP-based schemes consider the 3D points as vertices in a graph and transform the graph signals into frequency representations. They used graph Fourier transform (GFT) for energy compression of the graph signals and applied quantization over the GFT coefficients. Finally, the quantized GFT coefficients are entropy coded for compression. The entropy-coded bits are then passed to the transmission side. The transmission side uses channel coding for error resilience and determines the modulation format based on the measured channel quality.

A key challenge in existing graph-based schemes is the high-quality, low-delay delivery of graph signals over wireless

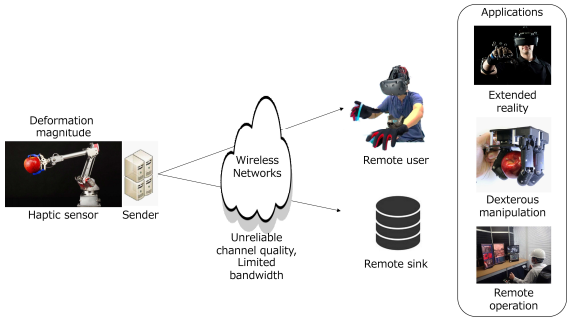
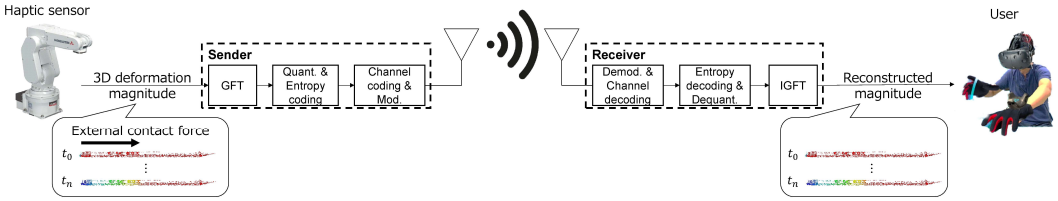


Fig. 1. Exemplification architecture of 3D haptic communication.

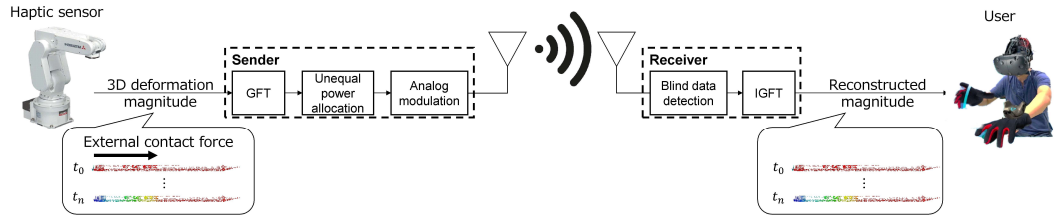
(a) Original 3D samples of the object (b) Deformation magnitude at time t_1

(c) Deformation magnitude at time t_2 (d) Deformation magnitude at time t_3

Fig. 2. Examples of deformation magnitude at each instant ($t_1 < t_2 < t_3$). Here, we consider that an external contact force is from the right side of the object. The 3D points with block color in (b) through (d) represent strong deformation magnitude at the 3D point.



(a) Graph-based digital joint source-channel coding (DJSCC)



(b) Proposed graph-based AJSCC scheme

Fig. 3. Framework of the graph-based DJSCC and proposed graph-based AJSCC schemes.

channels. Although the existing schemes require accurate channel estimation for proper source and channel coding, the channel conditions of wireless networks are unreliable. In particular, the channel conditions vary unexpectedly, resulting in inaccurate channel estimation and rate control for source and channel coding. Inaccurate channel estimation and rate control cause catastrophic quality degradation in the reconstructed deformation magnitude, i.e., the cliff and leveling off effects [7, 8]. A typical solution to prevent catastrophic quality degradation is retransmission. However, it is unsuitable for haptic communication due to the strict delay constraint of haptic applications.

As mentioned above, the existing graph-based schemes may cause critical issues in 3D haptic communication: 1) quality drop and 2) quality saturation. This paper proposes a novel graph-based analog joint source-channel coding (AJSCC) scheme to discuss the feasibility of 3D haptic communication. To overcome the above-mentioned issues, the proposed scheme integrates GFT for energy compaction, unequal power allocation for quality maximization, and analog modulation to prevent cliff and leveling-off effects. In addition, the proposed scheme adopts blind data detection [9] for decoding the transmitted deformation magnitudes with negligible communication overhead.

Evaluations using the simulated deformation magnitudes show that the proposed scheme reconstructs high-quality deformation magnitudes based on the instantaneous channel quality and available bandwidth. In addition, we discuss the effect of the graph shift operators and deformation objects on the reconstruction quality.

The contributions of this paper are three-fold:

- We consider 3D object deformations as graph signals to realize energy compression of irregular structure signals.
- We adopt GFT with various graph shift operators before analog modulation to find better energy compaction for 3D haptic communication.
- Although the existing AJSCC schemes cause large communication overhead for signal decoding, the proposed scheme can reconstruct the transmitted signals with only one constant value using blind data detection.

II. Related Work

A. Haptic Codec

Haptic codec solutions are mainly proposed for 1D vibrotactile signals to reduce traffic in haptic communication. The existing haptic codec solutions can be categorized into three groups: transform coding [10]–[12], analysis-by-synthesis coding [13, 14], and hybrid solutions [15, 16]. For example, PVC-SLP in [15, 16] is the state-of-the-art hybrid solution. It integrates Sparse Linear Prediction (SLP) and 1D DCT for traffic reduction.

The aforementioned solutions can reduce traffic, but they suffer from catastrophic quality degradation due to the all-or-nothing behavior of entropy and channel coding, as well as unrecoverable quantization errors.

B. GSP-Based Point Cloud Delivery

In the point cloud, a large number of 3D points are sampled non-uniformly in the 3D space to represent the real world environment. One of the key issues in point cloud coding is to compact the energy of the attributes of the 3D points. Some studies have used GFT for energy compaction of 3D coordinate and color attributes [17, 18]. In addition, a recent work [19] uses a graph neural network (GNN) architecture for energy reduction of the point cloud.

The proposed scheme uses GFT for 3D haptic data, i.e., deformation magnitudes in each 3D point, to discuss the impact of the graph-based solution on future 3D haptic communication. Although an effect of GFT-based decorrelation on 3D coordinate and color attributes is discussed in the existing

studies, we can see that GFT-based energy compression can have another great gain in 3D deformation magnitudes.

III. Proposed Scheme

Fig. 3 (b) shows the overview architecture of the proposed scheme. A sender obtains a time series of deformation magnitudes from the haptic sensors and then constructs graph signals from the measured deformation magnitudes. It performs a GFT over the graph signals for each instant to compress the signal energy. The GFT coefficients are then power-assigned to minimize the mean square error (MSE) between the original and reconstructed signals and mapped directly to the transmission symbols. At the receiver, blind data detection is used to decode the received signals with low communication overhead, and finally, inverse GFT (IGFT) is used on the filtered GFT coefficients to reconstruct the deformation magnitudes at each instant.

A. Energy Compaction

1) Graph Construction: The proposed scheme constructs graph signals from the measured deformation magnitudes in 3D space. Specifically, each deformation magnitude in the 3D space can be regarded as the vertex of a weighted and undirected graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{W})$, where $\boldsymbol{V}$ and $\boldsymbol{E}$ are the vertex and edge sets, respectively. The vertices of N samples have two attributes with 3D coordinates $\boldsymbol{p} = [\mathrm{x}, \mathrm{y}, \mathrm{z}]^{\mathrm{T}} \in \mathbb{R}^{3\times N}$ and deformation magnitudes $\boldsymbol{d} \in \mathbb{R}^{1\times N}$ in N points. $\boldsymbol{W}$ is an adjacency matrix, and each element $W_{i,j} \in \boldsymbol{W}$ represents the edge weight between vertices i and j. The proposed scheme considers the Gaussian kernel [20] for the calculation of each edge weight $W_{i,j}$.

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

where ϵ_p is a hyperparameter of the Gaussian kernel, such as the sample variance or standard deviation of 3D coordinate attributes.

2) GFT: The proposed scheme uses GFT on the constructed graph signals to obtain the frequency domain coefficients. GFT for the graph signals is defined by the parameterized graph shift operator $\boldsymbol{L}$ [21] based on the adjacency matrix $\boldsymbol{W}$, the degree matrix $\boldsymbol{D}$, and the parameter tuple $\boldsymbol{S} = (m_1, m_2, m_3, e_1, e_2, e_3)$ consisting of scalar multiplicative parameters m_1, m_2, m_3 and scalar exponential parameters e_1, e_2, e_3 as follows:

$$\boldsymbol{L} = m_1\boldsymbol{D}^{e_1} + m_2\boldsymbol{D}^{e_2}\boldsymbol{W}\boldsymbol{D}^{e_3} + m_3\boldsymbol{I}, \tag{2}$$

where $\boldsymbol{I}$ is the $N \times N$ identity matrix and $\boldsymbol{D}$ is the diagonal degree matrix of the adjacency matrix $\boldsymbol{W}$. The degree matrix can be derived as:

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

Table I lists the well-known graph shift operators. In Sec. IV-C, the effect of the graph shift operators on the reconstruction quality is discussed in detail.

TABLE I
Well-known graph shift operators based on parameter tuple

$\boldsymbol{S}$	Operator	
$(1, -1, 0, 1, 0, 0, 0)$	$D - W$	Regular graph Laplacian
$(1, 1, 0, 1, 0, 0, 0)$	$D + W$	Signless graph Laplacian
$(0, -1, 1, 0, -1, 0, 0)$	$I - D^{-1}W$	Random-walk graph Laplacian
$(0, 1, 0, 0, -1, 0, 0)$	$D^{-1}W$	Mean aggregation

The graph shift operator is a real symmetric matrix, and thus the operator will have a complete set of orthonormal eigenvectors and corresponding positive eigenvalues. The proposed scheme uses eigenvalue decomposition to obtain the eigenvectors $\boldsymbol{\Phi} \in \mathbb{R}^{N\times N}$ and eigenvalues $\boldsymbol{\Lambda}$ of the graph shift operator as $\boldsymbol{L} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{-1}$.

The GFT coefficients $\boldsymbol{f} \in \mathbb{R}^{1\times N}$ can be obtained by multiplying the eigenvectors by the deformation magnitudes $\boldsymbol{d}$ as $\boldsymbol{f} = \boldsymbol{d}\,\boldsymbol{\Phi}$, and the magnitudes can be reconstructed from the GFT coefficients as $\boldsymbol{d} = \boldsymbol{f}\,\boldsymbol{\Phi}^{-1}$.

B. Power Allocation

After the proposed scheme obtains GFT coefficients from deformation magnitudes, unequal transmission power is assigned to each GFT coefficient before transmission to minimize the MSE between the original and reconstructed deformation magnitudes. Specifically, power allocation for ith GFT coefficient can be written as:

$$x_i = g_i \cdot f_i. \tag{4}$$

This means that the ith transmission symbol x_i is the ith GFT coefficient scaled by the factor g_i. Here, the optimal scaling factor g_i is obtained by minimizing the MSE under the power constraint with an average power budget of P as follows:

$$\min \quad \mathrm{MSE} = E\left[(x_i - \hat{x}_i)^2\right] = \sum_{i}^{N} \frac{\sigma^2\lambda_i}{g_i^2\lambda_i + \sigma^2}, \tag{5}$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{i}^{N} g_i^2\lambda_i = P, \tag{6}$$

where $\hat{x}_i$ is an estimate at the receiver, λ_i is the power of the i-th GFT coefficient, and σ^2 is a noise variance. Using the method of Lagrange multipliers, the best g_i is obtained as [22]:

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

The power-assigned GFT coefficients are then mapped, two by two, to I (in-phase) and Q (quadrature-phase) components, i.e., analog modulation, for the coefficient transmissions.

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 considered in the noise variance. The received symbol y_i over wireless links can be modeled as $y_i = x_i + n_i$.

C. Metadata-Less Signal Decoding

At the receiver side, the received GFT coefficients are extracted from the I and Q components. A minimum MSE (MMSE) filter is a typical filter to reduce the communication noise from the received analog modulated coefficients. The MMSE filter is derived as follows [22]:

$$\hat{f}_i = \frac{g_i \lambda_i}{g_i^2 \lambda_i + \sigma^2} \cdot y_i. \tag{8}$$

However, the MMSE filter requires λ_i of all GFT coefficients at the receiver, so the sender must send λ_i as metadata. This metadata overhead will cause significant performance degradation and consume extra transmission power [23].

The proposed scheme uses blind data detection [9] to reconstruct high-quality GFT coefficients with negligible communication overhead. The proposed scheme scales the GFT coefficients with an optimal scaling factor according to Eq. (7). With $\lambda_i = |f_i|^2$, Eq. (7) can be rewritten as

$$g_i = \boldsymbol{c}|f_i|^{-1/2}. \tag{9}$$

In this case, the received symbol is

$$y_i = g_i \cdot f_i + n_i = \boldsymbol{c}|f_i|^{-1/2} f_i + n_i. \tag{10}$$

Here we can estimate the amplitude of f_i using a zero-forcing estimator and the sign of f_i using the sign of the received symbol. For example, for high quality wireless channels ($n_i \simeq 0$), we obtain an estimate of f_i as follows:

$$\hat{f}_i = (y_i/\boldsymbol{c})^2 \cdot \mathrm{sgn}(y_i). \tag{11}$$

In this case, the proposed scheme sends only the constant $\boldsymbol{c}$ as metadata for decoding the received symbols. Finally, the receiver reconstructs the deformation magnitudes $\hat{\boldsymbol{d}}$ by taking the IGFT for the filtered GFT coefficients $\hat{\boldsymbol{f}}$.

IV. Performance Evaluation

A. Simulation Settings

1) Performance Metric: We consider the MSE between the original and reconstructed deformation magnitudes as the quality of 3D haptic communication.

2) 3D Deformation Dataset: The following benchmark point clouds of 3D objects are used to simulate the deformation magnitudes: *pencil__10_0* with 2,731 points and *milk_color* with 13,704 points. *pencil__10_0* is first considered to discuss the baseline performance of the proposed and baseline schemes. *milk_color* is used to discuss the effect of a large 3D object on the reconstruction quality.

To simulate the deformation magnitudes, we consider a two-dimensional surface in 3D space. The four corners of the surface are ($y_{\min}$, $z_{\min}$), ($y_{\min}$, $z_{\max}$), ($y_{\max}$, $z_{\min}$), ($y_{\max}$, $z_{\max}$) where $y_{\min}$ and $z_{\min}$ are minimum coordinates in y-axis and z-axis and $y_{\max}$ and $z_{\max}$ are maximum coordinates in y-axis and z-axis, respectively. We consider the surface to be at $x_{\min}$ at t_0, shifted by Δx for each slice, and finally approaching $x_{\max}$ after T slices. For t-th slice, the deformation magnitude for i-th 3D point can be obtained as: $m_i = |x_i - x_t|/(x_{\max} - x_{\min})$, where x_t is x-coordinate at t-th slice. In this paper, the number of slices T is set to 100 and thus $\Delta x = (x_{\max} - x_{\min})/T$.

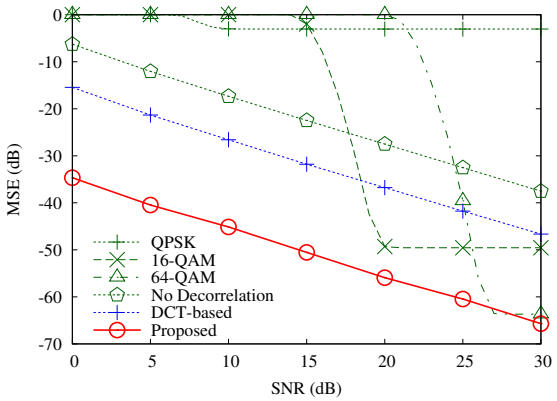


Fig. 4. MSE of deformation magnitude as a function of wireless channel SNRs for *pencil__10_0* ($N = 2731$). Here, the channel symbol rate at each slice is 1,365 symbols.

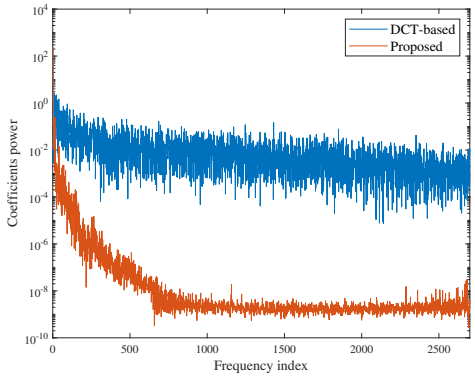


Fig. 5. Average power spectrum of the DCT coefficients and GFT coefficients for *pencil__10_0* across all the slices.

3) Wireless Settings: The digital-modulated and analog-modulated are impaired by an AWGN channel. For digital-based schemes, the digital modulation formats are either quadrature phase shift keying (QPSK), 16-ary quadrature-amplitude modulation (16-QAM), or 64-ary QAM (64-QAM). We also set the same channel symbol rate in all the comparison schemes.

B. Baseline Performance

This section evaluates the baseline performance of the conventional DJSCC and the proposed AJSCC schemes under channel quality fluctuations. For the comparison, we consider five baselines: graph-based DJSCC at the modulation formats of QPSK, 16-QAM, and 64-QAM, and AJSCC with/without decorrelation.

The DJSCC schemes use GFT with the regular graph Laplacian for the simulated deformation magnitudes to transform the magnitudes into the frequency representations, and uniformly

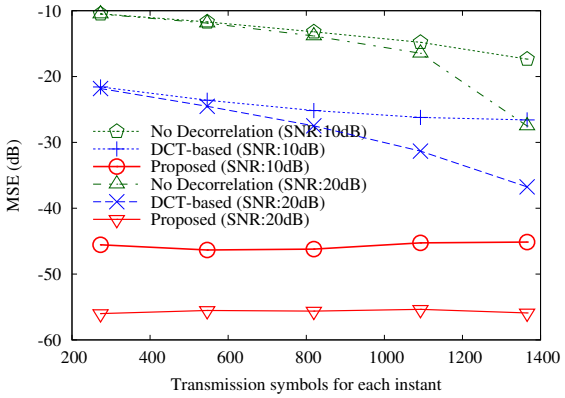


Fig. 6. MSE of deformation magnitudes as a function of channel symbol rates at each slice for *pencil__10_0* ($N = 2731$). Here, wireless channel SNRs are set to 10 dB and 20 dB, respectively.

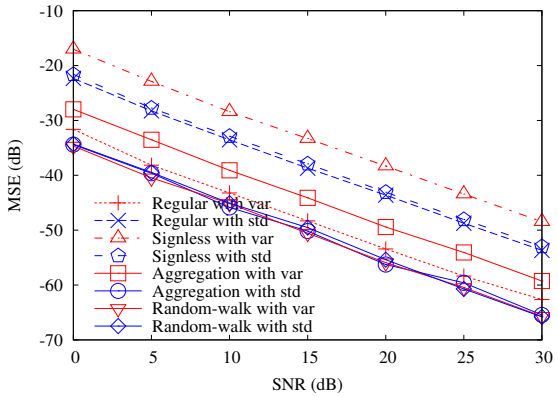


Fig. 7. MSE of deformation magnitude in the proposed scheme with different graph shift operators and hyperparameters for *pencil__10_0* ($N = 2731$). Here, the channel symbol rate at each slice is 1,365 symbols.

quantize the frequency representations followed by entropy coding. Finally, the digital modulation formats are used to transmit the bit stream. The AJSCC schemes consider the simulated deformation magnitudes as the ordered 1D signals. Here, the AJSCC with decorrelation scheme performs 1D-DCT (discrete cosine transform) on the signals, while the other scheme does not use any decorrelation techniques. Other operations, i.e., unequal power allocation, analog modulation, and blind data detection, are the same as the proposed scheme.

Fig. 4 shows the reconstruction quality of the deformation magnitudes as a function of the wireless channel SNRs. Here, the channel symbol rate at each slice is set to 1,365 symbols. In addition, the proposed scheme uses the random-walk graph Laplacian as the graph shift operator and the variance-based hyperparameter ϵ_p. From the evaluation results in Fig. 4, we can make the following observations:

- The proposed scheme gradually improves the reconstruction quality of the deformation magnitudes with the improvement of the wireless channel quality.
- The channel quality fluctuation causes a cliff effect in low-channel SNR regimes and a leveling-off effect in high-channel SNR regimes in the DJSCC schemes.
- The AJSCC schemes prevent cliff and leveling-off effects by removing quantization and entropy coding from the sender operations. However, the quality of the DCT-based AJSCC scheme is lower than that of the proposed scheme.

For example, the proposed scheme achieves 2.0 dB and 19.0 dB improvement at the channel SNR of 30 dB compared to the DJSCC with 64-QAM modulation format and the DCT-based AJSCC, respectively.

To discuss the advantages of the proposed scheme in detail, Fig. 5 shows the average power spectrum of DCT coefficients and GFT coefficients for *pencil__10_0* over all slices. The power values of the high frequency GFT coefficients are relatively lower than those of the high frequency DCT coefficients. For example, the sum of the square roots of the DCT coefficients and GFT coefficients in Fig. 5 is 292.4 and 33.6, respectively. As discussed in Eq. (5), the MSE of the AJSCC schemes decreases when the sum of the square roots of λ_i is small. As a result, the proposed scheme can achieve better performance compared to the DCT-based AJSCC scheme. We then evaluate the effect of the available bandwidth on the performance of the AJSCC schemes by decreasing the number of transmission symbols. Fig. 6 shows the quality of the AJSCC schemes as a function of channel symbol rates for each slice at wireless channel SNRs of 10 dB and 20 dB. We can see the following observations.

- The proposed scheme can maintain almost the same quality regardless of the number of symbols received, even though only 10% of the GFT coefficients can be received due to the bandwidth limitation.
- DCT-based and no decorrelation schemes gradually degrade the reconstruction quality as the number of received symbols decreases.

C. Effect of Graph Shift Operators and Hyperparameters

In the previous section, we evaluated the performance of the proposed scheme using the random-walk graph Laplacian and variance-based hyperparameters under heterogeneous channel quality and bandwidth availability. The proposed scheme can use different graph shift operators listed in Table I to encode and decode graph signals. In addition, the adjacency matrix $\boldsymbol{W}$ in Eq. (4) is highly dependent on the hyperparameter ϵ_p. In particular, the sample variance (var) or the standard deviation (std) of the point distances is often used for ϵ_p. In this section, we will discuss the effects of graph shift operators and hyperparameters on the quality of 3D haptic communication.

Fig. 7 shows the MSE of the deformation magnitudes in the proposed scheme with different graph shift operators and hyperparameters as a function of the wireless channel SNRs. We can see that the random walk graph Laplacian with the sample variance hyperparameter achieves the best performance. In this paper, we consider typical graph shift operators for graph signal coding and decoding. We leave the

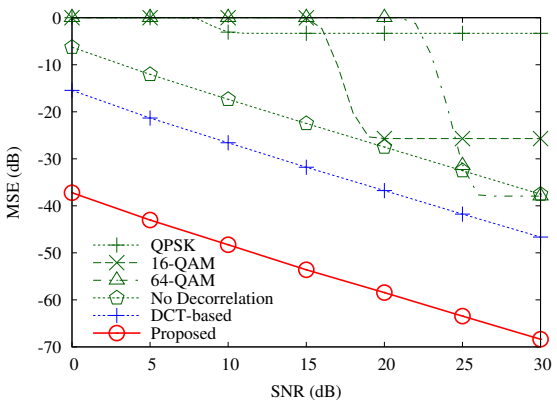


Fig. 8. MSE of deformation magnitude as a function of wireless channel SNRs for *milk_color* ($N = 13{,}704$). Here, the channel symbol rate at each slice is $6{,}852$ symbols.

optimization of graph shift operators to maximize the quality of 3D haptic communication as a future work.

D. Effect of Deformable Objects

In the previous sections, a relatively small 3D object with $2,731$ 3D points was used to demonstrate the advantage of the proposed scheme. In this section, we consider the simulated deformation magnitudes using a large 3D object to demonstrate the scalability of the proposed scheme. Fig. 8 shows the MSE of the deformation magnitudes for the point cloud of *milk_color* ($N = 13{,}704$).

Compared with a small 3D object in Fig. 4, the quality of the DJSCC schemes becomes lower for the same wireless channel quality. For example, the proposed scheme achieves 30.4 dB and 21.7 dB improvement at the channel SNR of 30 dB compared to the DJSCC with 64-QAM modulation format and DCT-based AJSCC coding, respectively.

V. Conclusion

In this paper, a graph-based AJSCC scheme was proposed for future 3D haptic communication. To transmit high-quality 3D deformation magnitudes over unreliable wireless channels, the proposed scheme integrates GFT-based energy compaction, MSE-minimized power allocation, and analog modulation. Evaluation results show that the proposed scheme achieves better quality than the existing DJSCC and AJSCC schemes, even under channel quality fluctuation and bandwidth limitation.

Acknowledgment

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

References

[1] S. Jain, R. Corcodel, and J. van Baar, "Visual 3d perception for interactive robotic tactile data acquisition," in *IEEE International Conference on Automation Science and Engineering*, 2021, pp. 1–7.

[2] Y. Wi, P. Florence, A. Zeng, and N. Fazeli, "Virdo: Visio-tactile implicit representations of deformable objects," in *IEEE International Conference on Robotics and Automation Workshop on Representing and Manipulating Deformable Objects*, 2022, pp. 1–8.

[3] J. Sanchez, C. M. Mateo, J. A. Corrales, B.-C. Bouzgarrou, and Y. Mezouar, "Online shape estimation based on tactile sensing and deformation modeling for robot manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 504–511.

[4] 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.

[5] 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.

[6] 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.

[7] T. Fujihashi, T. Koike-Akino, and T. Watanabe, "Soft delivery: Survey on a new paradigm for wireless and mobile multimedia streaming," *ACM Computing Surveys*, vol. 56, no. 2, 2023.

[8] T. Ogura, S. Kitamura, T. Fujihashi, K. Kizaki, S. Saruwatari, and T. Watanabe, "Semantic communication for multiple vibrotactile sensors," in *IEEE Global Communications Conference*, 2023, pp. 1–6.

[9] T. Zhang and S. Mao, "Metadata reduction for soft video delivery," *IEEE Networking Letters*, vol. 1, no. 2, pp. 84–88, 2019.

[10] E. Steinbach, M. Strese, M. Eid, X. Liu, A. Bhardwaj, Q. Liu, M. Al-Ja'afreh, T. Mahmoodi, R. Hassen, A. El Saddik, and O. Holland, "Haptic codecs for the tactile internet," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 447–470, 2019.

[11] A. Noll, B. Gülecyüz, A. Hofmann, and E. Steinbach, "A rate-scalable perceptual wavelet-based vibrotactile codec," in *IEEE Haptics Symposium*, 2020, pp. 854–859.

[12] O. Holland, E. Steinbach, R. V. Prasad, Q. Liu, Z. Dawy, A. Aijaz, N. Pappas, K. Chandra, V. S. Rao, S. Oteafy, M. Eid, M. Luden, A. Bhardwaj, X. Liu, J. Sachs, and J. Araújo, "The IEEE 1918.1 "tactile internet" standards working group and its standards," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 256–279, 2019.

[13] R. Chaudhari, B. Çizmeci, K. J. Kuchenbecker, S. Choi, and E. Steinbach, "Low bitrate source-filter model based compression of vibrotactile texture signals in haptic teleoperation," in *Proceedings of the 20th ACM International Conference on Multimedia*, 2012, p. 409–418.

[14] R. Chaudhari, C. Schuwerk, M. Danaei, and E. Steinbach, "Perceptual and bitrate-scalable coding of haptic surface texture signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 3, pp. 462–473, 2015.

[15] R. Hassen, B. Gülecyüz, and E. Steinbach, "PVC-SLP: perceptual vibrotactile-signal compression based-on sparse linear prediction," *IEEE Transactions on Multimedia*, vol. 23, pp. 4455–4468, 2021.

[16] R. Hassen and E. Steinbach, "Vibrotactile signal compression based on sparse linear prediction and human tactile sensitivity function," in *IEEE World Haptics Conference (WHC)*, 2019, pp. 301–306.

[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. 24, pp. 2179–2191, 2021.

[18] S. Ueno, T. Fujihashi, T. Koike-Akino, and T. Watanabe, "Point cloud soft multicast for untethered xr users," *IEEE Transactions on Multimedia*, vol. 25, pp. 7185–7195, 2023.

[19] 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.

[20] 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.

[21] G. Dasoulas, J. Lutzeyer, and M. Vazirgiannis, "Learning parametrised graph shift operators," in *International Conference on Learning Representations*, 2021, pp. 1–17.

[22] 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.

[23] 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, 2018.