High-Quality Reconstruction of Depth Maps From Graph-Based Non-Uniform Sampling

Jingyu Yang, *Senior Member, IEEE*, Wenqiang Xu, Yusen Hou, Xinchen Ye, *Member, IEEE*, Pascal Frossard, *Fellow, IEEE*, and Kun Li, *Member, IEEE*

***Abstract*—Depth sensing is essential for intelligent computer vision applications, but it often suffers from low range precision and spatial resolution. To address this problem, we propose a novel framework that combines non-uniform sampling and reconstruction based on graph theory. Our framework consists of two main components: (1) a graph Laplacian induced non-uniform sampling (GLINUS) scheme that samples depth signals more densely around edges and contours than in smooth regions, and (2) an ensemble of priors (EoP) model that reconstructs the high-quality depth map using adaptive dual-tree discrete wavelet packets (ADDWP) transform, graph total variation regularizer, and graph Laplacian regularizer with color guidance. We solve the reconstruction problem using the alternating direction method of multipliers (ADMM). Our experiments demonstrate that our framework can capture fine structures and global information in depth signals and produce superior depth reconstruction results.**

***Index Terms*—Graph signal processing, non-uniform sampling, depth reconstruction.**

I. Introduction

MANY applications, such as 3DTV, action recognition, and robot navigation, require high-quality depth images/videos. However, most existing depth acquisition devices using time-of-fly (ToF) or structured light can only provide low-resolution depth maps that are much lower than color image sensors. Therefore, effective depth super-resolution (SR) techniques are needed to produce high-resolution output from low-resolution input. Recent approaches can achieve high-quality results for $4\times$ SR and acceptable results for $8\times$ SR, but they would introduce significant artifacts for higher upsampling rates [1], [2].

Most depth sensors use rectangular patterns, which allow equal interval sampling on a uniform 2D grid without complex post-adaptation of the sampling grid. However, this may not be the optimal method for depth map reconstruction with limited sampling resources. From a signal processing perspective, related studies have shown that non-uniform sampling can achieve more efficient sensing and avoid frequency aliasing [3], [4]. Some works have also verified the benefits of non-uniform sampling on depth maps [5], [6]. Moreover, computational depth estimation methods, *e.g.*, stereo matching, can be regarded as a form of non-uniform sampling. They estimate depth values by computing the disparities between a pair of stereo images through their effective matching pixels/features distributed more densely around high-frequency regions than smooth regions.

As discussed, the sampling density and pattern are the two key factors in depth reconstruction. Instead of finding an appropriate regularization to invert the under-sampling problem, another approach for better recovery of depth signals is to use non-uniform sampling, which captures more prominent signal features with a set of sparse samples. As we observe, the main challenge in depth reconstruction is the ambiguity in locating the discontinuities such as edges and contours in depth maps. Depth maps can be reconstructed more accurately with the same sampling rate if more samples are allocated to areas around depth discontinuities. Given the limited related work [5], non-uniform sampling and reconstruction on depth maps deserve more investigation and evaluation.

This paper proposes a novel depth sampling and reconstruction framework based on graph signal processing [4], [7]. Our method has two main components: a graph Laplacian induced non-uniform sampling (GLINUS) scheme and an ensemble of priors (EoP) framework. The GLINUS scheme adapts the sampling rate to the graph filtering response of depth maps as band-limited graph signals. It preserves the spatial saliency of depth and avoids the rank reduction problem that general active sampling faces in the graph Fourier domain. We also provide theoretical justification for our sampling scheme. The EoP framework reconstructs the depth maps from the GLINUS samples using multiple priors: 1) an adaptive dual-tree discrete wavelet packets (ADDWP) transform to encode the depth maps and capture their multi-directional features, 2) a graph total variation regularizer to enforce piece-wise smoothness, and 3) a non-local fine-tuned graph Laplacian regularizer to exploit the structural correlation between the depth and auxiliary color images. To solve the reconstruction problem, we develop a numerical algorithm based on the alternating direction method of multipliers (ADMM) with variable splitting strategy. The main contributions are summarized as follows:

- We propose a graph Laplacian induced non-uniform sampling method based on the theory of graph signal processing, avoiding the rank reduction problem for more accurate and stable reconstruction.

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62231018 and Grant 62171317. (*Corresponding author: Kun Li*)

J. Yang, W. Xu, Y. Hou are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: yjy@tju.edu.cn, xwqtju@tju.edu.cn).

X. Ye is with the School of Information Science and Engineering, Dalian University of Technology, Dalian 116024, China (Email: yexch@dlut.edu.cn).

P. Frossard is with the Signal Processing Laboratory LTS4, Ecole Polytechnique Fédérale de Lausanne (EPFL),Lausanne 999034, Switzerland (Email: pascal.frossard@epfl.ch).

K. Li is with the College of Intelligence and Computing, Tianjin University, Tianjin 300350, China (Email: lik@tju.edu.cn).

- We propose a reconstruction framework based on ensemble of priors that leverages multiple sources of prior information, *i.e.*, the ADDWP transform, graph total variation, and the non-local fine-tuned graph Laplacian regularizer. We develop a numerical algorithm based on the ADMM framework.
- Experimental results show that our sampling scheme capture fine structure and efficient global information in depth signals, and our reconstruction algorithm outperforms the state-of-the-art methods for various depth maps at different sampling ratios.

II. Related Work

In this section, we review recent graph sampling methods, including spectral-domain and vertex-domain approaches. We also summarize popular depth reconstruction (DR) algorithms that are relevant to our proposed methods.

A. Graph Sampling Theory

In [8], a sampling theory for graph signals was developed to establish a sufficient condition for the unique reconstruction of signals from a given sampling set. Based on this condition, Narang *et al*. [9] derived a bound on the maximum bandwidth that a signal can possess to be uniquely recoverable from its samples on a subset of nodes. Anis *et al*. [10] improved this bound by considering a necessary and sufficient condition for sampling. They introduced a direct sampling set selection method that approximately maximizes this bound, enabling the unique reconstruction for a larger class of graph signals.

Previous methods for sampling set selection in graphs fall into two categories: spectral-domain methods and vertex-domain methods. We review them as follows.

1) Spectral-Domain Approaches: We classify the methods that rely on the explicit knowledge of a portion of the graph Fourier basis as spectral-domain approaches. These methods involve computing the spectrum of the variation operator (such as the graph Laplacian or adjacency matrix) [11]. Shomorony *et al*. [12] constructed a sampling set that guarantees a unique (but not necessarily stable) reconstruction for any signal spanned by the first r eigen-vectors of the graph Laplacian. Chen *et al*. [13] proposed a greedy algorithm that selects stable sampling sets for a given band-limited space by a spectral-domain criterion. Tsitsvero *et al*. [14] established a connection between the uncertainty principle for graph signals and sampling theory, and derived criteria for selecting sampling sets in the presence of sample noise.

2) Vertex-Domain Approaches: Instead of relying on graph spectral information, another category for sampling set selection uses vertex-domain characteristics. For instance, Narang *et al*. [15] and Nguyen *et al*. [16] selected sampling sets based on maximum graph cuts and spanning trees, respectively. However, these methods are more suitable for designing downsampling operators required in bipartite graph multi-resolution transforms [17], rather than optimizing the sampling sets to improve reconstruction quality. Unlike most existing methods that sample signal values at some nodes, Marques *et al*. [18] proposed to sample band-limited graph signals by aggregating local information at one node after successively applying graph-shift operator. However, this strategy requires a large neighborhood size equal to the dimension of the band-limited space for aggregating the sample values, which can cover a large portion of the graph.

Besides the deterministic methods of approximating optimal sampling sets, there is also a randomized sampling strategy [19] that guarantees a bound on the worst case reconstruction error in the presence of noise. This strategy samples nodes independently based on a carefully designed distribution over the nodes. However, it requires much more nodes than the dimension of the band-limited space to achieve the bound.

B. Depth Reconstruction From Subset of Samples

Depth reconstruction (DR) and depth super-resolution (DSR) are closely related problems. Many methods for enhancing depth maps can address both problems. The DR and DSR tasks are equivalent when the low-resolution (LR) depth map pixels are uniformly distributed on a regular high-resolution (HR) grid with a fixed spacing between observed pixels. The common goal is to estimate the missing depth values on the HR grid. In this section, we refer to them broadly as DR problems with uniform or non-uniform sampling patterns.

1) DR from Uniform Samples (DRUS): DRUS is a challenging task that has attracted considerable interests in recent years. A common strategy is to leverage the structural correlation between the low-resolution depth map and the corresponding high-resolution color image. This strategy relies on the assumption that the depth map and the color image share similar edge features, which is plausible as modern depth sensors can simultaneously capture registered depth maps and color images.

Many image filtering schemes with guidance [20], [21] have been adapted to reconstruct depth maps using their associated color images. These schemes predict the depth value at each missing pixel (linearly or non-linearly) from its neighboring depth values with weights derived from both depth and color information. Yang *et al*. [22] proposed an iterative joint bilateral filter (JBU) for depth SR. However, this method often introduces texture-copy artifacts in the reconstructed depth maps. Chan *et al*. [23] developed a noise aware filter that mitigates the texture-copy problem of JBU by dynamically adjusting between JBU and bilateral filter based on the noise level of the depth image. Ferstl *et al*. [24] constructed an anisotropic diffusion tensor based on the guidance image to preserve detailed structures. Liu *et al*. [25] estimated depth values of unknown pixels by finding the shortest geodesic paths to known pixels on the HR grid. Hua *et al*. [26] leveraged the structure and gradient information from the guidance image to generate a HR depth map.

In addition to the filtering-based methods, optimization-based methods have also been proposed for depth SR [1], [2], [27]–[29]. These methods use various priors, such as Markov random fields (MRF) [30]–[32], total variation model [24], truncated Huber penalty [33] , and autoregressive model [1], [34], to regularize the depth reconstruction problem. The priors are usually incorporated into minimization tasks, which

are solved by optimization techniques such as (weighted) least squares [1], [35], ℓ_1-norm minimization [2], [36], graph cut [32], half-quadratic optimization [33]. Yang *et al*. [34] analyzed the connection between the filtering-based methods and quadratic optimization methods, and proposed a fixed-point iteration algorithm to efficiently solve the large-scale quadratic programming, achieving state-of-the-art reconstruction performance. Under the framework of sparse representation, dictionary learning has been also introduced to learn efficient reconstruction of depth maps [6], [37].

With the advent of deep learning, many end-to-end neural network models have been proposed to enhance low-resolution (LR) depth maps using high-resolution (HR) intensity images or other priors. For example, Zuo *et al*. [38] used multi-scale global and local residual learning to exploit the guidance of HR intensity images at different scales. Liu *et al*. [33] introduced a novel truncated Huber penalty function that can adapt to different image smoothing preferences and achieve superior depth SR performance. Guo *et al*. [39] developed a residual U-Net architecture to generate HR depth maps by constructing an input pyramid to capture multiple level receptive fields. Ye *et al*. [40] proposed a progressive multi-branch aggregation network to recover the degraded depth maps gradually. Yao *et al*. [41] designed a texture-depth transformer that can learn the structural information of HR texture images and the corresponding interpolated depth maps for depth SR task. These learning-based methods rely heavily on data with some specific distribution characteristics or pre-trained models, which makes them lack flexibility and adaptability to different scenarios and data.

2) DR from Non-uniform Samples (DRNS): Many DRUS methods [1], [2], [34], [35] can cope with some degree of non-uniformity and sparsity in the samples. For example, the depth measurements warped to the color camera are not uniformly distributed near depth edges. However, these methods fail to handle more challenging cases, and there is a lack of literature that addresses these challenges. Liu *et al*. [5] made an initial attempt to study the sparse sampling and reconstruction of depth maps, and demonstrated that a carefully designed sampling scheme can greatly enhance the reconstruction quality.

Instead of using heuristics such as the two-stage sampling scheme [5], this paper investigates the non-uniform sampling of depth maps from a graph signal processing perspective (as the non-uniform samples can be naturally described on a graph), and proposes a depth reconstruction method to recover depth maps from severely-undersampled non-uniform samples (up to the sampling rate of 1% in our experiments). Preliminary results of our work were reported in a conference paper [42]. This paper further improves the sampling scheme by introducing a distribution bias towards smooth regions, which enhances the reconstruction stability. Moreover, it strengthens the reconstruction model by incorporating the more powerful EoP prior and provides a comprehensive analysis of both method development and experimental results.

III. Graph Laplacian Induced Non-uniform Sampling

This section introduces GSP preliminaries following the notations in [11], analyzes the characteristics of depth maps as band-limited graph signals, and presents the graph Laplacian induced non-uniform sampling (GLINUS) algorithm.

A. Preliminaries

An undirected graph $\mathcal{G} = (\mathcal{V}, \mathbf{W})$ consists of a set of nodes $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ and an adjacency matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, where w_{ij} denotes the weight between nodes v_i and v_j, if they are connected. The degree matrix $\mathbf{D}$ is a diagonal matrix, in which the diagonal element is the sum of each row of $\mathbf{W}$, *i.e.*, $d_{ii} = \sum_j w_{ij}$. The graph Laplacian matrix is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \tag{1}$$

which plays an important role in revealing underlying structures of the graph. Since $\mathbf{L}$ is a real, symmetric, and positive semi-definite matrix, it can be decomposed into a set of orthogonal eigen-vectors, denoted by $\{\mathbf{u}_i\}_{i=1}^N$, with the ascending order of associated real and non-negative eigen-values $0 = \alpha_1 \leq \alpha_2 \leq, \ldots, \leq, \alpha_N$. Concretely, the eigen-decomposition of $\mathbf{L}$ is written as

$$\mathbf{L} = \mathbf{U}\mathbf{A}\mathbf{U}^\top, \tag{2}$$

where the diagonal matrix $\mathbf{A}$ contains the eigen values $\{\alpha_i\}$ as its diagonal. $\mathbf{U} \triangleq (\mathbf{u}_1 \ldots \mathbf{u}_N)$ and $\mathbf{U}^\top$ are the matrices of inverse *graph Fourier transform* (GFT) and forward GFT, respectively.

A graph signal is a function, denoted by $f\colon \mathcal{V} \to \mathbb{R}$, defined on the vertices of the graph, where a scalar value is assigned to each vertex. For compact notation, the graph signal can be represented in a vector form, *i.e.*, $\mathbf{x} \in \mathbb{R}^N$ where x_i is the function value of v_i. Similarly to the standard Fourier transform, the graph signal $\mathbf{x}$ can be transformed to the graph frequency domain by

$$\mathbf{f} = \mathbf{U}^\top\mathbf{x}, \tag{3}$$

where $\mathbf{f}$ is the vector of GFT coefficients associated with generalized frequencies $\{\alpha_i\}$.

A signal $\mathbf{x} \in \mathbb{R}^N$ is ω-band-limited if its GFT coefficients are supported within the frequency range of $[0, \omega]$, *i.e.*, $\mathbf{f}_i = 0$ for all i with $\alpha_i > \omega$. Let $\mathcal{N} \triangleq \{1 \ldots N\}$, and denote by $\mathcal{R} = \{1 \ldots R\}$ the indices of eigen-values such that $\eta_i \leq \omega, i \in \mathcal{R}$. The bandwidth, denoted by $\omega(\mathbf{x})$, is defined as the largest eigen-values in $\{\alpha_i, i \in \mathcal{R}\}$, *i.e.*, $\omega(\mathbf{x}) = \alpha_R$. Denote by $\mathbf{U}_{\mathcal{N}\mathcal{R}}$ the sub-matrix of $\mathbf{U}$, whose rows and columns are indexed by $\mathcal{N}$ and $\mathcal{R}$, respectively, *i.e.*, $\mathbf{U}_{\mathcal{N}\mathcal{R}} \triangleq (\mathbf{u}_1 \ldots \mathbf{u}_R)$. An arbitrary ω-band-limited signal can be written as a linear combination of eigen-vectors:

$$\mathbf{x} = \mathbf{U}_{\mathcal{N}\mathcal{R}}\mathbf{f}_{\mathcal{R}}, \tag{4}$$

where $\mathbf{f}_{\mathcal{R}}$ is a sub-vector of $\mathbf{f}$ indexed by $\mathcal{R}$. Such ω-band-limited signals lie in the column space of $\mathbf{U}_{\mathcal{V}\mathcal{R}}$, known as the Paley-Wiener space denoted by $PW_\omega(\mathcal{G}) = \text{range}(\mathbf{U}_{\mathcal{V}\mathcal{R}})$ [8].

Assuming a linear observation model, sampling M measurements from signal $\mathbf{x}$ can be formulated as

$$\mathbf{y} \equiv \mathbf{x}_{\mathcal{M}} = \mathbf{\Psi}\mathbf{x}, \tag{5}$$

where $\mathbf{y}$ is the vector of samples, $\mathbf{\Psi} \in \{0,1\}^{M\times N}$ is the sampling operator, $\mathbf{x}_{\mathcal{M}}$ is the equivalent representation with the set of sample indices $\mathcal{M} = \{i_1 \dots i_M\}$. For a more general formulation, we use a probabilistic sampling setting. Concretely, $\mathbf{\Psi}$ is a linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$ defined as

$$\psi_{ij} = \begin{cases} 1, \text{ with probability } p_j, \\ 0, \text{ with probability } 1-p_j, \end{cases} \tag{6}$$

where $\{p_j\}_{j=1}^N$ is a sequence of probabilities for x_j to be sampled as y_i. Denote by ξ the sampling rate and assume the independent and identically distribution of samples. We have $\sum_{j=1}^N p_j = N\xi$.

A sampling scheme $\mathbf{\Psi}$ should be designed so that $\mathbf{x}$ can be reconstructed or approximated up to a tolerance from $\mathbf{y}$. For an (approximately) ω-band-limited signal $\mathbf{x}$, an intuitive way to reconstruct $\mathbf{x}$ from $\mathbf{y}$ is to minimize the squared error [4].

$$\hat{\mathbf{f}}_{\mathcal{R}} = \arg\min_{\mathbf{f}_{\mathcal{R}}} \|\mathbf{y} - \mathbf{U}_{\mathcal{M}\mathcal{R}}\mathbf{f}_{\mathcal{R}}\|_2^2, \tag{7}$$

where $\mathbf{U}_{\mathcal{M}\mathcal{R}} = \mathbf{\Psi}\mathbf{U}_{\mathcal{N}\mathcal{R}}$. Then, we have the reconstructed signal $\hat{\mathbf{x}} = \mathbf{U}_{\mathcal{N}\mathcal{R}}\hat{\mathbf{f}}_{\mathcal{R}} = \mathbf{U}_{\mathcal{N}\mathcal{R}}\left(\mathbf{U}_{\mathcal{M}\mathcal{R}}\right)^{\dagger}\mathbf{y}$, where $\left(\mathbf{U}_{\mathcal{M}\mathcal{R}}\right)^{\dagger} \triangleq \left(\mathbf{U}_{\mathcal{M}\mathcal{R}}^{\top}\mathbf{U}_{\mathcal{M}\mathcal{R}}\right)^{-1}\mathbf{U}_{\mathcal{M}\mathcal{R}}^{\top}$ is the pseudo-inverse of $\mathbf{U}_{\mathcal{M}\mathcal{R}}$. $\mathbf{U}_{\mathcal{N}\mathcal{R}}\left(\mathbf{U}_{\mathcal{M}\mathcal{R}}\right)^{\dagger}$ is the reconstruction matrix. To perfectly reconstruct $\mathbf{x} \in PW_{\omega}(\mathcal{G})$, there are two sufficient and necessary conditions: 1) the number of samples M is no less than the bandwidth R ($M \geq R$), and 2) the rank of $\mathbf{U}_{\mathcal{M}\mathcal{R}} = \mathbf{\Psi}^{\top}\mathbf{U}_{\mathcal{V}\mathcal{R}}$ is not less than R. While $\mathbf{U}_{\mathcal{V}\mathcal{R}}$ is of full column rank, condition (2) explicitly impose a constraint on the sampling scheme $\mathbf{\Psi}$ so that the sampled M rows from $\mathbf{U}_{\mathcal{V}\mathcal{R}}$ should have a largest linear independent group of R rows. This provides a upper bound on the required number of measurements and a distribution of samples on the graph $\mathcal{G}$ [4].

It should be noted that the derived sampling bound assumes a least square reconstruction. In practice, image reconstruction algorithms, including ours, often incorporate a variety of priors to effectively solve the reconstruction problem, which may reduce the number of samples needed. Nonetheless, this bound serves as a useful benchmark and provides valuable insights for designing efficient sampling schemes when images are represented as graphs.

B. Depth Maps as Approximately band-limited Graph Signals

Depth signals are a family of signals that can be tightly described by the piecewise-smooth image model: depth maps are usually smooth with small variations within each semantic object, and are non-smooth with large variations around object contours of different depths. In a graph representation $\mathcal{G} = (\mathcal{V}, \mathbf{W})$, the adjacency matrix $\mathbf{W}$ of the vertices $\mathcal{V}$ could capture significantly different characteristics between vertices in smooth regions and those around depth discontinuities. To encode the correlation between neighboring vertices, we define the adjacency matrix $\mathbf{W}$ similarly to the bilateral filter [22], considering both the spatial proximity and range similarity.

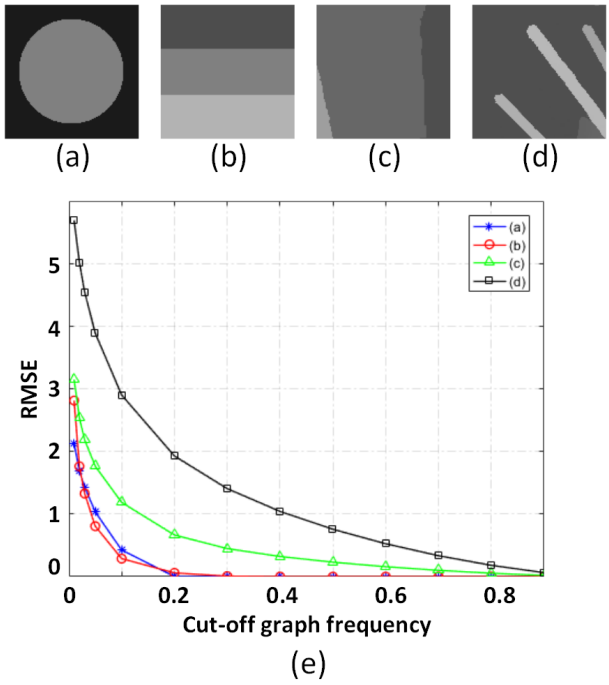

Fig. 1: The NLA results in RMSE by using different cut-off ratios (R/N) of graph frequency through the least squares reconstruction Eq. (7) for four typical depth signals.

Let $\mathcal{N}_i$ be the neighborhood of v_i. Then, the weight of the edge between v_i and v_j is defined as

$$W_{ij} = \exp\left(-\frac{\alpha\|p_i - p_j\|_2^2 + \|x_i - x_j\|_2^2}{\sigma_d^2}\right), j \in \mathcal{N}_i, \tag{8}$$

where p_i and x_i denote the coordinate and the depth value of pixel i, respectively, α is a constant to balance the weights between the spatial proximity and range similarity, and σ_d is a parameter to shape the mapping of spatial-depth proximity.

To verify the fitness of band-limited graph signal model for depth maps, we construct graphs for four typical depth patches of size 128×128. The depth signals are transformed into the GFT domain by Eq. (2) and Eq. (3), and then approximated by Eq. (4) with only the first R spectral components. Fig. 1 shows the nonlinear approximation (NLA) performance in root mean square error (RMSE) with respect to R/N. The NLA curves show a sharp decline at first, and level off after knee points, indicating that depth maps can be effectively represented by a few components in the graph frequency domain.

C. Graph Laplacian Induced Non-uniform Sampling

Since depth maps as approximately band-limited graph signals, they can be reconstructed from a small number of samples as discussed in Sec. III-B. One possible approach is to sample depth signals according to the graph frequency response as the example in Fig. 1. However, this would require the eigen-decomposition of the graph adjacency matrix and the selection of eigen-vectors, which is computational expensive in large signal-dependent graphs and impractical for depth sensing. To avoid this issue, we turn to design an

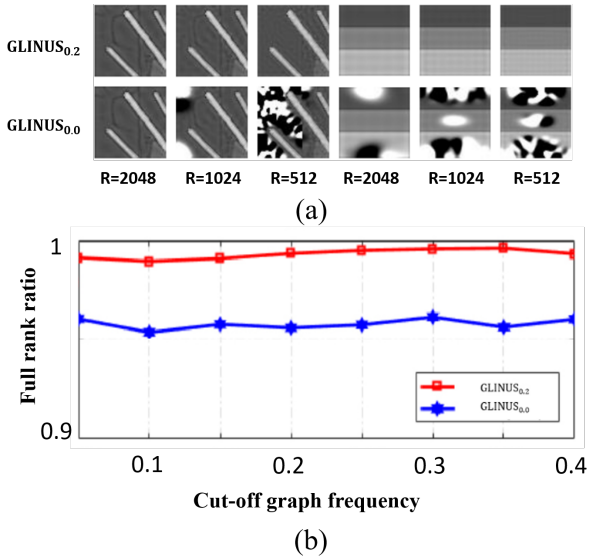

Fig. 2: (a) Reconstructed depth maps from measurements sampled by $\text{GLINUS}_{0.2}$ (top row) and $\text{GLINUS}_{0.0}$ (bottom row); (b) Normalized rank $rank(\mathbf{U}_K(\Psi^\top\Psi\mathbf{U}_K)^\dagger)/R$ curves with respect to normalized bandwidth (R/N) of depth signals on the graph frequency domain.

efficient non-uniform sampling scheme that can ensure reliable reconstruction.

Note that the major difficulty in depth reconstruction is the ambiguity in identifying the locations of discontinuities such as edges and contours given only a small fraction of samples. Under the same sampling rate, it would be easier to reconstruct the depth maps more accurately if more samples are assigned to areas around depth discontinuities. Thus, we assign higher probabilities for samples around discontinuities identified by a graph high-pass filter. The normalized graph Laplacian, denoted by $\hat{\mathbf{L}} = \mathbf{D}^{-1}\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{W}$, can be considered as a high-pass filter [43], which essentially computes the prediction error of the autoregressive filter $\mathbf{D}^{-1}\mathbf{W}$ defined on the graph $\mathcal{G}$. Let $\mathbf{e}$ be the response of $\hat{\mathbf{L}}$ on $\mathbf{x}$. The auto-regression error as the Laplacian response is defined as

$$\mathbf{e} = \mathbf{x} - \mathbf{D}^{-1}\mathbf{W}\mathbf{x}, \tag{9}$$

which can be used to design the non-uniform sampling. Specifically, depth value x_i associated with small prediction error e_i is able to be regressed from its connected neighbors on the graph, and thus should have a low sampling probability; otherwise, it should be assigned a high sampling probability. With the probabilistic sampling setting, the probability of pixel i being sampled, denoted by p_i, is defined by the normalized auto-regression error.

$$p_i = \frac{|e_i|}{\sum_i |e_i|} \tag{10}$$

Given a target sampling rate, the depth map $\mathbf{x}$ is sampled randomly according to the distribution $\{p_i\}$.

However, sampling $M \geq R$ measurements guided by the response of $\hat{\mathbf{L}}$ might not ensure the full column rank requirement on $\mathbf{U}_{\mathcal{MR}}$, which would affect the reconstruction quality as well as the numerical stability [11]. The sampling scheme in Eq. (9) focuses on only pixels near edges and contours of objects, but leaves large smooth areas severely undersampled. For stable reconstruction, we include depth samples also distributed in smooth areas. Specifically, we reshape the probability distribution by adding a constant bias to the response of the normalized graph Laplacian as follows.

$$\mathbf{e} = (1-\alpha)\left(\mathbf{x} - \mathbf{D}^{-1}\mathbf{W}\mathbf{x}\right) + \alpha\mathbf{1}, \tag{11}$$

where α is the bias and $\mathbf{1}$ is the all-one vector of the same size as $\mathbf{x}$. As a result, samples drawn according to the distribution defined by Eq. (10) and Eq. (11) can cover both smooth regions and texture regions.

To show the merit of such an improvement, Fig. 2(a) presents the visual comparison for two configurations: $\text{GLINUS}_{0.2}$ ($\alpha = 0.2$) and $\text{GLINUS}_{0.0}$ ($\alpha = 0.0$). The results show that $\text{GLINUS}_{0.0}$ cannot recover depth values for smooth regions, while $\text{GLINUS}_{0.2}$ is able to achieve successful depth reconstruction for both depth continuities and smooth areas. We also present normalized rank of the reconstruction matrix, *i.e.*, rank $\left(\mathbf{U}_{\mathcal{NR}}\left(\mathbf{U}_{\mathcal{MR}}\right)^\dagger\right)/R$, for the two sampling schemes. The results show that $\text{GLINUS}_{0.2}$ has higher ranks than $\text{GLINUS}_{0.0}$ for various cut-off graph frequencies, indicating more stable reconstruction. Note that other α values around 0.2 also yield similar results. A large α would reduce our sampling scheme to random sampling. Hence, we finally set $\alpha = 0.2$ in our implementation.

IV. Graph-Based Depth Reconstruction Algorithm

The reconstruction of a high-quality depth map from a sparse set of non-uniform samples is a severely ill-posed problem, and requires strong regularization. Previous depth reconstruction methods designed for (nearly) uniform sampling do not adapt well to heavily non-uniform undersampled depth measurements (verified in Sec. VI). To reconstruct high-quality depth map from sparse non-uniform samples, we propose an ensemble of priors (EoP) model to fully exploit signal correlation from multiple perspectives.

A. Depth Reconstruction Model with Ensemble of Priors

The observation model for depth sampling is described as

$$\mathbf{y} = \mathbf{\Psi}\mathbf{x} + \mathbf{n}, \tag{12}$$

where $\mathbf{x} \in \mathbb{R}^N$ denotes the vectorized form of the depth map with N pixels, $\mathbf{y} \in \mathbb{R}^M$ denotes the observed M samples, $\mathbf{\Psi} \in \mathbb{R}^{M\times N}$ represents the sampling matrix and $\mathbf{n} \in \mathbb{R}^M$ denotes additive noise. In depth reconstruction from highly sparse and non-uniform samples, the sampling matrix $\mathbf{\Psi}$ is flat, and makes our task highly ill-conditioned and challenging.

To tackle this problem, we propose the following depth reconstruction with an ensemble of priors (EoP):

$$\min_{\mathbf{x}} E_{\text{D}}\left(\mathbf{x}, \mathbf{y}\right) + E_{\text{EoP}}\left(\mathbf{x}\right), \tag{13}$$

where $E_{\text{D}}\left(\mathbf{x}, \mathbf{y}\right)$ is the data term to guarantee the observation consistency between $\mathbf{x}$ and $\mathbf{y}$. The ensemble of priors (EoP)

E_{EoP} is designed to exploit high-order correlation from three aspects. Specifically, the E_{EoP} term consists of three terms:

$$E_{\mathrm{EoP}}\left(\mathbf{x},\mathbf{y}\right)=\lambda_1 E_{\mathrm{tr}}\left(\mathbf{x}\right)+\lambda_2 E_{\mathrm{ls}}\left(\mathbf{x}\right)+\lambda_3 E_{\mathrm{nls}}\left(\mathbf{x}\right), \quad (14)$$

where the $E_{\mathrm{tr}}\left(\mathbf{x}\right)$ exploits the sparseness on transform domain, $E_{\mathrm{ls}}\left(\mathbf{x}\right)$ models local smoothness within the depth maps, and $E_{\mathrm{nls}}\left(\mathbf{x}\right)$ non-local smoothness across the color image and depth map. λ_1, λ_2, and λ_3 are the weighting parameters to balance the three terms. The design of the three priors are detailed in the following subsections.

B. Sparseness Prior on Multi-scale Directional Transform

Seeking effective signal representations is a well-studied subject in image processing, and wavelet transform has been proved to be an effective tool for image representation. Of various wavelet transforms, the dual-tree discrete wavelet transform (DDWT) [44], with directionality (six orientations: $\pm 75^\circ$, $\pm 15^\circ$, and $\pm 45^\circ$) and (approximate) shift invariance, provides efficient representation for directional features in images, such as edges and contours. By introducing adaptive anisotropic decomposition on DDWT subbands, adaptive dual-tree discrete wavelet packets (ADDWP) [45], [46] contains a much larger family of directional and elongated basis functions. We note that depth maps are piece-wise smooth with edges and contours as directional features, and thus can be efficiently represented by ADDWP. Therefore, we use ADDWP [46] in our transform-domain regularizer $E_{\mathrm{tr}}\left(\mathbf{x}\right)$. Since ADDWP is imposed as an analysis prior, the low-pass subbands is be excluded so that the regularizer focuses more on directional features in high-pass subbands.

Let $\mathbf{\Phi}\in\mathbb{R}^{M\times N}$ denote the partial transform matrix of ADDWP, excluding the basis functions of the low-pass subbands. The ADDWP prior is formulated as:

$$E_{\mathrm{ls}}\left(\mathbf{x}\right)=\left\|\mathbf{\Phi}\mathbf{x}\right\|_1, \quad (15)$$

where $\left\|\cdot\right\|_1$ is the ℓ_1 norm as the measurement of sparseness. Ablation results validating the effectiveness of this regularizer are presented in Sec. VI-D2.

C. Prior of Internal Local Smoothness

Depth reconstruction from only a sparse set of observed samples is severely ill-conditioned. Thus, it is desirable to further enforce internal local smoothness of the reconstructed depth map. Therefore, we introduce a total variation regularization on graph signals.

Let $\mathbf{S}$ denotes the shift operation defined on the graph [7] and a shifted version of $\mathbf{x}$ over the graph is denoted by $\mathbf{Sx}$. The normalized version of graph shift is defines as $\tilde{\mathbf{S}}:=\mathbf{S}/\left|\eta_{\max}\right|$, where $\eta_{\max}$ is the maximum eigen-value of $\mathbf{S}$. Similarly to the calculus on discrete signals that defines the discrete derivative, the derivative of a graph signal at the i^{th} vertex is defined as

$$\nabla_i(\mathbf{x})=\mathbf{x}_i-\sum_{j\in\mathcal{N}_i}\tilde{S}_{ij}\mathbf{x}_j \quad (16)$$

where the normalized shift operator $\tilde{\mathbf{S}}$ is to guarantee that the shifted signal is properly scaled for the comparison with the original signal.

While the local variation at the i^{th} pixel can be measured by $\nabla_i(\mathbf{x})$, the total variation of the graph signal $\mathbf{x}$ is the sum of variations for all vertices [7]:

$$\begin{aligned}E_{\mathrm{ls}}\left(\mathbf{x}\right)&=\sum_{i=1}^{N}\left|\mathbf{x}_i-\sum_{j\in\mathcal{N}_i}\tilde{S}_{ij}\mathbf{x}_j\right|\\&=\left\|\mathbf{x}-\tilde{\mathbf{S}}\mathbf{x}\right\|_1\\&=\left\|\mathbf{V}\mathbf{x}\right\|_1,\end{aligned} \quad (17)$$

where the variation operator $\mathbf{V}=\mathbf{I}-\tilde{\mathbf{S}}$ has the same eigenvectors as $\mathbf{S}$ with eigen-values $\xi_i=1-\eta_i/\left|\eta_{\max}\right|$ [7].

D. Prior of Color-guided Non-local Smoothness

In depth super-resolution, one important signal prior is the correlation between the depth map and the associated color image [1], [34], [47]. Although the textures of color image are usually much more complex than depth maps and could mislead depth reconstruction, depth-color pairs have strong correlation in terms of geometrical structures, *i.e.*, edges in the depth map often correspond to edges in color image [1], [47]. This implies that the locations of edges in depth maps can be inferred from the accompanying color images, and inspires us to exploit color-guided non-local graph-signal smoothness. For a depth map $\mathbf{x}$, in our method, a graph is constructed by connecting each pixel with its neighbours. Note that the adjacency matrix $\mathbf{W}$ cannot be calculated from the high quality depth map $\mathbf{x}$ to be reconstructed. Therefore, we estimated the adjacency matrix from the under-sampled depth information and its associated high-resolution color image.

Instead of simply taking a square neighborhood, we use the k-nearest-neighbourhood (kNN) method to determine a neighborhood for each pixel. Specifically, for each pixel, we first calculate its proximity to neighbouring pixels using the weighting scheme in Formula (8), and select those pixels with k largest weights as neighborhood (k = 10 in our implementation). In this way, the shapes of neighbourhoods, or equivalently the adjacency relationship, can more adapt to the content. To capture the cross depth-color correlation, we refine the weighting scheme (8) by incorporating high-resolution information from the associated color image. Without ambiguity, the adjacent matrix is formulated as

$$\begin{aligned}W_{i,j}=&\exp\left(\frac{-\|p_i-p_j\|^2}{2\sigma_1^2}\right)\exp\left(\frac{-\|\bar{x}_i-\bar{x}_j\|^2}{2\sigma_2^2}\right)\\&\exp\left(-\sum_{k\in\mathcal{K}}\|\mathbf{B}^{k,i}\circ(\mathcal{P}^{k,i}-\mathcal{P}^{k,j})\|^2/6\sigma_3^2\right),\end{aligned} \quad (18)$$

where $\bar{x}_i$ denotes the initially-estimated depth pixel (via bicubic interpolation in our implementation). In the color filter, $\mathcal{P}_i^c$ denotes an operator that extracts a $w\times w$ patch centered at pixel i from color image $\mathcal{C}=\{\mathbf{c}_k\}$, where $k\in\mathcal{K}\triangleq\{\mathrm{red},\mathrm{green},\mathrm{blue}\}$, $\circ$ represents the element-wise multiplication, and $\mathbf{B}$ denotes the bilateral filter kernel defined on the extracted $w\times w$ patch:

$$\begin{aligned}B_{m,n}^{k,i}=&\exp\left(-\|p_m-p_n\|^2/2\sigma_4^2\right)\\&\exp\left(-\|(\mathcal{P}^{k,m}-\mathcal{P}^{k,n})\|^2/6\sigma_5^2\right).\end{aligned} \quad (19)$$

In Eq. (18) and (19), $\{\sigma_i\}$ are decay rates of the range filter, depth filter or color filter.

Then, the graph Laplacian regularizer is formulated as

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} W_{i,j} \left(x_i - x_j\right)^2, \tag{20}$$

where the graph Laplacian is constructed from the adjacent matrix, *i.e.*, $\mathbf{L} = \mathbf{D} - \mathbf{W}$. However, the graph Laplacian $\mathbf{L}$ is un-normalized, and is sensitive to the vertex degrees of graph. As a remedy, we choose the random walk graph Laplacian [48] $\mathbf{L}_r = \mathbf{D}^{-1}\mathbf{L}$, which is essentially a high-pass filter. We therefore propose the graph Laplacian regularizer as follows.

$$E_{\text{lp}}\left(\mathbf{x}\right) = \left\|\mathbf{L}_r\mathbf{x}\right\|_2^2. \tag{21}$$

Note that the random walk graph Laplacian $\mathbf{L}_{\text{r}}$ estimated from the undersampled depth cannot accurately reflect the correlation among connected vertices. Therefore, we periodically update $\mathbf{L}_{\text{r}}$ as in Formula (18) by using the newly update depth map through the iterations, detailed in Algorithm 1.

V. Numerical Algorithm

By assembling the three priors, our depth reconstruction task is written as the following regularized optimization problem

$$\begin{aligned}\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2}\left\|\mathbf{\Psi}\mathbf{x} - \mathbf{y}\right\|_2^2 + \lambda_1 \left\|\mathbf{\Phi}\mathbf{x}\right\|_1 \\ +\lambda_2 \left\|\mathbf{V}\mathbf{x}\right\|_1 + \frac{\lambda_3}{2}\left\|\mathbf{L}_r\mathbf{x}\right\|_2^2.\end{aligned} \tag{22}$$

We solve the model via the alternating direction method of multipliers (ADMM) with Nesterov acceleration [49]. To handle the non-differentiable terms, *i.e.*, $\left\|\mathbf{V}\mathbf{x}\right\|_1$ and $\left\|\mathbf{\Phi}\mathbf{x}\right\|_1$, we introduce auxiliary variables, *i.e.*, ($\mathbf{u} = \mathbf{\Phi}\mathbf{x}$ and $\mathbf{v} = \mathbf{V}\mathbf{x}$), and obtain the following augmented Lagrangian function.

$$\begin{aligned}&\mathcal{L}\left(\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{z}\right) = \\ &\frac{1}{2}\left\|\mathbf{\Psi}\mathbf{x} - \mathbf{y}\right\|_2^2 + \lambda_1\left\|\mathbf{u}\right\|_1 + \lambda_2\left\|\mathbf{v}\right\|_1 + \frac{\lambda_3}{2}\left\|\mathbf{L}_r\mathbf{x}\right\|_2^2 \\ &+ \left\langle \mathbf{w}, \mathbf{u} - \mathbf{\Phi}\mathbf{x}\right\rangle + \left\langle \mathbf{z}, \mathbf{v} - \mathbf{V}\mathbf{x}\right\rangle \\ &+ \frac{\mu}{2}\left\|\mathbf{u} - \mathbf{\Phi}\mathbf{x}\right\|^2 + \frac{\rho}{2}\left\|\mathbf{v} - \mathbf{V}\mathbf{x}\right\|^2,\end{aligned} \tag{23}$$

where vector $\mathbf{w}$ and $\mathbf{y}$ are Lagrange multipliers, μ and ρ are half quadratic penalty factors. The Lagrangian (23) is minimized by alternatively solving the following sub-problems.

1) **x-subproblem**: We initially set $\mathbf{x}^{(0)}$ as the bicubic interpolation version of the input depth map $\mathbf{y}$. In the $(k+1)^{\text{th}}$ iteration, we update $\mathbf{x}^{(k+1)}$ by sovling the following $\mathbf{x}$-subproblem, given $\mathbf{u}^{(k)}$ and $\mathbf{v}^{(k)}$ from the last iteration.

$$\begin{aligned}\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}} \frac{1}{2}\left\|\mathbf{\Psi}\mathbf{x} - \mathbf{y}\right\|_2^2 + \frac{\lambda_3}{2}\left\|\mathbf{L}_r\mathbf{x}\right\|_2^2 \\ + \left\langle \mathbf{w}^{(k)}, \mathbf{u}^{(k)} - \mathbf{\Phi}\mathbf{x}\right\rangle + \left\langle \mathbf{z}^{(k)}, \mathbf{v}^{(k)} - \mathbf{V}\mathbf{x}\right\rangle \\ + \frac{\mu^{(k)}}{2}\left\|\mathbf{u}^{(k)} - \mathbf{\Phi}\mathbf{x}\right\|_2^2 + \frac{\rho^{(k)}}{2}\left\|\mathbf{v}^{(k)} - \mathbf{V}\mathbf{x}\right\|_2^2,\end{aligned} \tag{24}$$

which is quadratic in terms of $\mathbf{x}$, and is equivalent to solving the normal equation:

$$\begin{aligned}&\left(\mathbf{\Psi}^\top\mathbf{\Psi} + \mu^{(k)}\mathbf{\Phi}^\top\mathbf{\Phi} + \rho^{(k)}\mathbf{V}^\top\mathbf{V} + \lambda_3\mathbf{L}_r^\top\mathbf{L}_r\right)\mathbf{x}^{(k+1)} = \\ &\mathbf{S}^\top\mathbf{y} + \mathbf{\Phi}^\top(\mathbf{w}^{(k)} + \mu^{(k)}\mathbf{u}^{(k)}) + \mathbf{V}^\top(\mathbf{z}^{(k)} + \rho^{(k)}\mathbf{v}^{(k)}).\end{aligned} \tag{25}$$

TABLE I: Parameter Settings.

$\lambda_1 = 2\times10^{-4}$	$\lambda_2 = 1\times10^{-3}$	$\lambda_3 = 0.001 \sim 0.01$	$\mu = 0.001$	$\rho = 0.01$
$\sigma_1 = 50$	$\sigma_2 = 20$	$\sigma_3 = 20$	$\sigma_4 = 5$	$\sigma_5 = 4$

This can be efficiently solved via the preconditioned conjugate gradient method [50] instead of direct inversion.

2) **u-subproblem**: Ignoring irrelevant terms, the $\mathbf{u}$-subproblem is written as

$$\begin{aligned}\mathbf{u}^{(k+1)} = \arg\min_{\mathbf{u}} \lambda_1\left\|\mathbf{u}\right\|_1 + \left\langle \mathbf{w}^{(k)}, \mathbf{u} - \mathbf{\Phi}\mathbf{x}^{(k+1)}\right\rangle \\ + \frac{\mu^{(k)}}{2}\left\|\mathbf{u} - \mathbf{\Phi}\mathbf{x}^{(k+1)}\right\|_2^2.\end{aligned} \tag{26}$$

It has the following closed-form solution

$$\mathbf{u}^{(k+1)} = \text{soft}\left(\mathbf{\Phi}\mathbf{x}^{(k+1)} + \frac{\mathbf{w}^{(k)}}{\mu^{(k)}}, \frac{\lambda_1}{\mu^{(k)}}\right), \tag{27}$$

where $\text{soft}(\cdot)$ is the element-wise soft-thresholding function:

$$\text{soft}(\text{x}, \tau) = \begin{cases} x + \tau, & \text{if } x < -\tau, \\ 0 \quad , & \text{if } |x| \le \tau, \\ x - \tau, & \text{if } x > \tau. \end{cases} \tag{28}$$

3) **v-subproblem**: Similarly, the $\mathbf{v}$-subproblem is written as the following standard ℓ_1-norm minimization problem.

$$\begin{aligned}\mathbf{v}^{(k+1)} = \arg\min_{\mathbf{v}} \lambda_2\left\|\mathbf{v}\right\|_1 + \left\langle \mathbf{z}^{(k)}, \mathbf{v} - \mathbf{V}\mathbf{x}^{(k+1)}\right\rangle \\ + \frac{\rho^{(k)}}{2}\left\|\mathbf{v} - \mathbf{V}\mathbf{x}^{(k+1)}\right\|_2^2,\end{aligned} \tag{29}$$

which has the following closed solution

$$\mathbf{v}^{(k+1)} = \text{soft}\left(\mathbf{V}\mathbf{x}^{(k+1)} - \frac{\mathbf{z}^{(k)}}{\rho^{(k)}}, \frac{\lambda_2}{\rho^{(k)}}\right). \tag{30}$$

Other parameters are updated as

$$\begin{aligned}\mathbf{w}^{(k+1)} &= \mathbf{w}^{(k)} - \mu^{(k)}\left(\mathbf{u}^{(k+1)} - \mathbf{\Phi}\mathbf{x}^{(k+1)}\right), \\ \mathbf{z}^{(k+1)} &= \mathbf{z}^{(k)} - \rho^{(k)}\left(\mathbf{v}^{(k+1)} - \mathbf{V}\mathbf{x}^{(k+1)}\right), \\ \mu^{(k+1)} &= \gamma\mu^{(k)}, \\ \rho^{(k+1)} &= \gamma\rho^{(k)},\end{aligned} \tag{31}$$

where γ is set at 1.05 to ensure the increase of the penalty parameters.

We further accelerate the convergence of the algorithm by adding a Nesterov correction step [49] for each iteration.

$$\begin{aligned}\alpha_{k+1} &= \frac{1 + \sqrt{1 + 4\alpha_k^2}}{2}, \\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k+1)} + \frac{\alpha_k - 1}{\alpha_{k+1}}\left(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right), \\ \mathbf{u}^{(k+1)} &= \mathbf{u}^{(k+1)} + \frac{\alpha_k - 1}{\alpha_{k+1}}\left(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}\right), \\ \mathbf{v}^{(k+1)} &= \mathbf{v}^{(k+1)} + \frac{\alpha_k - 1}{\alpha_{k+1}}\left(\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\right),\end{aligned} \tag{32}$$

where the parameter α_0 is usually set to 1. The overall ADMM reconstruction algorithm is summarized in Algorithm 1.

TABLE II: Depth reconstruction results in RMSE (↓) / PE (%↓) recovered by the same algorithm to compare sampling schemes.

Method	Sampling Rates						
	1.56%	2.78%	4.00%	6.25%	11.10%	25.00%	average
uniform	5.24/2.53	2.95/0.70	2.70/0.57	2.47/0.46	2.17/0.34	1.88/0.23	2.90/0.805
random	5.75/3.61	3.15/0.84	2.86/0.65	2.56/0.52	2.18/0.36	1.70/0.22	3.03/1.033
IPCA [5]	4.81/2.42	2.52/0.50	1.99/0.31	1.46/0.19	0.76/0.056	0.32/0.010	1.98/0.581
$\text{GLINUS}_{0.0}$	2.41/1.16	1.90/0.80	1.67/0.62	1.45/0.55	1.37/0.50	1.19/0.45	1.67/0.680
$\text{GLINUS}_{0.2}$	**1.99/0.39**	**1.39/0.17**	**0.93/0.060**	**0.63/0.030**	**0.37/0.012**	**0.21/0.0091**	**0.92/0.112**

Algorithm 1 ADMM reconstruction algorithm

Input: Observed depth map $\mathbf{y}$, observation matrix $\mathbf{\Psi}$, aligned color image $\mathcal{C}$.

Output: Reconstruction depth $\mathbf{x}^*$

Construct the graph Laplacian $\mathbf{L}_r$;
Initialize $k = 0$,the maximum number of iterations K, and the stopping criterion tol;
Initialize $\mathbf{x}^{(0)} = \text{bicubic}(\mathbf{y})$, $\mathbf{u}^{(0)} = \mathbf{\Phi}\mathbf{x}^{(0)}, \mathbf{v}^{(0)} = \mathbf{V}\mathbf{x}^{(0)}$
while $k < K$ and $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{(k)}\|_2 \geq tol$ **do**
  Update $\mathbf{x}^{(k+1)}$ by solving Eq. (25)
  Update $\mathbf{u}^{(k+1)}$ by Eq. (27)
  Update $\mathbf{v}^{(k+1)}$ by Eq. (30)
  Update $\mathbf{w}^{(k+1)}$, $\mathbf{z}^{(k+1)}$ by Eq. (31)
  if $(k+1)\%10 == 0$ **then**
    Update $\mathbf{L}_r$ by Eq. (18) with $\mathbf{x}^{(k+1)}$
  end if
  Perform Nesterov acceleration by Eq. (32)
end while
return $\mathbf{x}^* \leftarrow \mathbf{x}^{(k+1)}$

VI. Experimental Results

This section evaluates the proposed sampling scheme in Sec. VI-B and depth reconstruction method in Sec. VI-C. To comprehensively investigate the behavior of the proposed algorithm, we presents ablation results in Sec. VI-D.

A. Settings

Testing depth maps in our experiments are from the Middlebury Stereo dataset [51]. Root mean square error (RMSE) is used to measure the reconstruction quality.

$$\text{RMSE}(\mathbf{x}, \mathbf{x}^{\text{gt}}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - x_i^{\text{gt}}\right)^2}, \tag{33}$$

where $\mathbf{x}$ and $\mathbf{x}^{\text{gt}}$ are the reconstructed image and ground truth, respectively.

We also use the percentage of bad pixels , denoted by $\text{PE}(\mathbf{x}, \mathbf{x}^{\text{gt}})$ to measure the depth quality [52]. Given a threshold $\tau > 0$, the percentage of bad pixels is defined as

$$\text{PE}(\mathbf{x}, \mathbf{x}^{\text{gt}}; \tau) = \frac{100}{N}\sum_{i=1}^{N}\Gamma\left(\max\left(\frac{x^{\text{gt}}}{x_i}, \frac{x_i}{x^{\text{gt}}}\right) > \tau\right), \tag{34}$$

where $\Gamma(\cdot)$ is an indication function to return one if the operand is true and return zero otherwise.

Parameters are set empirically in Table I. While other parameters are fixed, the regularization parameter λ_3 for the non-local smoothness term is chosen from the range of [0.001 $\sim$ 0.01] according the sampling rate. For low sampling rates, λ_3 is set to 0.01 for strong regularization, and is decreased up to 0.001 as more measurements are sampled at higher rates.

B. Results on Depth Sampling Schemes

Five sampling strategies are compared, including uniform sampling, random sampling, IPCA (improved principal component analysis) [5], and two variants of our proposed sampling methods: $\text{GLINUS}_{0.0}$ and $\text{GLINUS}_{0.2}$, where $\alpha = 0$ and $\alpha = 0.2$, respectively. For fair comparison, we use our depth reconstruction algorithm for all the tested sampling schemes.

Table II shows depth reconstruction results in RMSE and PE at six sampling rates averaged over the test depth images. IPCA and our GLINUS are significantly better than the uniform sampling and random sampling in both metrics, which suggests that adaptive sampling schemes is better than fixed ones. Among the two variants of our sampling scheme, $\text{GLINUS}_{0.2}$ outperforms $\text{GLINUS}_{0.0}$ in both metrics, which is in line with the stability results shown in Fig. 2. Our $\text{GLINUS}_{0.2}$ consistently outperforms IPCA, which shows the effectiveness of the proposed GLINUS scheme. For example, $\text{GLINUS}_{0.2}$ reduces average RMSE error by 53.54% (from 1.98 to 0.92) compared with IPCA. Our method also significantly reduces the proportion of depth pixels with large reconstruction errors (down to 0.112%). This would make our algorithm more appealing in depth-based rendering applications that are sensitive to large depth errors.

Fig. 3 shows depth maps reconstructed from sparse depth pixels with a sampling rate of 2.78% for these sampling schemes. The uniform and random sampling schemes produce significant artifacts around depth discontinuities. IPCA [5] produces less artifacts, but fails to preserve sharp depth edges. Our proposed GLINUS0.0 and GLINUS0.2 schemes achieve better results with more details, which demonstrate the effectiveness of our sampling scheme over other schemes.

C. Results on Depth Reconstruction

We further evaluate the performance of proposed EoP model for depth reconstruction. Ten methods are compared: bicubic interpolation, Wavelet-based method [5], the PFitDR method [53], the fast guided global interpolation (FGI) [27], robust color guided depth map restoration method (RCG) [28], edge-preserving and structure-preserving image smoothing (EPSP) [33], hierarchical features driven residual learning for depth map super-resolution (DepthSR) [39], progressive multi-branch aggregation network for scene depth super-resolution (PMBANet) [40], texture depth transformer network (TDTN)

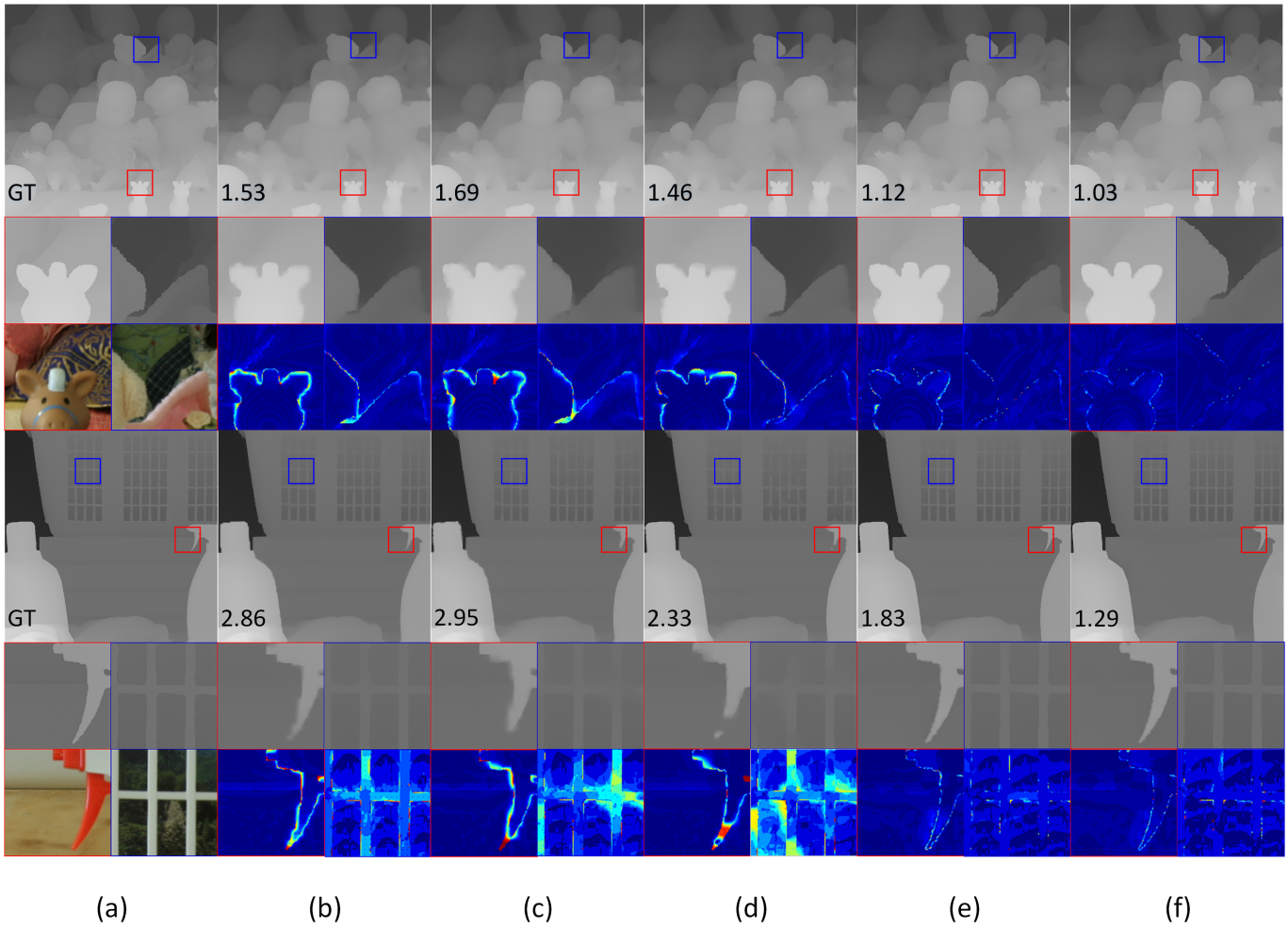

Fig. 3: Visual results of depth reconstruction for "Dolls" (top) and "Laundry"(bottom) with different sampling schemes at a ratio of 2.78%: (a) GT RGB-D pair, (b) uniform sampling, (c) random sampling, (d) IPCA [5], (e) our GLINUS$_{0.0}$, and (f) our GLINUS$_{0.2}$. For better visualization, we show two cropped patches and the associated error maps.

[41], and the primary version of our method (named GNDGRR) [42]. Depth samples are generated by our GLINUS$_{0.2}$ scheme as the input for algorithms with sparse depth input. Uniform sampling for compatible sampling rates is used for six super-resolution algorithms, including three optimization-based methods: FGI [27], RCG [28], and EPSP [33], and three deep learning based methods: DepthSR [39], PMBANet [40], and TDTN [41].

The RMSE results in Table III show that the proposed EoP model has much lower depth reconstruction error than Wavelet and PFitDR. The proposed model also significantly outperforms its primary version GNDGDR thanks to the elaborated design of the depth reconstruction algorithm. Note that Wavelet and PFitDR, which use non-uniform sampling, perform worse than generic optimization-based superresolution methods that use uniform sampling, such as FGI, RCG, and EPSP, for low sampling rates. This indicates that an effective depth reconstruction algorithm is also crucial for improving the overall reconstruction performance. For most cases of applicable sampling rates, our method outperforms deep learning based superresolution methods, *i.e.*, DepthSR, PMBANet, and TDTN. This verifies the effectiveness of our non-uniform sampling and reconstruction method.

Fig. 4 shows visual results of different depth reconstruction schemes at the sampling ratio of 1.56% for "*Art*", "*Reindeer*" and "*Aloe*". The visual results are consistent with the RMSE results reported in Table III. The two super-resolution methods, *i.e.*, FGI [27] and RCG [28], achieve acceptable reconstruction quality for smooth areas, but present annoying artifacts around depth boundaries. With the same non-uniform sampling scheme, Wavelet [5] and PFitDR [53] are able to recover sharp edges, but present significant artifacts in smooth areas: Wavelet introduces texture-copy artifacts while PFitDR presents cloud-like artifacts. In contrast, our EoP model can preserve sharp edges and boundaries with high accuracy while avoiding visible degradation in smooth areas.

D. Ablation Results

1) Algorithm Convergence: We compare the convergence performance of our EoP algorithm with Nesterov acceleration, the variant without Nesterov acceleration, and the wavelet-based method [5], which also uses an iterative optimization

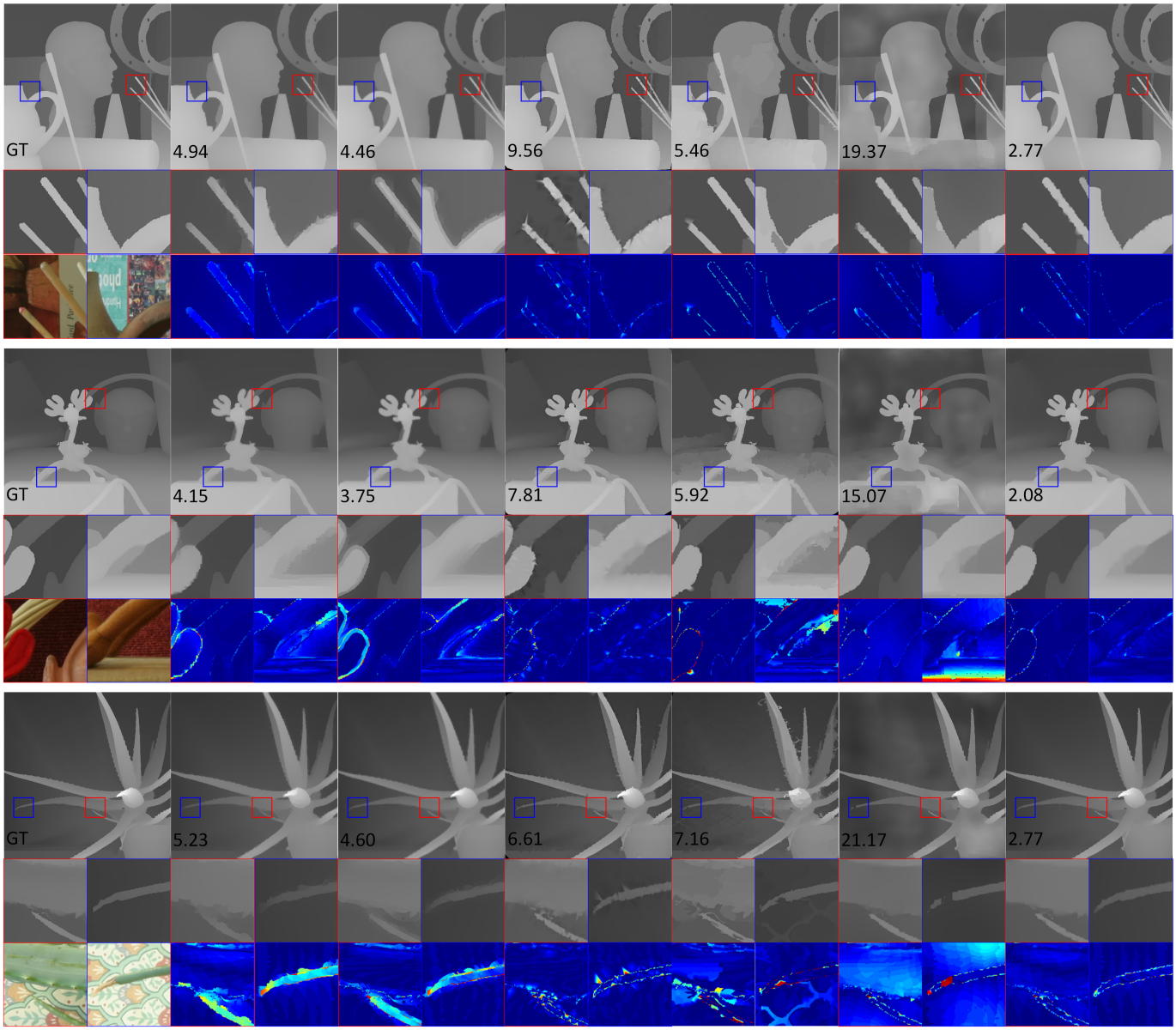

Fig. 4: Visual comparison results under different sampling schemes at the sampling ratio of 1.56% for image "*Art*"(top), "*Reindeer*"(middle) and "*Aloe*"(bottom). The results are generated by (a) GT, (b) FGI [27], (c) RCG [28], (d) Bicubic, (e) PFitDR [53], (f) Wavelet [5], and (g) our EoP model. For better visualization, for each image, we show two cropped patches as well as the associated error map in absolute difference.

TABLE III: Comparison results in RMSE with other methods. The (2^{nd}) best results are marked in (underline) bold.

Method	Percentage of samples						
	1.00%	1.56%	2.78%	4.00%	6.25%	11.1%	25.0%
FGI [27]	–	3.47	2.90	–	2.62	–	1.92
RCG [28]	–	3.04	2.97	–	2.27	–	1.56
EPSP [33]	–	2.97	2.63	–	2.21	1.89	1.50
DepthSR [39]	–	1.50	–	–	0.89	–	0.48
PMBANet [40]	–	2.19	–	–	1.21	–	0.43
TDTN [41]	–	1.40	–	–	0.79	–	0.32
Bicubic	10.64	7.60	4.49	3.79	2.76	1.60	1.05
PfitDR [53]	8.87	8.53	8.13	4.85	4.60	4.49	3.33
Wavelet [5]	8.77	5.35	3.49	2.55	1.71	0.92	0.42
GNSGDR [42]	5.53	3.31	2.75	1.74	1.08	1.25	0.48
Proposed EoP	**2.52**	1.86	**1.15**	**0.79**	**0.57**	**0.34**	**0.22**

framework. Fig. 5 shows the RMSE errors of the reconstructed depth maps compared to the ground truth. The convergence curves indicate that our EoP algorithm with Nesterov acceleration significantly reduces the reconstruction error in the early iterations, quickly reaches the knee point, and converges after sixty iterations. This demonstrates that our algorithm achieves faster convergence and more stable reconstruction than the other two algorithms.

2) Advantages of the ADDWP sparseness Prior: To evaluate the effectiveness of the sparseness prior, we compare three multiscale transforms: wavelet transform (WT), contourlet transform (CT), and the ADDWP. We use the Daubechies

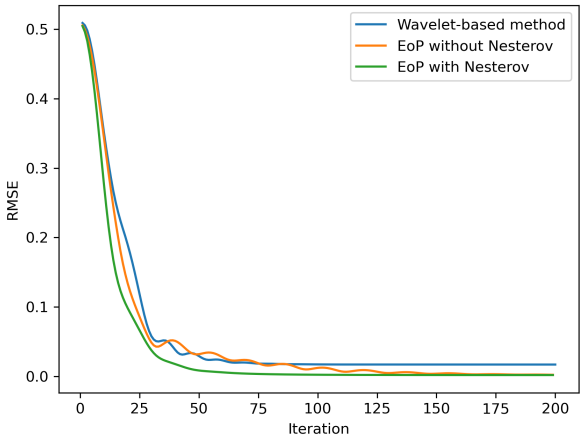

Fig. 5: Convergence curves of our EoP algorithm with Nesterov acceleration, the variant without acceleration, and the wavelet-based algorithm [5].

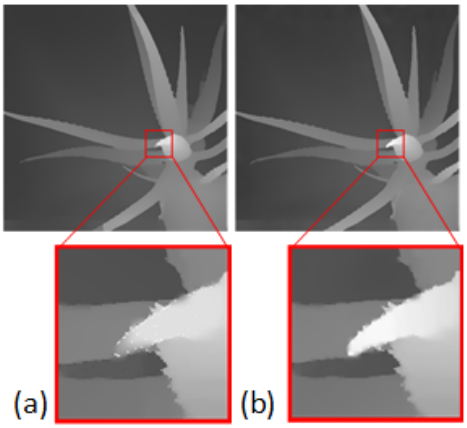

Fig. 6: Depth maps reconstructed by our algorithm with (a) a fixed graph Laplacian (RMSE = 4.34) and (b) an adaptive Laplacian (RMSE = 2.83).

TABLE IV: RMSE results of depth maps reconstructed by our EoP model with three different transforms.

Transform	Sampling Rates						
	1.56%	2.78%	4.00%	6.25%	11.10%	25.00%	average
WT	3.41	1.92	1.42	0.86	0.47	0.29	1.40
CT	3.37	1.89	1.40	0.84	0.47	0.28	1.38
ADDWP	**2.92**	**1.71**	**1.24**	**0.77**	**0.42**	**0.26**	**1.22**

DB4 for wavelet transform, the CDF 9/7 filters for pyramid decomposition in contourlet, and the filters of ADDWP as in [46]. For all the transforms, we use similar filter lengths and two-level decomposition. We also remove the external non-local smoothness (*i.e.*, $\lambda_3 = 0$) to reduce the influence of its strong regularization, and set $\lambda_2 = 1\times10^{-3}$. For the parameter of the transform regularizer, we set λ_1 to 4×10^{-5} for wavelet, to 2×10^{-4} for contourlet, and to 2×10^{-4} for ADDWP. Table IV shows that the ADDWP achieves the lowest RMSE results averaged over the test image set among the three transforms, thanks to the rich family of directional basis functions.

3) Parameter Adaptation: In our algorithm, the Laplacian $\mathbf{L}_r$ is periodically updated using the reconstructed depth map so that correlation among pixels can be better captured. Fig. 6 shows two depth maps recovered by our algorithm with a fixed graph Laplacian and the adaptive version, respectively. The adaptive Laplacian achieves lower reconstruction error with less artifacts than the fixed Laplacian, which verifies the effectiveness of our parameter adaptation.

E. Algorithm Complexity

The amount of computation in the proposed GLINUS sampling scheme mainly lies in the calculation of the adjacency matrix $\mathbf{W}$, which is a sparse matrix with non-zero elements proportional to the neighborhood size. Therefore, our GLINUS scheme has a computational complexity of $O(N)$.

For the reconstruction side, the computational cost is mainly dominated by $\mathbf{x}$-subproblem, as the other two subproblems are of $O(N)$ complexity by exploiting the spare structure of the matrices involved. Since Eq. (25) is a Toeplitz system, the preconditioned conjugate gradient method to solve the $\mathbf{x}$-subproblem has a complexity of $O(N\log^3 N)$ [54]. The overall complexity of the algorithm increase proportionally by a constant due to the number of iteration for convergence.

VII. Conclusions

This paper presents a novel framework for non-uniform sampling and reconstruction of depth maps based on graph signal analysis. We propose a GLINUS sampling scheme that allocates most of the sampling budget to pixels around depth discontinuities based on the response of the graph Laplacian filter, so that prominent features are sampled with higher probabilities. We also assign a small fraction of samples to other areas to ensure the successful reconstruction of the complete depth map. To reconstruct depth maps from highly undersampled measurements, we propose a depth reconstruction model with an ensemble of priors (EoP), which includes an ADDWP-based sparseness prior, a local smoothness prior with graph total variation, and a non-local smoothness prior with color-guided graph Laplacian regularizer. We solve the EoP model efficiently under the ADMM framework. Experimental results demonstrate that the GLINUS sampling scheme provides better depth reconstruction results than other sampling schemes when using the same reconstruction algorithm. Our EoP depth reconstruction algorithm also outperforms several state-of-the-art algorithms. The proposed framework offers a new paradigm for achieving more efficient sampling and reconstruction of depth maps.

References

[1] J. Yang, X. Ye, K. Li, C. Hou, and Y. Wang, "Color-guided depth recovery from RGB-D data using an adaptive autoregressive model," *IEEE TIP*, vol. 23, no. 8, pp. 3443–3458, 2014.

[2] Z. Jiang, Y. Hou, H. Yue, J. Yang, and C. Hou, "Depth super-resolution from RGB-D pairs with transform and spatial domain regularization," *IEEE TIP*, vol. 27, no. 5, pp. 2587–2602, 2018.

[3] H. G. Feichtinger, K. Gr, T. Strohmer *et al.*, "Efficient numerical methods in non-uniform sampling theory," *NM*, vol. 69, no. 4, pp. 423–440, 1995.

[4] S. Chen, R. Varma, A. Singh, and J. Kovačević, "Signal recovery on graphs: Fundamental limits of sampling strategies," *IEEE TSIPN*, vol. 2, no. 4, pp. 539–554, 2016.

[5] L. Liu, S. H. Chan, and T. Q. Nguyen, "Depth reconstruction from sparse samples: Representation, algorithm, and sampling," *IEEE TIP*, vol. 24, no. 6, pp. 1983–1996, June 2015.

[6] S. Mandal, A. Bhavsar, and A. K. Sao, "Depth map restoration from undersampled data," *IEEE TIP*, vol. 26, no. 1, pp. 119–134, Jan 2017.

[7] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *arXiv preprint arXiv:1211.0053*, 2012.

[8] I. Pesenson, "Sampling in paley-wiener spaces on combinatorial graphs," *TAMS*, vol. 360, no. 10, pp. 5603–5627, 2008.

[9] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *IEEE ICASSP*, 2013, pp. 5445–5449.

[10] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *IEEE ICASSP*, 2014, pp. 3864–3868.

[11] ——, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE TSP*, vol. 64, no. 14, pp. 3775–3789, 2016.

[12] H. Shomorony and A. S. Avestimehr, "Sampling large data on graphs," in *IEEE GlobalSIP*, 2014, pp. 933–936.

[13] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE TSP*, vol. 63, no. 24, pp. 6510–6523, Dec 2015.

[14] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE TSP*, vol. 64, no. 18, pp. 4845–4860, 2016.

[15] S. K. Narang and A. Ortega, "Local two-channel critically sampled filter-banks on graphs," in *IEEE ICIP*, 2010, pp. 333–336.

[16] H. Q. Nguyen and M. N. Do, "Downsampling of signals on graphs via maximum spanning trees," *IEEE TSP*, vol. 63, no. 1, pp. 182–191, 2014.

[17] S. K. Narang and A. Ortega, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE TSP*, vol. 61, no. 19, pp. 4673–4685, 2013.

[18] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE TSP*, vol. 64, no. 7, pp. 1832–1843, 2015.

[19] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst, "Random sampling of bandlimited signals on graphs," *Applied and Computational Harmonic Analysis*, vol. 44, no. 2, pp. 446–475, 2018.

[20] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE TPAMI*, vol. 35, no. 6, pp. 1397–1409, 2012.

[21] J. Lu, K. Shi, D. Min, L. Lin, and M. N. Do, "Cross-based local multipoint filtering," in *IEEE CVPR*, 2012, pp. 430–437.

[22] Q. Yang, R. Yang, J. Davis, and D. Nistér, "Spatial-depth super resolution for range images," in *IEEE CVPR*, 2007, pp. 1–8.

[23] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, "A noise-aware filter for real-time depth upsampling," 2008.

[24] D. Ferstl, C. Reinbacher, R. Ranftl, M. Rüther, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation," in *In Proc. IEEE ICCV*, 2013, pp. 993–1000.

[25] M.-Y. Liu, O. Tuzel, and Y. Taguchi, "Joint geodesic upsampling of depth images," in *In Proc. IEEE CVPR*, 2013, pp. 169–176.

[26] K.-L. Hua, K.-H. Lo, and Y.-C. F. F. Wang, "Extended guided filtering for depth map upsampling," *IEEE MultiMedia*, vol. 23, no. 2, pp. 72–83, 2015.

[27] Y. Li, D. Min, M. N. Do, and J. Lu, "Fast guided global interpolation for depth and motion," in *ECCV*. Springer, 2016, pp. 717–733.

[28] W. Liu, X. Chen, J. Yang, and Q. Wu, "Robust color guided depth map restoration," *IEEE TIP*, vol. 26, no. 1, pp. 315–327, 2016.

[29] X. Liu, D. Zhai, R. Chen, X. Ji, D. Zhao, and W. Gao, "Depth restoration from RGB-D data via joint adaptive regularization and thresholding on manifolds," *IEEE TIP*, vol. 28, no. 3, pp. 1068–1079, 2018.

[30] J. Diebel and S. Thrun, "An application of markov random fields to range sensing," in *Advances in neural information processing systems*, 2006, pp. 291–298.

[31] O. Mac Aodha, N. D. Campbell, A. Nair, and G. J. Brostow, "Patch based synthesis for single depth image super-resolution," in *ECCV*. Springer, 2012, pp. 71–84.

[32] J. Li, Z. Lu, G. Zeng, R. Gan, and H. Zha, "Similarity-aware patchwork assembly for depth image super-resolution," in *In Proc. IEEE CVPR*, 2014, pp. 3374–3381.

[33] W. Liu, P. Zhang, Y. Lei, X. Huang, J. Yang, and M. Ng, "A generalized framework for edge-preserving and structure-preserving image smoothing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6631–6648, 2021.

[34] J. Yang, X. Ye, and P. Frossard, "Global auto-regressive depth recovery via iterative non-local filtering," *IEEE TB*, vol. 65, no. 1, pp. 123–137, 2018.

[35] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. S. Kweon, "High-quality depth map upsampling and completion for RGB-D cameras," *IEEE TIP*, vol. 23, no. 12, pp. 5559–5572, 2014.

[36] J. Lu and D. Forsyth, "Sparse depth super-resolution," in *In Proc. IEEE CVPR*, 2015, pp. 2245–2253.

[37] M. Kiechle, S. Hawe, and M. Kleinsteuber, "A joint intensity and depth co-sparse analysis model for depth map super-resolution," in *In Proc. IEEE ICCV*, 2013, pp. 1545–1552.

[38] Y. Zuo, Q. Wu, Y. Fang, P. An, L. Huang, and Z. Chen, "Multi-scale frequency reconstruction for guided depth map super-resolution via deep residual network," *IEEE TCSVT*, vol. 30, no. 2, pp. 297–306, 2019.

[39] C. Guo, C. Li, J. Guo, R. Cong, H. Fu, and P. Han, "Hierarchical features driven residual learning for depth map super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2545–2557, 2018.

[40] X. Ye, B. Sun, Z. Wang, J. Yang, R. Xu, H. Li, and B. Li, "PMBANet: Progressive multi-branch aggregation network for scene depth super-resolution," *IEEE Transactions on Image Processing*, vol. 29, pp. 7427–7442, 2020.

[41] C. Yao, S. Zhang, M. Yang, M. Liu, and J. Qi, "Depth super-resolution by texture-depth transformer," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6.

[42] J. Yang, W. Xu, X. Ye, P. Frossard, and K. Li, "Graph based non-uniform sampling and reconstruction of depth maps," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 2324–2328.

[43] S. Chen, D. Tian, C. Feng, A. Vetro, and J. Kovačević, "Fast resampling of three-dimensional point clouds via graphs," *IEEE TSP*, vol. 66, no. 3, pp. 666–681, 2017.

[44] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and computational harmonic analysis*, vol. 10, no. 3, pp. 234–253, 2001.

[45] J. Yang, Y. Wang, W. Xu, and Q. Dai, "Image coding using dual-tree discrete wavelet transform," *IEEE TIP*, vol. 17, no. 9, pp. 1555–1569, Sep. 2008.

[46] J. Yang, W. Yao, W. Xu, and Q. Dai, "Image and video denoising using adaptive dual-tree discrete wavelet packets," *IEEE TCSVT*, vol. 19, no. 5, pp. 642–655, 2009.

[47] J. Yang, X. Ye, K. Li, and C. Hou, "Depth recovery using an adaptive color-guided auto-regressive model," in *ECCV*. Springer, 2012, pp. 158–171.

[48] X. Liu, G. Cheung, X. Wu, and D. Zhao, "Random walk graph laplacian-based smoothness prior for soft decoding of jpeg images," *IEEE TIP*, vol. 26, no. 2, pp. 509–524, 2016.

[49] V. Hryhorenko, D. Klyushin, and S. Lyashko, "Multiblock admm in machine learning," in *2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT)*. IEEE, 2019, pp. 461–464.

[50] S. F. Ashby and R. D. Falgout, "A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations," *NSE*, vol. 124, no. 1, pp. 145–159, 1996.

[51] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nešić, X. Wang, and P. Westling, "High-resolution stereo datasets with subpixel-accurate ground truth," in *GCPR*. Springer, 2014, pp. 31–42.

[52] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, no. 1-3, pp. 7–42, 2002.

[53] C. S. Balure, A. Bhavsar, and M. R. Kini, "Local segment-based dense depth reconstruction from very sparsely sampled data," in *IEEE Twenty-third NCC*, 2017, pp. 1–6.

[54] Y. Lu and C. M. Hurvich, "On the complexity of the preconditioned conjugate gradient algorithm for solving toeplitz systems with a fisher–hartwig singularity," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 3, pp. 638–653, 2005.