DECOUPLED GLOBAL-LOCAL IMPLICIT REPRESENTATION FOR CONTINUOUS NONLOCAL RECONSTRUCTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recovering signals from incomplete or noisy observations requires structural priors that remain effective across diverse sampling patterns. Nonlocal self-similarity (NSS) provides a powerful source of such prior knowledge by exploiting recurring local structures across spatially distant regions. Conventional NSS methods, however, typically rely on explicit neighborhood construction and grouping, making them less suitable for irregularly sampled observations and continuous coordinate queries. Implicit neural representations (INRs) offer a flexible alternative by modeling signals as continuous coordinate functions, but conventional INR formulations generally treat the signal as a single global implicit function and do not explicitly exploit recurring nonlocal structures. To bridge this gap, we propose a **Decoupled Global-Local Implicit Representation (D-GLIR)** for continuous nonlocal reconstruction. D-GLIR decomposes the implicit representation into shared local implicit bases and a coordinate-conditioned coefficient field. The local branch parameterizes reusable structural patterns over relative coordinates, while the global branch continuously maps spatial coordinates to the coefficients that combine these patterns at arbitrary query locations. Unlike conventional INRs that model the signal through a single global coordinate-to-value mapping, D-GLIR allows spatially distant regions to reuse the same local structural bases through different coefficient combinations. This introduces NSS as an intrinsic inductive bias of the representation while preserving its continuous coordinate-based nature, without explicit block matching, precomputed patch groups, or an additional NSS regularization term. Experiments on image inpainting, image denoising, off-grid function regression, point-cloud recovery, and weather data reconstruction demonstrate the effectiveness and flexibility of D-GLIR across diverse signal types, sampling patterns, and degradation conditions.

1 INTRODUCTION

Multidimensional signals, including images, point clouds, and geophysical observations, are sampled on regular grids or at irregular locations. Implicit neural representations (INRs) model these signals as continuous coordinate functions, enabling evaluation beyond observed samples (Xie et al., 2022). Advances in periodic activations, Fourier features, and multiresolution encodings (Sitzmann et al., 2020; Tancik et al., 2020; Müller et al., 2022), together with multiscale, localized, and adaptive frequency modeling (Lindell et al., 2022; Saragadam et al., 2023; Liu et al., 2024; Zhou et al., 2025), have improved their ability to represent fine-scale structures. INRs have consequently been applied to image reconstruction and three-dimensional shape modeling (Chen et al., 2021; Park et al., 2019; Mescheder et al., 2019). As illustrated in Fig. 1(a), conventional continuous representations directly fit the observations without a predefined reconstruction grid, but do not explicitly exploit recurring nonlocal structures.

Reconstruction from incomplete or noisy observations nevertheless requires effective structural priors. NeRP combines prior-image information with measurement physics (Shen et al., 2022); LRTFR extends low-rank modeling to continuous functions (Luo et al., 2023); and NeurTV characterizes local regularity through coordinate derivatives (Luo et al., 2025). Beyond external information, global

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

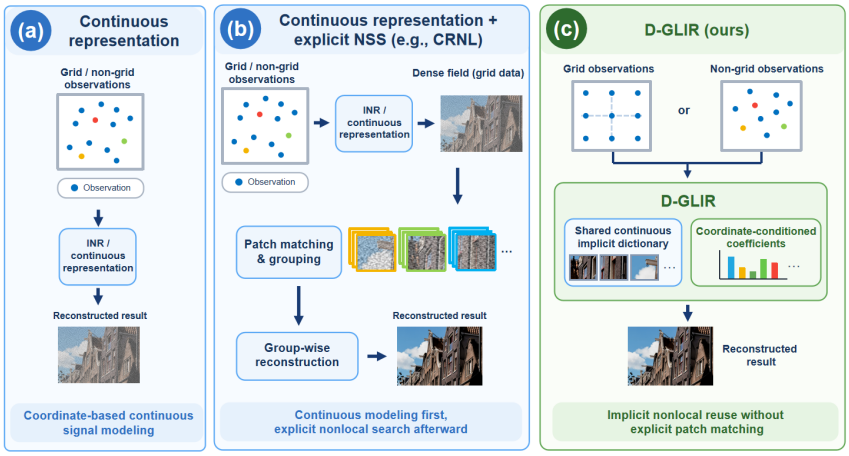


Figure 1: **Comparison of continuous reconstruction paradigms.** (a) Conventional continuous representations directly model the underlying signal from regularly or irregularly sampled observations, without explicitly exploiting nonlocal structural recurrence. (b) CRNL first constructs a continuous representation, converts it into a dense field for explicit patch matching and grouping, and then performs group-wise reconstruction. (c) D-GLIR directly learns a shared continuous implicit dictionary and coordinate-conditioned coefficients from the observations, enabling nonlocal structural reuse within the representation itself without explicit patch matching or group construction.

low-rank structure, and local smoothness, recurring local patterns provide a complementary prior: observations in one region can support the recovery of structurally similar regions elsewhere.

This type of recurring structure is commonly characterized by nonlocal self-similarity (NSS), a widely used prior in signal and image reconstruction. By relating local patterns across different spatial locations, NSS allows information from structurally similar regions to support the recovery of missing or corrupted content. Early methods such as nonlocal means exploit this redundancy through similarity-weighted aggregation (Buades et al., 2005), while BM3D groups similar patches for collaborative filtering in a transformed three-dimensional domain (Dabov et al., 2007). This idea was subsequently incorporated into sparse-representation models, including nonlocal sparse modeling, nonlocally centralized sparse representation (NCSR), and group-based sparse representation (GSR) (Mairal et al., 2009; Dong et al., 2012; Zhang et al., 2014). A related line of work exploits the low-rank structure of groups formed by similar patches, leading to methods based on nonlocal low-rank approximation, spectral priors, and weighted nuclear norm minimization (Dong et al., 2013; Wang et al., 2013; Gu et al., 2017). The same principle has also been extended beyond conventional image restoration to hyperspectral image super-resolution, video restoration, and multidimensional tensor completion (Dian et al., 2017; Wang et al., 2020; Xie et al., 2018; Zhang et al., 2019). Together, these studies establish NSS as an effective means of exploiting internal redundancy that is complementary to local smoothness and global low-rank structure.

Recent work also explores structural and nonlocal priors in continuous representations. The Implicit Neural Representation Regularizer (INRR) uses an additional INR to parameterize a learned Dirichlet-energy regularizer that captures structural correlations (Li et al., 2023). The Continuous Representation-based NonLocal method (CRNL) first fits an INR, partitions its representation into local cubes, explicitly matches and groups similar cubes, and reconstructs the groups through coupled low-rank function factorization (Luo et al., 2024). As illustrated in Fig. 1(b), these stages establish nonlocal correspondences through a separate search-and-group procedure rather than joint end-to-end optimization. Their computational overhead motivates embedding structural reuse directly into the representation, without intermediate reconstruction or explicit grouping.

Related efforts have sought to incorporate nonlocal interactions directly into learnable models. Nonlocal neural networks replace local-only processing with feature aggregation over spatially distant positions (Wang et al., 2018), while Neural Nearest Neighbors Networks (N3Net) introduce a continuous relaxation of nearest-neighbor selection to enable nonlocal matching within end-to-end train-

able architectures (Plötz & Roth, 2018). These developments alleviate some limitations of conventional hard correspondence selection, but still operate through explicit pairwise interactions or candidate-neighbor comparisons. The difficulty becomes more pronounced for continuous representations and irregularly sampled observations, where fixed patch layouts and point-to-point neighborhood alignment are no longer readily available. The central challenge is therefore not simply to make matching differentiable, but to enable local structures to be reused across distant locations as part of the continuous representation itself, while avoiding explicit correspondence search and retaining joint optimization directly from the observations.

To address this challenge, we propose a **Decoupled Global-Local Implicit Representation (D-GLIR)** for continuous nonlocal reconstruction. As shown in Fig. 1(c), the local branch maps relative coordinates to shared implicit bases, while the global branch maps absolute coordinates to coefficients that combine these bases at each location. Sharing the bases allows spatially distant regions to reuse structural components without explicit correspondence search. Both branches are jointly optimized from observations, embedding structural reuse directly into the representation without precomputed patch groups or an additional NSS regularizer. This coordinate-based formulation supports both regularly gridded and irregularly sampled observations.

The main contributions of this work are summarized as follows:

- We propose D-GLIR, a decoupled continuous representation that separates reusable local structures from their spatially varying expressions through shared local implicit bases and a coordinate-conditioned coefficient field.
- We embed nonlocal structural reuse directly into the representation and jointly optimize its global and local branches from observations, eliminating the need for explicit block matching, precomputed patch groups, or an additional NSS regularization term.
- We evaluate D-GLIR on image inpainting, image denoising, off-grid function regression, point-cloud recovery, and weather data reconstruction, demonstrating its effectiveness across diverse signal types, sampling patterns, and degradation conditions.

2 METHODOLOGY

This section presents the formulation of the proposed Decoupled Global-Local Implicit Representation (D-GLIR). We first introduce a continuous representation of local structures and then develop a global-local factorization to parameterize this representation. Finally, we describe the joint learning objective and the implementation details of the proposed framework.

2.1 PROBLEM FORMULATION

Let $f^\star : \Omega \to \mathbb{R}^C$ denote an unknown clean signal defined over a coordinate domain $\Omega \subset \mathbb{R}^d$, where d is the coordinate dimension and C is the number of signal channels. Both regularly gridded and irregularly sampled data can be viewed as observations of this continuous function at a finite set of coordinates. Suppose that only a degraded set of observations is available:

$$\mathcal{D}_{\text{obs}} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N}, \qquad \mathbf{y}_i = f^\star(\mathbf{x}_i) + \boldsymbol{\varepsilon}_i, \tag{1}$$

where $\mathbf{x}_i \in \Omega$ denotes an observed coordinate and $\boldsymbol{\varepsilon}_i$ represents measurement noise. For missing-data and sparse-sampling problems, the observed coordinates form only a subset of the full signal domain. The objective is to recover a continuous reconstruction function $\widehat{f} : \Omega \to \mathbb{R}^C$ from $\mathcal{D}_{\text{obs}}$ such that

$$\widehat{f}(\mathbf{x}) \approx f^\star(\mathbf{x}), \qquad \mathbf{x} \in \Omega. \tag{2}$$

Once learned, $\widehat{f}$ can be evaluated at both observed and unobserved coordinates, allowing reconstruction on regular grids as well as arbitrary query locations.

2.2 DECOUPLED GLOBAL-LOCAL IMPLICIT REPRESENTATION

2.2.1 CONTINUOUS REPRESENTATION OF LOCAL STRUCTURES

Conventional nonlocal methods exploit recurring local structures through explicit patch construction and matching, which typically relies on neighborhoods that can be consistently extracted and aligned

on a shared sampling grid. In contrast, D-GLIR models local structures directly in a continuous coordinate domain. It complements the conventional global coordinate-to-value mapping with a local-coordinate formulation, allowing recurring local patterns to be represented and reused across different spatial locations.

For an anchor coordinate $\mathbf{s} \in \Omega$, let $\mathbf{r} \in \mathcal{R}$ denote a relative coordinate within its local neighborhood, where $\mathcal{R}$ defines a bounded relative-coordinate domain centered at the anchor. We define the continuous local field centered at $\mathbf{s}$ as

$$g^\star(\mathbf{s}, \mathbf{r}) = f^\star(\mathbf{s} + \mathbf{r}), \qquad \mathbf{r} \in \mathcal{R}, \tag{3}$$

where $\mathbf{s}$ specifies the spatial location of the local neighborhood and $\mathbf{r}$ specifies its internal coordinate.

This formulation provides a local-coordinate view of the original signal: for each anchor $\mathbf{s}$, the function $g^\star(\mathbf{s}, \cdot)$ describes the signal variation within its neighborhood. Rather than directly parameterizing the signal $f^\star(\mathbf{x})$ with a single coordinate-to-value function, D-GLIR learns a continuous representation of this local field, allowing local structures at different locations to be jointly modeled through shared components.

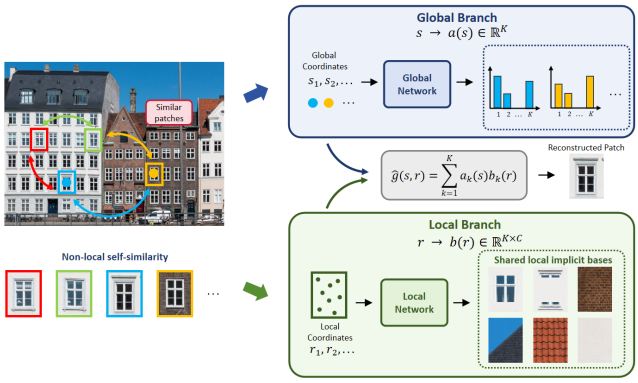


Figure 2: Overview of the proposed D-GLIR for continuous nonlocal reconstruction. Instead of explicit patch matching, recurring local structures across spatial locations are represented in a shared continuous basis space.

2.2.2 Global-Local Decomposition of the Continuous Local Field

To parameterize the continuous local field defined in Eq. 3, D-GLIR decomposes its dependence on the anchor coordinate $\mathbf{s}$ and the relative coordinate $\mathbf{r}$:

$$\widehat{g}(\mathbf{s}, \mathbf{r}) = \sum_{k=1}^{K} a_k(\mathbf{s}) \mathbf{b}_k(\mathbf{r}), \tag{4}$$

where K denotes the number of shared components, $\mathbf{b}_k(\mathbf{r}) \in \mathbb{R}^C$ is the k-th local basis function, and $a_k(\mathbf{s})$ is its location-dependent coefficient.

Here, $\widehat{g}$ is the continuous implicit representation learned by D-GLIR to approximate $g^\star$. The basis functions $\mathbf{b}_k(\mathbf{r})$ characterize reusable local structural components over the relative-coordinate domain, while $a_k(\mathbf{s})$ specifies how these components are combined at each spatial location.

Since the same basis functions are shared over the entire signal domain, local regions at different positions are represented in a common structural space. Regions with similar local patterns can therefore activate related combinations of shared bases, allowing nonlocal structural reuse to emerge within the representation itself.

Figure 2 illustrates the overall formulation. Instead of explicitly searching for similar neighborhoods, local structures throughout the signal domain are represented in a shared continuous basis space, while their spatial variations are encoded by location-dependent coefficients.

2.2.3 CONTINUOUS PARAMETERIZATION

Both components in Eq. 4 are parameterized as continuous functions:

$$\mathbf{a}(\mathbf{s}) = \mathcal{G}_{\theta_g}(\mathbf{s}), \qquad \mathbf{b}(\mathbf{r}) = \mathcal{F}_{\theta_l}(\mathbf{r}), \tag{5}$$

where $\mathbf{a}(\mathbf{s}) = [a_1(\mathbf{s}), \ldots, a_K(\mathbf{s})]^{\mathrm{T}}$ and $\mathbf{b}(\mathbf{r}) = [\mathbf{b}_1(\mathbf{r}), \ldots, \mathbf{b}_K(\mathbf{r})]$.

This coordinate-based parameterization enables continuous querying without requiring a fixed sampling lattice, making D-GLIR applicable to both regularly gridded and irregularly sampled observations.

D-GLIR is also independent of a particular function approximator. The functions $\mathcal{G}_{\theta_g}$ and $\mathcal{F}_{\theta_l}$ may be instantiated using different differentiable coordinate-based models according to the dimensionality, sampling structure, and computational requirements of the reconstruction task.

2.3 JOINT LEARNING FROM OBSERVATIONS

Unlike conventional NSS methods, D-GLIR does not require nonlocal groups to be constructed before reconstruction. The shared local bases and their spatial coefficients are learned jointly and directly from the observations.

The anchor coordinate $\mathbf{s}$ defines a reference location for representing a local structure and is not restricted to an observed sample. In practice, a finite set of anchor locations is selected according to the sampling pattern and reconstruction task. For an anchor $\mathbf{s}$ and a relative coordinate $\mathbf{r} \in \mathcal{R}$, the corresponding signal coordinate is

$$\mathbf{x} = \mathbf{s} + \mathbf{r}. \tag{6}$$

Therefore, an observed sample can contribute to multiple local representations associated with different anchor locations. Each observation thus constrains both the signal value and the shared local structural space.

Let

$$\mathcal{D}_{\mathrm{local}} = \{(\mathbf{s}_j, \mathbf{r}_j, \mathbf{y}_j)\}_{j=1}^{M} \tag{7}$$

denote the set of local training tuples constructed from available observations. Each tuple satisfies $\mathbf{s}_j + \mathbf{r}_j = \mathbf{x}_i$, where $\mathbf{x}_i$ is an observed coordinate and $\mathbf{s}_j$ is a selected reference location.

The selection of anchor locations only affects the construction of local training tuples, while the learned representation remains continuous over both anchor and relative-coordinate domains.

The reconstruction loss is defined as

$$\mathcal{L}_{\mathrm{data}} = \sum_{(\mathbf{s},\mathbf{r},\mathbf{y}) \in \mathcal{D}_{\mathrm{local}}} \|\widehat{g}(\mathbf{s}, \mathbf{r}) - \mathbf{y}\|_2^2 . \tag{8}$$

For missing-data and sparse-sampling tasks, $\mathcal{L}_{\mathrm{data}}$ is evaluated only at observed coordinates. For noisy measurements, additional regularization can be introduced to prevent overfitting. The overall objective is formulated as

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \mathcal{L}_{\mathrm{reg}}, \tag{9}$$

where $\mathcal{L}_{\mathrm{data}}$ enforces consistency with the observed measurements and $\mathcal{L}_{\mathrm{reg}}$ denotes optional task-dependent regularization.

Importantly, these task-dependent regularizations are not responsible for the nonlocal behavior of D-GLIR. Nonlocal structural reuse is induced by the shared basis representation in Eq. 4, rather than by an additional NSS loss or explicitly constructed patch groups.

2.4 CONTINUOUS SIGNAL QUERY

D-GLIR models the local signal over the entire relative-coordinate domain $\mathcal{R}$, while the reconstructed value at a target coordinate can be obtained through a center query. Given $\mathbf{x} \in \Omega$, we set the anchor to $\mathbf{s} = \mathbf{x}$ and the relative coordinate to $\mathbf{r} = \mathbf{0}$:

$$\widehat{f}(\mathbf{x}) = \widehat{g}(\mathbf{x}, \mathbf{0}) = \sum_{k=1}^{K} a_k(\mathbf{x}) \mathbf{b}_k(\mathbf{0}). \tag{10}$$

Although only the center value is queried for final reconstruction, the complete basis functions $\mathbf{b}_k(\mathbf{r})$ over $\mathcal{R}$ are essential during learning. Observations at nonzero relative coordinates jointly constrain the shape of each basis and enable local structures surrounding different anchors to share information.

This center-query strategy also avoids the need to average multiple overlapping patch predictions or introduce manually designed patch-fusion weights. Once optimized, the reconstructed signal can be queried directly at any desired coordinate.

2.5 Efficient Instantiation Used in Experiments

The formulation above defines D-GLIR at the functional level and is independent of a particular parameterization of $\mathcal{F}_{\theta_l}$ and $\mathcal{G}_{\theta_g}$. For the main experiments, we adopt a lightweight implementation that preserves the global-local decomposition while reducing the computational cost of optimization.

2.5.1 Local Branch

The shared local basis functions are parameterized by learnable control lattices and evaluated continuously through tensor-product cubic interpolation. For a d-dimensional signal domain, we define the learnable control parameters as

$$\mathbf{F} \in \mathbb{R}^{K \times \underbrace{p \times \cdots \times p}_{d \text{ dimensions}} \times C}, \tag{11}$$

where K is the number of shared bases, p denotes the resolution of the control lattice along each spatial dimension, and C is the signal dimensionality.

Given a relative coordinate $\mathbf{u} = (u_1, \ldots, u_d)$, the value of the k-th basis in channel c is obtained by

$$B_k(\mathbf{u}, c) = \sum_{n_1=1}^{p} \cdots \sum_{n_d=1}^{p} F_{k,n_1,\ldots,n_d,c} \prod_{r=1}^{d} \phi\Big(u_r - \xi_{n_r}^{(r)}\Big), \tag{12}$$

where $\xi_{n_r}^{(r)}$ denotes the control coordinate along the r-th spatial dimension and $\phi(\cdot)$ is the cubic interpolation kernel:

$$\phi(t) = \begin{cases} 1.5|t|^3 - 2.5|t|^2 + 1, & |t| \le 1, \\ -0.5|t|^3 + 2.5|t|^2 - 4|t| + 2, & 1 < |t| \le 2, \\ 0, & |t| > 2. \end{cases} \tag{13}$$

This formulation provides continuous evaluation over the local coordinate domain while avoiding the cost of repeatedly evaluating a deep network. It naturally accommodates signals of different spatial dimensionalities: $d = 2$ yields bicubic interpolation for images and other two-dimensional signals, whereas $d = 3$ yields the corresponding tricubic formulation for point-cloud data. When suitable observed local structures are available, the control lattices are initialized from randomly selected local observations, providing meaningful structural patterns at the beginning of optimization and improving convergence stability.

2.5.2 Global Branch

The global branch associates each anchor $\mathbf{s}_j$ with a coefficient vector that combines the shared local structural bases. The anchors are arranged on a d-dimensional grid and need not coincide with the observed samples. Depending on the data size and computational requirements, we either optimize the grid coefficients directly or generate them from a learnable low-resolution latent tensor:

$$\widetilde{\mathbf{A}} = \mathcal{H}_d\left(\mathrm{Upsample}(\mathbf{Z})\right), \tag{14}$$

where $\mathbf{Z}$ denotes the learnable latent tensor and $\mathcal{H}_d$ is an optional lightweight convolutional refinement module. We use 3×3 convolutions for two-dimensional domains and $3 \times 3 \times 3$ convolutions for three-dimensional domains. When convolutional refinement is omitted, $\mathcal{H}_d$ is the identity mapping. The low-resolution parameterization reduces the number of latent variables, while upsampling

and convolutional refinement introduce spatial coupling among neighboring coefficients. For direct parameterization, $\widetilde{\mathbf{A}}$ itself is optimized as a learnable coefficient tensor.

We further apply a learnable soft-thresholding operator to encourage compact combinations of the shared bases:

$$A_{j,k} = \mathcal{S}_{\lambda_k}\left(\widetilde{A}_{j,k}\right), \qquad \mathcal{S}_\lambda(x) = \operatorname{sign}(x)\max(|x| - \lambda, 0), \tag{15}$$

where j indexes a grid anchor and λ_k is a learnable nonnegative threshold associated with the k-th basis. This encourages each anchor to represent its local structure using a compact subset of the shared bases.

To evaluate the global coefficients at a query location, we interpolate the thresholded coefficient tensor $\mathbf{A}$ using the same tensor-product cubic construction as in Eqs. 12 and 13. The query coordinate is first mapped to grid-index coordinates, and the interpolation is applied independently to each coefficient channel. This defines a continuous coefficient field over the grid domain, using bicubic interpolation in two dimensions and tricubic interpolation in three dimensions, for the center-query reconstruction.

To examine the effect of branch parameterization, we consider three variants within the same global-local formulation: SIREN-global replaces only the global branch with a SIREN, SIREN-local replaces only the local branch, and SIREN-both uses SIREN for both branches. Unchanged branches retain the parameterizations described above. These variants are compared in Appendix B.6; additional per-sampling-rate results for SIREN-global are provided in Appendix Table 5.

3 Experiments

We evaluate D-GLIR on both regular meshgrid and irregular coordinate-based data. The on-meshgrid experiments include image inpainting and denoising, whereas the beyond-meshgrid experiments cover synthetic function regression, weather recovery and denoising, and point-cloud color regression and denoising. These tasks jointly examine the robustness of D-GLIR to missing or noisy observations under regular and irregular sampling. Additional experimental results and implementation details are provided in Appendix A and B.

3.1 Experimental Settings

D-GLIR is optimized independently for each degraded observation without requiring external paired training data. The anchor configuration is determined according to the sampling geometry and observation pattern of each dataset. For regularly gridded data, anchors are naturally associated with the underlying spatial lattice. For irregularly sampled data, anchors are uniformly placed on a regular grid spanning the coordinate domain. The grid resolution is chosen according to the total number of data points, yielding approximately the same number of anchors as data points. All reported metrics are averaged over five independent trials. Further details on the datasets, compared methods, evaluation metrics, implementation, and optimization settings are provided in Appendix A.

3.2 Image Inpainting Results

We evaluate D-GLIR on RGB image and MSI inpainting. Table 1 reports the quantitative results averaged over sampling rates of 5%, 10%, 15%, and 20%. Representative and additional qualitative comparisons are provided in Appendix B.

As shown in Table 1, D-GLIR achieves the best overall performance on both RGB images and MSIs. On RGB images, it obtains the best average results across all three metrics, demonstrating its ability to preserve recurring local structures under sparse observations. Its improvement over SIREN further indicates the benefit of explicitly modeling reusable local patterns beyond direct coordinate fitting. The advantage is more pronounced on MSIs, where D-GLIR achieves the best average results across all three metrics. These results support the effectiveness of jointly modeling shared local structures and their spatially varying combinations. The effects of the individual branch parameterizations are examined in Appendix B.6.

Table 1: Quantitative comparison on RGB image and MSI inpainting. Results are averaged over sampling rates of 5%, 10%, 15%, and 20%, with each rate evaluated using five independent random samplings. The best and second-best results are highlighted in bold and underlined, respectively.

Method	Color images			MSIs		
	PSNR↑	SSIM↑	NRMSE↓	PSNR↑	SSIM↑	NRMSE↓
Observed	5.1450	0.0265	0.9350	14.5650	0.2098	0.9350
TRLRF	20.3000	0.4035	0.1773	28.5525	0.7255	0.1885
SIREN	24.3000	0.7223	0.1103	26.4275	0.7158	0.2358
FTNN	19.8500	0.4230	0.1988	35.3700	0.9228	0.0990
HLRTF	23.1450	0.5440	0.1280	38.3450	0.9480	0.0710
t-ϵ-LogDet	20.9150	0.3778	0.1603	32.2825	0.8420	0.1308
TCTV	<u>26.3425</u>	0.7490	<u>0.0850</u>	34.2875	0.9145	0.1040
CRNL	25.7725	<u>0.7688</u>	0.0918	<u>39.0125</u>	<u>0.9668</u>	<u>0.0663</u>
Ours	**26.5300**	**0.7883**	**0.0848**	**39.8525**	**0.9708**	**0.0573**

Table 2: Quantitative comparison on off-meshgrid multivariate regression. Results are averaged over sampling rates of 10%, 15%, and 20%, with each sampling rate evaluated using five independent random samplings.

Method	Explicit functions			Weather data		
	f_1 NRMSE / R^2	f_2 NRMSE / R^2	f_3 NRMSE / R^2	55°N 117°W NRMSE / R^2	53°N 111°W NRMSE / R^2	47°N 100°W NRMSE / R^2
SVR	0.2497 / 0.8337	0.2787 / 0.8297	0.1393 / 0.6307	0.0180 / 0.9437	0.0217 / 0.6880	0.0237 / 0.7863
KNR	0.0637 / 0.9890	0.0683 / 0.9893	0.0467 / 0.9567	0.0147 / 0.9617	0.0190 / 0.7617	0.0197 / 0.8493
DT	0.1330 / 0.9517	0.1623 / 0.9410	0.0683 / 0.9100	0.0203 / 0.9287	0.0223 / 0.6687	0.0240 / 0.7787
RF	0.0743 / 0.9847	0.0820 / 0.9843	0.0397 / 0.9690	0.0147 / 0.9623	<u>0.0180 / 0.7837</u>	0.0187 / 0.8643
FSA-HTF	0.0140 / 0.9993	0.0187 / 0.9990	0.0197 / 0.9927	<u>0.0137 / 0.9670</u>	0.0190 / 0.7490	<u>0.0177 / 0.8760</u>
CRNL	<u>0.0090 / 0.9997</u>	<u>0.0133 / 0.9997</u>	<u>0.0073 / 0.9983</u>	0.0140 / 0.9617	0.0207 / 0.6833	0.0183 / 0.8607
Ours	**0.0063 / 1.0000**	**0.0043 / 1.0000**	**0.0033 / 1.0000**	**0.0127 / 0.9713**	**0.0160 / 0.8287**	**0.0153 / 0.9077**

3.3 Off-Meshgrid Multivariate Regression

We further evaluate D-GLIR on off-meshgrid multivariate regression, where observations are sparsely and irregularly distributed in the coordinate domain. Table 2 reports the results averaged over the 10%, 15%, and 20% sampling rates.

As shown in Table 2, D-GLIR achieves the best overall performance on both the explicit functions and the real-world weather data. For the three explicit functions, it attains R^2 values rounded to 1.0000, demonstrating accurate recovery from sparse and irregular observations without requiring a predefined regular grid. This result highlights the effectiveness of jointly modeling reusable local variations and their spatially adaptive combinations. D-GLIR also consistently outperforms the compared methods on all three weather datasets. These results indicate that the learned representation can capture not only smooth analytical functions but also heterogeneous spatial variations, demonstrating its applicability to irregular coordinate domains.

3.4 Point-Cloud Color Regression

We further evaluate D-GLIR on point-cloud color regression, where RGB attributes are recovered from sparsely observed points on irregular 3D surfaces. Table 3 reports the results averaged over the 10%, 15%, and 20% sampling rates.

As shown in Table 3, D-GLIR achieves the best average performance on all five point clouds, consistently obtaining lower NRMSE and higher R^2 than the competing methods. These results demonstrate its effectiveness in recovering RGB attributes from sparse observations on irregular 3D surfaces.

D-GLIR also consistently outperforms CRNL across all samples, indicating that the proposed global-local representation more effectively captures spatially varying color patterns in point clouds.

Table 3: Quantitative comparison on point-cloud color regression. Results are averaged over sampling rates of 10%, 15%, and 20%, with each rate evaluated using five independent random samplings.

Method	Doll NRMSE / R^2	Duck NRMSE / R^2	Frog NRMSE / R^2	Mario NRMSE / R^2	PeterRabbit NRMSE / R^2
SVR	0.2950 / 0.0930	0.1437 / 0.1720	0.2123 / 0.2117	0.2783 / 0.0163	0.1520 / 0.0037
KNR	0.1140 / 0.8640	0.0513 / 0.8930	0.0493 / 0.9567	0.0850 / 0.9077	0.0587 / 0.8527
DT	0.1147 / 0.8613	0.0547 / 0.8803	0.0523 / 0.9517	0.0917 / 0.8927	0.0707 / 0.7853
RF	0.1250 / 0.8377	0.0553 / 0.8770	0.0553 / 0.9467	0.1027 / 0.8660	0.0757 / 0.7550
FSA-HTF	0.1423 / 0.7877	0.0570 / 0.8700	0.0570 / 0.9433	0.0957 / 0.8833	0.0673 / 0.8023
HLRTF	0.1490 / 0.7683	0.0827 / 0.7270	0.0720 / 0.9097	0.1243 / 0.8040	0.0800 / 0.7253
CRNL	0.1053 / 0.8830	0.0410 / 0.9320	0.0420 / 0.9693	0.0690 / 0.9387	0.0530 / 0.8773
Ours	**0.1030 / 0.8883**	**0.0393 / 0.9377**	**0.0407 / 0.9710**	**0.0677 / 0.9420**	**0.0517 / 0.8837**

The qualitative comparison in Appendix Fig. 7 further shows improved color fidelity and fewer reconstruction artifacts.

3.5 Denoising on Heterogeneous Data

We evaluate D-GLIR on denoising across three heterogeneous data types: images, weather fields, and point clouds. Table 4 reports representative results under the strongest noise level, $\sigma = 0.25$. Complete results under all evaluated noise levels are provided in Appendix B.4.

Table 4: Quantitative comparison of denoising performance at $\sigma = 0.25$. Results are averaged over five independent trials.

Image denoising			Weather denoising			Point-cloud denoising		
Method	PSNR↑	NRMSE↓	Method	R^2 ↑	NRMSE↓	Method	R^2 ↑	NRMSE↓
Observed	13.1238	0.3767	Observed	0.937925	0.013603	Observed	-0.584794	0.230291
LRTDTV	20.8169	0.1568	KNN-Mean	0.936236	0.012149	KNN-Median	0.790142	0.086457
HLRTF	18.1657	0.2140	Gaussian	0.942738	0.011507	KNN-Mean	0.846685	0.076767
CRNL	21.2728	0.1487	IDW	0.938115	0.013571	Gaussian-KNN	0.848379	0.076162
RCTV	23.8374	0.1117	Local-Linear	0.947085	0.011085	Bilateral-KNN	-0.354299	0.212831
LTDL	23.4070	0.1175	FSA	0.934754	0.012433	HLRTF	0.747807	0.106195
Ours	**24.1136**	**0.1072**	CRNL	0.958689	0.010471	FSA-HTF	0.835998	0.083599
			Ours	**0.963303**	**0.010262**	CRNL	0.857728	0.075647
						Ours	**0.875267**	**0.070767**

Under severe noise, D-GLIR achieves the best performance across all three denoising tasks. For image denoising, it obtains the highest PSNR and the lowest NRMSE, demonstrating effective noise removal while retaining the underlying image content.

D-GLIR also achieves the best R^2 and NRMSE on both weather fields and point clouds. These results demonstrate that the proposed global-local representation can effectively recover underlying signals across regular image grids, irregular spatial fields, and 3D point sets, indicating its robustness to heterogeneous data structures.

4 Conclusion

We presented D-GLIR, a decoupled global-local implicit representation for continuous nonlocal reconstruction. By combining shared local bases with spatially varying coefficients, D-GLIR enables structural reuse across distant regions without explicit block matching, patch grouping, or an additional NSS regularizer. Experiments demonstrate its effectiveness across diverse signal types, sampling patterns, and noise levels, while computational comparisons show a favorable balance between reconstruction quality and efficiency. Future work will focus on scaling the framework to larger datasets and incorporating task-specific priors for more challenging reconstruction problems.

AI USE STATEMENT

Generative AI tools were used to improve the readability of the manuscript. It was not used to generate the research idea, design the algorithm, or run the experiments. All AI-assisted changes were reviewed by the authors. The authors take responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

To support reproducibility, Section 3.1 summarizes the observation-only optimization protocol, anchor-selection strategy, and repeated-trial procedure used in our experiments. Appendix A provides complete dataset descriptions, implementation and optimization details, baseline methods, evaluation metrics, anchor construction rules, and computational-cost settings. Appendix B reports complete per-setting quantitative results and additional qualitative comparisons. The theoretical analysis appendix states the assumptions underlying the theoretical results and provides the corresponding proofs.

REFERENCES

Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 2, pp. 60–65. Ieee, 2005.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8628–8638, 2021.

Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16 (8):2080–2095, 2007.

Renwei Dian, Leyuan Fang, and Shutao Li. Hyperspectral image super-resolution via non-local sparse tensor factorization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5344–5353, 2017.

Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li. Nonlocally centralized sparse representation for image restoration. *IEEE transactions on Image Processing*, 22(4):1620–1630, 2012.

Weisheng Dong, Guangming Shi, and Xin Li. Nonlocal image restoration with bilateral variance estimation: A low-rank approach. *IEEE Transactions on Image Processing*, 22(2):700–711, 2013. doi: 10.1109/TIP.2012.2221729.

Xiao Gong, Wei Chen, and Jie Chen. A low-rank tensor dictionary learning method for hyperspectral image denoising. *IEEE Transactions on Signal Processing*, 68:1168–1180, 2020. doi: 10.1109/TSP.2020.2971441.

Shuhang Gu, Qi Xie, Deyu Meng, Wangmeng Zuo, Xiangchu Feng, and Lei Zhang. Weighted nuclear norm minimization and its applications to low level vision. *International journal of computer vision*, 121(2):183–208, 2017.

Tai-Xiang Jiang, Michael K Ng, Xi-Le Zhao, and Ting-Zhu Huang. Framelet representation of tensor nuclear norm for third-order tensor completion. *IEEE Transactions on Image Processing*, 29:7233–7244, 2020.

Nikos Kargas and Nicholas D Sidiropoulos. Supervised learning and canonical decomposition of multivariate functions. *IEEE Transactions on Signal Processing*, 69:1097–1107, 2021.

Seungyong Lee, George Wolberg, and Sung Yong Shin. Scattered data interpolation with multilevel b-splines. *IEEE transactions on visualization and computer graphics*, 3(3):228–244, 2002.

Zhemin Li, Hongxia Wang, and Deyu Meng. Regularize implicit neural representation by itself. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10280–10288. IEEE, 2023.

David B Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. Bacon: Band-limited coordinate networks for multiscale scene representation. In *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 16231–16241. IEEE, 2022.

Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. Finer: Flexible spectral-bias tuning in implicit neural representation by variableperiodic activation functions. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2713–2722. IEEE, 2024.

Yisi Luo, Xi-Le Zhao, Deyu Meng, and Tai-Xiang Jiang. Hlrtf: Hierarchical low-rank tensor factorization for inverse problems in multi-dimensional imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19303–19312, 2022.

Yisi Luo, Xile Zhao, Zhemin Li, Michael K Ng, and Deyu Meng. Low-rank tensor function representation for multi-dimensional data recovery. *IEEE transactions on pattern analysis and machine intelligence*, 46(5):3351–3369, 2023.

Yisi Luo, Xile Zhao, and Deyu Meng. Revisiting nonlocal self-similarity from continuous representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(1):450–468, 2024.

Yisi Luo, Xile Zhao, Kai Ye, and Deyu Meng. Neurtv: Total variation on the neural domain. *SIAM Journal on Imaging Sciences*, 18(2):1101–1140, 2025.

Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. Non-local sparse models for image restoration. In *2009 IEEE 12th international conference on computer vision*, pp. 2272–2279. IEEE, 2009.

Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4460–4470, 2019.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, 41(4):1–15, 2022.

Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 165–174, 2019.

Jiangjun Peng, Hailin Wang, Xiangyong Cao, Xinling Liu, Xiangyu Rui, and Deyu Meng. Fast noise removal in hyperspectral images via representative coefficient total variation. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–17, 2022.

Tobias Plötz and Stefan Roth. Neural nearest neighbors networks. *Advances in Neural information processing systems*, 31, 2018.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. Wire: Wavelet implicit neural representations. In *2023 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 18507–18516. IEEE, 2023.

Liyue Shen, John Pauly, and Lei Xing. Nerp: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. *IEEE transactions on neural networks and learning systems*, 35(1):770–782, 2022.

Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33:7462–7473, 2020.

Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in neural information processing systems*, 33:7537–7547, 2020.

Hailin Wang, Jiangjun Peng, Wenjin Qin, Jianjun Wang, and Deyu Meng. Guaranteed tensor recovery fused low-rankness and smoothness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9):10990–11007, 2023.

Shenlong Wang, Lei Zhang, and Yan Liang. Nonlocal spectral prior model for low-level vision. In *Computer Vision – ACCV 2012*, volume 7726, pp. 231–244. Springer, 2013. doi: 10.1007/978-3-642-37431-9_18.

Xiaolong Wang, Ross Girshick, Abhinav Gupta, and Kaiming He. Non-local neural networks. In *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 7794–7803. IEEE, 2018.

Yao Wang, Jiangjun Peng, Qian Zhao, Yee Leung, Xi-Le Zhao, and Deyu Meng. Hyperspectral image restoration via total variation regularized low-rank tensor decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(4):1227–1243, 2017.

Yugang Wang, Ting-Zhu Huang, Xi-Le Zhao, and Tai-Xiang Jiang. Video deraining via nonlocal low-rank regularization. *Applied Mathematical Modelling*, 79:896–913, 2020.

Ting Xie, Shutao Li, Leyuan Fang, and Licheng Liu. Tensor completion via nonlocal low-rank regularization. *IEEE transactions on cybernetics*, 49(6):2344–2354, 2018.

Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. Neural fields in visual computing and beyond. In *Computer graphics forum*, volume 41, pp. 641–676. Wiley Online Library, 2022.

Ming Yang, Qilun Luo, Wen Li, and Mingqing Xiao. 3-d array image data completion by tensor decomposition and nonconvex regularization approach. *IEEE Transactions on Signal Processing*, 70:4291–4304, 2022.

Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE transactions on image processing*, 19(9):2241–2253, 2010.

Longhao Yuan, Chao Li, Danilo Mandic, Jianting Cao, and Qibin Zhao. Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pp. 9151–9158, 2019.

Jian Zhang, Debin Zhao, and Wen Gao. Group-based sparse representation for image restoration. *IEEE transactions on image processing*, 23(8):3336–3351, 2014.

Lefei Zhang, Liangchen Song, Bo Du, and Yipeng Zhang. Nonlocal low-rank tensor completion for visual data. *IEEE transactions on cybernetics*, 51(2):673–685, 2019.

Ting-Wei Zhou, Xi-Le Zhao, Wei-Hao Wu, Jian-Li Wang, and Yi-Si Luo. Frequency-aware implicit neural representation for multi-dimensional data recovery. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(11):10862–10874, 2025.

Decoupled Global-Local Implicit Representation for Continuous Nonlocal Reconstruction

Supplementary Material

A Additional Experimental Details

A.1 Tasks and Datasets

On-meshgrid restoration. Image inpainting is evaluated on four RGB images, namely *Peppers*, *Plane*, *Sailboat*, and *House*, and on multispectral images from the CAVE dataset (Yasuma et al., 2010). For each image, 5%, 10%, 15%, or 20% of the pixels are randomly retained as observations. Image denoising is evaluated under additive Gaussian noise with different standard deviations, following common multidimensional restoration settings (Wang et al., 2017; Gong et al., 2020; Peng et al., 2022; Luo et al., 2024).

Beyond-meshgrid restoration. The coordinate-based experiments use synthetic functions from the scattered-data interpolation benchmark (Lee et al., 2002), North American precipitation data[1], and colored point clouds obtained from the SHOT dataset website[2]. Each dataset is represented as a collection of coordinate-value pairs, and the objective is to predict scalar values or point attributes at unobserved coordinates.

Synthetic functions. We randomly sample 3000 points from $[0, 1]^2$ and evaluate the following three functions:

$$f_1(x, y) = \frac{1.25 + \cos(5.4y)}{6 + 6(3x - 1)^2}, \tag{16}$$

$$f_2(x, y) = \frac{1}{9}\left[\tanh(9 - 9x - 9y) + 1\right], \tag{17}$$

$$f_3(x, y) = \frac{\sqrt{64 - 81\left[(x - 0.5)^2 + (y - 0.5)^2\right]}}{9} - 0.5. \tag{18}$$

Weather data. We use three spatial precipitation datasets centered at $(55°\mathrm{N}, 117°\mathrm{W})$, $(53°\mathrm{N}, 111°\mathrm{W})$, and $(47°\mathrm{N}, 100°\mathrm{W})$. These datasets contain irregularly distributed spatial observations with nonuniform local variations, providing more realistic regression benchmarks than the explicit functions.

Point-cloud data. We further evaluate multi-channel attribute regression on five colored point clouds: *Doll*, *Duck*, *Frog*, *Mario*, and *PeterRabbit*. Each sample is represented by a 3D coordinate and its associated RGB attribute, and the task is to infer the colors of unobserved points from sparse observations. Unlike image inpainting, these samples lie on irregular 3D surfaces and do not possess a predefined lattice neighborhood.

For all synthetic, weather, and point-cloud datasets, 10%, 15%, and 20% of the samples are retained for model fitting. The reported NRMSE and R^2 values at each sampling rate are averaged over five independent random samplings.

Denoising settings. Additive Gaussian noise with standard deviations $\sigma = 0.10$, 0.15, 0.20, and 0.25 is applied to the clean data. Results for images, weather fields, and point clouds are reported in Appendix B.4; additional MSI denoising results at $\sigma = 0.25$ are reported in Appendix B.5.

Computational-cost setting. Computational efficiency is evaluated on `house.mat` ($512 \times 512 \times 3$) at SR $= 0.10$. The results are reported in Appendix B.7.

[1] `https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=2130`

[2] `http://www.vision.deis.unibo.it/research/80-shot`

A.2 Implementation Details

Training protocol and anchor selection. D-GLIR is independently optimized for each degraded observation without external paired training data. All models are implemented in PyTorch and optimized with Adam on an NVIDIA RTX 3090 GPU.

For RGB image inpainting, D-GLIR uses 5×5 local patches and 32 local bases. For MSI inpainting, the same patch size and 64 bases are used. For image denoising, we use 9×9 patches and 192 local bases. The optimization objective combines reconstruction fidelity, total-variation regularization, coefficient sparsity, and noise separation followed the previous works (Luo et al., 2024; 2025).

For irregularly sampled data, spatial coordinates are normalized, and anchors are uniformly placed on a regular grid spanning the coordinate domain. The grid resolution is chosen according to the total number of data points, yielding approximately the same number of anchors as data points. Local training neighborhoods are constructed using nearest observed samples according to Euclidean distances in the normalized coordinate space. The global coefficients are parameterized on the anchor grid as described in the main paper and evaluated at query coordinates using tensor-product cubic interpolation. The anchor grid provides an internal parameterization and does not require the observations to lie on a regular lattice. The coefficients and shared local bases are jointly optimized using only observed values, without external paired training data.

Compared methods and evaluation metrics. For image inpainting, we compare D-GLIR with TRLRF (Yuan et al., 2019), SIREN (Sitzmann et al., 2020), FTNN (Jiang et al., 2020), HLRTF (Luo et al., 2022), t-ϵ-LogDet (Yang et al., 2022), TCTV (Wang et al., 2023), and CRNL (Luo et al., 2024). The SIREN-global variant replaces only the global branch with SIREN and retains the interpolation-based local bases. For denoising, the compared methods additionally include LRTDTV (Wang et al., 2017), RCTV (Peng et al., 2022), and LTDL (Gong et al., 2020). For coordinate-based regression and point-cloud recovery, we compare with SVR, KNR, DT, RF, FSA-HTF (Kargas & Sidiropoulos, 2021), HLRTF, and CRNL.

For image-like data, reconstruction quality is evaluated using PSNR, SSIM, and NRMSE. For synthetic functions, weather data, and point-cloud attributes, we report NRMSE and R^2. Higher PSNR, SSIM, and R^2, together with lower NRMSE, indicate better performance.

B Complete Experimental Results

B.1 Image Inpainting Results

Table 5 presents the detailed RGB image and MSI inpainting results at sampling rates of 5%, 10%, 15%, and 20%, complementing the averaged results reported in the main paper. Qualitative comparisons on Plane and additional RGB/MSI examples are provided in Figs. 3 and 4, respectively.

Table 5: Quantitative comparison of RGB image and MSI inpainting at different sampling rates. The best and second-best values are highlighted.

Sampling rate (SR)		0.05			0.10			0.15			0.20		
Dataset	Method	PSNR	SSIM	NRMSE	PSNR	SSIM	NRMSE	PSNR	SSIM	NRMSE	PSNR	SSIM	NRMSE
Color images 512×512×3	Observed	4.78	0.015	0.975	5.01	0.023	0.949	5.26	0.030	0.922	5.53	0.038	0.894
	TRLRF	15.85	0.148	0.282	20.14	0.380	0.171	22.13	0.510	0.135	23.08	0.576	0.121
	SIREN	20.81	0.583	0.158	24.19	0.723	0.108	25.62	0.775	0.092	26.58	0.808	0.083
	FTNN	13.57	0.148	0.361	19.69	0.394	0.183	22.19	0.530	0.138	23.95	0.620	0.113
	HLRTF	19.16	0.353	0.194	22.75	0.534	0.127	24.74	0.623	0.102	25.93	0.666	0.089
	t-ϵ-LogDet	17.60	0.208	0.227	20.23	0.339	0.167	22.14	0.441	0.134	23.69	0.523	0.113
	TCTV	24.00	0.670	0.109	25.87	0.738	0.088	27.20	0.779	**0.076**	28.30	0.809	**0.067**
	CRNL	23.24	0.670	0.119	25.45	0.761	0.093	26.78	0.809	0.081	27.62	0.835	0.074
	SIREN-global	23.46	0.689	0.117	25.70	0.773	0.091	26.78	0.805	0.081	27.46	0.826	0.075
	Ours	**24.34**	**0.721**	**0.106**	**26.11**	**0.778**	**0.087**	**27.34**	**0.813**	0.077	**28.33**	**0.841**	0.069
MSIs 256×256×31	Observed	14.20	0.159	0.975	14.43	0.193	0.949	14.68	0.227	0.922	14.95	0.260	0.894
	TRLRF	26.12	0.591	0.241	28.78	0.743	0.183	29.45	0.774	0.169	29.86	0.794	0.161
	SIREN	24.34	0.627	0.291	26.00	0.701	0.244	27.19	0.750	0.215	28.18	0.785	0.193
	FTNN	30.23	0.844	0.156	34.42	0.925	0.101	37.28	0.954	0.077	39.55	0.968	0.062
	HLRTF	31.92	0.872	0.128	37.44	0.957	0.070	40.69	0.977	0.049	43.33	0.986	0.037
	t-ϵ-LogDet	27.28	0.690	0.212	31.45	0.842	0.133	34.15	0.902	0.099	36.25	0.934	0.079
	TCTV	29.89	0.836	0.158	33.29	0.912	0.110	35.89	0.946	0.083	38.08	0.964	0.065
	CRNL	33.16	0.926	0.110	37.73	0.969	0.068	41.24	0.983	0.049	43.92	0.989	0.038
	SIREN-global	32.09	0.875	0.134	34.51	0.920	0.101	35.33	0.932	0.091	36.23	0.943	0.083
	Ours	**34.72**	**0.938**	**0.093**	**38.62**	**0.971**	**0.060**	**41.80**	**0.984**	**0.043**	**44.27**	**0.990**	**0.033**

As shown in Table 5, for RGB image inpainting, D-GLIR achieves the best PSNR and SSIM at all sampling rates and remains competitive in NRMSE. At the challenging sampling rate of 5%, it obtains a PSNR of 24.34 dB and an SSIM of 0.721. Compared with SIREN, D-GLIR improves PSNR by 3.53 dB and 1.75 dB at sampling rates of 5% and 20%, respectively. SIREN-global also outperforms the original SIREN in the reported RGB results. Since SIREN-global retains the same interpolation-based local bases as the full model, their comparison concerns the global parameterization. The local-branch comparisons are reported in Appendix B.6.

In Table 5, the improvements are more pronounced on MSIs, where D-GLIR achieves the best results across all three metrics at every sampling rate. Its PSNR increases from 34.72 dB at a sampling rate of 5% to 44.27 dB at 20%. At the 5% sampling rate, D-GLIR improves PSNR over CRNL by 1.56 dB and reduces NRMSE from 0.110 to 0.093. At the 20% sampling rate, it further reduces NRMSE from 0.038 to 0.033. These results demonstrate that D-GLIR effectively captures recurring spatial-spectral structures across different levels of observation sparsity.

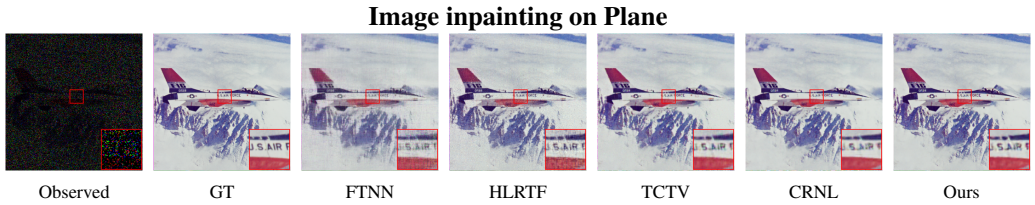


Figure 3: Qualitative comparison of image inpainting on *Plane*. From left to right: Observed, GT, FTNN, HLRTF, TCTV, CRNL, and Ours.

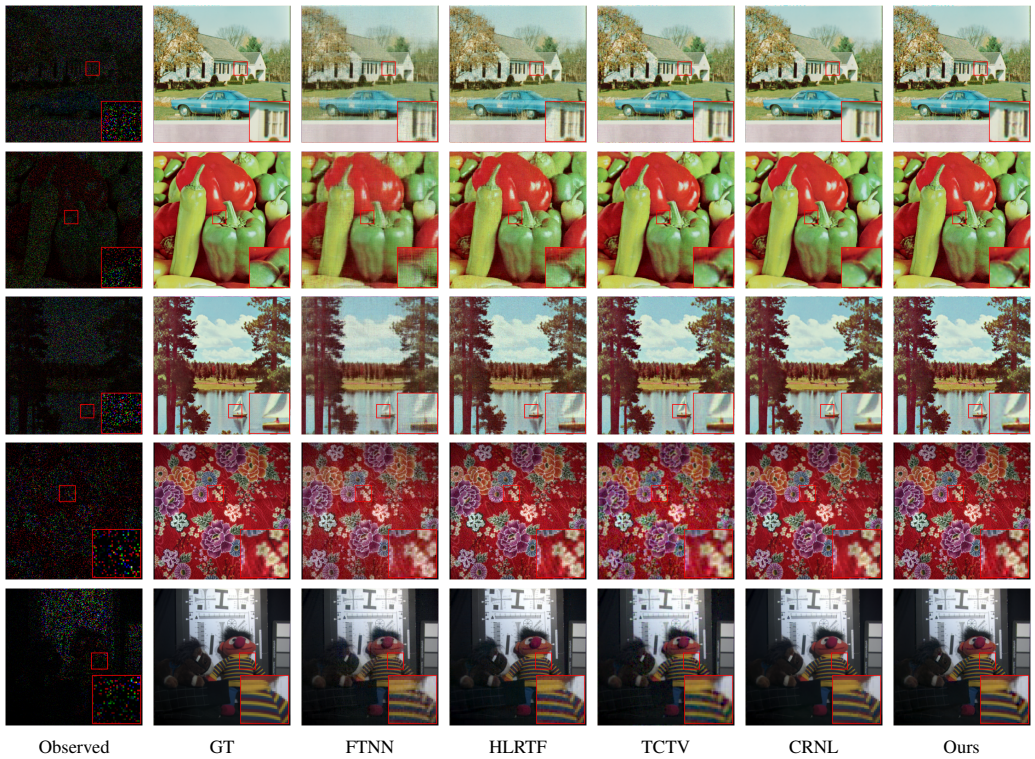


Figure 4: Qualitative comparison of inpainting results at a sampling rate of 10% on three RGB images and two MSIs. From top to bottom: *House*, *Peppers*, *Sailboat*, *Cloth*, and *Toys*. From left to right: Observed, GT, FTNN, HLRTF, TCTV, CRNL, and Ours.

B.2 Off-Meshgrid Multivariate Regression

Table 6 presents the detailed off-meshgrid regression results at sampling rates of 10%, 15%, and 20%, complementing the averaged results reported in the main paper. The synthetic functions and sampling setup are specified in Appendix A.1, Eqs. equation 16–equation 18. Additional visual comparisons on the explicit functions and weather fields at a sampling rate of 10% are provided in Figs. 5 and 6, respectively.

Table 6: Quantitative comparison of off-meshgrid multivariate regression at different sampling rates. Performance is measured by NRMSE↓ and R^2 ↑. The best and second-best results are highlighted in bold and underlined, respectively.

Rate	Method	Explicit functions			Weather data		
		f_1 NRMSE / R^2	f_2 NRMSE / R^2	f_3 NRMSE / R^2	55°N 117°W NRMSE / R^2	53°N 111°W NRMSE / R^2	47°N 100°W NRMSE / R^2
10%	SVR	0.253 / 0.830	0.282 / 0.827	0.138 / 0.642	0.019 / 0.937	0.022 / 0.671	0.025 / 0.765
	KNR	0.073 / 0.986	0.084 / 0.985	0.060 / 0.931	0.017 / 0.950	0.021 / 0.711	0.022 / 0.811
	DT	0.156 / 0.935	0.187 / 0.923	0.081 / 0.877	0.023 / 0.906	0.024 / 0.621	0.028 / 0.699
	RF	0.088 / 0.979	0.110 / 0.973	0.050 / 0.953	0.017 / 0.950	0.020 / 0.743	0.021 / 0.825
	FSA-HTF	0.015 / 0.999	0.022 / 0.999	0.022 / 0.991	0.016 / 0.956	0.024 / 0.610	0.023 / 0.797
	CRNL	0.010 / 0.999	0.015 / 0.999	0.009 / 0.998	0.018 / 0.940	0.023 / 0.612	0.024 / 0.778
	Ours	**0.007 / 1.000**	**0.005 / 1.000**	**0.004 / 1.000**	**0.016 / 0.956**	**0.018 / 0.779**	**0.018 / 0.872**
15%	SVR	0.249 / 0.834	0.278 / 0.831	0.140 / 0.624	0.018 / 0.945	0.022 / 0.692	0.024 / 0.784
	KNR	0.063 / 0.989	0.066 / 0.990	0.042 / 0.966	0.014 / 0.965	0.019 / 0.771	0.020 / 0.844
	DT	0.132 / 0.953	0.163 / 0.942	0.065 / 0.919	0.020 / 0.934	0.023 / 0.659	0.023 / 0.798
	RF	0.072 / 0.986	0.075 / 0.988	0.037 / 0.974	0.014 / 0.965	0.018 / 0.787	0.019 / 0.868
	FSA-HTF	0.014 / 0.999	0.017 / 0.999	0.019 / 0.993	0.013 / 0.970	0.017 / 0.804	0.015 / 0.912
	CRNL	0.008 / 1.000	0.013 / 1.000	0.005 / 0.999	0.012 / 0.974	0.021 / 0.665	0.017 / 0.882
	Ours	**0.006 / 1.000**	**0.005 / 1.000**	**0.003 / 1.000**	**0.012 / 0.975**	**0.016 / 0.835**	**0.015 / 0.913**
20%	SVR	0.247 / 0.837	0.276 / 0.831	0.140 / 0.626	0.017 / 0.949	0.021 / 0.701	0.022 / 0.810
	KNR	0.055 / 0.992	0.055 / 0.993	0.038 / 0.973	0.013 / 0.970	0.017 / 0.803	0.017 / 0.893
	DT	0.111 / 0.967	0.137 / 0.958	0.059 / 0.934	0.018 / 0.946	0.020 / 0.726	0.021 / 0.839
	RF	0.063 / 0.989	0.061 / 0.992	0.032 / 0.980	0.013 / 0.972	0.016 / 0.821	0.016 / 0.900
	FSA-HTF	0.013 / 1.000	0.017 / 0.999	0.018 / 0.994	0.012 / 0.975	0.016 / 0.833	0.015 / 0.919
	CRNL	0.009 / 1.000	0.012 / 1.000	0.008 / 0.998	0.012 / 0.971	0.018 / 0.773	0.014 / 0.922
	Ours	**0.006 / 1.000**	**0.003 / 1.000**	**0.003 / 1.000**	**0.010 / 0.983**	**0.014 / 0.872**	**0.013 / 0.938**

In Table 6, D-GLIR consistently achieves the lowest NRMSE and an R^2 close to or equal to 1.000 across all sampling rates. It also outperforms CRNL on every function, with particularly clear improvements at the most challenging sampling rate of 10%. These results demonstrate the effectiveness of D-GLIR in recovering continuous functions from sparse, irregular observations.

Table 6 also shows best or tied-best performance across the three weather regions. Its advantage is particularly evident in the second and third regions, where it consistently obtains lower NRMSE and higher R^2 than the competing methods. As shown in Figs. 5 and 6, D-GLIR produces reconstructions that more closely match the ground truth while retaining both global trends and local variations. Overall, the results confirm its effectiveness across different functions, spatial fields, and sampling rates.

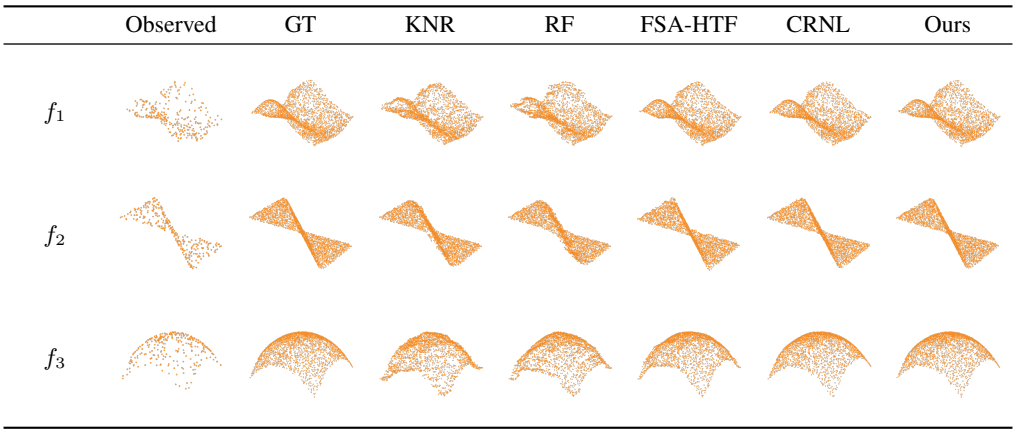


Figure 5: Qualitative comparison of off-meshgrid regression on three explicit functions at a sampling rate of 10%. Each row corresponds to one function. From left to right: Observed, GT, KNR, RF, FSA-HTF, CRNL, and Ours.

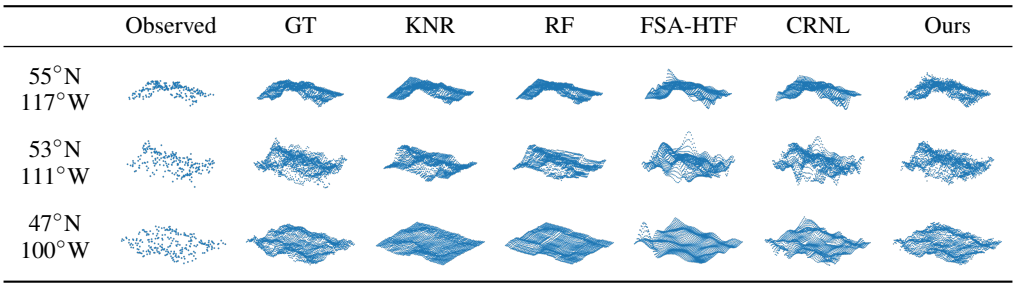


Figure 6: Qualitative comparison of off-meshgrid regression on three weather regions at a sampling rate of 10%. Each row corresponds to one region. From left to right: Observed, GT, KNR, RF, FSA-HTF, CRNL, and Ours.

B.3 POINT-CLOUD COLOR REGRESSION

Table 7 reports the detailed results for point-cloud color regression at sampling rates of 10%, 15%, and 20%, complementing the averaged results presented in the main paper. Qualitative results are shown in Figs. 7 and 8. Results cover five point clouds: *Doll*, *Duck*, *Frog*, *Mario*, and *PeterRabbit*.

Table 7: Quantitative comparison of point-cloud color regression at different sampling rates. Performance is measured by NRMSE↓ and R^2 ↑. The best and second-best results are highlighted in bold and underlined, respectively.

Rate	Method	Point-cloud data				
		Doll NRMSE / R^2	Duck NRMSE / R^2	Frog NRMSE / R^2	Mario NRMSE / R^2	PeterRabbit NRMSE / R^2
10%	SVR	0.297 / 0.082	0.144 / 0.168	0.223 / 0.131	0.280 / 0.005	0.153 / -0.006
	KNR	0.128 / 0.829	0.057 / 0.868	0.055 / 0.947	0.095 / 0.886	0.065 / 0.820
	DT	0.131 / 0.822	0.060 / 0.856	0.058 / 0.941	0.103 / 0.865	0.078 / 0.740
	RF	0.132 / 0.819	0.058 / 0.864	0.058 / 0.942	0.107 / 0.856	0.078 / 0.738
	FSA-HTF	0.155 / 0.749	0.058 / 0.864	0.058 / 0.941	0.101 / 0.870	0.073 / 0.766
	HLRTF	0.150 / 0.764	0.083 / 0.725	0.072 / 0.910	0.125 / 0.802	0.080 / 0.722
	CRNL	<u>0.119 / 0.852</u>	<u>0.046 / 0.915</u>	<u>0.047 / 0.962</u>	<u>0.078 / 0.922</u>	<u>0.058 / 0.853</u>
	Ours	**0.116 / 0.860**	**0.044 / 0.922**	**0.045 / 0.964**	**0.076 / 0.927**	**0.057 / 0.859**
15%	SVR	0.295 / 0.094	0.144 / 0.172	0.211 / 0.220	0.278 / 0.018	0.152 / 0.005
	KNR	0.113 / 0.868	0.051 / 0.896	0.049 / 0.958	0.084 / 0.910	0.058 / 0.857
	DT	0.112 / 0.868	0.054 / 0.884	0.052 / 0.952	0.091 / 0.896	0.070 / 0.791
	RF	0.123 / 0.843	0.055 / 0.880	0.055 / 0.948	0.101 / 0.870	0.075 / 0.760
	FSA-HTF	0.138 / 0.800	0.057 / 0.871	0.057 / 0.944	0.094 / 0.887	0.065 / 0.816
	HLRTF	0.148 / 0.772	0.083 / 0.727	0.073 / 0.908	0.123 / 0.807	0.080 / 0.727
	CRNL	<u>0.104 / 0.888</u>	<u>0.040 / 0.935</u>	<u>0.041 / 0.971</u>	<u>0.068 / 0.941</u>	<u>0.053 / 0.880</u>
	Ours	**0.101 / 0.893**	**0.039 / 0.940**	**0.040 / 0.972**	**0.067 / 0.944**	**0.051 / 0.887**
20%	SVR	0.293 / 0.103	0.143 / 0.176	0.203 / 0.284	0.277 / 0.026	0.151 / 0.012
	KNR	0.101 / 0.895	0.046 / 0.915	0.044 / 0.965	0.076 / 0.927	0.053 / 0.881
	DT	0.101 / 0.894	0.050 / 0.901	0.047 / 0.962	0.081 / 0.917	0.064 / 0.825
	RF	0.120 / 0.851	0.053 / 0.887	0.053 / 0.950	0.100 / 0.872	0.074 / 0.767
	FSA-HTF	0.134 / 0.814	0.056 / 0.875	0.056 / 0.945	0.092 / 0.893	0.064 / 0.825
	HLRTF	0.149 / 0.769	0.082 / 0.729	0.071 / 0.911	0.125 / 0.803	0.080 / 0.727
	CRNL	<u>0.093 / 0.909</u>	<u>0.037 / 0.946</u>	<u>0.038 / 0.975</u>	<u>0.061 / 0.953</u>	<u>0.048 / 0.899</u>
	Ours	**0.092 / 0.912**	**0.035 / 0.951**	**0.037 / 0.977**	**0.060 / 0.955**	**0.047 / 0.905**

As shown in Table 7, D-GLIR achieves the lowest NRMSE and highest R^2 on all five point clouds at every sampling rate. It consistently outperforms CRNL under the particularly sparse sampling rate of 10%, and maintains this advantage as the sampling rate increases to 15% and 20%. Although the numerical improvements are moderate, their consistency across different objects and sampling conditions demonstrates the robustness of D-GLIR for point-cloud attribute recovery.

The qualitative comparisons in Figs. 7 and 8 further show that D-GLIR recovers color distributions that more closely resemble the ground truth, with fewer visible artifacts. Together, these results demonstrate the effectiveness of the proposed global-local representation in recovering RGB attributes from sparse observations on irregular 3D surfaces.

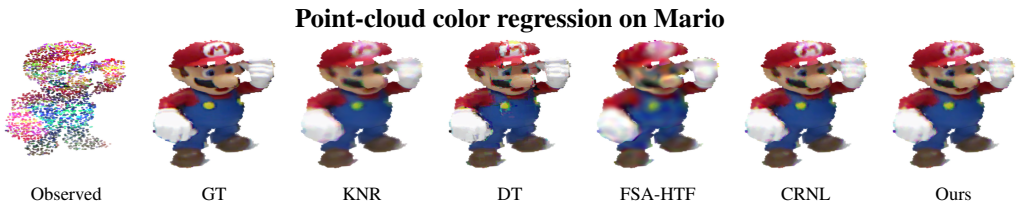


Figure 7: Qualitative comparison of point-cloud color regression on *Mario*. From left to right: Observed, GT, KNR, DT, FSA-HTF, CRNL, and Ours.

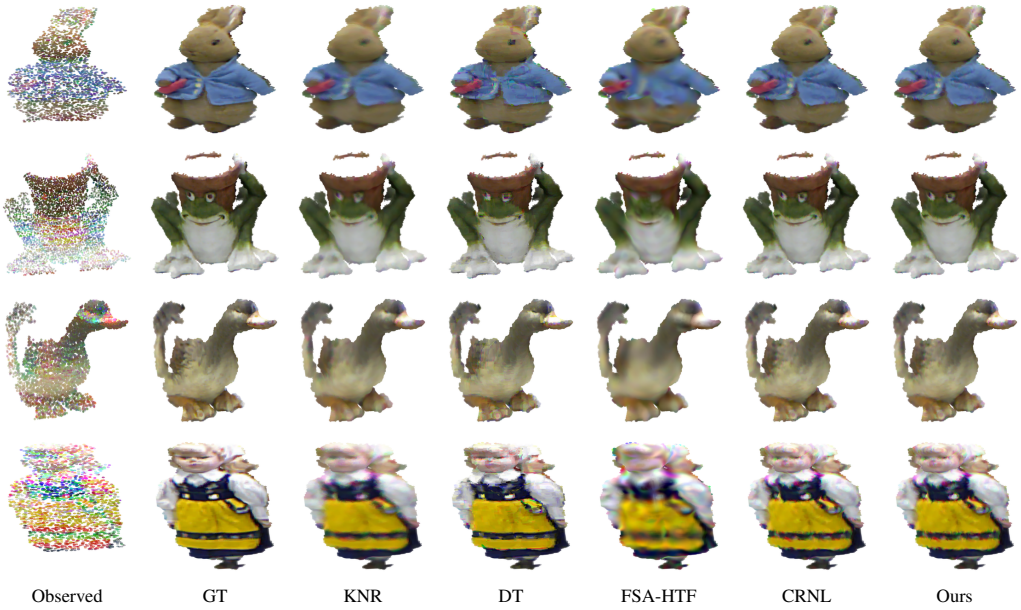


Figure 8: Qualitative comparison of point-cloud color regression on *PeterRabbit*, *Frog*, *Duck*, and *Doll*. From left to right: Observed, GT, KNR, DT, FSA-HTF, CRNL, and Ours.

B.4 DENOISING ON HETEROGENEOUS DATA

We report denoising results under the noise settings specified in Appendix A. Table 4 summarizes representative results at the strongest noise level, while Table 8 presents the detailed results at all evaluated noise levels.

Table 8: Quantitative comparison of denoising performance on images, point clouds, and weather fields under different noise levels. The best and second-best results under each setting are highlighted in bold and underlined, respectively.

Method	$\sigma = 0.10$		$\sigma = 0.15$		$\sigma = 0.20$		$\sigma = 0.25$	
	PSNR↑	NRMSE↓	PSNR↑	NRMSE↓	PSNR↑	NRMSE↓	PSNR↑	NRMSE↓
Image denoising								
Observed	20.1888	0.1669	16.9144	0.2434	14.7312	0.3130	13.1238	0.3767
LRTDTV	22.5752	0.1311	22.4521	0.1323	21.8740	0.1401	20.8169	0.1568
HLRTF	20.8138	0.1669	19.9654	0.1782	19.0678	0.1946	18.1657	0.2140
CRNL	27.0579	0.0762	24.9842	0.0965	23.0706	0.1206	21.2728	0.1487
RCTV	27.3757	0.0754	25.9767	0.0881	<u>24.7900</u>	<u>0.1004</u>	<u>23.8374</u>	<u>0.1117</u>
LTDL	<u>28.1440</u>	<u>0.0680</u>	<u>26.1279</u>	<u>0.0859</u>	24.6932	0.1013	23.4070	0.1175
Ours	**29.0235**	**0.0615**	**26.5762**	**0.0810**	**25.0943**	**0.0958**	**24.1136**	**0.1072**
Point-cloud denoising								
Method	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑
Observed	0.100970	0.698749	0.147853	0.347596	0.191155	-0.094112	0.230291	-0.584794
KNN-Median	0.048460	0.938978	0.060134	0.902738	0.073032	0.852968	0.086457	0.790142
KNN-Mean	0.046347	0.946237	0.055302	0.921450	0.065650	0.888004	0.076767	0.846685
Gaussian-KNN	<u>0.044602</u>	<u>0.949913</u>	**0.054013**	<u>0.924575</u>	<u>0.064743</u>	<u>0.890446</u>	0.076162	0.848379
Bilateral-KNN	0.068450	0.863442	0.120098	0.570044	0.169136	0.142888	0.212831	-0.354299
HLRTF	0.100571	0.769239	0.100821	0.771137	0.102641	0.762892	0.106195	0.747807
FSA-HTF	0.072908	0.876653	0.075013	0.868836	0.078793	0.854236	0.083599	0.835998
CRNL	0.057941	0.916587	0.062352	0.903016	0.068320	0.883558	<u>0.075647</u>	<u>0.857728</u>
Ours	**0.043736**	**0.950687**	<u>0.054117</u>	**0.925571**	**0.062766**	**0.900149**	**0.070767**	**0.875267**
Weather denoising								
Method	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑	NRMSE↓	R^2 ↑
Observed	0.005437	0.990069	<u>0.008200</u>	<u>0.977454</u>	0.010902	0.960133	0.013603	0.937925
KNN-Mean	0.011857	0.938009	0.011945	0.937437	0.012059	0.936722	0.012149	0.936236
Gaussian	0.011192	0.944580	0.011288	0.943989	0.011416	0.943209	0.011507	0.942738
IDW	<u>0.005437</u>	<u>0.990081</u>	0.008218	0.977429	0.010892	0.960156	0.013571	0.938115
Local-Linear	0.010702	0.949327	0.010811	0.948695	0.010948	0.947847	0.011085	0.947085
FSA	0.012128	0.936980	0.012202	0.936376	0.012305	0.935669	0.012433	0.934754
CRNL	0.007785	0.977074	0.008213	0.974234	<u>0.008990</u>	<u>0.969005</u>	<u>0.010471</u>	<u>0.958689</u>
Ours	**0.004715**	**0.991920**	**0.006159**	**0.986457**	**0.008044**	**0.977155**	**0.010262**	**0.963303**

As shown in Table 8, D-GLIR achieves the highest PSNR and lowest NRMSE at all noise levels. Compared with the second-best method, it improves PSNR by 0.88 dB, 0.45 dB, 0.30 dB, and 0.28 dB as σ increases from 0.10 to 0.25, respectively. As illustrated in Fig. 9, D-GLIR effectively suppresses Gaussian noise while preserving image structures and visual details.

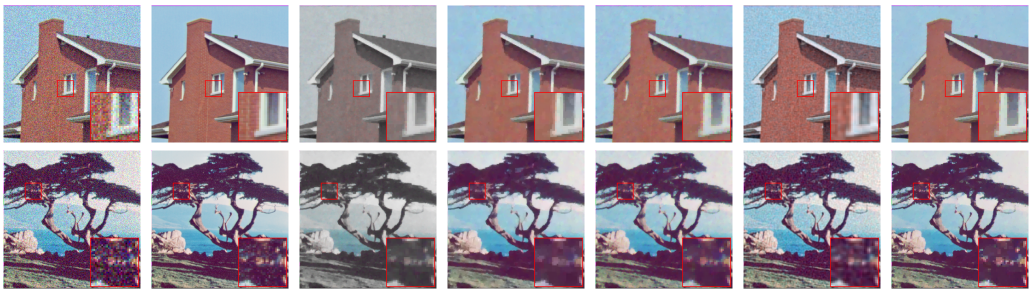


Figure 9: Qualitative comparison of image denoising at $\sigma = 0.10$. The top and bottom rows show results on *House-II* and *Tree*, respectively. From left to right: Observed, GT, LRTDTV, RCTV, LTDL, CRNL, and Ours.

For point-cloud denoising, Table 8 shows that D-GLIR obtains the highest R^2 at every noise level and the lowest NRMSE in three of the four settings. At $\sigma = 0.15$, its NRMSE is only slightly higher than that of Gaussian-KNN, while its R^2 remains the highest. The consistent performance across noise levels demonstrates the effectiveness of D-GLIR in recovering RGB attributes on irregular 3D surfaces.

For weather denoising, Table 8 shows that D-GLIR consistently achieves the lowest NRMSE and highest R^2 under all evaluated noise levels. These results indicate that the proposed representation can recover spatially varying fields from noisy and irregular observations. Overall, the experiments demonstrate that D-GLIR provides robust denoising performance across regular image grids, irregular point sets, and spatial weather fields within a unified framework.

B.5 MSI Denoising Results and Limitations

We further examine the limitations of D-GLIR on MSI denoising. Compared with conventional images, MSIs contain substantially more spectral channels and exhibit strong spatial–spectral dependencies. As shown in Table 9, D-GLIR achieves competitive performance under Gaussian noise with $\sigma = 0.25$, but remains inferior to the specialized MSI denoising methods LTDL and SDeCNN.

Table 9: Quantitative comparison on MSI denoising under Gaussian noise with $\sigma = 0.25$.

Method	PSNR ↑	SSIM ↑	NRMSE ↓
HSID-CNN	29.90	0.762	0.209
LRTDTV	30.27	0.654	0.166
HLRTF	31.49	0.760	0.159
LTDL	**34.83**	**0.925**	**0.098**
RCTV	30.58	0.721	0.159
SDeCNN	<u>33.92</u>	<u>0.864</u>	<u>0.114</u>
WNLRATV	28.63	0.550	0.215
CRNL	33.13	0.820	0.122
Ours	33.21	0.845	0.118

The performance gap in Table 9 indicates that MSI denoising benefits substantially from dedicated spatial–spectral modeling. For example, LTDL explicitly exploits the low-rank and correlated structures of spectral data, while supervised methods such as SDeCNN learn data-driven denoising priors from external training samples. In contrast, D-GLIR is designed as a general self-supervised continuous representation and incorporates neither MSI-specific spectral priors nor external paired training data. Nevertheless, it achieves the third-best results across all three metrics and remains competitive with representative optimization-based and continuous reconstruction methods.

The results in Table 9 also suggest a clear direction for future development. Incorporating adaptive spectral bases, stronger cross-channel modeling, or data-driven MSI priors into the global-local representation may further improve denoising performance while retaining the generality and flexibility of D-GLIR.

B.6 Effect of Branch Parameterization

We compare the full model with SIREN-global, SIREN-local, and SIREN-both on RGB image inpainting to examine the effects of the global and local parameterizations. Table 10 summarizes results over the four sampling rates. Comparing the full model with SIREN-global retains the interpolation-based local bases while changing the global branch. Comparing the full model with SIREN-local, or SIREN-global with SIREN-both, retains the respective global parameterization while changing the local branch.

Table 10: Branch-parameterization ablation on RGB image inpainting. Param denotes the corresponding branch parameterization used in the full model. Values are equally weighted averages of the means at sampling rates of 5%, 10%, 15%, and 20%.

Variant	Global	Local	PSNR↑	SSIM↑	NRMSE↓
D-GLIR	Param	Param	26.5300	0.7883	0.0848
SIREN-global	SIREN	Param	25.8500	0.7733	0.0910
SIREN-local	Param	SIREN	19.4241	0.5300	0.2002
SIREN-both	SIREN	SIREN	19.8204	0.5589	0.1879

As shown in Table 10, the full model attains the best average reconstruction quality. With the interpolation-based local bases retained, replacing the global branch with SIREN yields an average PSNR of 25.85 dB, compared with 26.53 dB for the full model. Replacing the local bases with SIREN yields lower reconstruction quality under both global parameterizations. These results favor the interpolation-based local parameterization in the tested setting. They compare the resulting implementations rather than isolate whether the differences arise from representation, initialization, or optimization.

B.7 Computational Cost Analysis

The computational-cost results in Table 11 follow the evaluation setting specified in Appendix A. As shown in Table 11, D-GLIR achieves the best reconstruction performance while maintaining the second-shortest running time. Compared with CRNL, it reduces the running time and parameter count by 63.8% and 69.4%, respectively, while improving PSNR by 1.72 dB. It also outperforms TCTV in reconstruction quality while requiring less than one-sixth of its running time. These results demonstrate that D-GLIR provides a favorable trade-off between reconstruction accuracy and computational efficiency. Note that methods without trainable parameters may still involve costly iterative optimization and therefore do not necessarily incur lower computational costs.

Table 11: Computational efficiency and reconstruction performance on `house.mat` ($512\times512\times3$) at SR $= 0.10$.

Method	Time (s) ↓	PSNR ↑	SSIM ↑	NRMSE ↓	Params
TRLRF	**7.60**	20.9450	0.4451	0.1277	20.56K
FTNN	23.10	16.2082	0.3708	0.2177	0
HLRTF	27.31	24.2728	0.6129	0.0867	3.134M
t-ϵ-LogDet	35.59	21.7773	0.4018	0.1116	0
SIREN	38.27	25.2955	0.7834	0.0768	264.45K
TCTV	119.59	<u>27.3248</u>	0.8253	<u>0.0604</u>	0
CRNL	51.20	27.0618	<u>0.8520</u>	0.0624	6.905M
Ours	<u>18.51</u>	**28.7819**	**0.8849**	**0.0522**	2.114M

C THEORETICAL ANALYSIS OF D-GLIR

This appendix provides a theoretical interpretation of D-GLIR from two perspectives: the global-local decomposition of continuous patch functions, and the relation between coefficient similarity and nonlocal self-similarity.

C.1 GLOBAL-LOCAL DECOMPOSITION OF PATCH FUNCTIONS

Let $\mathcal{S}$ denote the domain of spatial anchors and let $\mathcal{R}$ denote the relative-coordinate domain. Given an underlying signal $f^\star : \Omega \to \mathbb{R}^C$, we define its associated patch function as

$$g^\star(\mathbf{s}, \mathbf{r}) = f^\star(\mathbf{s} + \mathbf{r}), \qquad (\mathbf{s}, \mathbf{r}) \in \mathcal{S} \times \mathcal{R}. \tag{19}$$

D-GLIR approximates this function using the finite global-local expansion

$$\widehat{g}_K(\mathbf{s}, \mathbf{r}) = \sum_{k=1}^{K} a_k(\mathbf{s}) \mathbf{b}_k(\mathbf{r}). \tag{20}$$

Proposition 1. Assume that $g^\star \in L^2(\mathcal{S} \times \mathcal{R}; \mathbb{R}^C)$. Then there exist singular values $\{\sigma_k\}_{k=1}^{\infty}$, an orthonormal set of spatial functions $\{u_k\}_{k=1}^{\infty} \subset L^2(\mathcal{S})$, and an orthonormal set of local basis functions $\{\mathbf{v}_k\}_{k=1}^{\infty} \subset L^2(\mathcal{R}; \mathbb{R}^C)$ such that

$$g^\star(\mathbf{s}, \mathbf{r}) = \sum_{k=1}^{\infty} \sigma_k u_k(\mathbf{s}) \mathbf{v}_k(\mathbf{r}). \tag{21}$$

Furthermore, the truncated expansion $g_K^\star(\mathbf{s}, \mathbf{r}) = \sum_{k=1}^{K} \sigma_k u_k(\mathbf{s}) \mathbf{v}_k(\mathbf{r})$ is the optimal rank-K separable approximation, satisfying:

$$\|g^\star - g_K^\star\|^2_{L^2(\mathcal{S} \times \mathcal{R})} = \sum_{k > K} \sigma_k^2. \tag{22}$$

Proof. We interpret $g^\star(\mathbf{s}, \mathbf{r})$ as the kernel of a linear integral operator $T : L^2(\mathcal{S}) \to L^2(\mathcal{R}; \mathbb{R}^C)$ defined by $(T\phi)(\mathbf{r}) = \int_{\mathcal{S}} g^\star(\mathbf{s}, \mathbf{r}) \phi(\mathbf{s})\, d\mathbf{s}$. Since $g^\star \in L^2(\mathcal{S} \times \mathcal{R}; \mathbb{R}^C)$, the operator T is a compact Hilbert–Schmidt operator. The expansion in Eq. equation 21 is a direct application of the Singular Value Decomposition (SVD) theorem for compact operators on Hilbert spaces. The optimality of the rank-K truncation and the associated error bound are guaranteed by the Eckart–Young–Mirsky theorem, which states that the sum of squared singular values of the discarded components yields the minimum L^2 reconstruction error. □

Interpretation. Proposition 1 provides an approximation-theoretic foundation for D-GLIR, showing that the global-local decomposition is mathematically equivalent to a low-rank separable approximation of the continuous patch function. This decomposition is highly effective when the signal exhibits structural redundancy; in such cases, the singular values σ_k decay rapidly, allowing a small number of components (K) to capture the vast majority of the patch function's energy. It is important to note that the learned global and local branches in D-GLIR do not need to explicitly match the theoretical singular functions $\{u_k, \mathbf{v}_k\}$; instead, they are optimized to construct a functionally equivalent subspace that best reconstructs the signal from degraded observations.

C.2 IMPLICIT NONLOCAL SELF-SIMILARITY

D-GLIR represents all local patch functions using a common set of local bases and location-dependent coefficient vectors. This formulation establishes a direct relationship between the similarity of two reconstructed patches and the similarity of their coefficient vectors.

For two anchors $\mathbf{s}_i$ and $\mathbf{s}_j$, the reconstructed patch difference is

$$\widehat{g}_K(\mathbf{s}_i, \mathbf{r}) - \widehat{g}_K(\mathbf{s}_j, \mathbf{r}) = \sum_{k=1}^{K} \left[a_k(\mathbf{s}_i) - a_k(\mathbf{s}_j) \right] \mathbf{b}_k(\mathbf{r}). \tag{23}$$

Defining the Gram matrix of the local bases as

$$G_{k\ell} = \langle \mathbf{b}_k, \mathbf{b}_\ell \rangle_{L^2(\mathcal{R})}, \tag{24}$$

and the coefficient vector as

$$\mathbf{a}(\mathbf{s}) = [a_1(\mathbf{s}), \ldots, a_K(\mathbf{s})]^{\mathrm{T}}. \tag{25}$$

The squared distance between the two reconstructed patches is then

$$\begin{aligned} &\|\widehat{g}_K(\mathbf{s}_i, \cdot) - \widehat{g}_K(\mathbf{s}_j, \cdot)\|^2_{L^2(\mathcal{R})} \\ &\quad = [\mathbf{a}(\mathbf{s}_i) - \mathbf{a}(\mathbf{s}_j)]^{\mathrm{T}} \mathbf{G} [\mathbf{a}(\mathbf{s}_i) - \mathbf{a}(\mathbf{s}_j)]. \end{aligned} \tag{26}$$

If the local bases are orthonormal, then $\mathbf{G} = \mathbf{I}$ and Eq. equation 26 simplifies to the Euclidean distance in the coefficient space:

$$\|\widehat{g}_K(\mathbf{s}_i, \cdot) - \widehat{g}_K(\mathbf{s}_j, \cdot)\|_{L^2(\mathcal{R})} = \|\mathbf{a}(\mathbf{s}_i) - \mathbf{a}(\mathbf{s}_j)\|_2. \tag{27}$$

More generally, if $\mathbf{G}$ is positive definite with eigenvalues $0 < \lambda_{\min}(\mathbf{G}) \leq \lambda_{\max}(\mathbf{G})$, then:

$$\begin{aligned} \lambda_{\min}(\mathbf{G}) \|\Delta \mathbf{a}_{ij}\|_2^2 &\leq \|\widehat{g}_K(\mathbf{s}_i, \cdot) - \widehat{g}_K(\mathbf{s}_j, \cdot)\|^2_{L^2(\mathcal{R})} \\ &\leq \lambda_{\max}(\mathbf{G}) \|\Delta \mathbf{a}_{ij}\|_2^2, \end{aligned} \tag{28}$$

where

$$\Delta \mathbf{a}_{ij} = \mathbf{a}(\mathbf{s}_i) - \mathbf{a}(\mathbf{s}_j).$$

Interpretation. Equations equation 26–equation 28 demonstrate that patch similarity is characterized by a metric in the coefficient space defined by $\mathbf{G}$. Specifically, local structures with similar coefficient vectors necessarily produce similar reconstructed patches. This provides a formal mechanism for nonlocal self-similarity: spatially distant regions share structural information through the common local basis set, effectively coupling their representations without explicit patch matching or grouping.