ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Self-supervised migration image restoration constrained by spatially variant point-spread functions

Jingwen Zeng[1], Hao Wang[1]*, Yingzhe Bai[1], and Lu Xu[1,2]

[1]SINOPEC Geophysical Research Institute Co., Ltd., Nanjing, Jiangsu, China

[2]School of Geophysics and Information Technology, China University of Geosciences (Beijing), Beijing, China

Abstract

Least-squares migration can improve image resolution and compensate for uneven illumination, but data-domain implementations require repeated modeling and migration. Image-domain methods based on point-spread functions (PSFs) reduce this cost by approximating the local Hessian response, yet their performance depends on the representativeness of local PSFs in structurally complex areas. We propose SPIRE-Net, a self-supervised migration image restoration framework constrained by spatially variant PSFs. The network predicts a restored image patch, which is then filtered by the PSF assigned to the corresponding spatial location. Training minimizes the mismatch between this PSF-filtered prediction and the input image patch without requiring supervised target images. Structure-aware adaptive PSF-paired sampling and a structure-enhanced consistency loss strengthen the constraint in complex structural regions. The Marmousi experiment demonstrates the ability of SPIRE-Net to solve the prescribed PSF-based image-domain inverse problem through self-supervised learning. Field-data comparisons show improved local reflector continuity and event separation relative to image-domain LSM in the examined regions, while preserving the main reflection geometry. These results suggest that spatially variant PSFs can serve as physical constraints for self-supervised migration image restoration.

Keywords: Spatially variant point-spread function; Migration image restoration; Self-supervised learning; Hessian compensation; Seismic imaging

***Corresponding author:**
Hao Wang
(wanghao.swty@sinopec.com)

Citation: Zeng J, Wang H, Bai Y, Xu L. Self-supervised migration image restoration constrained by spatially variant point-spread functions. *J Seismic Explor.* doi: 10.36922/JSE026260115

Received: June 24, 2026

Revised: September 18, 2026

Accepted: September 22, 2026

Published online: October 29, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Seismic imaging uses surface-recorded wavefields to construct subsurface images that reveal geological interfaces and structural geometries. It is a fundamental step in resource exploration and geological interpretation. Reverse time migration (RTM) maps seismic records into the subsurface image domain by back-propagating the observed wavefield and applying an imaging condition, and has been widely used for imaging complex media.[1,2] However, conventional migration is essentially an adjoint operation rather than a true inverse solution. The resulting migrated image is therefore affected by limited aperture, uneven illumination, finite bandwidth, and the non-orthogonality of the imaging operator. These factors may lead to amplitude distortion, limited resolution,

and residual local artifacts.

Least-squares migration (LSM) formulates seismic imaging as a data-fitting inverse problem, in which the subsurface image is updated by minimizing the mismatch between predicted and observed data.[3,4] Compared with conventional migration, LSM can more systematically compensate for uneven illumination and band-limited acquisition, thereby improving image resolution and reflector focusing.[5] Least-squares reverse time migration (LSRTM), as a representative data-domain implementation of least-squares imaging, usually uses Born modeling as the linearized forward operator and RTM as its adjoint. It can improve illumination compensation, amplitude recovery, and spatial resolution. Nevertheless, each gradient update in LSRTM usually requires forward modeling and adjoint migration over many shots, resulting in substantial computational and storage costs. This limitation restricts its application to large-scale field data.[6-8]

Image-domain least-squares methods reduce this cost by transferring part of the inversion compensation from the data domain to the image domain. Previous studies have shown that a target-oriented Hessian approximation can describe the degradation relationship between a conventional migrated image and a high-resolution image-domain representation.[9,10] Under the local stationarity assumption, the local response of the Hessian can be approximated by a point-spread function (PSF). Accordingly, compensation for migration-induced blurring can be formulated as an image-domain restoration problem constrained by local PSFs. Non-iterative LSRTM and PSF-based image-domain deconvolution avoid repeated data-domain modeling by using sampled or approximated Hessian responses in the image domain. These methods provide an efficient way to sharpen migrated events and compensate for illumination-related blurring.[11,12] However, when PSFs vary spatially, local inverse problems still need to be defined, regularization parameters must be selected, and compensation is often performed region by region. The restoration quality also depends strongly on whether the selected PSF is representative of the local migrated-image patch.

Existing PSF-based image-domain compensation methods can be broadly divided into explicit deconvolution and data-driven correction. Local PSF modeling, spatially variant deconvolution, and efficient PSF deconvolution have been used to enhance reflector focusing and reduce illumination imbalance.[12] Deep neural networks have also been introduced to deblur migrated images, indicating the potential of learning-based methods for efficient migration image restoration.[13] Related seismic learning studies have used global and multiscale feature fusion for low-frequency extrapolation and hybrid bilateral attention for seismic image super-resolution, further demonstrating the ability of deep networks to recover information across scales and frequency bands.[14,15] For example, Zhou *et al.*[16] combined cost-effective PSF construction with a data-driven imaging correction network, using PSFs as an efficient Hessian approximation for image-domain high-resolution imaging. These studies indicate that PSFs can provide physically meaningful image-domain priors for learning-based migration image restoration. Nevertheless, many existing workflows still rely on target images, pseudo-labels, or auxiliary paired samples for supervised training. For field data, such supervised targets are often difficult to obtain. In addition, the physical consistency between the network prediction and the corresponding local PSF response in the target area is not always explicitly enforced.

Physics-constrained self-supervised learning provides a natural way to reduce the dependence on labeled targets. Instead of learning only from paired input–target samples, a known physical operator can be embedded into the training process to test whether the network prediction can reproduce the observation. Physics-informed neural networks introduce differential-equation constraints into network training and provide a general framework for physics-constrained learning.[17,18] In seismic inversion, physics-guided networks incorporate forward operators, wave-equation constraints, or data residuals into the training objective, so that network predictions are constrained by observed physical relationships rather than by data-driven mappings alone.[19,20] Recent bidirectional physics-constrained full-waveform inversion further combines neural representations, residual learning, and wave-equation gradients to reduce dependence on manually constructed training data.[21] Semi-supervised initial model building and probabilistic physics-guided inversion further extend neural-network training under limited labels.[22-24] For migration image restoration, local PSFs provide a direct image-domain degradation operator. They characterize how a locally restored image is blurred into the corresponding migrated image. This relationship can be used to construct a self-supervised consistency constraint between the PSF-filtered prediction and the input migrated-image patch, with the network output serving as the latent restored image.

Motivated by the above limitations, we develop SPIRE-Net, a self-supervised migration image restoration framework constrained by spatially variant PSFs. Its central contribution is to reformulate local PSF-based image-domain compensation as a self-supervised learning problem: the network predicts an image-domain model, the paired local PSF maps that prediction back to the migrated-

image domain, and the reconstruction mismatch provides the training signal without reflectivity labels. This differs from existing PSF-based image-domain restoration, which typically relies on iterative deconvolution at inference, and from supervised restoration approaches, which typically require paired high-quality target images. To make this constraint more effective in complex structural areas, we introduce a structure-aware PSF-paired sampling strategy that increases the coverage of high-gradient image features. We also incorporate a structure-enhanced migrated-image-domain consistency loss to emphasize local boundaries and reflector variations during optimization. The numerical example evaluates operator-consistent recovery against a known reference and an image-domain LSM baseline, whereas the field-data example evaluates the practical restoration of an independently obtained RTM image.

2. Methods

2.1. Least-squares migration imaging

Least-squares migration is commonly formulated under a given background velocity model and the Born approximation. Under this approximation, the scattered data can be related to the reflectivity perturbation through a linearized modeling equation:[25]

$$d = Lm + e, \tag{1}$$

where d denotes the scattered data generated by the reflectivity perturbation, m is the reflectivity model to be estimated, L is the Born modeling operator, and e represents noise and linearization errors. The corresponding data-domain least-squares objective can be written as:

$$J(m) = \frac{1}{2}\left\|Lm - d_{obs}\right\|_2^2 + \lambda R(m), \tag{2}$$

where d_{obs} denotes the observed scattered data, $R(m)$ is a regularization term, and λ is the regularization weight. This formulation updates the reflectivity model by minimizing the mismatch between predicted and observed data. Compared with conventional migration, LSM can partially compensate for uneven illumination, limited bandwidth, and the non-orthogonality of the imaging operator. However, data-domain LSM requires repeated Born modeling and adjoint migration, which leads to high computational and storage costs, especially for large-scale field data.[26,27]

To derive the image-domain least-squares formulation, the regularization and noise terms are first omitted. Taking the gradient of **Equation 2** with respect to m and setting it to zero gives:

$$L^{\mathrm{T}}\left(Lm - d_{obs}\right) = 0, \tag{3}$$

or equivalently:

$$L^{\mathrm{T}}Lm = L^{\mathrm{T}}d_{obs}. \tag{4}$$

Let:

$$H = L^{\mathrm{T}}L, \qquad y = L^{\mathrm{T}}d_{obs}, \tag{5}$$

where H denotes the Hessian operator and y is the migrated image obtained by conventional adjoint migration. The least-squares problem can then be rewritten, in the sense of the normal equation, as an image-domain inverse problem:[28]

$$Hm = y. \tag{6}$$

In the presence of noise, modeling inaccuracies, operator approximation errors, and Hessian estimation errors, this relationship is typically regularized in the image domain as:

$$m^* = \arg\min_m \frac{1}{2}\left\|Hm - y\right\|_2^2 + \mu\Phi(m), \tag{7}$$

where $\Phi(m)$ is an image-domain regularization term and μ is its weighting parameter. **Equations 6** and **7** indicate that a migrated image can be interpreted as the response of the reflectivity model after passing through the imaging system. Compared with data-domain LSM, this formulation treats the migrated image as the observation and performs compensation directly in the image domain, thereby reducing the problem dimension and avoiding repeated wave-equation modeling during restoration.

The representation of the Hessian H plays a key role in image-domain least-squares imaging. If only its diagonal terms are used, the inversion mainly performs illumination compensation. When off-diagonal terms are included, interactions between neighboring image points can be captured, which helps mitigate local blurring and improves spatial resolution.[29] Physically, each column of the Hessian corresponds to the response of a point scatterer after modeling and migration. This response is commonly referred to as the PSF. Therefore, the full Hessian describes a global blurring operator, whereas a PSF represents its localized behavior. The Hessian formulation is suitable for global analysis, while PSFs are more convenient for local modeling and numerical implementation.[30,31]

In continuous form, the relationship between the migrated image and the reflectivity model can be written as:

$$y(r) = \int_\Omega h(r,\xi)m(\xi)\mathrm{d}(\xi) + \eta(r), \tag{8}$$

where $h(r,\xi)$ denotes the contribution at image point r from a unit reflector located at ξ, and $\eta(r)$ represents noise and modeling errors. The kernel $h(r,\xi)$ incorporates the effects of illumination, acquisition aperture, finite bandwidth, velocity-model errors, and operator non-orthogonality. Because the subsurface medium and acquisition geometry are generally spatially varying, this

kernel cannot, in general, be represented using a globally shift-invariant convolution.

When velocity, illumination, and structural variations are mild within a local region, a local stationarity approximation can be adopted. Let P_k denote the operator that extracts the k-th patch from the global migrated image y, and let $x_k = P_k y$. **Equation 8** can then be approximated in the k-th region using a locally shift-invariant convolution [32,33]:

$$x_k = h_k \otimes m_k + \eta_k, \tag{9}$$

where x_k is the local migrated-image patch, m_k is the corresponding local image-domain model, h_k is the local PSF, and $\otimes$ denotes convolution. Under this approximation, **Equation 7** can be interpreted as a set of local deconvolution problems constrained by spatially variant PSFs:

$$m_k^* = \arg\min_{m_k} \frac{1}{2}\|h_k \otimes m_k - x_k\|_2^2 + \alpha\Psi(m_k), \tag{10}$$

where $\Psi(m_k)$ is a local regularization term and α is its weighting parameter. **Equations 9 and 10** provide the physical basis of this study. Specifically, the input patch x_k is treated as a degraded response of the local image m_k after filtering using the position-dependent PSF h_k. Therefore, migration image restoration can be regarded as compensation for the local imaging response.

This relationship is illustrated using a horizontally layered model. Figure 1 shows the velocity model, and Figure 2 presents the corresponding PSF-affected image. The reflected events exhibit vertical broadening and sidelobe energy, which reduce interface resolution and cause local energy leakage. Figure 3 shows the corresponding image-domain least-squares result for the same model. Compared with the input image, the compensated result shows improved reflector focusing and weaker sidelobe energy. This example illustrates the degradation–compensation mechanism assumed in the following PSF-constrained restoration framework.

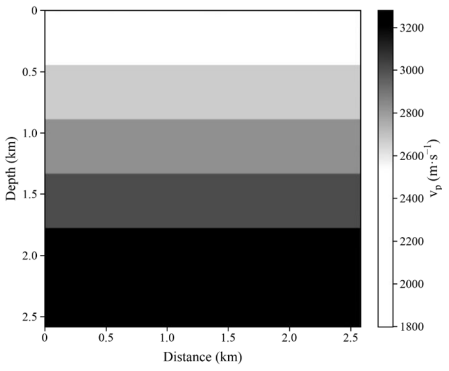


Figure 1. Horizontally layered velocity model

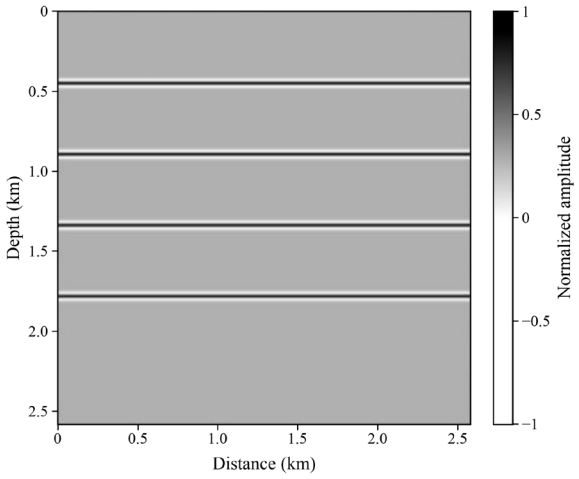


Figure 2. Point-spread function–degraded image for the horizontally layered model

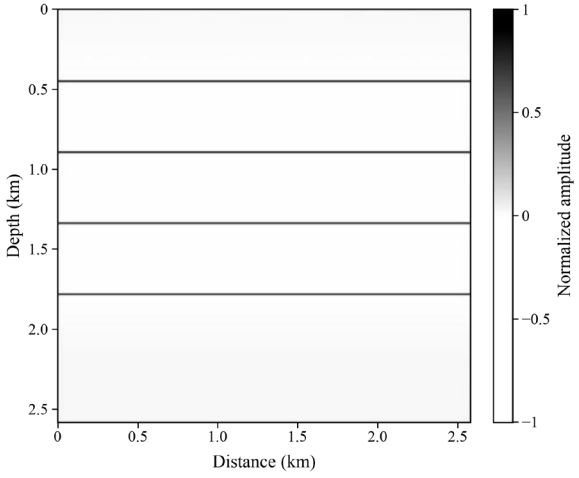


Figure 3. Image-domain least-squares migration result for the horizontally layered model

2.2. SPIRE-Net framework

Based on the local PSF degradation model, we developed SPIRE-Net, a self-supervised migration image restoration framework constrained by spatially variant PSFs. The proposed method parameterizes local patch-wise image-domain compensation using a neural network. A single network learns a shared restoration mapping from local image patches, while the PSF associated with each spatial position imposes a position-dependent physical constraint on the network output.

Let y denote the global input image, and let $x_k = P_k y$ be the k-th local input patch extracted from y. The restored local image patch predicted by the network is written as:

$$\hat{m}_k = f_\theta(x_k), \tag{11}$$

where $f_\theta(\cdot)$ denotes the restoration network with parameters θ, and $\hat{m}_k$ is the estimated local image-domain model. In this study, the term input image was used in a general sense. It may be obtained from conventional migration, RTM, or other seismic imaging workflows. In the synthetic experiment, the input image was generated by degrading the reference reflectivity with local PSFs. In the field-data experiment, the input image was an RTM image.

In field applications, high-quality target images or supervised reflectivity labels are usually unavailable, making it difficult to construct a reliable supervised training set. In this study, local PSFs associated with the input image were assumed to be available. The training constraint was constructed from the image-domain physical relationship between the restored local image patch and its local PSF response. Specifically, the corresponding local PSF was applied to the network output to synthesize its migrated-image-domain response:

$$\hat{x}_k = h_k \otimes \hat{m}_k = h_k \otimes f_\theta(x_k). \tag{12}$$

where $\hat{x}_k$ denotes the PSF-filtered prediction obtained from the restored local image patch. If $\hat{m}_k$ is a physically consistent restored image-domain patch, filtering it with the corresponding local PSF h_k should reproduce the input migrated-image patch x_k. Therefore, the network can be trained in a self-supervised manner by enforcing consistency between $\hat{x}_k$ and x_k.

Figure 4 illustrates the overall SPIRE-Net framework. The left block represents the input and PSF-paired sampling module, which extracts the input migrated-image patch x_k and assigns its corresponding local PSF h_k. The middle block is the image-domain model prediction branch, where the restoration network $f_\theta(\cdot)$ estimates the restored local image patch $\hat{m}_k$. The lower block is the PSF-consistency branch, in which local PSF filtering is used to generate the PSF-filtered prediction $\hat{x}_k$. The right block represents the self-supervised optimization module. The network parameters are updated by minimizing the migrated-image-domain consistency error between the PSF-filtered prediction and the input migrated-image patch.

2.3. Point-spread function–constrained self-supervised optimization

In SPIRE-Net, the training signal is derived from the physical consistency between the input migrated-image patch and the corresponding local PSF response. For the i-th sample, let x_i denote the input image patch, h_i denote the local PSF, and $\hat{m}_i = f_\theta(x_i)$ denote the restored local image patch predicted by the network. The corresponding PSF-filtered prediction is given by:

$$\hat{x}_i = h_i \otimes \hat{m}_i. \tag{13}$$

The basic physical consistency loss is defined as:

$$L_{mig} = \frac{1}{N}\sum_{i=1}^{N} \left\| \hat{x}_i - x_i \right\|_2^2, \tag{14}$$

where N is the number of training samples. This loss corresponds to the migrated-image-domain consistency term in the image-domain least-squares objective. It requires the restored local image patch to reproduce the input migrated-image patch after being filtered by the corresponding local PSF. If the prediction contains structures, amplitudes, or frequency components that are inconsistent with the current PSF, residuals will appear in the reconstructed image patch and guide the update of the network parameters through back-propagation.

A uniformly weighted amplitude reconstruction error provides limited constraints on faults, steeply dipping events, and thin layers with rapid local variations. Structure weights are constructed from the local gradient of the input migrated-image patch to increase the contribution of complex structural regions to the training objective. Let:

$$g_i(p) = \| \nabla x_i(p) \|_2, \tag{15}$$

where $g_i(p)$ denotes the local gradient magnitude of the i-th sample at pixel position p. The weight is defined as:

$$w_i(p) = 1 + \alpha \frac{g_i(p)}{\overline{g_i} + \epsilon}, \tag{16}$$

where $\overline{g_i}$ is the mean gradient magnitude within the sample, ϵ is a stabilizing constant, and α controls the weight increase in high-gradient regions. In implementation, the gradient-based weight is capped at $1+4\alpha$. Normalization by the mean gradient reduces sensitivity to overall patch amplitude, while the cap bounds the weight assigned to isolated large gradients.

The corresponding structure-weighted migrated-image-domain consistency loss is written as:

$$L_{\text{wmig}} = \frac{\sum_{i,p} w_i(p)\left(\hat{x}_i(p) - x_i(p)\right)^2}{\sum_{i,p} w_i(p)}. \tag{17}$$

This term assigns larger weights to positions with strong local variations, such as fault boundaries, thin-layer interfaces, and steep reflectors. To further constrain local boundaries and structural variations in the reconstructed image patches, a migrated-image-domain gradient-consistency loss is incorporated:

$$L_{\text{grad}} = \frac{1}{N}\sum_{i=1}^{N} \left(\left\| \nabla_z \hat{x}_i - \nabla_z x_i \right\|_1 + \left\| \nabla_x \hat{x}_i - \nabla_x x_i \right\|_1 \right) \tag{18}$$

where ∇_z and ∇_x denote the vertical and horizontal gradient operators, respectively. The final training objective is:

$$L = L_{\text{wmig}} + \lambda_g L_{\text{grad}}, \tag{19}$$

where λ_g is the weight of the gradient consistency term.

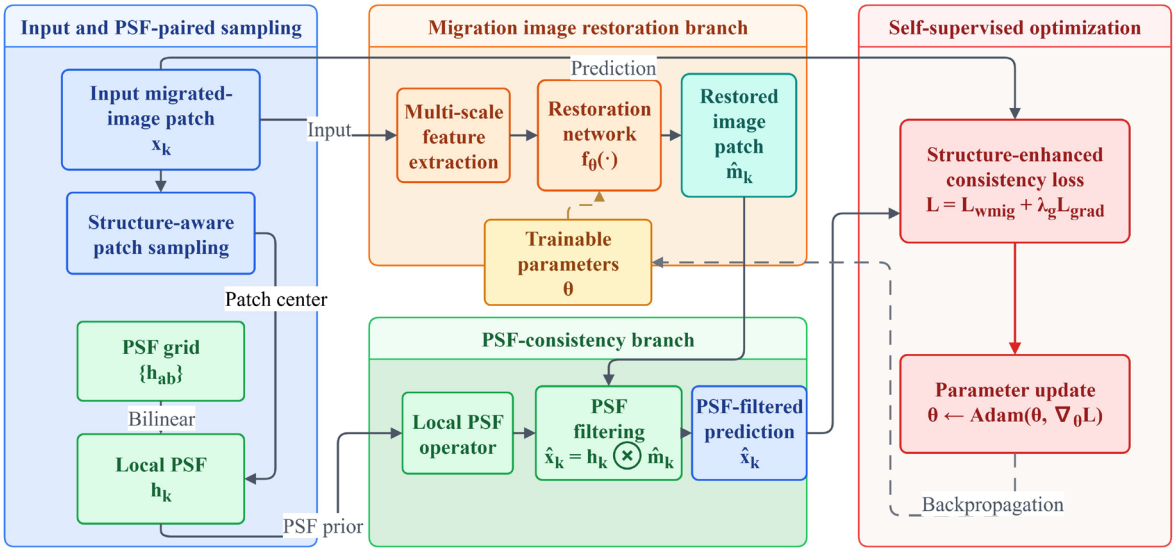


Figure 4. Overall SPIRE-Net framework
Abbreviation: PSF: Point-spread function.

This training strategy does not require explicit target labels. Reference reflectivity models in synthetic tests and image-domain LSM results were used only for post-training evaluation and were not involved in network training, early stopping, or model selection.

During training, all local samples shared the same network parameters. Therefore, the same network was required to explain input patches from different spatial positions together with their corresponding PSFs. The model learns restoration behavior that can be shared among local regions within the current section, while each local PSF defines the physical relationship that the corresponding sample must satisfy between the restored local image patch and the input image. This PSF-constrained self-supervised optimization follows the local image-domain degradation model and helps preserve the physical consistency and interpretability of the restored result.

2.4. Structure-aware adaptive point-spread function–paired sampling

Under spatially variant PSFs, each training sample should contain a local migrated-image patch and a local PSF associated with the same spatial position. Fixed-position samples are first extracted on a predefined PSF grid. The input image and the corresponding PSF grid are divided according to the same spatial grid, so that each image patch is paired with the local PSF at the corresponding position. The k-th fixed-grid sample is defined as:

$$s_k = \{x_k, h_k\}, \tag{20}$$

where x_k is the input migrated-image patch and h_k is the local PSF at the corresponding grid position. Fixed-grid samples cover the main spatial positions, but their sampling density is limited by the PSF grid spacing. In regions with faults, steeply dipping reflectors, and rapidly varying thin layers, key structures may lie between neighboring PSF grid points. If only fixed-grid samples are used, complex structural areas may be underrepresented in the training set, and the corresponding physical constraints may be relatively weak.

To improve sample coverage in complex structural areas, we used a structure-aware adaptive PSF-paired sampling strategy. The method first computes the local gradient magnitude of the input image:

$$g(z,x) = \sqrt{\left(\partial_z y(z,x)\right)^2 + \left(\partial_x y(z,x)\right)^2}, \tag{21}$$

where z and x denote the vertical and horizontal coordinates, respectively. Image gradients are evaluated by finite differences on the sample grid, using central differences in the interior and one-sided differences at the image boundaries. High-gradient positions often correspond to fault boundaries, steeply dipping events, thin-layer interfaces, or regions with rapid local amplitude variations. Positions with gradient magnitudes at or above a prescribed quantile threshold are selected as candidate centers for adaptive sampling. Small random perturbations are then added around these centers to generate additional sample centers q_j and the corresponding training patches x_{q_j}.

This strategy draws additional samples from structurally active regions while preventing the extra patches from being excessively concentrated around a few local extrema. Candidate centers are restricted to nonzero input samples whose gradient magnitudes are at or above the prescribed quantile. The candidates are randomly permuted using a fixed seed and perturbed within a bounded integer range. Patch origins are clipped to keep the complete patch inside the image, and duplicate origins are removed.

An adaptive sample center q_j generally does not exactly coincide with the original PSF grid points. Therefore, the local PSF matched to this position is obtained by bilinear interpolation using neighboring PSFs. Let h_{11}, h_{12}, h_{21}, and h_{22} denote the four local PSFs surrounding q_j, and let ω_{ab} be the corresponding bilinear interpolation weights. The interpolated local PSF for the adaptive sample is written as:

$$h_{q_j} = \sum_{a=1}^{2}\sum_{b=1}^{2} \omega_{ab} h_{ab}, \qquad \sum_{a=1}^{2}\sum_{b=1}^{2} \omega_{ab} = 1. \tag{22}$$

The final training set is the union of the fixed-grid samples and the adaptive samples:

$$D = D_{\text{grid}} \cup D_{\text{adaptive}}, \quad D_{\text{grid}} = \left\{(x_k, h_k)\right\}_{k=1}^{K},$$
$$D_{\text{adaptive}} = \left\{(x_{q_j}, h_{q_j})\right\}_{j=1}^{M}. \tag{23}$$

where D_{grid} is the fixed-grid sample set, D_{adaptive} is the adaptive sample set supplemented from high-gradient regions, and K and M are the numbers of fixed-grid and adaptive samples, respectively. In this way, each additional patch is paired with a local PSF, and the added samples remain consistent with their spatially matched local imaging responses.

Figure 5 illustrates the structure-aware adaptive PSF-paired sampling procedure. By combining fixed-grid sampling with supplementary sampling in high-gradient areas, the training set includes PSF-paired samples from both the main spatial regions and complex structural zones.

2.5. Training implementation and inference

The restoration mapping f_θ is implemented using a U-Net-type encoder–decoder architecture. This architecture is suitable for same-size mapping from local migrated-image patches to restored image-domain model patches. Its multiscale feature extraction and skip connections help recover large-scale structural features while preserving local boundaries and thin-layer details.[34] In our implementation, the input is a single-channel local migrated-image patch, and the output is a single-channel restored image patch with the same spatial size. The encoder channel numbers are 64, 128, and 256, and the bottleneck contains 512 channels. The decoder uses bilinear up-sampling to recover spatial resolution, followed by a final convolution layer to generate the restored image patch. During training, this prediction is filtered using the corresponding local PSF. The network parameters are updated according to the self-supervised objective defined in Section 2.3.

Because different samples in the same mini-batch correspond to different PSFs, loss computation must use the local convolution kernel assigned to each sample. Batch-parallel PSF convolution is adopted to implement this sample-wise PSF filtering process. For a mini-batch with N samples, let the input patches, PSFs, and network predictions be $\{x_k\}_{k=1}^{N}$, $\{h_k\}_{k=1}^{N}$, and $\{\hat{m}_k\}_{k=1}^{N}$, respectively. For each mini-batch sample, the PSF-filtered prediction is computed as:

$$\hat{x}_k = h_k \otimes \hat{m}_k, \qquad k = 1, \ldots, N. \tag{24}$$

In implementation, these N sample-wise filtering operations are performed in parallel in the Fourier domain. For the field implementation, linear convolution uses zero-extension at the boundaries and preserves the sampled PSF origin when cropping the output to the prediction size. Each predicted patch is filtered only with its own sample-specific PSF, including interpolated PSFs for off-grid adaptive samples. The resulting PSF-filtered predictions are used in the losses defined in **Equations 17–19**.

After training, the learned restoration mapping f_θ is applied to the full input migrated section. Inference is performed using a sliding-window strategy. Overlapping predictions are averaged to reduce block-boundary artifacts and maintain structural continuity. The window size, fixed-grid sampling positions, number of adaptive samples, PSF interpolation method, loss weights, training parameters, and inference settings are specified in the corresponding numerical and field-data example.

3. Numerical examples

3.1. Experimental setup

To evaluate whether SPIRE-Net can recover reflectivity using the local PSF constraint as the training signal, we designed a controlled experiment on the two-dimensional Marmousi model. A spatially variant PSF operator was applied to the reference reflectivity to generate the synthetic input image. This provides a known image-domain degradation relationship for evaluating the network reconstruction. The original model size was 602 × 1,134. Because imaging blur and its directionality varied spatially, training samples were constructed from local patches extracted from the PSF-degraded synthetic image,

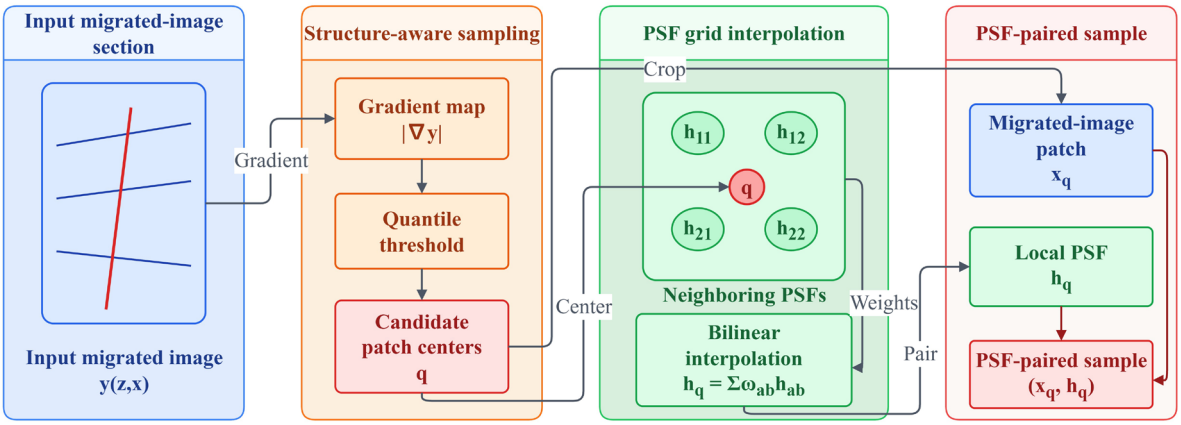


Figure 5. Structure-aware adaptive point-spread function (PSF)-paired sample construction

each paired with the local PSF at the corresponding spatial position. The PSF grid contained 14 sampling positions in the vertical direction and 18 sampling positions in the horizontal direction, resulting in 252 fixed-grid PSF-paired samples. Each fixed-grid sample consisted of a PSF-degraded synthetic image patch and the local PSF at the same spatial position.

To improve the coverage of complex structural regions, high-gradient areas were automatically selected from the input image. An additional 512 adaptive samples were extracted from these regions. Specifically, adaptive sampling used the 75th-percentile gradient threshold, a perturbation range of ± 12 samples, and seed 42. For the off-grid adaptive samples, the corresponding local PSFs were obtained using bilinear interpolation from neighboring PSF grid points. The original size of each extracted patch was 43×63, and zero-padding was applied to obtain 64×64 patches for network input.

The network was trained using the Adam optimizer.[35] The batch size was 64, and the initial learning rate was 1.0×10^{-3}. The weighting coefficient for high-gradient regions in the structure-weighted migrated-image-domain consistency loss was 2.0. The weight of the migrated-image-domain gradient-consistency loss was 0.2. Full-section inference was performed using 64×64 local windows with a sliding step of 8×8 grid points. Window predictions were placed back at their original spatial positions, and overlapping regions were averaged to reduce block-boundary artifacts. Quantitative evaluation was performed in the valid imaging region. True reflectivity was used only for evaluation after training.

Training on four graphics processing units (GPUs) took 248 s for 387 Marmousi epochs. Image-domain LSM required 90 s. Network inference required 23 s. Figure 6 shows the training loss curve for the Marmousi model. The loss decreased rapidly at the early stage and then gradually stabilized, indicating that the local PSF-based consistency constraint provides a stable self-supervised optimization signal.

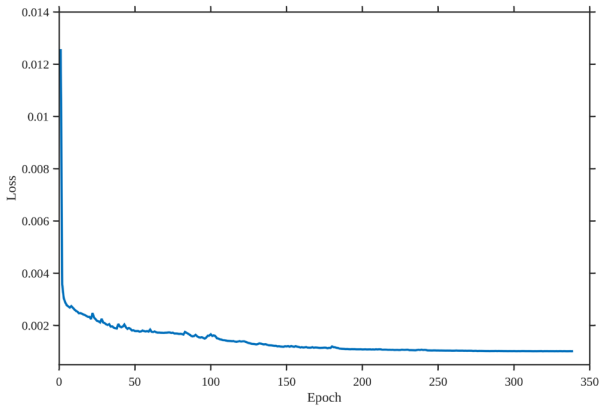


Figure 6. Loss curve for the Marmousi model

3.2. Results analysis

Figure 7 compares the Marmousi velocity model, the reference reflectivity, the PSF-degraded synthetic image, the image-domain LSM result, and the SPIRE-Net result. The PSF-degraded synthetic image in Figure 7C retains the main subsurface structures but exhibits reflector broadening, sidelobe interference, and local energy spreading. The image-domain LSM result in Figure 7D compensates for these effects and provides a reconstruction closer to the reference reflectivity. The SPIRE-Net result in Figure 7E preserves the principal reflector positions and reproduces

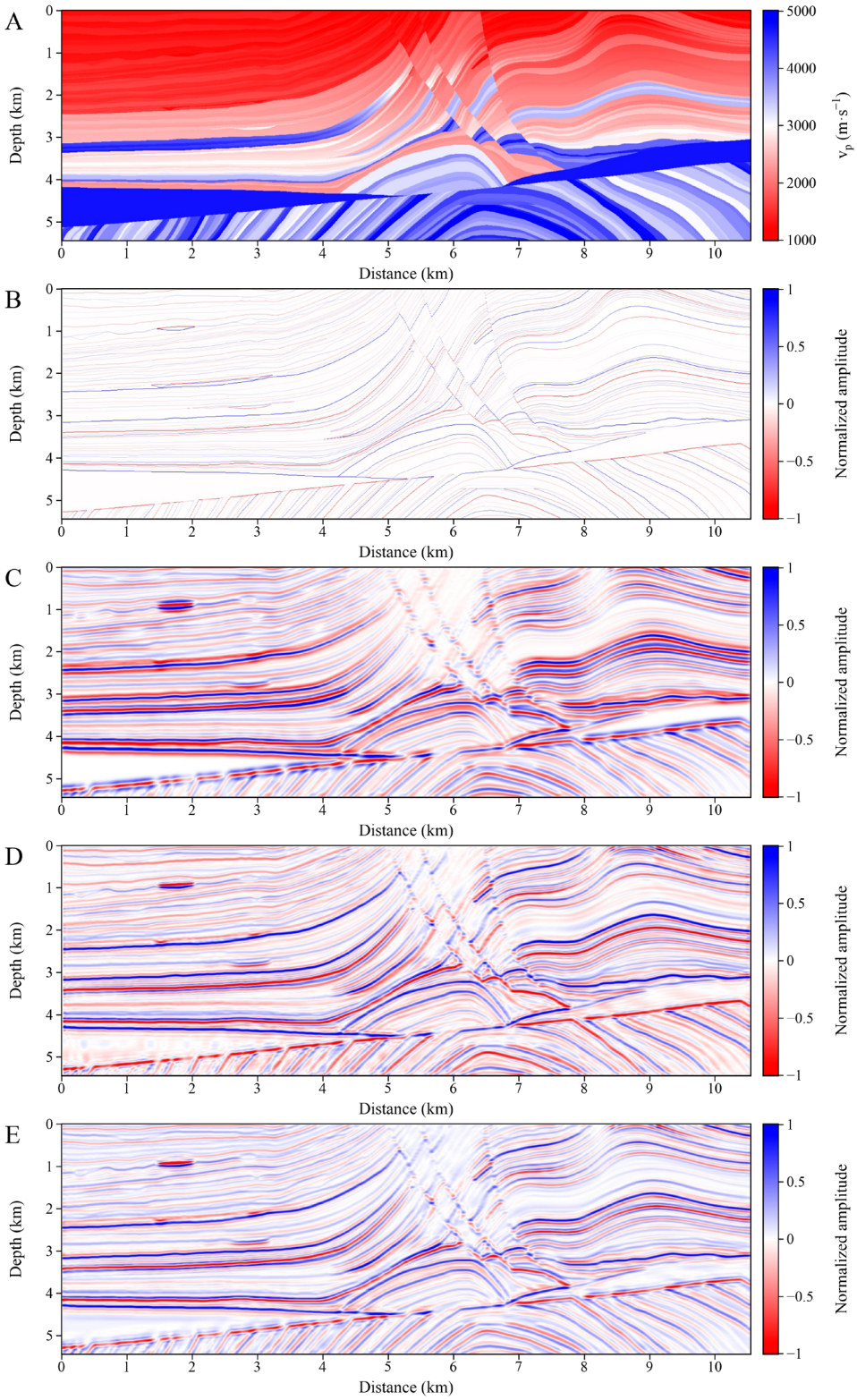


Figure 7. Comparison of Marmousi imaging results. (A) Velocity model. (B) True reflectivity. (C)Point-spread function–degraded synthetic image. (D) Image-domain least-squares migration result. (E) SPIRE-Net result.

the main focusing and continuity characteristics of the image-domain LSM result. This controlled comparison therefore evaluates whether the self-supervised network can recover the known reflectivity and approximate the compensation produced using image-domain LSM under the same PSF operator.

Figure 8 provides a local comparison between the image-domain LSM result and the SPIRE-Net result in a region containing curved reflectors and local dip variations. The two results show consistent reflector positions, layer boundaries, and local event geometry. SPIRE-Net reproduced the principal focusing and continuity characteristics of image-domain LSM, with only minor local differences. This enlarged view supports the consistency of the two restored results at the local structural scale.

For quantitative evaluation of imaging quality, mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM) were computed for different methods relative to the reference reflectivity within the valid imaging region. MSE measures the overall difference between the recovered image and the reference reflectivity. PSNR reflects overall signal fidelity, whereas SSIM evaluates structural similarity. Table 1 shows that SPIRE-Net and image-domain LSM achieved comparable metrics, with both improving the PSF-degraded synthetic image. Together with the structural comparisons in Figures 7 and 8, these results support the ability of SPIRE-Net to recover reflectivity through the prescribed PSF-based self-supervised constraint.

3.3. Ablation study

An ablation study on the Marmousi model was performed to evaluate the contributions of adaptive PSF-paired sampling and the structure-enhanced migrated-image-domain consistency loss. All ablation experiments used the same input image, PSF grid, restoration network, optimizer, and inference parameters. Only the sample construction strategy and the loss formulation were changed, so that the different configurations can be compared under consistent experimental conditions.

Table 1. Quantitative comparison of different methods relative to the reference reflectivity on the Marmousi model

Method	MSE	PSNR (dB)	SSIM
PSF-degraded synthetic image	1.337×10^{-2}	18.738	0.161
Image-domain LSM	1.196×10^{-2}	19.224	0.323
SPIRE-Net	1.156×10^{-2}	19.368	0.367

Abbreviations: LSM: Least squares migration; MSE: Mean squared error; PSF: Point-spread function; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure.

Four configurations were compared. The basic configuration used fixed-grid PSF-paired samples and the basic migrated-image-domain consistency loss. The second configuration added adaptive PSF-paired samples while keeping the basic loss unchanged. The third configuration used fixed-grid samples but replaced the basic loss with the structure-enhanced migrated-image-domain consistency loss. The full method combined adaptive PSF-paired sampling with the structure-enhanced loss.

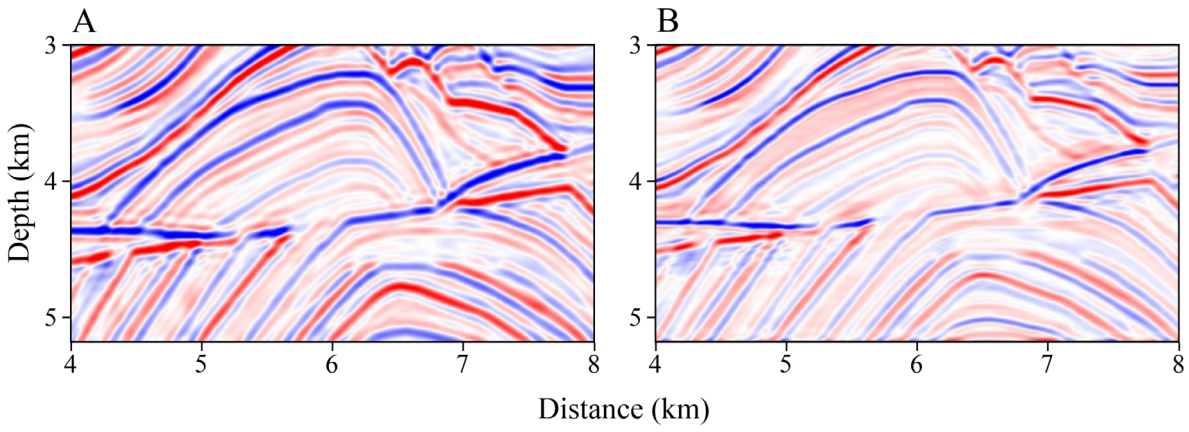


Figure 8. Representative local comparison on the Marmousi model. (A) Image-domain least-squares migration result. (B) SPIRE-Net result.

Table 2 lists the quantitative metrics of the four configurations relative to the reference reflectivity, and Figure 9 shows the corresponding metric comparison. Adding adaptive PSF-paired samples alone did not improve the metrics over the basic configuration, indicating that additional high-gradient samples are not sufficient without a matched loss weighting. Adding the structure-enhanced loss alone gave results close to the basic configuration. One possible explanation is that the training samples remain restricted to the fixed PSF grid. The full method achieved the best MSE, PSNR, and SSIM, suggesting that adaptive PSF-paired sampling and the structure-enhanced loss are most effective when used together.

The two components play complementary roles rather than producing independent gains under all configurations. Adaptive sampling alone shifts the training distribution toward fault, thin-layer, and high-gradient patches, which are more difficult to restore; consequently, the globally averaged metrics can decrease when the loss still weights all image samples in the same manner. Conversely, the structure-enhanced loss alone assigns greater importance to high-gradient pixels, but the fixed PSF grid contains a relatively limited number of complex-structure samples on which this weighting can act. When the two components are combined, adaptive sampling provides more structurally informative PSF-paired examples, while the structure-enhanced loss increases their contribution during optimization. The ablation results therefore support a complementary interaction between the two components under the tested configuration rather than an independent improvement from either component alone.

4. Field data

4.1. Experimental setup

The Marmousi example demonstrates the basic restoration capability of SPIRE-Net under controlled synthetic conditions. To further evaluate its applicability beyond synthetic examples, the proposed method was applied to field data. The field experiment used an RTM image from a survey area, spatially variant PSFs, and an image-domain LSM result. The RTM image served as the input image, and the local PSFs were used to construct the migrated-image-domain consistency constraint. The image-domain LSM result was used only as a comparison baseline and was not involved in network training.

Table 2. Ablation study results

Methods	MSE	PSNR (dB)	SSIM
Basic	1.175×10^{-2}	19.299	0.341
+ Adaptive samples	1.231×10^{-2}	19.099	0.319
+ Structural loss	1.182×10^{-2}	19.274	0.336
Full method	1.156×10^{-2}	19.368	0.367

Abbreviations: MSE: Mean squared error; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure.

The field RTM and PSF workflow used a three-dimensional tilted transversely isotropic acoustic-wave implementation. The computational image intervals were 12.5 m horizontally, 25 m in the crossline direction, and 10 m vertically. PSF construction used a selected subset of 1,671 survey shots and a 2-km migration aperture on each horizontal side. Sparse point responses were calculated on a 24 × 3 × 11 three-dimensional grid. The field PSFs were generated using the PSF module of the image-domain least-squares processing workflow. Let D denotes the

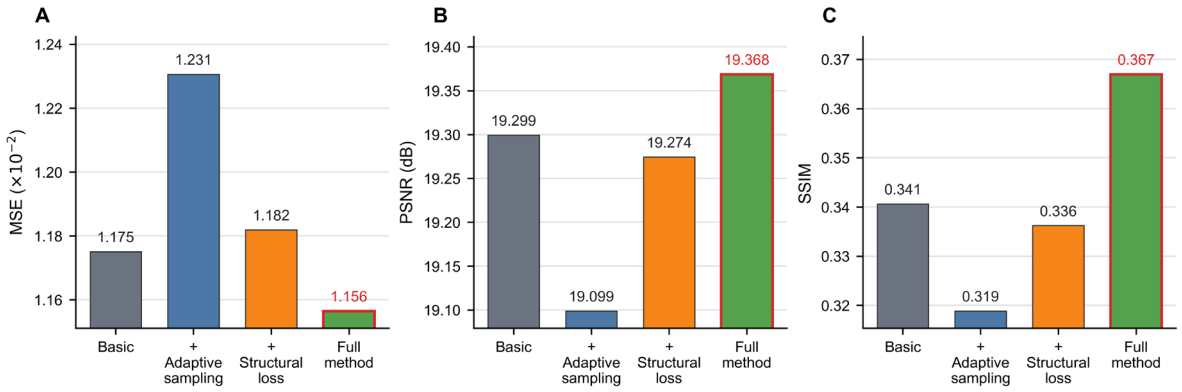


Figure 9. Quantitative comparison of ablation-study configurations. (A) Mean squared error (MSE). (B) Peak signal-to-noise ratio (PSNR). (C) Structural similarity index measure (SSIM).

demigration operator defined by the survey acquisition geometry and migration background model, and let M denotes the corresponding RTM operator. A sparse point-scatterer model:

$$p = \sum_j a_j e_j \tag{25}$$

was first demigrated to generate synthetic receiver data, after which RTM was applied:

$$d_{\text{psf}} = Dp, \qquad y_{\text{psf}} = Md_{\text{psf}} = MDp. \tag{26}$$

The spatially separated response around the j-th point was extracted as the corresponding local PSF. When M is the adjoint counterpart of D, MD approximates the Hessian $L^T L$ in **Equation 5**.

The size of the field-data section was 1,201 × 2,501. The PSF grid contained 24 sampling positions in the horizontal direction and 11 sampling positions in the vertical direction. Each local PSF had a size of 100 × 100, and the spacing between adjacent sampling positions was 100 grid points. The regular grid therefore provided 264 PSF-paired RTM samples. Using zero-based indices, the window origins started at 50 in both directions and advanced in steps of 100 samples, giving 24 horizontal and 11 vertical positions. To improve the coverage of complex structures and strong reflectors, 256 additional high-gradient positions were selected from the RTM image during training. Adaptive sampling used an 85th-percentile gradient threshold, a perturbation range of ±8 samples, and seed 42. The corresponding local PSFs for these off-grid samples were obtained using bilinear interpolation from neighboring regular-grid PSFs. Before network training, the PSF-patch collection and RTM-patch collection were separately normalized by their respective maximum absolute amplitudes. The same PSF bank was used for the image-domain LSM baseline and SPIRE-Net.

The training process followed the PSF-constrained self-supervised framework described in Sections 2.2–2.4. For the field data, the high-gradient-region weighting coefficient was set to 1.0, and the weight of the migrated-image-domain gradient-consistency loss was set to 0.1. Both values were lower than those used in the Marmousi experiment to reduce the influence of field-data noise and locally unstable structures during optimization. The lower coefficients reduce the relative contribution of high-gradient features while retaining the amplitude-consistency constraint. This choice accounts for the greater ambiguity between geological edges and artifacts in field RTM images. Gradient-based weighting can also emphasize artifacts, so the reduced coefficients control their contribution without identifying them as geological structure. The network was trained using the Adam optimizer. The batch size was 32, the initial learning rate was 1.0×10^{-5}, and the maximum number of epochs was 116. During inference, a 128×128 sliding window was used with a step of 4×4 grid points. Field-data training took 875.5 s. Inference took 89.4 s on four accelerators with 128 × 128 windows, a stride of 4 samples, and a batch size of 256.

Figure 10 shows the training loss curve for the field-data experiment. The loss decreased rapidly during the early training stage and then gradually stabilized, indicating that the local PSF-based consistency constraint still provides a stable self-supervised optimization signal under field-data conditions. Field data usually contain noise, velocity-model uncertainty, and PSF approximation errors. The stable loss curve indicates that the training objective remains numerically stable in this field example.

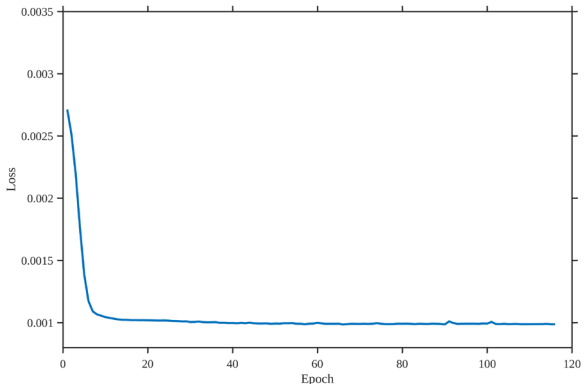


Figure 10. Loss curve for the field-data experiment

4.2. Results analysis

Because no reference reflectivity model is available for the field data, this section focuses on visual comparisons of local imaging results. The analysis mainly examined reflector continuity, local focusing, and the preservation of structural geometry. Figures 11 and 12 show two representative local regions from the field-data section.

Figure 11 compares the RTM, image-domain LSM, and SPIRE-Net results for field data. All three results preserved the main reflection geometry. In the green-boxed region, SPIRE-Net provided a clearer and more continuous delineation of the fault and its adjacent events than the RTM and image-domain LSM results. In the yellow-boxed region, SPIRE-Net made weak reflection events more distinct.

Figure 12 shows the RTM, image-domain LSM, and SPIRE-Net results for field data in another region of the field-data section. In the green-boxed region, compared with the RTM and image-domain LSM results, SPIRE-Net

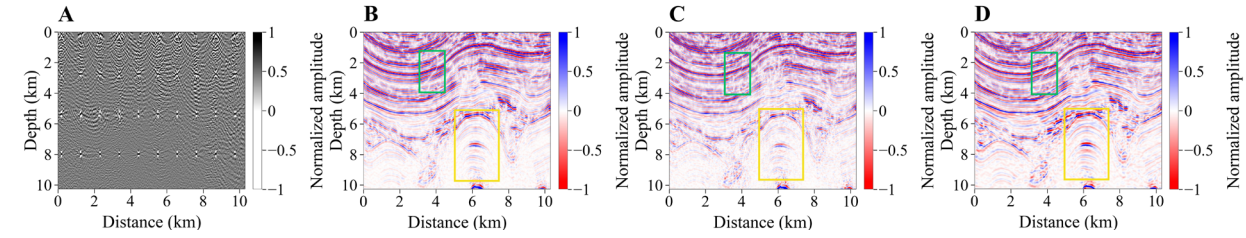


Figure 11. Field-data local region 1. (A) Local point-spread function. (B) Reverse time migration image at the corresponding position. (C) Image-domain least-squares migration result. (D) SPIRE-Net result.

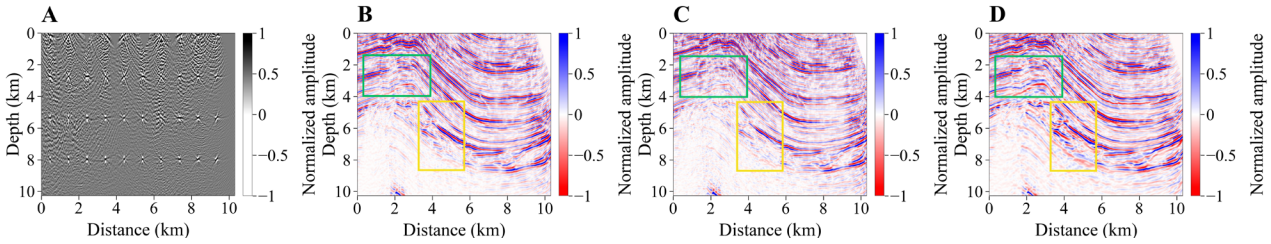


Figure 12. Field-data local region 2. (A) Local point-spread function. (B) Reverse time migration image at the corresponding position. (C) Image-domain least-squares migration result. (D) SPIRE-Net result.

made several weak curved reflections clearer and more laterally continuous. In the yellow-boxed region, SPIRE-Net enhanced weak dipping reflections.

5. Discussion

The reliability of SPIRE-Net depends on the accuracy and spatial representativeness of the local PSFs. The local approximation can weaken where velocity, illumination, or wave propagation varies rapidly within a patch. Bilinear interpolation introduces additional uncertainty when PSF phase, orientation, or sidelobe patterns change between sampled locations. These effects can lead to biased or incomplete compensation and motivate denser or spatially adapted PSF sampling in strongly heterogeneous regions.

The PSF-consistency constraint also leaves ambiguity in poorly resolved spatial-frequency components. Shared network parameters, the multiscale architecture, and the stopping rule provide implicit regularization, but they do not guarantee a unique reconstruction or exclude high-frequency amplification. This limitation should be considered when interpreting sharpened reflectors, particularly where the input contains coherent artifacts.

The field-data assessment was based on local comparisons with RTM and image-domain LSM. Well logs, interpreted horizons, or local gather consistency would strengthen the geological validation of the restored structures. The current application is also restricted to two-dimensional sections; extending it to three-dimensional volumes requires more efficient PSF storage, local filtering, and inference.

6. Conclusion

This study presents SPIRE-Net, a self-supervised migration image restoration method constrained by spatially variant PSFs. Adaptive PSF-paired sampling and a structure-enhanced consistency loss incorporate local structural information into the training objective. The Marmousi experiment verifies the ability of the network to solve the prescribed PSF-based image-domain inverse problem. In the field-data example, SPIRE-Net improves local reflector continuity and event separation relative to image-domain LSM in the examined regions while preserving the main reflection geometry. These findings suggest the use of local PSF constraints to learn migration image restoration directly from the input image.

Acknowledgments

None.

Funding

This work was supported by the SINOPEC Science and Technology Department under Grant No. P26032 ("Geophysical Prospective Technology Research Program 2026"), Grant No. P26011_2, and Grant No. P24165; and the National Natural Science Foundation of China under Grant No. U24B6001.

Conflict of interest

All authors are employees of SINOPEC Geophysical Research Institute Co., Ltd.; however, they were not involved in any activities that could constitute a conflict of interest in relation to this study. The authors declare they have no competing interests.

Author contributions

Conceptualization: Jingwen Zeng
Methodology: Jingwen Zeng
Investigation: Jingwen Zeng, Lu Xu
Software: Jingwen Zeng, Yingzhe Bai
Validation: Lu Xu
Visualization: Hao Wang, Jingwen Zeng
Writing–original draft: Jingwen Zeng
Writing–review & editing: Hao Wang

Availability of data

The data that support the findings of this study are from the SINOPEC Geophysical Research Institute, but were used under license for the current study and are not publicly available. Data are available from the corresponding author upon reasonable request and with permission of the institute.

References

1. Baysal E, Kosloff DD, Sherwood JWC. Reverse time migration. *Geophysics*. 1983;48(11):1514-1524.

   doi: 10.1190/1.1441434

2. McMechan GA. Migration by extrapolation of time-dependent boundary values. *Geophys Prospect*. 1983;31(3):413-420.

   doi: 10.1111/j.1365-2478.1983.tb01060.x

3. Schuster GT. Least-squares cross-well migration. In: *SEG Technical Program Expanded Abstracts 1993*. Houston, TX: Society of Exploration Geophysicists; 1993:110-113.

   doi: 10.1190/1.1822308

4. Nemeth T, Wu C, Schuster GT. Least-squares migration of incomplete reflection data. *Geophysics*. 1999;64(1):208-221.

   doi: 10.1190/1.1444517

5. Chavent G, Plessix RE. An optimal true-amplitude least-squares prestack depth-migration operator. *Geophysics*. 1999;64(2):508-515.

   doi: 10.1190/1.1444557

6. Dai W, Fowler P, Schuster GT. Multi-source least-squares reverse time migration. *Geophys Prospect*. 2012;60(4):681-695.

   doi: 10.1111/j.1365-2478.2012.01092.x

7. Dong S, Cai J, Guo M, *et al*. Least-squares reverse time migration: towards true amplitude imaging and improving the resolution. In: *SEG Technical Program Expanded Abstracts 2012*. Houston, TX: Society of Exploration Geophysicists; 2012:1-5.

   doi: 10.1190/segam2012-1488.1

8. Zeng C, Dong S, Wang B. Least-squares reverse time migration: inversion-based imaging toward true reflectivity. *Lead Edge*. 2014;33(9):962-968.

   doi: 10.1190/tle33090962.1

9. Tang Y. Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian. *Geophysics*. 2009;74(6):WCA95-WCA107.

   doi: 10.1190/1.3204768

10. Fletcher RP, Nichols D, Bloor R, Coates RT. Least-squares migration—data domain versus image domain using point spread functions. *Lead Edge*. 2016;35(2):157-162.

    doi: 10.1190/tle35020157.1

11. Xu P, Feng B, Wang H, Liu S. Noniterative least-squares reverse time migration based on an efficient asymptotic Hessian/point-spread function estimation. *Geophysics*. 2022;87(4):S169-S184.

    doi: 10.1190/geo2021-0401.1

12. Yang J, Huang J, Zhu H, McMechan G, Li Z. An efficient and stable high-resolution seismic imaging method: point-spread function deconvolution. *J Geophys Res Solid Earth*. 2022;127(7):e2021JB023281.

    doi: 10.1029/2021JB023281

13. Liu C, Sun M, Dai N, *et al*. Deep learning-based point-spread function deconvolution for migration image deblurring. *Geophysics*. 2022;87(4):S249-S265.

    doi: 10.1190/geo2020-0904.1

14. Dong S, Dong X, Zhang R, Cong Z, Zhong T, Wang H. Global-feature-fusion and multiscale network for low-frequency extrapolation. *IEEE Trans Geosci Remote Sens*. 2024;62:5917714.

    doi: 10.1109/TGRS.2024.3408949

15. Zhong T, Yang F, Dong X, Dong S, Luo Y. SHBGAN: hybrid bilateral attention GAN for seismic image super-resolution reconstruction. *IEEE Trans Geosci Remote Sens*. 2024;62:5934312.

    doi: 10.1109/TGRS.2024.3492142

16. Zhou Z, Liu S, Ni W, Yan Z, Gu H, Zhang B. A data-driven high-resolution imaging based on the cost-effective construction of point-spread functions. *Geophys Prospect*. 2025;73(8):e70086.

    doi: 10.1111/1365-2478.70086

17. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed

neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys*. 2019;378:686-707.

doi: 10.1016/j.jcp.2018.10.045

18. Rasht-Behesht M, Huber C, Shukla K, Karniadakis GE. Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions. *J Geophys Res Solid Earth*. 2022;127(5):e2021JB023120.

doi: 10.1029/2021JB023120

19. Biswas R, Sen MK, Das V, Mukerji T. Prestack and poststack inversion using a physics-guided convolutional neural network. *Interpretation*. 2019;7(3):SE161-SE174.

doi: 10.1190/INT-2018-0236.1

20. Zhang Y, Singh S, Thanoon D, Devarakota P, Jin L, Tsvankin I. Physics-guided unsupervised deep-learning seismic inversion with uncertainty quantification. *J Seism Explor*. 2023;32(3):257-270.

21. Dong X, Yuan Z, Lin J, Dong S, Tong X, Li Y. Bidirectional physics-constrained full-waveform inversion: reducing seismic data dependence in velocity model building. *Geophys J Int*. 2025;244(2):ggaf466.

doi: 10.1093/gji/ggaf466

22. Wu B, Meng D, Zhao H. Semi-supervised learning for seismic impedance inversion using generative adversarial networks. *Remote Sens (Basel)*. 2021;13(5):909.

doi: 10.3390/rs13050909

23. Sun Q, Zong Z. Building initial model for seismic inversion based on semi-supervised learning. *Geophys Prospect*. 2024;72(5):1800-1815.

doi: 10.1111/1365-2478.13491

24. Sun QH, Zong ZY, Li X. Probabilistic seismic inversion based on physics-guided deep mixture density network. *Pet Sci*. 2024;21(3):1611-1631.

doi: 10.1016/j.petsci.2023.12.015

25. Kühl H, Sacchi MD. Least-squares wave-equation migration for AVP/AVA inversion. *Geophysics*. 2003;68(1):262-273.

doi: 10.1190/1.1543212

26. Xue Z, Chen Y, Fomel S, Sun J. Seismic imaging of incomplete data and simultaneous-source data using least-squares reverse time migration with shaping regularization. *Geophysics*. 2016;81(1):S11-S20.

doi: 10.1190/geo2014-0524.1

27. Shoja A, Van Der Neut J, Wapenaar K. Target-oriented least-squares reverse time migration with marchenko redatuming and double focusing: field data application. *Geophysics*. 2024;89(3):S267-S274.

doi: 10.1190/geo2023-0374.1

28. Rickett JE. Illumination-based normalization for wave-equation depth migration. *Geophysics*. 2003;68(4):1371-1379.

doi: 10.1190/1.1598130

29. Lecomte I. Resolution and illumination analyses in PSDM: a ray-based approach. *Lead Edge*. 2008;27(5):650-663.

doi: 10.1190/1.2919584

30. Cavalca M, Fletcher RP, Nichols D. Image-domain least-squares migration in rapidly-varying media: practical considerations. In: *First EAGE/SBGf Workshop on Least-Squares Migration*. Houten, Netherlands: European Association of Geoscientists & Engineers; 2018.

doi: 10.3997/2214-4609.201803057

31. Osorio LN, Pereira-Dias B, Bulcão A, Landau L. Migration deconvolution using domain decomposition. *Geophysics*. 2021;86(3):S247-S256.

doi: 10.1190/geo2020-0352.1

32. Li B, Sun M, Xiang C, Bai Y. Least-squares reverse time migration in imaging domain based on global space-varying deconvolution. *Appl Sci*. 2022;12(5):2361.

doi: 10.3390/app12052361

33. Zhang W, Guo X, Ravasi M, Gao J, Sun W. Angle-dependent image-domain least-squares migration through analytical point spread functions. *Geophysics*. 2024;89(5):S339-S360.

doi: 10.1190/geo2023-0499.1

34. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (Lecture Notes in Computer Science)*. Cham, Switzerland: Springer International Publishing; 2015:234-241.

doi: 10.1007/978-3-319-24574-4_28

35. Kingma DP, Ba J. Adam: a method for stochastic optimization. Preprint. Posted online December 22, 2014. arXiv.

doi: 10.48550/arXiv.1412.6980