

# Hyperspectral Image Reconstruction Algorithms for Broadband Spectral Modulation: A Comprehensive Review

Fangning Li,<sup>a</sup> Heyan Meng,<sup>a</sup> Jiajie Zhang,<sup>a</sup> Xianye Li,<sup>b,c</sup> Zhangjun Wang,<sup>d</sup> Guoliang Shentu,<sup>e</sup> Baoqing Sun,<sup>a\*</sup> Yupeng Wang<sup>a\*</sup>

<sup>a</sup>School of Information Science and Engineering, Shandong University, Qingdao, 266237, China

<sup>b</sup>School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China

<sup>c</sup>Shandong Key Laboratory of Intelligent Electronic Packaging Testing and Application, Weihai 264209, China

<sup>d</sup>State Key Laboratory of Physical Oceanography, Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266100, China

<sup>e</sup>Shandong Guoyao Quantum Lidar technology Co. Ltd., Jinan 250101, China

**Abstract.** Spectral modulation-based hyperspectral imaging is an important approach for acquiring high-dimensional hyperspectral data via computational reconstruction. This method addresses limitations of traditional scanning systems in terms of hardware complexity, acquisition speed, and data storage. This review systematically summarizes the development and current status of reconstruction algorithms for spectral modulation-based hyperspectral imaging. Existing methods are categorized into three main types: prior-constraint approaches based on mathematical models, learning-based methods using deep neural networks, and hybrid frameworks that combine model-based priors with data-driven strategies. Prior-constraint methods, such as those using sparsity, low-rank, or regularization constraints, offer solid theoretical foundations and interpretability, but often face challenges related to parameter selection and computational efficiency. Deep learning approaches, including convolutional neural networks and transformer architectures, can achieve high reconstruction accuracy and real-time performance, but typically require large amounts of training data and may have limited generalization to diverse imaging conditions. Hybrid methods, especially deep unfolding networks, seek to balance theoretical interpretability and adaptive performance. Key challenges include balancing reconstruction accuracy and computational cost, integrating physical knowledge into network design, and establishing standardized evaluation frameworks. This review aims to provide a reference for the development and application of spectral modulation-based hyperspectral imaging in areas such as remote sensing, biomedical diagnostics, and industrial inspection.

**Keywords:** hyperspectral imaging; compressive sensing; spectral modulation; image reconstruction; deep learning.

\*Seventh Author, E-mail: baoqing.sun@sdu.edu.cn

\*Eighth Author, E-mail: yupeng\_wang@mail.sdu.edu.cn

## 1 Introduction

Hyperspectral imaging has undergone a remarkable transformation over the past three decades, evolving from bulky scanning systems to sophisticated computational frameworks that fundamentally reimagine how spectral information is captured and processed. Traditional hyperspectral imaging systems, which mechanically scan scenes across spatial or spectral dimensions, have long been constrained by the fundamental trade-off between spatial resolution, spectral resolution, and temporal resolution. These scanning-based approaches, while providing

high-quality spectral data, suffer from prolonged acquisition times, substantial data storage requirements, and limited applicability to dynamic scenes.

The emergence of computational spectral imaging in the early 2000s marked a paradigm shift in hyperspectral data acquisition. By leveraging the mathematical framework of **CS (compressive sensing)**, researchers demonstrated that hyperspectral datacubes could be reconstructed from significantly fewer measurements than dictated by the Nyquist-Shannon theorem. This breakthrough catalyzed the development of snapshot hyperspectral imaging systems that encode three-dimensional spectral-spatial information onto two-dimensional detector through optical modulation. Current state-of-the-art systems achieve compression ratios exceeding 100:1 while maintaining reconstruction quality suitable for quantitative analysis. Key technological milestones include the introduction of coded aperture snapshot spectral imaging (CASSI), spatial-spectral encoded compressive imaging, and more recently, metasurface-based spectral modulators that enable unprecedented miniaturization and integration capabilities.

Broadband spectral modulation offers an alternative approach to hyperspectral imaging that addresses certain limitations of narrowband filtering techniques. Unlike sequential band capture or fixed narrowband filters, broadband modulation encodes information across the entire spectral range through parallel processing. This approach provides several advantages: improved light throughput and signal-to-noise characteristics, flexible post-acquisition spectral resolution adjustment through computational reconstruction, and enhanced performance under low-light conditions. By distributing spectral information across detector pixels through designed modulation patterns, these systems can achieve light collection efficiency improvements compared to scanning approaches, extending hyperspectral imaging capabilities to challenging scenarios in biomedical diagnostics and remote sensing applications.

Despite advances in computational hyperspectral imaging, current methods still face limitations that restrict their practical application. Most existing systems suffer from a fundamental disconnect between hardware design and reconstruction algorithms, where optical encoding schemes are developed independently from the computational methods used to recover spectral information. This separation often results in suboptimal overall system performance, as the encoding patterns may not align with the assumptions underlying reconstruction algorithms. Current reconstruction methods predominantly rely on generic mathematical priors such as sparsity or low-rank constraints, which fail to fully exploit the rich physical properties and domain-specific characteristics of hyperspectral data. The computational burden of iterative reconstruction algorithms remains prohibitive for real-time applications, while learning-based approaches struggle with generalization across different scene types and imaging conditions. Moreover, existing systems lack standardized evaluation frameworks and benchmarks, making objective comparison and systematic improvement challenging. The field urgently requires integrated design approaches that jointly optimize optical encoding and computational reconstruction, adaptive algorithms that can adjust to scene-specific properties, and efficient implementations suitable for embedded systems and real-time processing.

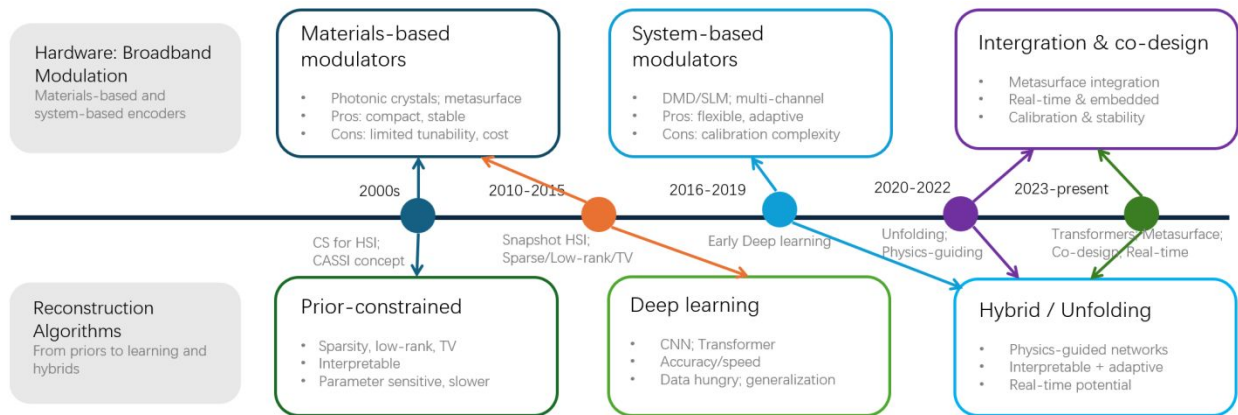
Material-based approaches to broadband spectral modulation exploit the intrinsic optical properties of advanced materials to achieve passive spectral encoding. These methods primarily utilize dispersive elements, photonic crystals, and engineered metasurfaces that modulate incident light based on wavelength-dependent interactions. Recent developments in nanofabrication have enabled metasurfaces with customized spectral responses, offering compact alternatives to traditional bulk optics. The main advantages include mechanical stability, absence of moving parts, and potential for extreme miniaturization. However, material-based solutions face significant

challenges: the fixed spectral response limits adaptability to different imaging scenarios, fabrication costs remain high for large-aperture systems, and wavelength-dependent efficiency variations can complicate reconstruction algorithms. Despite these limitations, material-based modulators remain attractive for applications requiring robust, maintenance-free operation.

System-based approaches integrate multiple optical and electronic components to achieve flexible broadband spectral modulation. These methods combine spatial light modulators, coded apertures, and multi-channel detection schemes within unified architectures. By treating spectral modulation as a system-level challenge rather than a component-level problem, these approaches enable adaptive encoding strategies and real-time optimization based on scene content. The primary benefits include operational flexibility, the ability to incorporate feedback mechanisms, and superior performance through hardware-software co-optimization. However, system complexity introduces calibration challenges, increased cost, and potential reliability issues from multiple interdependent components. Additionally, the need for precise alignment and synchronization limits portability. Despite these constraints, system-based methods offer the best performance for applications demanding high accuracy and adaptability.

Advanced reconstruction algorithms are fundamental to broadband spectral modulation systems, transforming raw encoded measurements into hyperspectral datacubes with enhanced spectral accuracy and spatial resolution. Modern reconstruction algorithms leverage diverse mathematical frameworks, from traditional optimization-based methods incorporating physical priors to deep learning architectures that learn implicit spectral-spatial representations from data. These computational approaches effectively compensate for hardware limitations and enable the extraction of spectral information that would be impossible to obtain through optical means alone. The evolution from iterative optimization algorithms to end-to-end learning frameworks has

dramatically improved reconstruction speed and quality. Deep unfolding networks, which embed domain knowledge into learnable architectures, achieve superior performance by combining the interpretability of model-based methods with the adaptability of data-driven approaches. Transformer-based architecture has recently demonstrated exceptional capability in modeling long-range spectral correlations, while physics-informed neural networks ensure reconstructions remain consistent with fundamental imaging principles. The primary advantage of algorithmic solutions lies in their flexibility and continuous improvement potential through software updates. However, challenges remain in balancing reconstruction accuracy with computational efficiency, ensuring generalization across diverse imaging conditions, and maintaining interpretability for scientific applications requiring uncertainty quantification.



**Fig. 1** Development Roadmap of Broadband Spectral Modulation-based HSI.

As discussed, the field of broadband spectral modulation has advanced along two parallel yet interconnected tracks: the evolution of modulation hardware from materials-based to system-based, and the innovation in reconstruction algorithms from prior-constrained to deep learning and hybrid methods. Figure 1 provides a visual roadmap of this evolution, illustrating the progression from early concepts like CS and CASSI to the current trends of physics-guided deep networks and

hardware-software "co-design." This increasing integration of hardware encoding and computational reconstruction is a core driver of the field and a key theme of this review.

This review provides a comprehensive analysis of hyperspectral reconstruction algorithms for spectral modulation-based imaging systems. Chapter 2 establishes the theoretical foundations necessary for understanding reconstruction methodologies. Building on these foundations, Chapter 3 examines model-based reconstruction algorithms that leverage mathematical priors and constraints. As the field has evolved, Chapter 4 discusses hybrid and advanced reconstruction frameworks that integrate multiple algorithmic strategies to overcome individual method limitations. Moving toward modern approaches, Chapter 5 explores learning-based reconstruction algorithms that harness data-driven techniques for improved performance. Throughout each chapter, we analyze the theoretical foundations, implementation details, and performance characteristics of the presented methods. The review concludes in Chapter 6 with current challenges and future research directions in this rapidly evolving field.

## **2 Fundamental of Hyperspectral Reconstruction Algorithms**

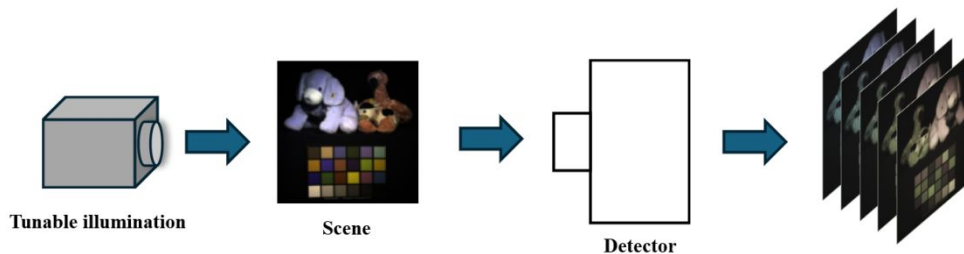
### *2.1 Hyperspectral Imaging Model*

Broadband spectral modulation techniques play an important role in hyperspectral imaging. The core principle involves encoding both spectral and spatial information of incident light to achieve compressed sampling of hyperspectral data. This chapter derives a general linear observation model from the physical principles of broadband spectral modulation, establishing a unified mathematical framework applicable to various broadband modulation techniques and providing theoretical foundation for subsequent reconstruction algorithms.

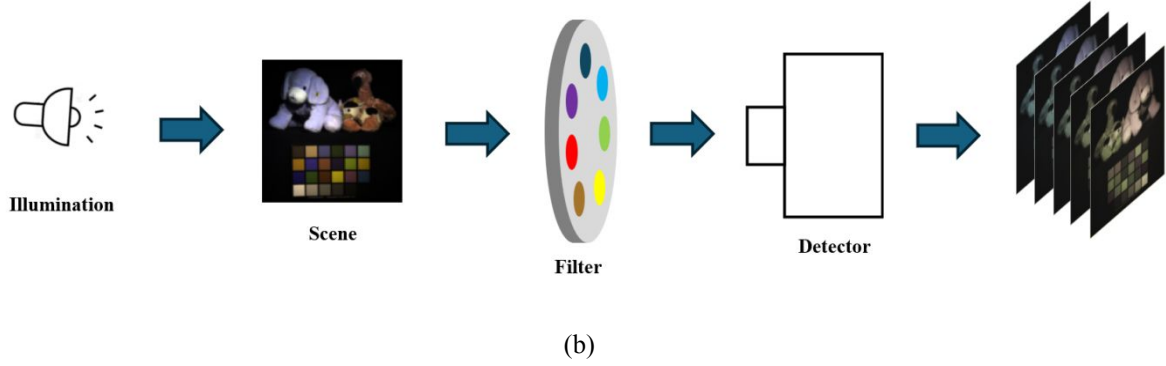
Broadband spectral modulation techniques can be categorized into two fundamental approaches based on their spectral selection strategy: active modulation and passive modulation. Active modulation systems employ tunable illumination sources to provide wavelength-specific excitation of target scenes. These systems typically utilize tunable lasers<sup>1</sup>, LED arrays<sup>2</sup>, or acousto-optic modulators<sup>3</sup> to sequentially illuminate targets at discrete wavelengths. Passive modulation systems rely on wavelength-selective optical elements to extract spectral information from broadband incident light. Common implementations include filter-wheel cameras that sequentially position bandpass filters in the optical path, coded aperture systems that combine spatial masks with dispersive elements, and snapshot hyperspectral cameras that simultaneously encode spectral and spatial information through specialized optical architectures<sup>4</sup>. Both active and passive systems can employ either single-pixel detectors or area-array cameras depending on the spatial sampling requirements. Single-pixel configurations require spatial scanning mechanisms to acquire complete spatial information, while area-array configurations enable simultaneous multi-pixel acquisition across the spatial field.

The physical implementation differs between active and passive modulation approaches in their optical configurations and signal processing methods. Active modulation systems integrate tunable sources with illumination optics to provide broadband spectral modulation of target scenes. In single-pixel implementations, imaging optics focus the target scene onto spatial light modulators such as digital micromirror devices (DMDs), where spatial encoding occurs before the modulated light reaches a single-pixel detector. Spatial scanning is achieved through programmable modulation patterns. In area-array implementations, imaging optics directly focus the reflected or transmitted light onto the detector array, capturing full spatial information simultaneously across the modulated spectral range. Spectral modulation occurs at the illumination stage in both

configurations, allowing direct control over the incident spectral characteristics and intensity distribution. Passive modulation systems operate by filtering broadband incident light through wavelength-selective components. Filter-wheel implementations mechanically rotate different optical filters into the optical path, enabling temporal sampling of spectral bands while maintaining spatial resolution through conventional detector arrays. Coded aperture systems employ structured masks combined with dispersive elements, where spectral information is recovered through computational reconstruction from multiple measurements. Snapshot systems integrate dispersive and encoding elements to simultaneously capture spectral and spatial information in a single exposure. Figure 2 illustrates these two general approaches. Figure 2(a) shows an active system using tunable illumination. Figure 2(b) provides a simplified conceptual diagram of a passive system, where the 'Filter' component is representative of the broader category of wavelength-selective elements—such as optical filters, prisms, or gratings—that are characteristic of this approach and are discussed in the text. The detector in both configurations can be either a single-pixel sensor or an area-array camera.



(a)



**Fig. 2** (a) An active spectral modulation system using tunable laser illumination. (b) A passive spectral modulation system employing optical filters.

Despite the diversity in physical implementations, all these systems share a common mathematical structure. This review focuses on elastic light–matter interactions for hyperspectral imaging—namely reflectance, transmittance, and absorption measurements—in which detected photons are wavelength-preserving with respect to the scene interaction. Consequently, inelastic mechanisms that induce wavelength shifts (e.g., Raman scattering and fluorescence) fall outside the scope of this mathematical framework and the subsequent discussion. Under these assumptions, without cross-wavelength coupling, we establish a unified observation model. A general continuous-domain forward model is given by:

$$y_i = \iiint H(u, v, \lambda) \cdot M_i(u, v, \lambda) \cdot D(u, v, \lambda) du dv d\lambda \quad (1)$$

where  $M_i(u, v, \lambda)$  represents the joint spatial-spectral modulation function for the  $i$ -th measurement configuration,  $D(u, v, \lambda)$  represents the wavelength- and pixel-dependent detector responsivity under linear operating conditions (after dark/flat-field calibration), and  $y_i$  is the measurement response for the  $i$ -th modulation configuration.

In many practical systems, the joint modulation function can be separated into independent spatial and spectral components, such that  $M_i(u, v, \lambda) = T_s^i(\lambda) \cdot T_t^i(u, v)$ . This separability condition

holds when the physical implementation ensures that spectral modulation, spatial modulation, and detector response operate as independent multiplicative components in the measurement process. Under this assumption, the spectral modulation function  $T_s^i(\lambda)$  characterizes the wavelength-dependent transmission or filtering properties, while the spatial modulation function  $T_t^i(u, v)$  describes the position-dependent encoding patterns applied to the incoming light field. The detector response function  $D(u, v, \lambda)$  is a critical component that accounts for both the spatial sensitivity variations across pixels (a function of  $u, v$ ) and the wavelength-dependent quantum efficiency of the sensing array (a function of  $\lambda$ ). While  $D(u, v, \lambda)$  is a general form, it is often the case in well-calibrated array detectors that spatial uniformity is better than spectral uniformity. In such scenarios, the detector response can be reasonably approximated as a product of a spatially varying sensitivity map and a dominant spectral response curve.

To convert the continuous-domain model into a discrete form suitable for numerical computation, we discretize each dimension appropriately. The spectral dimension is sampled at  $L$  wavelengths, the spatial dimension at  $N$  positions, with  $M$  different modulation patterns. The hyperspectral data cube is vectorized as  $\mathbf{x} \in \mathbb{R}^{LN}$ , where element  $x_j$  represents the radiance intensity at a specific wavelength-position combination.

Based on the discretized representation, each measurement can be expressed as:

$$y_i = \sum_{j=1}^{LN} \phi_{i,j} h_j \quad (2)$$

where  $\phi_{i,j}$  represents the combined modulation effect at the  $j$ -th spectral-spatial position for the  $i$ -th measurement.

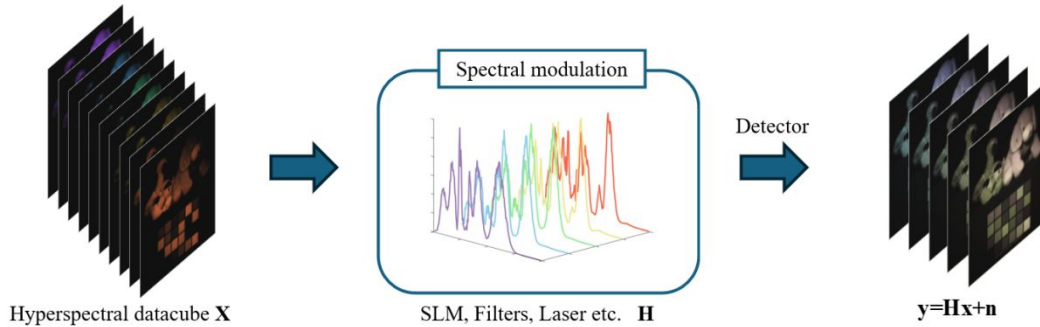
Combining all  $M$  measurements forms the complete observation vector  $\mathbf{y} = [y_1, y_2, \dots, y_M]^T$ , with the corresponding system matrix  $\mathbf{H} \in \mathbb{R}^{M \times LN}$ . Considering noise effects in actual measurement processes, we obtain the unified linear observation model:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (3)$$

Linearity assumption and detector operating regime. Equation (3) models detection as a linear mapping, which presumes the sensor operates within its specified linear response range such that the expected digital output is proportional to the incident irradiance<sup>5, 6</sup>. This assumption is standard in compressive hyperspectral imaging and is supported by characterization studies of scientific-grade CCD/CMOS and InGaAs sensors, which typically exhibit <0.5–1% nonlinearity over the central portion of their dynamic range when operated below saturation and after radiometric calibration<sup>7-9</sup>. In our experiments/scope, exposure time and illumination are controlled to keep the detector within this linear regime, and routine preprocessing (dark subtraction, flat-fielding, optional nonlinearity correction) is applied. Residual nonlinearities are therefore negligible for the purposes of our analysis and are absorbed into the noise term  $\mathbf{n}$ . For detectors known to exhibit stronger nonlinear behavior (e.g., uncooled microbolometers or some HgCdTe arrays), the model can be extended to  $\mathbf{y} = f(\mathbf{H}\mathbf{x}) + \mathbf{n}$  with a per-pixel monotonic response  $f$  calibrated offline<sup>10</sup>; a systematic treatment of such nonlinear models is beyond the scope of this review.

In this model, the observation vector  $\mathbf{y}$  contains all measurements; the system matrix  $\mathbf{H}$  encodes the combined effects of spectral modulation, spatial modulation, and detector response; the signal vector  $\mathbf{x}$  is the vectorized hyperspectral scene; and the noise vector  $\mathbf{n}$  represents measurement errors. Figure 3 illustrates the hyperspectral data processing flow based on spectral

modulation, showing the transformation from input hyperspectral datacube through spectral modulation to the final detector output.



**Fig. 3** Hyperspectral data process based on spectral modulation.

This unified linear observation model provides a consistent mathematical framework that accommodates the diverse physical implementations described above. The model naturally reduces to established formulations for specific system types while maintaining generality across different modulation strategies. This mathematical consistency enables systematic algorithm development and performance comparison across various broadband modulation techniques, facilitating both theoretical analysis and practical implementation of hyperspectral reconstruction methods.

## 2.2 Overview of Hyperspectral Reconstruction Algorithms

The theoretical foundation of this approach rests on CS principles, which demonstrate that sparse signals can be accurately reconstructed from significantly fewer measurements than dictated by the Nyquist-Shannon sampling theorem<sup>11, 12</sup>. Hyperspectral data exhibits natural compatibility with the CS framework due to its inherent spatiotemporal correlations and spectral redundancies, manifesting as sparse representations in appropriate transform domains or low-rank structures across spatial and spectral dimensions.

CS-based hyperspectral imaging systems comprise two fundamental components: optical encoding modulation and computational reconstruction. The optical encoding stage implements spectral-spatial modulation patterns to compress high-dimensional hyperspectral information into lower-dimensional measurements. Subsequently, computational reconstruction algorithms exploit the underlying signal structure to recover the complete hyperspectral datacube from these compressed observations.

Evaluating the performance of hyperspectral reconstruction algorithms requires a comprehensive set of quantitative metrics to assess the quality of the recovered datacube against a ground-truth reference. These metrics are crucial for objective comparison and can be broadly categorized based on whether they measure spatial reconstruction quality, assess spectral fidelity, or provide a global assessment of the entire datacube. A robust evaluation relies on a combination of metrics from these categories to capture the multi-faceted nature of reconstruction quality<sup>13</sup>.

The spatial quality of a reconstructed hyperspectral image is fundamental, and it is commonly evaluated using metrics such as PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). PSNR is a widely adopted metric that quantifies pixel-wise accuracy based on MSE (Mean Squared Error) between the reconstructed and ground-truth images, typically calculated for each spectral band and then averaged<sup>14, 15</sup>. The PSNR is defined as:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (4)$$

where  $\text{MAX}_I$  is the maximum possible pixel value. While PSNR provides a straightforward measure of error, it does not always align with human visual perception. To address this, SSIM was developed to evaluate the similarity of structure, luminance, and contrast between two images,

offering a more perceptually relevant assessment<sup>16</sup>. A higher PSNR value and an SSIM value closer to 1 both indicate superior spatial reconstruction<sup>17, 18</sup>.

Beyond spatial accuracy, preserving the integrity of spectral information is paramount in hyperspectral imaging. SAM (Spectral Angle Mapper) is the primary metric for this purpose, designed specifically to assess spectral fidelity. SAM measures the spectral similarity by calculating the angle between the spectral vectors of the same pixel in the reconstructed and ground-truth images, effectively quantifying the shape similarity of the spectral signatures irrespective of their absolute intensity. For a single pixel at spatial location  $(i, j)$ , the SAM is defined as:

$$\text{SAM}(i, j) = \arccos\left(\frac{\mathbf{x}_{i,j}^T \mathbf{x}_{i,j}}{\|\mathbf{x}_{i,j}\|_2 \|\mathbf{x}_{i,j}\|_2}\right) \quad (5)$$

The final SAM score is averaged over all spatial pixels, with a smaller value (in degrees or radians) signifying higher spectral fidelity and a more accurate preservation of material-specific spectral characteristics<sup>19, 20</sup>.

To provide a holistic assessment of the entire datacube, global quality indices that integrate both spatial and spectral errors are employed. Among these, ERGAS (Erreur Relative Globale Adimensionnelle de Synthèse) is particularly valuable. ERGAS quantifies the overall distortion by considering the root mean squared error in each band relative to its mean radiance, thus providing a single figure of merit for the entire reconstructed image<sup>21</sup>. Another important global metric is MRAE (Mean Relative Absolute Error), which measures the average relative error across all pixels and bands, making it less sensitive to the absolute scale of pixel intensities<sup>22</sup>. The formulas are given by:

$$\mathbf{ERGAS} = 100 \frac{h}{l} \sqrt{\frac{1}{N_\lambda} \sum_{k=1}^{N_\lambda} \left( \frac{\mathbf{RMSE}(k)}{\mu(k)} \right)^2} \quad (6)$$

$$\mathbf{MRAE} = \frac{1}{N_x N_y N_\lambda} \sum_{i,j,k} \frac{|X_{i,j,k} - \hat{X}_{i,j,k}|}{X_{i,j,k} + \delta} \quad (7)$$

In summary, the evaluation of HSI reconstruction is a multi-dimensional task that cannot rely on a single metric. A comprehensive and fair comparison of different algorithms necessitates a combined application of complementary metrics. This typically involves assessing spatial quality (e.g., PSNR, SSIM), spectral fidelity (e.g., SAM), and overall datacube integrity (e.g., ERGAS, MRAE). By reporting a suite of these metrics, researchers can provide a holistic and objective assessment of an algorithm's performance, capturing its strengths and weaknesses across different aspects of reconstruction quality<sup>23-25</sup>.

Section 2.2 has established the theoretical foundation for spectral reconstruction algorithms based on CS principles. The analysis encompasses both theoretical foundations and practical performance characteristics, providing readers with the necessary knowledge to select appropriate reconstruction algorithms for specific application requirements. The subsequent chapters provide detailed analysis of these algorithmic categories, beginning with model-based methods in Chapter 3, followed by hybrid frameworks in Chapter 4, and learning-based approaches in Chapter 5. Figure 4 presents a comprehensive classification of spectral reconstruction algorithms, categorizing them into prior-constraint methods, hybrid constraints, and learning-based approaches, which provides a structural overview of the algorithmic landscape discussed in the following chapters.

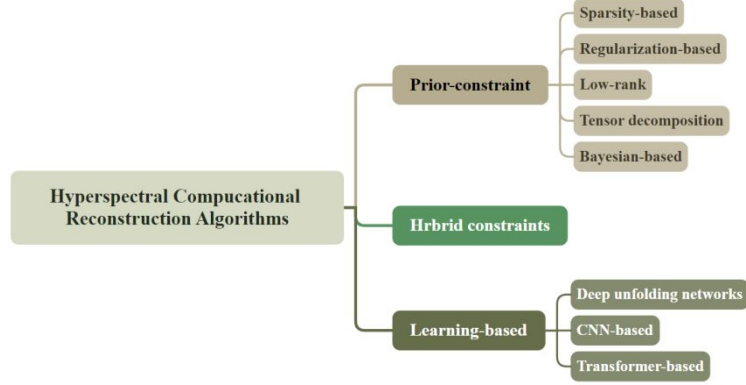


Fig. 4 Hyperspectral reconstruction algorithm classification.

### 3 Model-Based Reconstruction Algorithms

Prior constraint-based methods constitute a fundamental paradigm in computational hyperspectral reconstruction from spectral modulation measurements. These approaches address the inherently ill-posed inverse problem<sup>26</sup> by incorporating mathematical representations of hyperspectral data characteristics into the reconstruction framework. The main idea is that, although the number of measurements is much less than the data dimensions, we can still achieve accurate reconstruction by properly utilizing the inherent structural properties of the measurements.

Building upon the theoretical foundation and linear observation model established in Chapter 2, this chapter examines model-based reconstruction methods that incorporate explicit mathematical priors to address the inherently ill-posed inverse problem<sup>8</sup>. These approaches leverage mathematical representations of hyperspectral data characteristics, transforming the underdetermined system into well-posed optimization problems through regularization.

The reconstruction problem can be formulated within two equivalent optimization frameworks:

$$\min R(x) \quad \text{subject to} \quad \|Hx - y\|_2^2 \leq \varepsilon \quad (8)$$

$$\min \|Hx - y\|_2^2 + \mu R(x) \quad (9)$$

where  $\|Hx - y\|_2^2$  quantifies data fidelity,  $R(x)$  encodes prior constraints reflecting structural assumptions about hyperspectral data, and the parameters  $\varepsilon$  or  $\mu$  balance measurement consistency against prior enforcement. The choice of  $R(x)$  distinguishes different reconstruction methodologies and embodies distinct assumptions regarding hyperspectral data structure.

Hyperspectral imagery exhibits several exploitable characteristics that inform prior design. Spatial domains typically feature piecewise-smooth regions delineated by material boundaries, motivating gradient-based regularizers such as total variation. Spectral dimensions demonstrate strong inter-band correlations, manifesting as low-rank structures in appropriately matricized representations. Transform-domain sparsity emerges when hyperspectral data is projected onto suitable bases. The inherent three-dimensional structure naturally admits tensor decompositions that simultaneously capture spatial-spectral correlations. Furthermore, statistical properties enable probabilistic modeling within Bayesian frameworks that explicitly quantify reconstruction uncertainties.

These intrinsic properties have motivated five principal categories of prior-based reconstruction methods. Sparsity-based approaches exploit transform-domain representations, typically employing  $l_1$ -norm minimization. Spatial regularization methods primarily utilize total variation and its variants to enforce piecewise smoothness while preserving critical material boundaries. Low-rank methods capitalize on spectral redundancy through rank constraints applied to matricized data representations, exploiting correlation structures across spectral bands. Tensor decomposition approaches directly model the multidimensional structure, leveraging correlations across all modes simultaneously. Bayesian methods adopt a probabilistic perspective, encoding

prior knowledge through carefully designed probability distributions and deriving solutions via principled statistical inference.

The evolution of these methodologies reflects concurrent advances in optimization theory and deepening understanding of hyperspectral data characteristics. Early formulations employed classical regularizers such as Tikhonov penalties. The advent of compressed sensing theory catalyzed the development of sparsity-promoting algorithms, while breakthroughs in matrix completion enabled sophisticated low-rank approaches. Recent advances in multilinear algebra and hierarchical Bayesian modeling have further enriched the methodological landscape, enabling increasingly sophisticated reconstruction algorithms.

Although presented as distinct categories for clarity, these methods share foundational principles in regularized inverse problems and often exhibit complementary strengths. Contemporary state-of-the-art algorithms frequently integrate multiple prior types, creating hybrid formulations that simultaneously exploit diverse structural properties. The subsequent sections provide comprehensive analyses of each methodological category, detailing theoretical foundations, algorithmic implementations, and comparative performance characteristics.

### *3.1 Sparsity-Based Reconstruction Algorithms*

Sparsity-based methods constitute a cornerstone of hyperspectral reconstruction, exploiting the fundamental property that hyperspectral data admits sparse representations in appropriately chosen transform domains. These approaches derive their theoretical foundation from compressed sensing (CS) theory, which establishes that sparse signals can be accurately recovered from sub-Nyquist measurements, provided the sensing mechanism satisfies specific mathematical conditions.

A signal  $x \in \mathbb{R}^n$  is considered  $k$ -sparse if it contains at most  $k$  non-zero elements, where  $k \leq n$ .

While hyperspectral data rarely exhibits direct sparsity, it demonstrates remarkable sparsity when projected onto suitable transform bases:

$$x = \Psi \alpha \quad (10)$$

where  $\Psi$  represents a transform basis (e.g., wavelet, DCT, or learned dictionary) and  $\alpha$  is the sparse coefficient vector with  $\|\alpha\|_0 < k$ . Compressed sensing theory guarantees that when the measurement matrix  $H$  satisfies the restricted isometry property (RIP) or exhibits sufficient incoherence with  $\Psi$ , the original signal can be recovered by solving:

$$\min \|\alpha\|_0 \quad \text{subject to } y = H\Psi\alpha + n \quad (11)$$

Since  $\ell_0$ -norm minimization is NP-hard, practical algorithms employ the convex  $\ell_1$  relaxation:

$$\min \|\alpha\|_1 \quad \text{subject to } \|y - H\Psi\alpha\|_2 \leq \varepsilon \quad (12)$$

Under appropriate conditions on  $H$  and  $\Psi$ , this convex relaxation properly yields identical solutions to the original combinatorial problem.

Hyperspectral data manifests three distinct forms of sparsity that inform algorithm design. Spectral sparsity emerges because material reflectance spectra typically comprise linear combinations of few fundamental signatures determined by molecular composition and surface properties. Spatial sparsity arises from scene homogeneity—natural images contain piecewise-smooth regions with abrupt transitions at material boundaries, yielding sparse gradient representations. Joint spatial-spectral sparsity exploits the coupled nature of spatial and spectral

dimensions, where correlations across both domains create structured sparsity patterns in appropriate multidimensional transforms.

Building on these theoretical insights and the identified sparsity structures, the evolution of sparsity-based methods in hyperspectral compressive imaging represents a development from simple mathematical models to sophisticated data-driven approaches. Early methods relied on fixed analytical dictionaries like DCT and wavelets with standard  $\ell_1$  minimization, achieving compression at 40-50% sampling rates. While these pioneering approaches proved the viability of sparse representation, their reliance on predetermined bases limited their adaptability to diverse hyperspectral scenes.

This limitation sparked a paradigm shift toward dictionary learning methods, exemplified by K-SVD variants that learn optimal bases directly from data. By formulating the problem as basis pursuit denoising (BPDN), these methods achieved impressive reconstruction quality at 25% sampling rates<sup>27</sup>. However, the computational burden of dictionary learning and the challenge of capturing complex spectral-spatial correlations remained significant obstacles.

Researchers soon recognized that hyperspectral data possesses rich structure beyond simple sparsity. The introduction of spatial-spectral encoded approaches<sup>27</sup> marked a crucial insight: rather than treating spectral and spatial dimensions separately, joint encoding could exploit their inherent correlations. This realization led to innovative hardware designs, including filters with random transmittance patterns<sup>28</sup> that improved spectral resolution through stochastic sampling—a clever exploitation of randomness to overcome deterministic limitations.

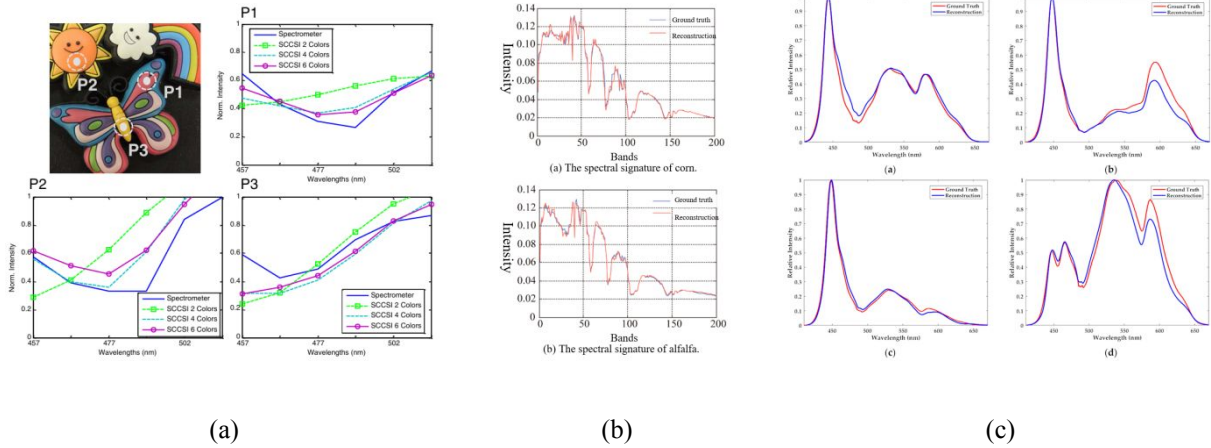
The field's maturation is evident in recent structured sparsity methods that elegantly combine multiple priors. Modern approaches like SSHBCS employ hierarchical dictionary structures with reweighted Laplacian regularization<sup>29</sup>, effectively balancing smoothness preservation with edge

sharpness. These methods demonstrate how theoretical advances in optimization—particularly maximum a posteriori frameworks—can be practically implemented to decompose complex joint reconstruction-dictionary learning problems<sup>30</sup>. Figure 5 illustrates the comparative performance of these sparse representation-based reconstruction algorithms, showing both reconstruction quality metrics and spectral fidelity for different materials like corn and alfalfa across various wavelength bands.

A particularly promising direction involves tensor-based formulations that naturally represent the three-dimensional structure of hyperspectral datacubes. By moving beyond matrix representations to embrace the data's inherent tensor structure, these methods achieve superior reconstruction through nonlinear compressed sensing<sup>31</sup>, though at the cost of increased computational complexity.

The ultimate test of these theoretical advances lies in their practical implementation. Snapshot imaging systems<sup>32</sup> and miniaturized spectrometers<sup>33</sup> represent the culmination of algorithmic sophistication meeting engineering constraints. These systems employ adaptive sparse coding strategies that must balance reconstruction quality against real-time processing requirements, demonstrating that the most elegant mathematical solutions must ultimately yield to practical considerations of speed, size, and power consumption.

Throughout this evolution, a clear trend emerges: the most successful methods are those that deeply understand and exploit the physical properties of hyperspectral data rather than treating it as generic high-dimensional signals. The progression from generic sparsity to structured, physically-informed models illustrates how domain knowledge increasingly drives algorithmic innovation in this field.



**Fig. 5** Comparison of hyperspectral compressive imaging reconstruction algorithms based on sparse representation. (a) Comparison of the spectral signature for the highlighted spatial points, measured with a commercial spectrometer and the SCCSI reconstructions using two, four, and six colors<sup>32</sup>. (b) Spectral Sparsity-Based Spectral Reconstruction and Curve Removal for Corn and Alfalfa<sup>34</sup>. (c) Comparison of Ground Truth and Reconstructed LED Spectra Using Dictionary Learning with L1-Norm Minimization<sup>33</sup>.

The methodological progression from predefined dictionaries to fully adaptive, structured sparsity models represents a fundamental paradigm shift in hyperspectral reconstruction philosophy. Modern approaches integrate multiple constraint types—including non-negativity, spatial smoothness, and spectral consistency—within unified optimization frameworks that better reflect the physical properties of hyperspectral data.

Despite remarkable advances, several fundamental challenges persist in sparsity-based hyperspectral reconstruction. Parameter sensitivity remains problematic, particularly in balancing multiple regularization terms within unified optimization frameworks. Dictionary learning approaches, while significantly improving representation quality, substantially increase computational complexity and memory requirements. The fundamental reliance on sparsity assumptions may prove limiting in scenarios involving complex spectral mixtures, high noise levels, or materials with intrinsically non-sparse spectral signatures.

Future research directions include developing robust, data-driven parameter selection strategies, exploring deep dictionary learning architectures that combine traditional sparse coding with neural network representations, extending theoretical guarantees to more realistic measurement models that account for noise correlation and system nonlinearities, and optimizing algorithms for embedded systems and real-time processing applications. Additionally, emerging research explores the integration of sparsity-based methods with complementary imaging modalities and the development of adaptive algorithms that automatically adjust sparsity assumptions based on scene characteristics.

### *3.2 Regularization-Based Reconstruction Algorithms*

Regularization-based methods address the ill-posed nature of hyperspectral reconstruction by incorporating structural priors into the optimization framework. These approaches transform underdetermined systems into well-posed optimization problems by imposing penalty terms that encode prior knowledge about desired solution properties, including spatial smoothness, edge preservation, and spectral coherence.

The general regularization framework follows the variational formulation:

$$\min_x \frac{1}{2} \|y - Hx\|_2^2 + \mu R(x) \quad (13)$$

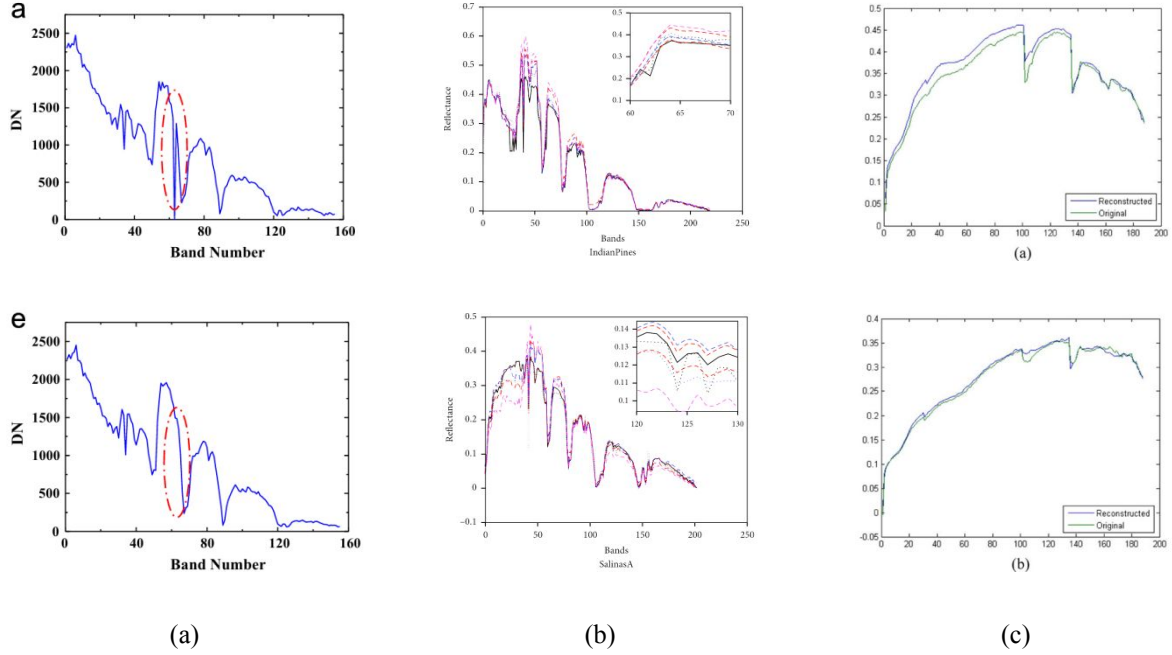
where the first term  $\|y - Hx\|_2^2$  ensures consistency with measurements,  $R(x)$  encodes structural priors through regularization functionals, and  $\mu$  governs the trade-off between measurement fitting and prior enforcement. Within the Bayesian framework, this formulation corresponds to maximum a posteriori (MAP) estimation, where  $R(x)$  represents the negative log-prior probability.

Total variation (TV)<sup>35</sup> regularization has emerged as the preeminent approach for hyperspectral reconstruction, predicated on the observation that natural images exhibit piecewise smoothness with sparse gradient distributions. The TV functional is defined as:

$$R_{TV}(x) = \sum_{i,j} \sqrt{(\nabla_x x_{i,j})^2 + (\nabla_y x_{i,j})^2} \quad (14)$$

This formulation promotes gradient sparsity while preserving discontinuities at material boundaries—a critical requirement for hyperspectral imagery where sharp spectral transitions delineate different materials.

Building on the success of standard TV regularization, researchers have developed several specialized variants to address specific reconstruction challenges. Anisotropic TV<sup>36</sup> decouples horizontal and vertical gradients, sacrificing rotational invariance for computational efficiency and simplified optimization<sup>37</sup>. Higher-order TV incorporates second-order derivatives to mitigate staircase artifacts and better preserve textural information<sup>38</sup>. Non-local TV<sup>39, 40</sup> extends the regularization beyond local neighborhoods, exploiting image self-similarity through patch-based comparisons to enhance reconstruction of repetitive structures. Weighted TV<sup>41</sup> introduces spatially-adaptive regularization strengths, preserving fine details in textured regions while strongly regularizing homogeneous areas. To leverage the complementary advantages of these variants, researchers have developed multi-TV collaboration approaches that combine different regularization strategies for enhanced reconstruction performance<sup>37</sup>. Figure 6 demonstrates the comparative effectiveness of these regularization-based reconstruction algorithms, showing spectral fidelity measurements and reconstruction accuracy across different band ranges, with close agreement between reconstructed and original hyperspectral signatures. These developments collectively establish TV-based regularization as a cornerstone methodology in hyperspectral compressed sensing reconstruction.



**Fig. 6** Comparison of hyperspectral compressive imaging reconstruction algorithms based on regularization. (a) Reconstruction results in Ref. 18. (b) Reconstruction results in Ref. 15. (c) Reconstruction results in Ref. 23.

Beyond TV-based approaches, the regularization landscape encompasses diverse methodologies exploiting different structural assumptions. Table 1 provides a comprehensive comparison of these regularization functionals, highlighting their mathematical formulations, key properties, computational complexities, and typical application domains. Tikhonov regularization<sup>42</sup> employs  $l_2$ -norm penalties to constrain solution energy, offering computational tractability at the expense of edge preservation. Graph-based regularization<sup>43</sup> constructs a dual-edge graph structure where vertices represent sparse coefficients and edges encode both intra-channel spatial similarity and inter-channel spectral correlation, preserving the joint spatial-spectral correlation structure of hyperspectral data through normalized Laplacian constraints. Spectral regularization specifically targets inter-band correlations, imposing smoothness or sparsity constraints along the spectral dimension to exploit the continuity of material reflectance spectra.

The three-dimensional nature of hyperspectral datacubes necessitates sophisticated multi-dimensional regularization strategies that respect the distinct statistical properties of spatial and spectral dimensions. Tensor TV<sup>44</sup> extends vectorial TV to third-order tensors, simultaneously regularizing spatial gradients and spectral variations through:

$$R_{TTV}(\mathcal{X}) = \sum_{i,j,k} \sqrt{\|\nabla_s \mathcal{X}_{i,j,k}\|_2^2 + |\nabla_\lambda \mathcal{X}_{i,j,k}|^2} \quad (15)$$

where  $\nabla_s$  and  $\nabla_\lambda$  denote spatial and spectral gradient operators, respectively. Hybrid approaches apply dimension-specific regularization<sup>37</sup>, combining spatial TV with spectral smoothness constraints to leverage the complementary nature of spatial and spectral priors. Adaptive frameworks dynamically adjust regularization parameters based on local data characteristics, enabling spatially-variant and spectrally-variant regularization that responds to scene complexity.

**Table 1** Comparison of Regularization Functionals.

Regularization Type	Mathematical Form	Key Properties	Computational Complexity	Typical Applications
Tikhonov <sup>45</sup>	$\ \nabla x\ _2^2$	Smooth solutions, easy optimization	$O(n)$	Denosing, smooth scenes
Total Variation <sup>40</sup>	$\ \nabla x\ _1 = \sum_i  \nabla_i x $	Edge preservation, gradient sparsity	$O(n \log n)$	Natural images, material boundaries
Non-local TV <sup>39</sup>	$\sum_{i,j} w_{i,j} \ x_i - x_j\ _2^2$	Texture preservation	$O(n^2)$	Repetitive structure
Graph-based <sup>46</sup>	$x^T Lx$	Manifold structure	$O(n^2)$	Spectral unmixing
Tensor TV <sup>44</sup>	$\ \nabla_s \mathcal{X}\ _1 + \ \nabla_{sp} \mathcal{X}\ _1$	3D structure preservation	$O(n^3)$	Full datacube reconstruction

Many regularization functionals, especially TV-based methods, are non-differentiable and require specialized optimization algorithms. Table 2 summarizes the key optimization algorithms employed for regularization-based reconstruction, comparing their convergence rates, memory requirements, and computational advantages. The Alternating Direction Method of Multipliers<sup>47</sup> (ADMM) has emerged as the workhorse algorithm, decomposing non-smooth problems into

sequences of simpler subproblems through variable splitting and augmented Lagrangian techniques. Proximal gradient methods<sup>48</sup>, including the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA), exploit the proximal operator structure of regularization terms while incorporating momentum-based acceleration. Split Bregman iteration<sup>49, 50</sup> provides efficient solutions for  $l_1$ -regularized problems by reformulating them as sequences of  $l_2$  subproblems. Primal-dual algorithms<sup>51</sup>, exemplified by the Chambolle-Pock method, simultaneously update primal and dual variables, avoiding computationally expensive matrix inversions while maintaining convergence guarantees.

**Table 2** Optimization Algorithms for Regularization.

<b>Regularization Type</b>	<b>Suitable for</b>	<b>Convergence Rate</b>	<b>Memory Requirements</b>	<b>Key Advantages</b>
ADMM <sup>47</sup>	Non-smooth, variable splitting	$O(1/k)$	Moderate	Parallelizable
FISTA <sup>52</sup>	Proximal operators	$O(1/k^2)$	Low	Fast convergence
Split Bregman <sup>49</sup>	$l_1$ -problems	$O(1/k)$	Moderate	Efficient for TV
Chambolle-Pock <sup>53</sup>	Primal-dual	$O(1/k)$	Low	No matrix inversion

Recent advances have introduced several algorithmic innovations in regularization-based reconstruction. The spatial-spectral non-local means regularization framework proposed by Meza et al.<sup>39</sup> exploits multi-scale self-similarity through split Bregman optimization, achieving superior noise suppression while preserving fine spectral details. The SAMNLTV algorithm extends non-local regularization to multiple dimensions, constructing joint spatial-spectral gradient operators that capture cross-dimensional correlations through alternating minimization schemes<sup>54</sup>.

A particularly noteworthy advance is the extension of Generalized Alternating Projection (GAP) algorithms<sup>55</sup> from transform domains to direct spatial domain TV minimization. This reformulation eliminates the complexity of basis selection and coefficient grouping inherent in

traditional wavelet or DCT-based GAP implementations, enabling direct optimization in the signal space while maintaining theoretical convergence guarantees.

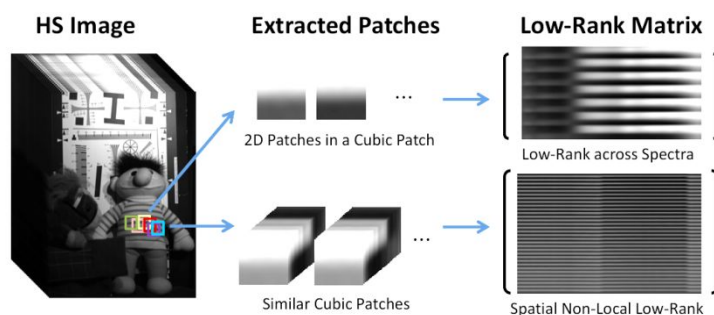
The spectral layer adaptive reconstruction method introduces automated parameter selection through TV subgradient boundedness theory, achieving up to 305% computational acceleration via the PANCoS algorithm while maintaining reconstruction fidelity<sup>56</sup>. This approach exemplifies the trend toward self-tuning regularization frameworks that reduce the burden of manual parameter selection.

Regularization-based methods offer meaningful advantages: strong theoretical foundations with convergence guarantees, effective noise suppression capabilities, interpretable results aligned with physical image properties, and robustness to measurement degradation. However, significant challenges remain. Parameter sensitivity necessitates careful tuning of regularization weights, often requiring problem-specific expertise. Certain regularizers may inadequately capture complex structural patterns or introduce characteristic artifacts—TV's staircase effect being a prime example. Computational demands scale unfavorably with data dimensionality, limiting applicability to large-scale hyperspectral datasets.

Current research directions address these limitations through data-adaptive regularization schemes that learn parameters or even functional forms from the data itself. The integration of regularization principles with deep learning architecture represents a particularly promising avenue, combining the theoretical rigor and interpretability of variational methods with the representational power of neural networks. These hybrid approaches may ultimately provide the flexibility and efficiency required for next-generation hyperspectral imaging systems.

### 3.3 Low-rank Reconstruction Algorithms

Low-rank methods constitute a fundamental paradigm in hyperspectral reconstruction, exploiting the intrinsic spectral correlations and redundancies that characterize hyperspectral data. These approaches are predicated on the observation that hyperspectral datacubes, when appropriately matricized, exhibit pronounced low-rank structures arising from the limited diversity of material spectra and strong inter-band correlations within natural scenes. As illustrated in Fig. 7, this low-rank property can be exploited through patch extraction strategies, where similar cubic patches from the hyperspectral image are organized into matrices that demonstrate clear low-rank characteristics both across spectral dimensions and through spatial non-local similarities.



**Fig. 7** Low-rank matrix formation from hyperspectral images via patch extraction, showing both spectral and spatial non-local low-rank characteristics utilized in reconstruction algorithms<sup>57</sup>.

The mathematical foundation rests on the insight that a three-dimensional hyperspectral cube  $\chi \in \mathbb{R}^{n_x \times n_y \times n_\lambda}$  can be reshaped into a matrix  $\mathbf{X} \in \mathbb{R}^{n_s \times n_\lambda}$  (where  $n_s = n_x \times n_y$ ), which admits a low-rank approximation despite its high ambient dimension. This property emerges from the physical constraint that material reflectance spectra are determined by molecular composition, resulting in spectral signatures representable as linear combinations of relatively few fundamental endmember spectra.

The canonical low-rank reconstruction problem seeks to minimize matrix rank while ensuring measurement consistency:

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad \text{subject to} \quad \|\mathbf{H}\text{vec}(\mathbf{X}) - \mathbf{y}\|^2 \leq \varepsilon \quad (16)$$

where  $\mathbf{H}$  encodes the measurement process,  $\mathbf{y}$  represents acquired measurements, and  $\varepsilon$  bounds the noise level. Since rank minimization is NP-hard, practical algorithms employ the nuclear norm as a convex surrogate:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathbf{H}\text{vec}(\mathbf{X}) - \mathbf{y}\|^2 \leq \varepsilon \quad (17)$$

where  $\|\mathbf{X}\|_* = \sum_i \sigma_i(\mathbf{X})$  denotes the sum of singular values. This relaxation enjoys strong theoretical guarantees under restricted isometry conditions, paralleling the  $\| \cdot \|_1 - \| \cdot \|_0$  relationship in compressed sensing.

**Table 3** Comparison of Low-rank Method Variants.

Method Type	Formulation	Advantages	Limitations	Computational Complexity
Global Low-rank <sup>58</sup>	$\min \ X\ _* + \lambda \ HX - y\ _2^2$	Simple, efficient, strong theory	May oversmooth local features	$O(mnp)$ per iteration
Local Low-rank <sup>59</sup>	$\min \sum \ X\ _*$	Preserves local variations	Blocking artifacts, parameter tuning	$O(Nmnr)$ , $N = \#$ patch
Weighted Nuclear <sup>60</sup>	$\min \ X\ _{w,*}$	Adaptive regularization	Weight selection complexity	$O(mnp)$ + weight computation
Tensor-based <sup>61</sup>	Tensor nuclear form	Preserves 3D structure	Higher complexity	$O(mnp^2)$ or worse

The low-rank framework has spawned diverse methodological variants tailored to different aspects of hyperspectral reconstruction. Table 3 provides a comparative analysis of these low-rank method variants, outlining their formulations, advantages, limitations, and computational complexities. Global Low-rank Methods operate on the entire matricized datacube, imposing a single low-rank constraint that captures scene-wide spectral correlations. While computationally

efficient for moderate-sized problems, these methods may oversimplify heterogeneous scenes containing diverse materials. Local Low-rank Methods<sup>59</sup> partition the hyperspectral cube into overlapping spatial patches, applying low-rank constraints to each patch independently before aggregating results. This approach better captures spatially-varying spectral characteristics at the cost of increased computational complexity and potential blocking artifacts. Structured Low-rank Approaches augment basic low-rank constraints with additional structural priors. These include spectral smoothness regularization that enforces continuity in reflectance spectra, and spatial coherence terms that promote consistency among neighboring pixels sharing similar materials. Weighted Nuclear Norm Methods replace uniform singular value penalization with adaptive weighting schemes, enabling differential treatment of signal and noise subspaces. The weighted nuclear norm  $\|\mathbf{X}\|_{w,*} = \sum_i w_i \sigma_i(\mathbf{X})$  allows preservation of significant spectral components while aggressively suppressing noise-dominated dimensions.

The non-smooth nuclear norm minimization requires specialized algorithms centered on the Singular Value Thresholding (SVT) operator:

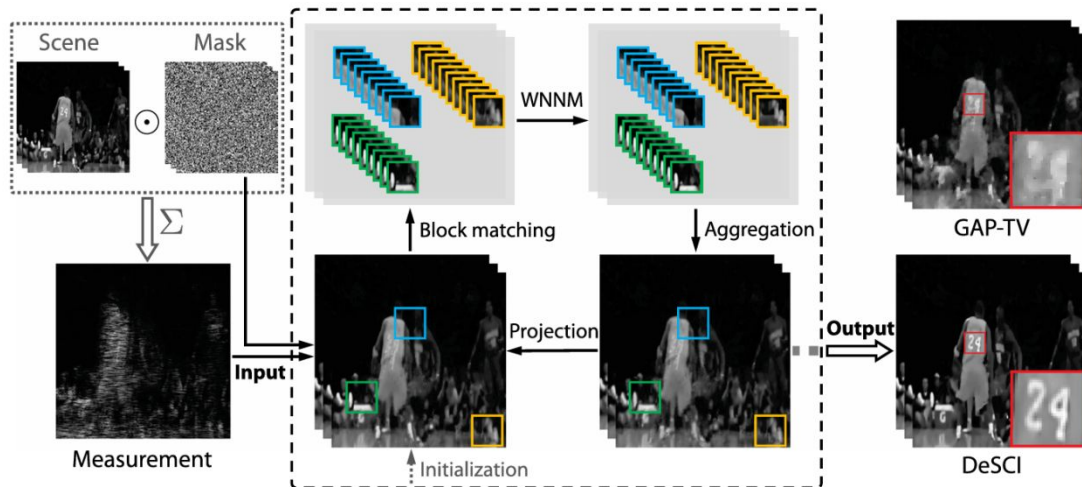
$$\text{SVT}_\tau(\mathbf{X}) = \mathbf{U} \text{diag}(\max(\boldsymbol{\sigma} - \tau, 0)) \mathbf{V}^T \quad (18)$$

where  $\mathbf{X} = \mathbf{U} \text{diag}(\boldsymbol{\sigma}) \mathbf{V}^T$  is the singular value decomposition.

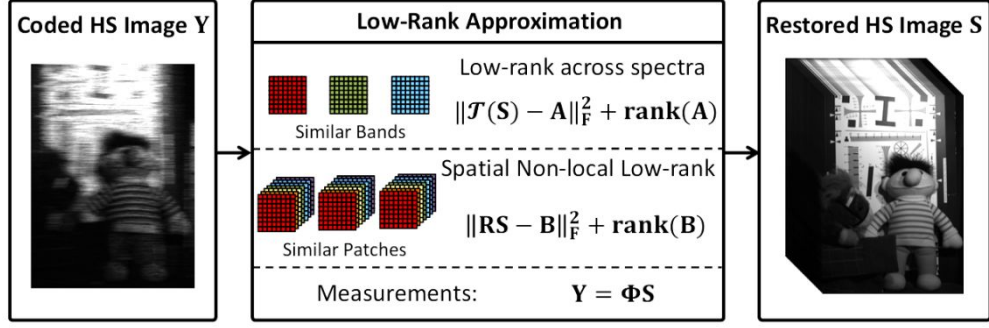
Augmented Lagrangian methods, particularly ADMM formulations, decompose the problem into alternating X-minimization and multiplier update steps, each admitting closed-form solutions via SVT. Accelerated proximal gradient algorithms adapt FISTA to matrix variables, achieving  $O(1/k^2)$  convergence rates through momentum-based acceleration. For large-scale problems,

randomized SVD approximations reduce computational complexity from  $O(mn \min(m, n))$  to  $O(mnr)$ , where  $r$  is the target rank.

Contemporary research has yielded sophisticated low-rank reconstruction algorithms addressing specific challenges in computational spectral imaging. The Weighted Nuclear Norm Minimization (WNNM) framework adaptively shrinks singular values through self-learned weights, demonstrating superior preservation of spectral details compared to uniform penalization<sup>62</sup>. As illustrated in Figure 8, the WNNM-based reconstruction algorithm employs an iterative process involving block matching, weighted nuclear norm minimization, and aggregation to achieve enhanced reconstruction performance. Dual low-rank regularization approaches simultaneously exploit spectral correlations through global matricization and spatial self-similarity via local patch grouping, achieving state-of-the-art reconstruction quality<sup>57</sup>. Figure 9 demonstrates this dual strategy, where low-rank approximations are applied both across spectral dimensions using similar bands and spatially through non-local patch similarities.



**Fig. 8** Dual low-rank approximation framework demonstrating spectral band correlation and spatial non-local patch similarity exploitation in hyperspectral image reconstruction<sup>62</sup>.



**Fig. 9** The flowchart of dual low-rank reconstruction algorithm for coded hyperspectral image restoration by using spectral low-rank and spatial non-local lowrank regularizations<sup>57</sup>.

Multi-level reconstruction frameworks employ gradient-boosted sampling strategies to enhance prior learning across different frequency bands, utilizing local low-rank properties within coefficient matrices for targeted denoising<sup>59</sup>. The Low-rank Matrix Recovery (LRMR) approach, grounded in linear spectral mixing theory, decomposes observations as  $\mathbf{D} = \mathbf{L} + \mathbf{S} + \mathbf{N}$ , jointly estimating low-rank clean signals, sparse corruptions, and Gaussian noise through principled optimization<sup>63</sup>.

Low-rank methods offer notable advantages: effective global correlation modeling, strong theoretical foundations in matrix completion theory, robustness to missing data and noise, and favorable sampling complexity compared to pixel-wise approaches. The framework naturally aligns with the physical properties of hyperspectral data, where spectral redundancy is inherent to material properties.

However, significant challenges persist in practical applications. Global low-rank constraints may over-smooth unique spectral signatures critical for material discrimination. The matricization process disrupts spatial relationships, potentially degrading edge preservation. Computational scaling remains problematic for high-resolution datasets, with SVD complexity limiting real-time

applications. Rank selection and regularization parameter tuning require careful calibration, often demanding problem-specific expertise.

Current research directions address these limitations through adaptive rank estimation schemes that dynamically adjust model complexity based on local data characteristics. The integration of low-rank priors within deep learning architectures represents a particularly promising avenue, where neural networks implicitly learn low-dimensional representations while maintaining the flexibility to capture nonlinear spectral mixing phenomena. These hybrid approaches may ultimately provide the adaptability and computational efficiency required for next-generation hyperspectral imaging systems.

### *3.4 Tensor-Based Reconstruction Algorithms*

Tensor decomposition methods represent an advanced paradigm for hyperspectral reconstruction that preserves and exploits the intrinsic multi-dimensional structure of hyperspectral data. Unlike matrix-based approaches requiring destructive reshaping operations, tensor methods maintain the native three-dimensional organization of hyperspectral cubes, treating them as third-order tensors  $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , where  $I_1$ ,  $I_2$  denote spatial dimensions and  $I_3$  represents the spectral dimension. This natural representation enables simultaneous modeling of correlations across all modes, providing a mathematically principled framework for capturing the complex interdependencies characterizing hyperspectral imagery.

While matrix-based approaches such as sparse representation and low-rank matrix recovery have proven effective for hyperspectral reconstruction, they require reshaping the three-dimensional data cube into two-dimensional matrices, inevitably disrupting the intrinsic multi-dimensional correlations. This vectorization process decouples spatial-spectral relationships, treats

neighboring pixels independently, and increases the ambiguity of the inverse problem. In contrast, tensor methods preserve the native three-dimensional structure throughout processing, enabling simultaneous modeling of correlations across all modes. This structural preservation becomes particularly advantageous under severe undersampling conditions (below 10% sampling rate), where tensor methods typically achieve 2-4 dB higher PSNR than matrix approaches. Furthermore, tensor frameworks naturally accommodate multiple priors—combining low-rank constraints with sparsity or smoothness regularization—and excel in scenarios involving spatially varying spectral signatures, non-uniform sampling patterns, and applications requiring accurate spectral signature preservation. The choice between tensor and matrix methods ultimately depends on the specific reconstruction requirements, with tensor approaches offering superior performance when maintaining multi-dimensional structure is critical for the application.

The mathematical underpinnings of tensor decomposition draw from multilinear algebra, with several canonical models forming the basis for reconstruction algorithms:

CANDECOMP/PARAFAC (CP) Decomposition<sup>64</sup> approximates the tensor as a sum of rank-one components(Fig. 10 (a)):

$$\mathbf{X} \approx \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \quad (19)$$

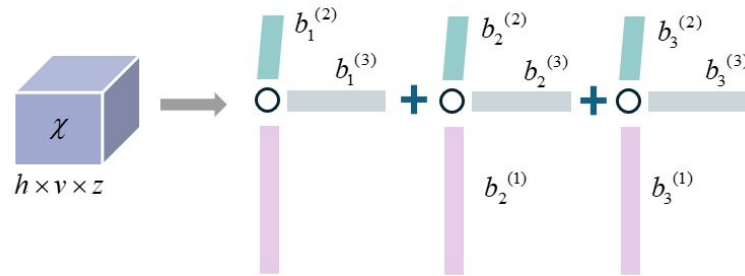
where  $\mathbf{a}_r, \mathbf{b}_r \in \mathbb{R}^{j_1, j_2}$  capture spatial distributions,  $\mathbf{c}_r \in \mathbb{R}^{j_3}$  represents spectral signatures, and  $\circ$  denotes the outer product. This decomposition directly models the linear mixing of endmember spectra, aligning with physical imaging processes.

Tucker Decomposition<sup>65</sup> provides a higher-order generalization of matrix SVD(Fig. 10 (b)):

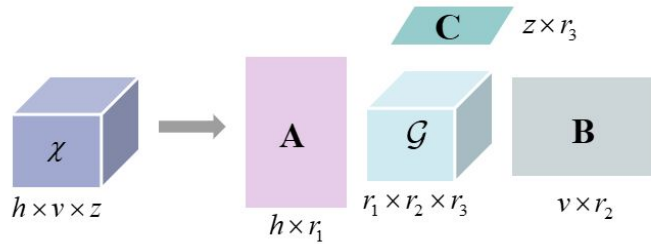
$$\mathbf{X} \approx \mathbf{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \quad (20)$$

where  $\mathbf{G} \in \mathbb{J}^{R_1 \times R_2 \times R_3}$  is the core tensor encoding component interactions, and  $\mathbf{A} \in \mathbb{J}^{I_1 \times R_1}$ ,  $\mathbf{B} \in \mathbb{J}^{I_2 \times R_2}$ ,  $\mathbf{C} \in \mathbb{J}^{I_3 \times R_3}$  are factor matrices. This model offers greater expressiveness for complex spatial-spectral interactions at the cost of increased parametric complexity. The Tucker model has inspired numerous algorithmic innovations, including structured sparse variants where the core tensor  $\mathbf{G}$  is constrained to be sparse ( $\|\mathbf{G}\|_1 \leq \epsilon$ ), enabling automatic rank adaptation and improved reconstruction under limited measurements<sup>66</sup>.

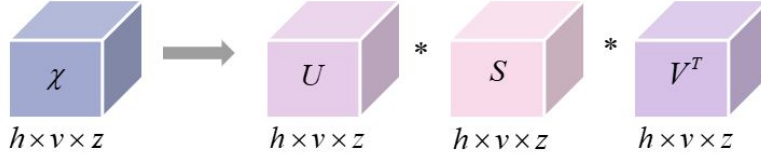
Tensor-SVD (t-SVD)<sup>30</sup> introduces a novel tensor-tensor product based on circular convolution (Fig. 10 (c)), yielding decompositions particularly suited to preserving spectral continuity. Tensor Train (TT) decomposition represents high-order tensors as chains of three-dimensional cores, providing scalable representations for large-scale hyperspectral data.



(a)



(b)



(c)

**Fig. 10** Tensor Decomposition of a Third-Order Tensor<sup>67</sup>. (a) CP decomposition. (b) Tucker decomposition. (c) t-SVD.

Tensor-based reconstruction problems typically adopt the variational formulation:

$$\min_{\mathbf{X}} R(\mathbf{X}) \quad \text{subject to} \quad \|H(\mathbf{X}) - \mathbf{y}\| \leq \delta \quad (21)$$

where  $R(\mathbf{X})$  promotes low-rank tensor structure,  $H$  represents the measurement operator, and

$\delta$  models noise. The specific form of  $R(\mathbf{X})$  depends on the chosen decomposition:

- For CP:  $R(\mathbf{X}) = \min_R \sum_{i=1}^R \|\mathbf{a}_i\|_2 \|\mathbf{b}_i\|_2 \|\mathbf{c}_i\|_2$  subject to the CP constraint;
- For Tucker:  $R(\mathbf{X}) = \sum_{n=1}^3 \alpha_n \|\mathbf{X}_{(n)}\|_*$  where  $\mathbf{X}_{(n)}$  denotes mode- $n$  unfolding;
- For t-SVD:  $R(\mathbf{X}) = \|\mathbf{X}\|_{TNN}$ , the tensor nuclear norm;

**Table 4** Computational complexity comparison of tensor decomposition methods.

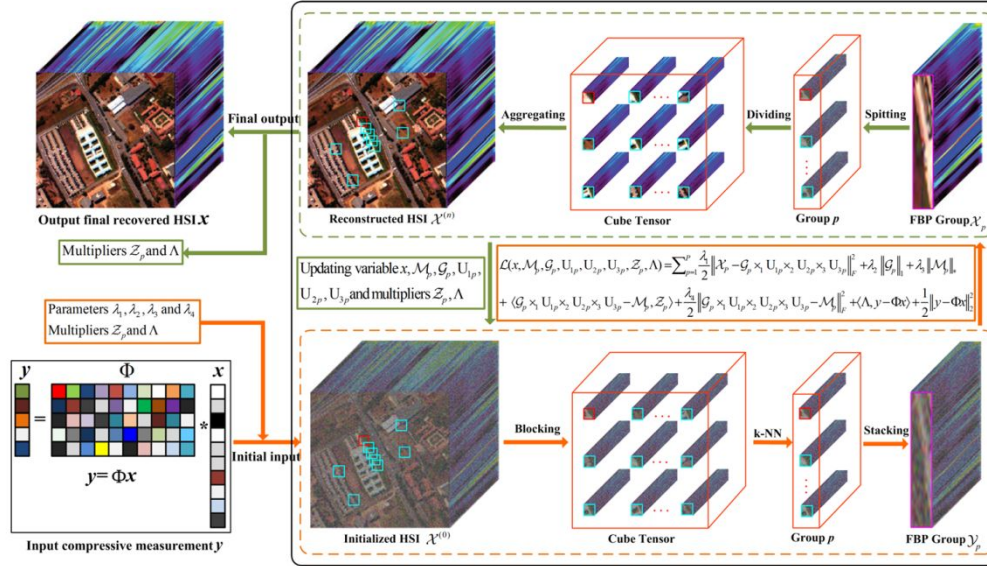
Method	Time Complexity	Space Complexity	Parallelizable
Matrix SVD <sup>68</sup>	$O(mnp \cdot \min(m, np))$	$O(mnp)$	Moderate
CP <sup>69</sup>	$O(RI \ mnp)$	$O(R(m+n+p))$	High
Tucker <sup>70</sup>	$O(mnpR^3 + R^3(m+n+p))$	$O(R^3 + R^3(m+n+p))$	Moderate
t-SVD <sup>51</sup>	$O(mnp \log p)$	$O(mnp)$	High
Tensor Train <sup>69</sup>	$O(R^2 mnp)$	$O(R^2(m+n+p))$	Moderate
Tensor Ring <sup>71</sup>	$O(R^3 mnp)$	$O(R^2(m+n+p))$	High

The computational efficiency of tensor decomposition methods varies significantly with algorithm choice and data structure. Table 4 provides a detailed comparison of the computational complexities for major tensor decomposition methods, including their time and space requirements and parallelization capabilities. Matrix SVD, requiring  $O(mnp \cdot \min(m, np))$  operations, serves as a baseline but scales poorly with data size and destroys the multi-dimensional structure of hyperspectral cubes. Among tensor methods, CP decomposition offers the best efficiency with  $O(RI mnp)$  complexity, where  $I$  denotes the number of iterations and  $R$  is the rank. Despite its apparent higher complexity, CP's linear scaling in rank and high parallelizability often result in fast practical implementation. Tucker decomposition's  $O(mnpR^3 + R^3(m + n + p))$  complexity becomes prohibitive for ranks above 10-20 due to cubic scaling, limiting its use despite superior modeling flexibility. The t-SVD method achieves  $O(mnp \log p)$  complexity through FFT operations, providing excellent performance for hyperspectral data with hundreds of spectral bands. Tensor Train offers balanced  $O(R^2 mnp)$  scaling suitable for large-scale problems, while Tensor Ring's  $O(R^3 mnp)$  complexity restricts it to small-rank applications. For typical hyperspectral datasets ( $256 \times 256 \times 200$  voxels), practical runtime differences are substantial: CP with rank 20 requires  $\sim 10^7$  operations per iteration versus  $\sim 10^9$  for Tucker decomposition. Modern acceleration techniques including GPU parallelization and randomized algorithms can reduce computation time by 1-2 orders of magnitude. The optimal method selection depends on the specific trade-off between reconstruction accuracy, computational resources, and the target application's real-time requirements.

Variable splitting strategies often reformulate these problems for computational tractability:

$$\min_{\mathbf{X}, \mathbf{Z}} R(\mathbf{Z}) + \frac{\mu}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2 \quad \text{subject to} \quad \mathbf{y} = H(\mathbf{X}) + \mathbf{d} \quad (22)$$

The non-convexity of tensor rank minimization necessitates specialized algorithms. Alternating optimization schemes decompose the problem into manageable subproblems: CP-ALS cyclically updates factor matrices, while Tucker-HOOI generalizes power iteration to tensor spaces. Convex relaxation approaches employ sum of nuclear norms (SNN) as tractable surrogates, enabling principled optimization with convergence guarantees. Accelerated algorithms adapt FISTA to tensor domains, achieving  $O(1/k^2)$  convergence through momentum-based updates. Practical implementations often employ ADMM for handling the non-convex constraints. ADMM decomposes complex optimization problems into tractable subproblems through auxiliary variable introduction and augmented Lagrangian formulations, as demonstrated in methods like NTSRLR<sup>66</sup> where the tensor sparse and low-rank constraints are handled separately. Figure 11 illustrates the NTSRLR algorithm workflow, showing how the method alternates between tensor blocking/aggregating operations and singular value thresholding (SVT) to simultaneously enforce sparsity and low-rank structure in the reconstruction process.

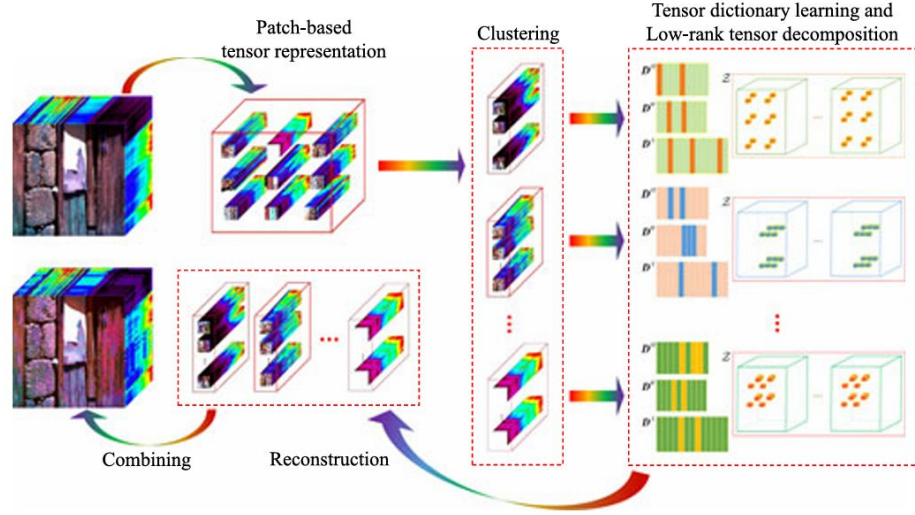


**Fig. 11** The flowchart of the NTSRLR algorithms<sup>66</sup>, which consists of two steps: sensing and reconstruction. First, it acquires the compressive measurement  $y$  by a random sampling matrix  $H$ . Second, NTSRLR recovers an HSI from the measurements  $y=Hx$ .

Contemporary tensor-based reconstruction algorithms demonstrate increasingly sophisticated integration of structural priors with computational efficiency. The Collaborative Tucker3 framework exemplifies patch-based strategies, segmenting hyperspectral cubes into overlapping volumetric blocks that are reorganized into fourth-order tensors capturing non-local similarities<sup>70</sup>. By applying Tucker decomposition to these augmented tensors, the method achieves superior reconstruction under extreme compression ratios while preserving spectral fidelity through collaborative filtering across similar patches. Low-rank Tensor Recovery (LRTR) methods formulate reconstruction directly as tensor completion problems, employing tensor nuclear norm regularization  $\min_x \|X\|_{TNN} + \frac{\mu}{2} \|P_\Omega(X) - Y\|_F^2$ , where  $P_\Omega$  represents the sampling operator<sup>72</sup>.

The key innovation lies in efficient t-SVD computation through FFT transformations, reducing

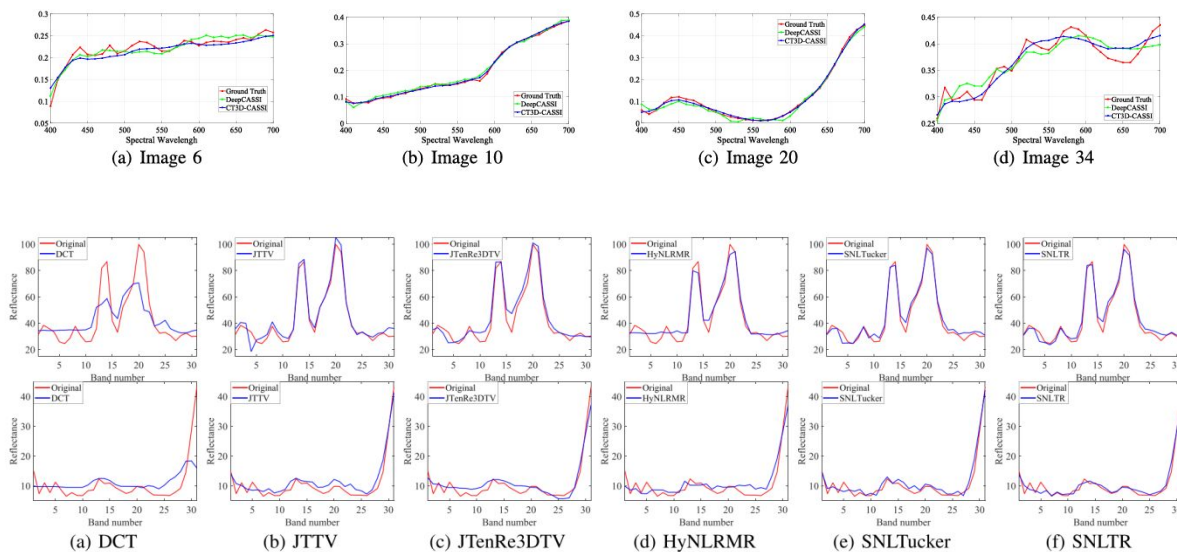
complexity from  $O(I_1 I_2 I_3^2)$  to  $O(I_1 I_2 I_3 \log I_3)$  and enabling scalable processing of large-scale datacubes.



**Fig. 12** The flowchart of the PLTD algorithm<sup>73</sup>, illustrating patch-based tensor representation, nonlocal similarity clustering, and low-rank tensor decomposition.

More advanced approaches integrate multiple structural assumptions within unified frameworks. The Patch-based Low-rank Tensor Decomposition<sup>73</sup> (PLTD) algorithm combines local and non-local processing by first clustering similar patches via K-means, then constructing separate fourth-order tensors for each cluster. As illustrated in Fig. 12, the PLTD workflow systematically progresses from patch-based tensor representation through clustering to tensor dictionary learning and low-rank decomposition for each cluster group. Adaptive Tucker decomposition applied to each cluster tensor preserves local spatial details while exploiting both intra-patch and inter-patch correlations, with final reconstruction achieved through weighted aggregation of overlapping patches. The state-of-the-art Subspace-based Nonlocal Tensor Ring<sup>71</sup> (SNTR) decomposition replaces traditional Tucker models with tensor ring representations, expressing each frontal slice as products of latent cores:  $\mathbf{X}(:, :, k) = \mathbf{G}_1^{(k)} \times \mathbf{G}_2^{(k)} \times \mathbf{L} \times \mathbf{G}_N^{(k)}$ . This

formulation offers enhanced parameter efficiency  $O(I_1I_2R)$  vs.  $O(I_1I_2I_3R)$  for Tucker, invariance to mode permutations, and superior handling of irregular sampling patterns. Figure 13 demonstrates the practical effectiveness of these tensor decomposition approaches, showing convergence behavior across different test images and detailed spectral reconstruction comparisons between methods. Comprehensive evaluations across diverse datasets consistently demonstrate 3-5 dB PSNR improvements over matrix-based methods, particularly in scenarios with sampling rates below 10%, validating the fundamental advantage of preserving multi-dimensional structure throughout the reconstruction process.



**Fig. 13** Comparison of hyperspectral compressive imaging reconstruction algorithms based on tensor decomposition<sup>71</sup>.

Tensor decomposition methods offer compelling advantages for hyperspectral reconstruction. They preserve multi-dimensional correlations throughout processing, provide compact yet expressive representations with clear physical interpretability, and demonstrate robust performance under severe undersampling. The framework naturally accommodates diverse constraints reflecting domain knowledge, from spectral smoothness to spatial coherence.

However, practical implementation faces challenges. Computational complexity scales with tensor dimensions, requiring careful algorithm selection based on datacube size. Non-convexity introduces sensitivity to initialization, necessitating robust strategies like multi-start optimization. Rank selection remains problem-specific, though recent adaptive schemes show promise. The mathematical sophistication may impede adoption by practitioners without multilinear algebra expertise.

Current research addresses these limitations through several promising directions. Neural tensor networks integrate decomposition within deep architectures, learning optimal representations end-to-end. Automated rank selection employs reinforcement learning or meta-learning to adapt ranks dynamically. Quantum-inspired tensor networks borrow concepts from quantum physics to develop ultra-efficient decomposition. These advances promise to unlock the full potential of tensor methods for next-generation hyperspectral imaging systems.

#### **4 Hybrid and Advanced Reconstruction Frameworks**

The model-based reconstruction algorithms presented in Chapter 3 have demonstrated remarkable success in exploiting specific structural properties of hyperspectral data. Sparsity-based methods (Section 3.1) effectively capture transform-domain representations, regularization approaches (Section 3.2) preserve spatial smoothness while maintaining edge fidelity, low-rank methods (Section 3.3) exploit spectral correlations, and tensor-based algorithms (Section 3.4) maintain the intrinsic multi-dimensional structure. However, real-world hyperspectral imaging scenarios present complexities that often exceed the modeling capacity of any single prior constraint.

Consider a typical remote sensing scene containing urban areas, vegetation, water bodies, and bare soil. Such scenes simultaneously exhibit sparse spectral signatures from man-made materials, smooth spatial variations within homogeneous regions, low-rank structure from natural material

spectra, and complex tensor correlations across all dimensions. No single mathematical model can adequately capture this diversity of structural properties. Furthermore, each individual method faces inherent limitations: sparsity-based approaches may struggle with dense spectral mixtures, regularization methods can over-smooth fine spectral details, low-rank constraints may fail to preserve unique spectral signatures, and tensor methods suffer from computational complexity at scale.

These fundamental limitations have motivated a paradigm shift toward hybrid reconstruction frameworks that synergistically combine multiple complementary priors. The evolution from single-constraint to multi-constraint optimization represents more than a simple aggregation of methods—it reflects a deeper understanding of hyperspectral data complexity and the need for adaptive, context-aware reconstruction strategies. By leveraging the collective strengths of different approaches while mitigating individual weaknesses, hybrid methods achieve superior reconstruction performance across diverse imaging scenarios.

The general hybrid reconstruction problem can be formulated as a multi-objective optimization:

$$\min_x \frac{1}{2} \|Hx - y\|_2^2 + \sum_{i=1}^N \mu_i R_i(x) \quad (23)$$

where  $\frac{1}{2} \|Hx - y\|_2^2$  enforces measurement consistency, and each  $R_i(x)$  imposes a distinct structural constraint weighted by  $\mu_i$ . This formulation encompasses diverse hybrid strategies, from straightforward dual-constraint combinations to sophisticated multi-prior frameworks operating across different transform domains and scales. The key challenge lies not merely in combining constraints, but in designing optimization algorithms that effectively balance competing objectives while maintaining computational tractability.

This chapter explores the theoretical foundations and practical implementations of hybrid reconstruction frameworks. We begin by examining the mathematical formulation of multi-constraint optimization problems and the computational challenges they present. We then categorize hybrid methods based on their constraint combinations: sparsity-regularization hybrids that balance transform-domain sparsity with spatial smoothness, sparsity-low-rank decompositions that separate structured backgrounds from sparse anomalies, low-rank-regularization approaches that combine global spectral modeling with local structure preservation, and tensor-enhanced hybrids that augment multi-dimensional decomposition with additional constraints. Throughout our analysis, we emphasize how these hybrid frameworks address the specific limitations of their constituent methods while introducing new challenges in parameter selection and computational efficiency.

Hybrid prior methods combine multiple constraints to address different aspects of hyperspectral data structure. This section classifies these methods based on the types of priors being integrated. The classification includes four main categories: (1) Sparsity-Regularization Hybrids combine transform-domain sparsity with spatial smoothness constraints; (2) Sparsity-Low-rank Decomposition models data as structured background plus sparse anomalies; (3) Low-rank-Regularization Approaches integrate spectral correlation modeling with spatial structure preservation; (4) Tensor-Enhanced Hybrids augment tensor decomposition with additional constraints. Methods can be further divided into single-modal approaches (using only hyperspectral data) and multi-modal fusion approaches (incorporating external data sources). Table 5 provides a comprehensive comparison of these hybrid reconstruction algorithms, summarizing their mathematical formulations, main advantages, key limitations, and

computational characteristics. This framework provides a basis for understanding method characteristics and selecting appropriate algorithms for specific reconstruction tasks.

**Table 5** Comparison of hybrid reconstruction algorithms for spectral modulation-based hyperspectral imaging

Method type	Main Priors Combined	Mathematical Formulation	Advantages	Limitations
Sparsity-Regularization <sup>74, 75</sup>	Sparsity + Spatial Smoothness	$l_1$ +TV	Preserves edges, robust to noise	Parameter selection, may over-smooth
Sparsity-Low-rank <sup>72</sup>	Sparsity + Low-rank	$l_1$ +Nuclear Norm	Separates background/target, anomaly det	Sensitive to rank/sparsity, complex scenes
Low-rank-Regularization <sup>32</sup>	Low-rank + Total Variation	Nuclear Norm+TV	Global consistency, local smoothness	May lose fine details, computation
Tensor-Enhanced Hybrid <sup>70</sup>	Tensor Decomposition + Sparsity/TV	Tensor Norm+ $l_1$ /TV	Multi-dimensional structure, detail pres	High computation, parameter tuning
Multi-modal Fusion <sup>1</sup>	HSI + MSI/LiDAR, etc.	Coupled Subspace/Sparsity	Use extra info, improves robustness	Needs extra data, model design complexity

Sparsity-Regularization Hybrids<sup>74, 75</sup> merge transform-domain sparsity with spatial smoothness constraints, typically formulated as:

$$\min_{\mathbf{x}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \mu_1 \|\Psi\mathbf{x}\|_1 + \mu_2 \mathbf{TV}(\mathbf{x}) \quad (24)$$

This combination captures both localized spectral features through wavelet or learned dictionaries while preserving spatial coherence through total variation regularization. The dual constraints prove particularly effective for scenes containing both smooth regions and sharp material boundaries with distinct spectral signatures.

Sparsity-Low-rank Decompositions model<sup>72</sup> hyperspectral data as superpositions of structured and anomalous components:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{A}(\mathbf{L} + \mathbf{S}) - \mathbf{y}\|_2^2 + \mu_1 \|\mathbf{L}\|_* + \mu_2 \|\mathbf{S}\|_1 \quad (25)$$

where  $\mathbf{L}$  captures the low-rank background corresponding to dominant materials, while  $\mathbf{S}$  represents sparse anomalies or rare signatures. This separation strategy proves especially valuable for anomaly detection and target identification applications. The approach assumes that background materials exhibit low-rank spectral behavior, which may not hold in highly complex or heterogeneous scenes. Despite this limitation, the method provides effective automatic background-foreground separation capabilities.

Low-rank-Regularization Approaches<sup>32</sup> simultaneously exploit spectral correlation and spatial structure:

$$\min_{\mathbf{X}} \|\mathbf{H}\mathbf{vec}(\mathbf{X}) - \mathbf{y}\|_2^2 + \mu_1 \|\mathbf{X}\|_* + \mu_2 \sum_{i,j} \|\nabla_{i,j} \mathbf{X}\|_2 \quad (26)$$

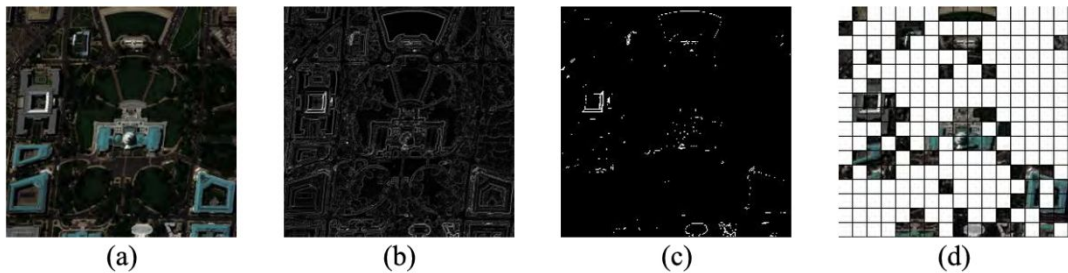
This formulation balances global spectral modeling through nuclear norm minimization with edge-preserving spatial regularization. The framework is versatile, allowing various spatial regularization choices including total variation and non-local filtering. These methods typically achieve consistent reconstruction quality across different scene types and compression ratios. The primary limitation is potential over-smoothing of fine details when regularization weights are not properly calibrated.

Tensor-Enhanced Hybrids augment tensor decomposition with complementary constraints:

$$\min_{\mathbf{X}} \|\mathbf{H}(\mathbf{X}) - \mathbf{y}\|_2^2 + \mu_1 \mathbf{L}_{\text{tensor}}(\mathbf{X}) + \mu_2 \|\mathbf{X}\|_{1,1,2} \quad (27)$$

where  $\mathbf{L}_{\text{tensor}}$  promotes low-rank tensor structure and the mixed norm encourages group sparsity across spatial locations. This approach excels in complex scenes with multiple materials and varying conditions, though it requires more computational resources and careful parameter tuning. The method naturally accommodates the spatial-spatial-spectral organization of hyperspectral data but comes with increased computational complexity.

Contemporary hybrid reconstruction algorithms demonstrate sophisticated integration of multiple priors within unified optimization frameworks. The 3D Weighted Total Variation with Adaptive Low-rank (3DWTV-ALR)<sup>76</sup> method exemplifies intelligent constraint adaptation, applying nuclear norm regularization selectively based on local spatial characteristics while maintaining global spatial-spectral smoothness through weighted TV. Figure 14 illustrates the adaptive processing workflow, showing the original hyperspectral scene, gradient computation across spectral bands, identification of high-gradient regions, and the resulting block-based regularization scheme that enables spatially-adaptive constraint application. The algorithm constructs sparse dictionaries from observed pixels, avoiding spectral detail loss inherent in purely global low-rank models. Its adaptive weighting scheme modulates constraint strength based on local texture metrics, preserving fine details in complex regions while enforcing stronger regularization in homogeneous areas.



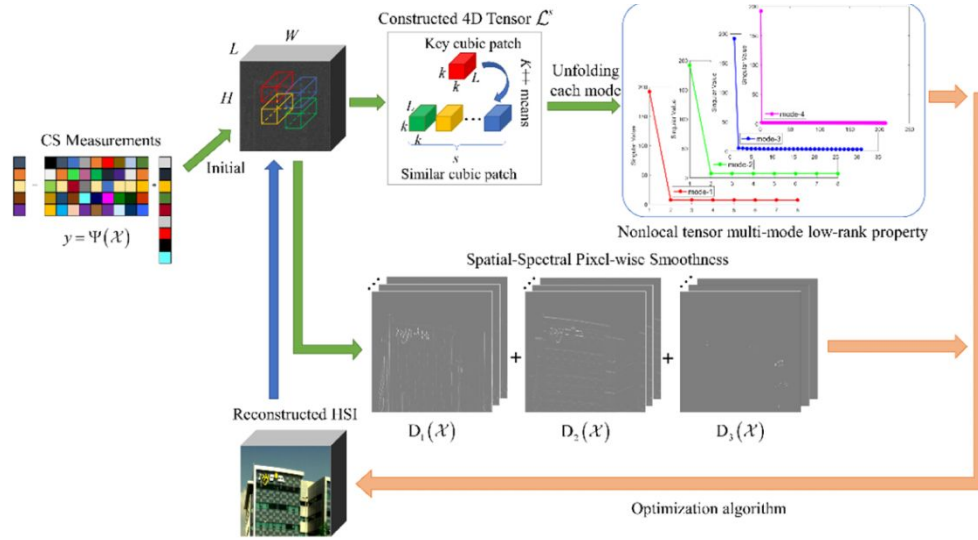
**Fig. 14** (a) False color image of the original hyperspectral image<sup>76</sup>. (b) The sum of gradients at each pixel position across all spectral bands. (c) The top 1% pixels with maximum gradients. (d) Blocks using 3DWTV regularization.

Tucker-based Multispectral Fusion<sup>77</sup> frameworks demonstrate the power of combining tensor structure with multiple regularizers. These methods apply logarithmic sum penalties to factor matrices for enhanced low-rank promotion, incorporate TV constraints on spectral factors to enforce smoothness, and impose  $l_1$  sparsity on core tensors to identify dominant multilinear

components. The resulting optimization problem, though non-convex, admits efficient solutions through careful algorithm design exploiting the multi-block structure.

The Low-rank Tensor Ring with Group Sparsity (LRTRDGS)<sup>37</sup> algorithm represents a sophisticated marriage of global and local modeling. By combining tensor ring decomposition for efficient global structure representation with weighted group sparse regularization for local feature preservation, this approach handles mixed noise scenarios effectively. The ADMM framework enables tractable optimization despite the complexity of multiple constraints, with each subproblem admitting closed-form or efficient iterative solutions.

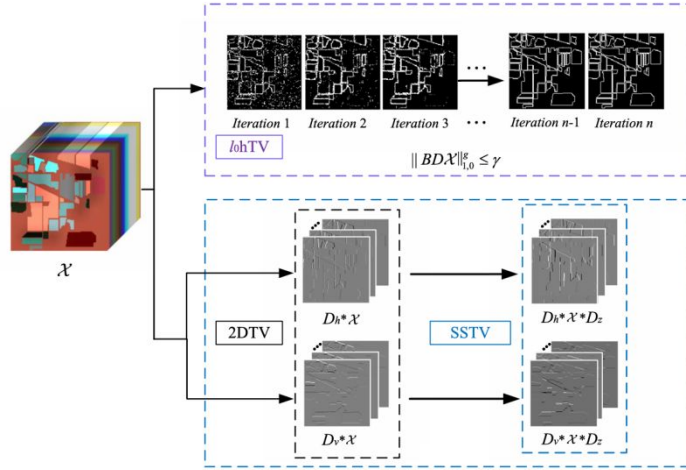
Recent innovations push the boundaries of constraint combination strategies. The Multimode Low-rank Tensor with Anisotropic Smoothness (MLRTAS)<sup>78</sup> method constructs fourth-order tensors to preserve structural information while approximating multimode low-rank constraints through weighted Schatten-p norms. As depicted in Fig. 15, the MLRTAS algorithm systematically combines nonlocal tensor multi-mode low-rank properties with spatial-spectral pixel-wise smoothness constraints, constructing 4D tensors from similar cubic patches to exploit both local similarities and global structural relationships. This non-convex relaxation provides superior low-rank modeling compared to nuclear norm approaches. Combined with 3D anisotropic TV regularization, the method achieves a remarkable balance between global structure preservation and local detail retention.



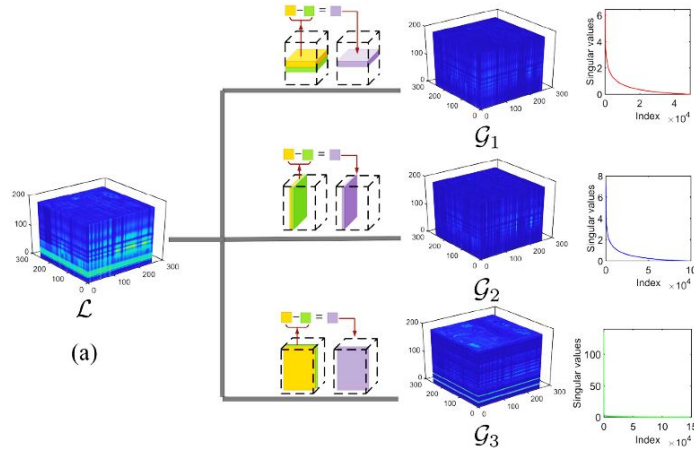
**Fig. 15** The flowchart of the MLRTAS Algorithm for HSI compressive sensing reconstruction<sup>78</sup>, 4D tensor construction from similar cubic patches with nonlocal tensor multi-mode low-rank properties and spatial-spectral pixel-wise smoothness constraints for HSI compressive sensing reconstruction.

The  $l_0 - l_1$  Hybrid Total Variation<sup>79</sup> approach addresses the fundamental tension between edge preservation and noise suppression by combining  $l_0$  gradient constraints for sharp edges with  $l_1$  spatial-spectral TV for global smoothness. As illustrated in Fig. 16, this hybrid regularization strategy systematically processes boundary information through alternating BDTV and SSTV operations to achieve precise edge localization while maintaining overall smoothness. This formulation directly controls the number of non-zero gradients, enabling precise edge localization while maintaining overall smoothness. Similarly, Low-rank Guided Spatial-Spectral TV<sup>80</sup> methods model gradient maps as simultaneously sparse and low-rank, capturing the dual nature of natural image gradients through joint  $l_1$  and tensor nuclear norm constraints. Figure 17 demonstrates how hyperspectral image gradient domains exhibit both sparsity and low-rank

properties across different gradient directions, providing the theoretical foundation for these joint regularization approaches.



**Fig. 16** Illustration of the l0-l1 hybrid total variation regularization<sup>79</sup>, alternating BDTV and SSTV operations for balancing edge preservation and noise suppression.



**Fig. 17** Sparsity and low-rank characteristics of hyperspectral image gradient domains<sup>80</sup>: evidence for joint regularization approaches combining  $\ell_1$  and tensor nuclear norm constraints.

Beyond mathematical prior integration, hybrid approaches increasingly exploit complementary information from multiple data sources. Joint reconstruction frameworks process

related datasets—hyperspectral with multispectral guides, hyperspectral with LiDAR, or multi-temporal acquisitions—through coupled optimization:

$$\min_{\mathbf{X}_1, \mathbf{X}_2} \sum_{k=1}^2 [\mathbf{D}_k(\mathbf{A}_k \mathbf{X}_k, \mathbf{Y}_k) + \lambda_k \mathbf{R}_k(\mathbf{X}_k)] + \mu \mathbf{C}(\mathbf{X}_1, \mathbf{X}_2) \quad (28)$$

where  $\mathbf{C}(\cdot, \cdot)$  enforces cross-modal consistency through shared subspace constraints, joint sparsity patterns, or coupled tensor decompositions. This paradigm extends hybrid constraints from purely mathematical priors to data-driven constraints, enabling mutually beneficial reconstruction particularly valuable in remote sensing applications where multi-modal acquisition is standard.

Hybrid constraint methods offer significant advantages: enhanced robustness to noise and artifacts through redundant constraints, improved resilience to model misspecification via complementary priors, superior preservation of both global structure and local details, more effective utilization of limited measurements through diverse information extraction, and better generalization across scenes with varying characteristics. The framework naturally accommodates physically interpretable decompositions, with different constraints aligning with meaningful data components.

However, these benefits incur costs. Computational complexity increases substantially, with optimization requiring sophisticated algorithms to handle multiple non-smooth terms. Parameter selection becomes increasingly challenging as the number of hyperparameters grows multiplicatively. Competing constraints may create optimization landscapes with poor local minima or slow convergence. Theoretical analysis becomes more difficult, limiting performance guarantees.

Current research addresses these challenges through adaptive parameter selection using bilevel optimization or machine learning approaches, efficient optimization algorithms exploiting problem structure through operator splitting or proximal methods, and theoretical frameworks for analyzing multi-constraint problems. The integration of hybrid constraints with deep learning architectures represents a particularly promising direction, where neural networks learn optimal constraint combinations and parameters directly from data while maintaining the interpretability and reliability of model-based approaches. These advances position hybrid constraint methods at the forefront of hyperspectral reconstruction, offering the flexibility and performance required for next-generation imaging systems.

## **5 Learning-Based Reconstruction Algorithms**

The model-based approaches examined in Chapter 3 and hybrid frameworks discussed in Chapter 4 have established a solid theoretical foundation for hyperspectral reconstruction. These methods leverage explicit mathematical priors—sparsity, low-rank structures, smoothness constraints, and their combinations—to transform ill-posed inverse problems into tractable optimization tasks. While achieving notable success in controlled scenarios, these traditional approaches increasingly encounter fundamental limitations when confronted with the complexity and diversity of real-world hyperspectral imaging applications.

The limitations of model-based methods stem from several inherent challenges. First, the hand-crafted priors, despite their mathematical elegance, often fail to capture the full complexity of natural hyperspectral scenes. Real-world data exhibits intricate spatial-spectral correlations, non-linear mixing phenomena, and scene-specific characteristics that resist simple mathematical modeling. Second, the performance of these methods critically depends on careful parameter tuning—regularization weights, sparsity levels, rank estimates—which typically requires domain

expertise and extensive trial-and-error. Third, the iterative nature of optimization-based reconstruction imposes significant computational burdens, particularly problematic for real-time applications or large-scale datasets. Finally, these methods struggle to adapt to varying imaging conditions, sensor characteristics, and degradation patterns without manual intervention.

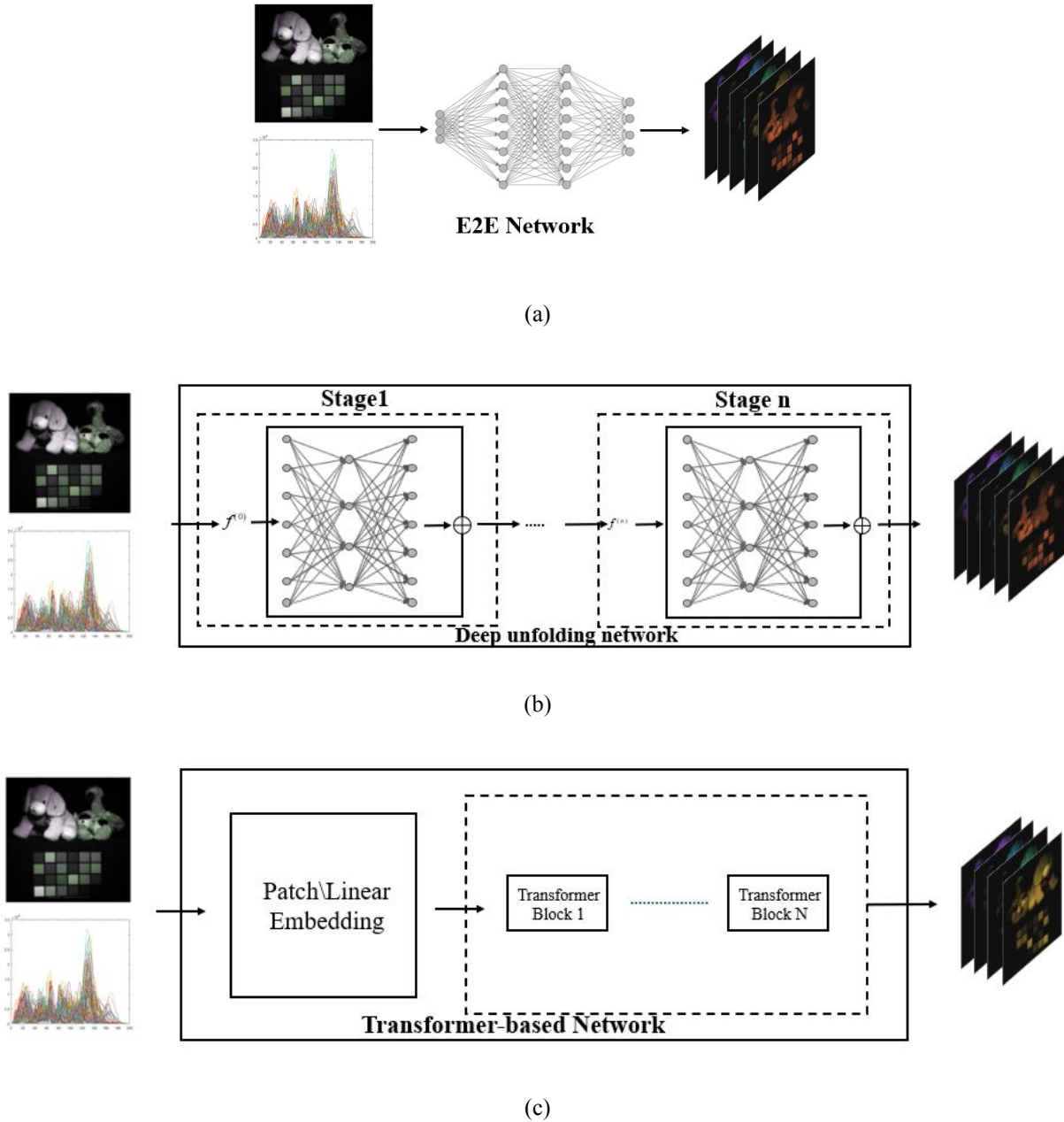
The emergence of deep learning has opened transformative possibilities for addressing these limitations. Unlike traditional approaches that rely on predefined mathematical models, learning-based methods can automatically discover optimal representations and reconstruction strategies directly from data. This paradigm shift—from hand-crafted to learned priors—enables several key advantages: (1) the ability to capture complex, non-linear relationships in hyperspectral data that elude mathematical description; (2) adaptive performance across diverse scenes and imaging conditions through training on varied datasets; (3) potential for real-time reconstruction by shifting computational burden from test-time optimization to offline training; and (4) end-to-end optimization that jointly learns all components of the reconstruction pipeline

Learning-based hyperspectral reconstruction fundamentally involves training neural networks to approximate the inverse mapping from degraded measurements to pristine hyperspectral cubes. Formally, we seek to learn a parameterized function  $f_{\theta} : \mathbb{I}^M \rightarrow \mathbb{I}^N$  that maps compressed measurements  $\mathbf{y} \in \mathbb{I}^M$  to reconstructed hyperspectral images  $\hat{\mathbf{x}} \in \mathbb{I}^N$ :

$$\hat{\mathbf{x}} = f_{\theta}(\mathbf{y}) \tag{29}$$

where  $\theta$  represents network parameters optimized to minimize expected reconstruction loss  $L$  over the data distribution. This formulation shifts computational complexity from test-time optimization to offline training, enabling real-time reconstruction capabilities essential for

dynamic imaging scenarios. The choice of architecture  $f_\theta$ , loss function  $L$ , and training strategy profoundly impacts reconstruction quality, generalization capability, and computational efficiency.



**Fig. 18** Architectures of partial reconstruction network. (a) E2E network. (b) Deep unfolding network. (c) Transformer-based network.

This chapter explores three complementary paradigms in learning-based hyperspectral reconstruction. Figure 18 illustrates the key architectural approaches: (a) end-to-end networks that

directly map measurements to reconstructions, (b) deep unfolding networks that preserve optimization structure through iterative stages, and (c) transformer-based networks that employ attention mechanisms for global spatial-spectral modeling. We begin with model-driven learning approaches (Section 5.1), exemplified by Deep Unfolding Networks that preserve the interpretability of optimization algorithms while enabling data-driven parameter learning. These methods represent a natural bridge between traditional and learning-based approaches, maintaining theoretical guarantees while adapting to data characteristics. We then examine purely data-driven methods (Sections 5.2-5.3), including CNN-based and Transformer-based architectures that leverage the full representational power of deep learning. These end-to-end approaches demonstrate how architectural innovations—from local convolutions to global attention mechanisms—can effectively model the complex spatial-spectral relationships in hyperspectral data. Throughout our analysis, we emphasize the trade-offs between reconstruction accuracy, computational efficiency, interpretability, and generalization capability, providing practical guidance for method selection in diverse application contexts.

### *5.1 Deep Unfolding Networks*

Deep unfolding networks represent an elegant synthesis of model-based optimization and data-driven learning for hyperspectral reconstruction, transforming classical iterative algorithms into trainable neural architectures. The fundamental principle involves "unrolling" a fixed number of optimization iterations into sequential network layers, where traditionally fixed parameters and operations become learnable components optimized through backpropagation<sup>81</sup>. This architectural philosophy preserves the interpretability and theoretical guarantees of optimization-based methods while harnessing deep learning's capacity to adapt to data-specific characteristics. By maintaining algorithmic structure while enabling parameter learning, deep unfolding networks address a

critical limitation of classical methods—their reliance on hand-tuned parameters that may be suboptimal for diverse imaging conditions.

The transformation from iterative optimization to deep unfolding begins with the canonical hyperspectral reconstruction problem:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \mu \mathbf{R}(\mathbf{x}) \quad (30)$$

Classical algorithms like ISTA<sup>82</sup> (Iterative Shrinkage-Thresholding Algorithm) solve this through iterative updates:

$$\mathbf{x}^{(k+1)} = \mathbf{S}_{\lambda/L} \left( \mathbf{x}^{(k)} - \frac{1}{L} \mathbf{H}^T (\mathbf{H}\mathbf{x}^{(k)} - \mathbf{y}) \right) \quad (31)$$

where  $\mathbf{S}_{\lambda/L}$  denotes the soft-thresholding operator and  $L$  is the Lipschitz constant. Deep unfolding transforms each iteration into a network layer with learnable components:

$$\mathbf{x}^{(k+1)} = \mathbf{F}_{\theta_k} \left( \mathbf{x}^{(k)} - \mathbf{W}_k^T (\mathbf{W}_k \mathbf{x}^{(k)} - \mathbf{y}) \right) \quad (32)$$

Here, the linear operators  $\mathbf{H}, \mathbf{H}^T$  become learnable transformations  $\mathbf{W}_k, \mathbf{W}_k^T$ , the step size  $\frac{1}{L}$

becomes a learnable parameter, and the shrinkage function  $\mathbf{S}$  transforms into a parameterized activation  $\mathbf{F}_{\theta_k}$ . Crucially, parameters can vary across layers, enabling stage-specific adaptations that outperform fixed-parameter algorithms.

The distinctive properties of hyperspectral images—high spectral dimensionality, strong spectral correlations, and spatial-spectral joint structures—require specialized unfolding architectures tailored for hyperspectral reconstruction tasks.

ADMM-Net<sup>25</sup> unfolds the Alternating Direction Method of Multipliers, particularly suited for hyperspectral reconstruction with multiple constraints. The network learns penalty parameters  $\rho_k$ , projection operators for different constraint sets, and proximal mappings that preserve spectral consistency. Each ADMM iteration translates to network modules:

- Variable update:  $\mathbf{x}^{(k+1)} = \text{Conv3D}_{\theta_1}(\mathbf{z}^{(k)} - \mathbf{u}^{(k)})$
- Auxiliary update:  $\mathbf{z}^{(k+1)} = \text{ProxNet}_{\theta_2}(\mathbf{x}^{(k+1)} + \mathbf{u}^{(k)})$
- Dual update:  $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \gamma_k(\mathbf{x}^{(k+1)} - \mathbf{z}^{(k+1)})$

where Conv3D processes spatial-spectral volumes and ProxNet implements learnable proximal operators.

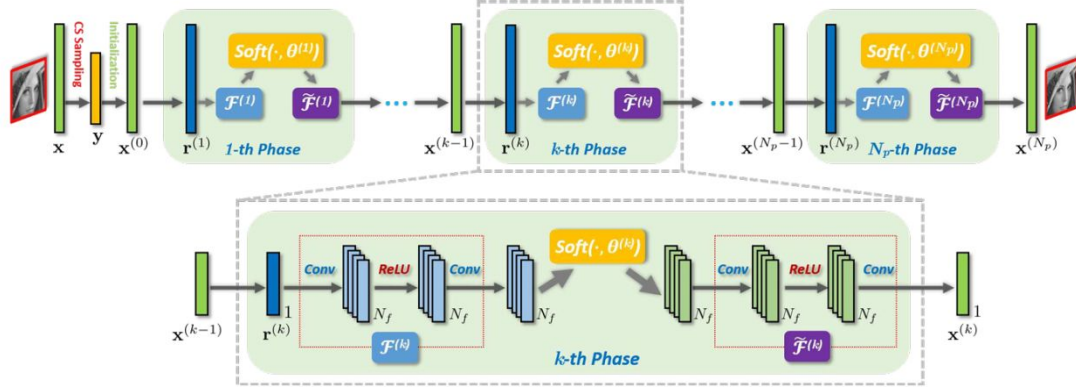
HSI-SpectralDiff-Net incorporates spectral difference modules within the unfolding framework, explicitly modeling relationships between adjacent bands through learnable filters:

$$\mathbf{D}_{\text{spectral}} = \sum_{i=1}^S w_i \text{CircShift}(\mathbf{x}, i) - \mathbf{x} \quad (33)$$

This architecture captures spectral smoothness prior more effectively than generic regularizers, leading to superior preservation of spectral signatures.

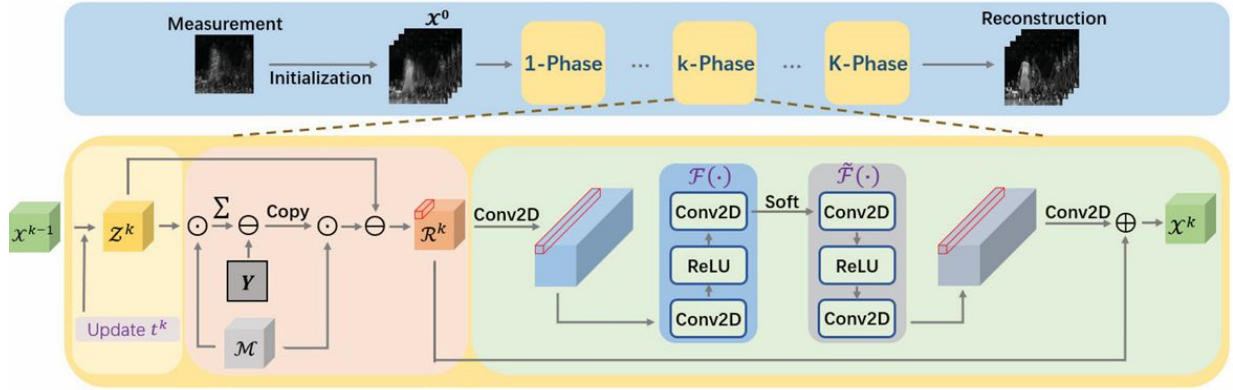
Building upon the ISTA framework introduced earlier (Equation 27-29), several practical implementations have emerged for hyperspectral compressive imaging. ISTA-Net<sup>82</sup> pioneered this application by introducing learnable nonlinear transforms  $F(\cdot)$  that adapt to hyperspectral data characteristics. As shown in Fig. 19, the ISTA-Net architecture unfolds the iterative ISTA algorithm into a deep network structure, where each phase corresponds to one ISTA iteration with learnable soft-thresholding operations and convolutional transforms. The network reformulates the

proximal mapping as  $x^{(k)} = \mathcal{P}(\text{soft}(F(r^{(k)}), \theta))$ , where both  $F$  and  $\mathcal{P}$  are implemented as convolutional layers specifically designed for spectral-spatial feature extraction.



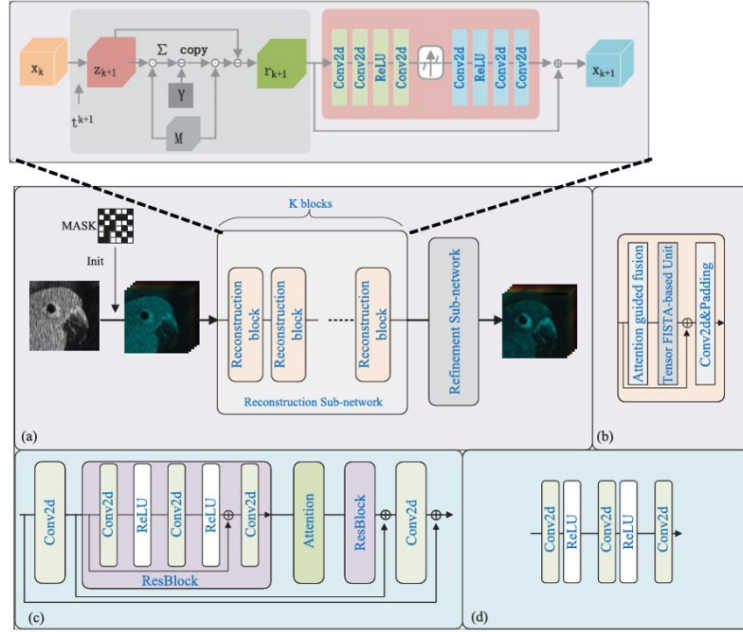
**Fig. 19** The architecture of ISTA-Net ISTA-Net architecture for hyperspectral compressive imaging: deep unfolding of iterative ISTA algorithm with learnable nonlinear transforms and soft-thresholding operations across multiple phases<sup>82</sup>.

Tensor FISTA-Net<sup>83</sup> advances this concept by exploiting the inherent tensor structure of hyperspectral data. Unlike standard ISTA-Net that processes vectorized data, this architecture maintains the 3D tensor format throughout iterations, preserving crucial spatial-spectral correlations. As illustrated in Fig. 20, the Tensor FISTA-Net architecture systematically processes hyperspectral data through multiple phases while preserving the tensor structure, with each phase incorporating tensorized  $F(\cdot)$  and  $\mathcal{P}(\cdot)$  operations that utilize Conv2D layers and soft-thresholding specifically designed for 3D data. The tensorized operations significantly reduce information loss between unfolding stages, a critical improvement for high-dimensional hyperspectral reconstruction.



**Fig. 20** The architecture of Tensor FISTA-Net, maintaining 3D tensor format across iterative phases to preserve spatial-spectral correlations and reduce information loss in high-dimensional hyperspectral reconstruction<sup>83</sup>.

Recent architectures like ReAttFISTA-Net<sup>52</sup> demonstrates how deep unfolding can be enhanced through hybrid architectures. As depicted in Fig. 21, the ReAttFISTA-Net architecture showcases a sophisticated multi-component design that integrates reconstruction sub-networks, attention-guided fusion mechanisms, and finer FISTA-based residual blocks, demonstrating the evolution from simple algorithm unfolding to complex hybrid frameworks. While maintaining the Tensor FISTA-Net as its optimization backbone, it incorporates attention-guided fusion mechanisms and learnable refinement modules. This design exemplifies the evolution from pure algorithm unfolding to sophisticated hybrid frameworks that balance theoretical guarantees with data-driven performance gains. These ISTA-based variants collectively illustrate how the theoretical framework presented in Equations 27-29 translates into practical architectures, each addressing specific challenges in hyperspectral compressive imaging through targeted innovations.



**Fig. 21** The architecture of ReAttFISTA-Net with attention-guided fusion and multi-component reconstruction design<sup>52</sup>.

The training of deep unfolding networks for hyperspectral reconstruction requires careful consideration of the interplay between algorithmic structure and learning flexibility. A particularly effective strategy involves physics-informed initialization<sup>84</sup>, where network parameters are initialized to mimic the original algorithm's behavior. For instance, convolutional layers replacing  $\mathbf{H}^T\mathbf{H}$  are initialized using the system's point spread function, while activation functions approximate classical shrinkage operators. This initialization ensures the network begins from a theoretically grounded starting point before data-driven refinement.

Deep unfolding networks provide several advantages over both classical optimization and black-box deep learning approaches. Compared to iterative algorithms, they provide deterministic runtime through fixed-depth architectures, eliminating convergence checks and enabling real-time deployment. The learned parameters adapt to data statistics, often outperforming hand-tuned values across diverse scenes. Against conventional neural networks, unfolding architectures offer

enhanced interpretability—each layer corresponds to an algorithmic step with analyzable parameters. They demonstrate superior sample efficiency, requiring significantly less training data due to embedded domain knowledge, and exhibit robust generalization to out-of-distribution scenarios through algorithmic inductive bias.

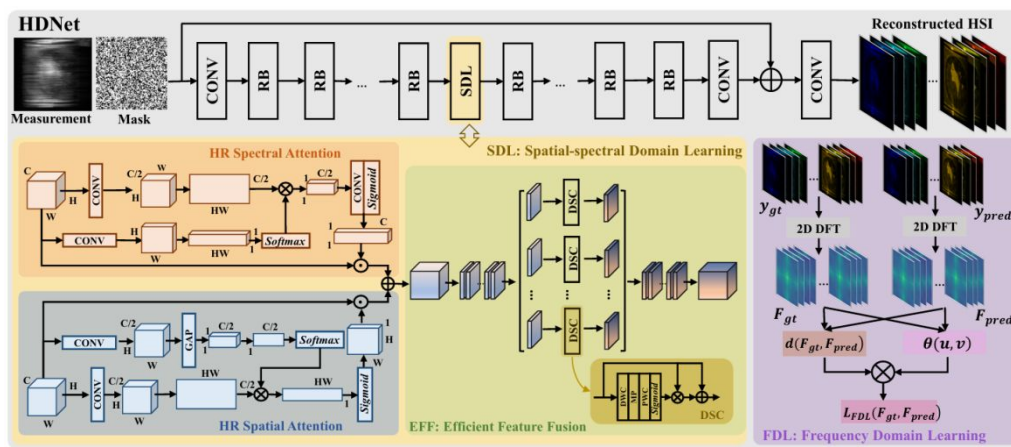
Deep unfolding networks face several important limitations that affect their practical deployment. First, their performance depends heavily on whether the underlying algorithmic assumptions match the target problem. When optimization priors like sparsity or smoothness poorly represent the actual data characteristics, these methods may underperform pure data-driven approaches. Second, architectural design requires careful balancing of algorithmic structure and learning flexibility. Poor design choices can cause training instabilities or loss of theoretical guarantees. Third, computational demands often exceed those of classical methods, particularly during training and when processing high-dimensional hyperspectral data. Fourth, generalization remains limited by training data distributions, with performance degrading when test conditions differ from training scenarios. Finally, as architectural modifications increase to improve performance, the interpretability advantage may diminish, potentially reducing the transparency that motivates the unfolding approach. These constraints highlight the need for careful evaluation when selecting deep unfolding networks over alternative reconstruction methods.

### *5.2 CNN-Based Reconstruction Algorithms*

Convolutional Neural Networks (CNNs) have emerged as the dominant architecture in early end-to-end hyperspectral image reconstruction methods, primarily due to their exceptional capability to capture local spatial structures and spectral correlations through hierarchical feature learning. The fundamental design of CNN-based reconstruction networks typically employs stacked convolutional layers that progressively extract multi-scale features from compressed

measurements, where small receptive fields enable fine-grained spatial pattern recognition while deeper layers aggregate broader contextual information to enhance reconstruction fidelity.

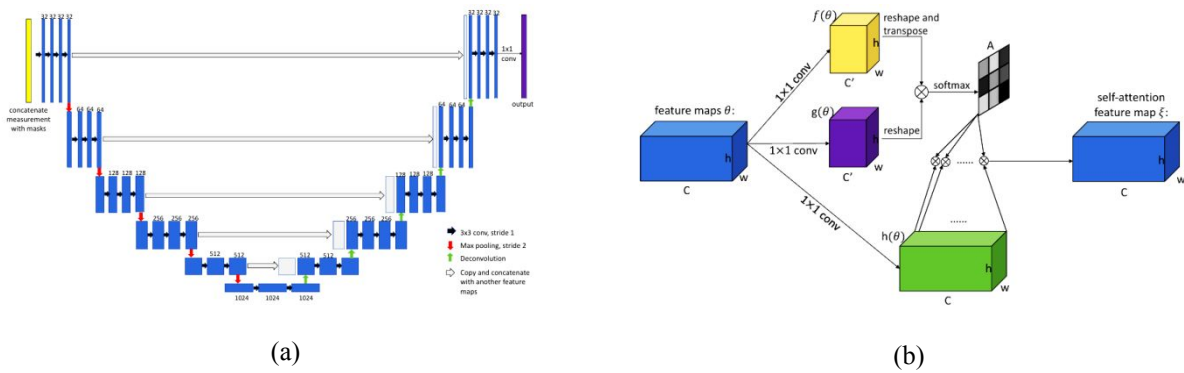
The concept of dual-domain processing has proven particularly effective in addressing the inherent trade-off between spatial resolution and spectral accuracy in compressive spectral imaging. Hu<sup>85</sup> pioneered this approach with HDNet, which implements a high-resolution dual-domain learning strategy that jointly operates in both spatial and spectral domains. As illustrated in Fig. 22, the HDNet architecture integrates multiple specialized components including HR spectral attention, HR spatial attention, spatial-spectral domain learning (SDL), efficient feature fusion (EFF), and frequency domain learning (FDL) modules, systematically addressing different aspects of hyperspectral reconstruction through coordinated multi-domain processing. Through specialized convolutional blocks and multi-scale feature extraction mechanisms, HDNet successfully preserves fine spatial details while maintaining spectral fidelity across the entire wavelength range.



**Fig. 22** The architecture of HDNet<sup>85</sup>, a dual-domain learning network for spectral compressive imaging.

Building upon the multi-stage reconstruction paradigm, Miao et al. developed  $\lambda$ -net<sup>25</sup>, a dual-stage generative model that decouples the reconstruction process into initial generation and refinement phases. The first stage employs a self-attention generator integrated with a hierarchical

channel reconstruction (HCR) strategy to produce an initial hyperspectral estimate from 2D snapshot measurements. Figure 23 illustrates the key architectural components of  $\lambda$ -Net: (a) the U-Net architecture employed in the reconstruction stage featuring encoder-decoder structure with skip connections for progressive refinement, and (b) the attention module that enables adaptive feature weighting through self-attention mechanisms. Subsequently, a refinement stage comprising a compact U-Net architecture with residual learning connections further enhances the reconstruction quality, demonstrating the effectiveness of progressive refinement in hyperspectral reconstruction.



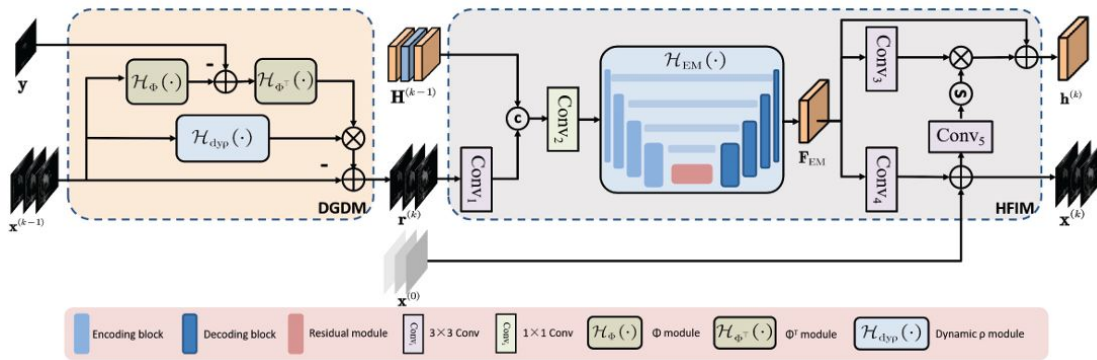
**Fig. 23** (a) U-net architecture used in the reconstruction stage of  $\lambda$ -Net. (b) The attention module of  $\lambda$ -Net<sup>25</sup>.

Recognizing the importance of jointly modeling spatial and spectral information, several studies have focused on developing CNN architectures that effectively capture both local spatial context and long-range spectral dependencies. Wang et al.<sup>19</sup> introduced a learned deep spatial-spectral prior that integrates spatial and spectral cues throughout the feature extraction process. This approach enables the network to resolve complex variations inherent in hyperspectral data, resulting in more accurate reconstructions compared to methods that process spatial and spectral information independently.

The integration of attention mechanisms has further enhanced the capability of CNN-based methods to model spatial-spectral relationships. Meng et al.<sup>17</sup> embedded spatial-spectral self-

attention modules within the CNN framework, enabling dynamic focus on the most informative spatial and spectral features. This attention-guided approach not only improves reconstruction accuracy but also prioritizes computational efficiency and hardware compatibility, making it particularly suitable for low-cost, end-to-end compressive spectral imaging systems.

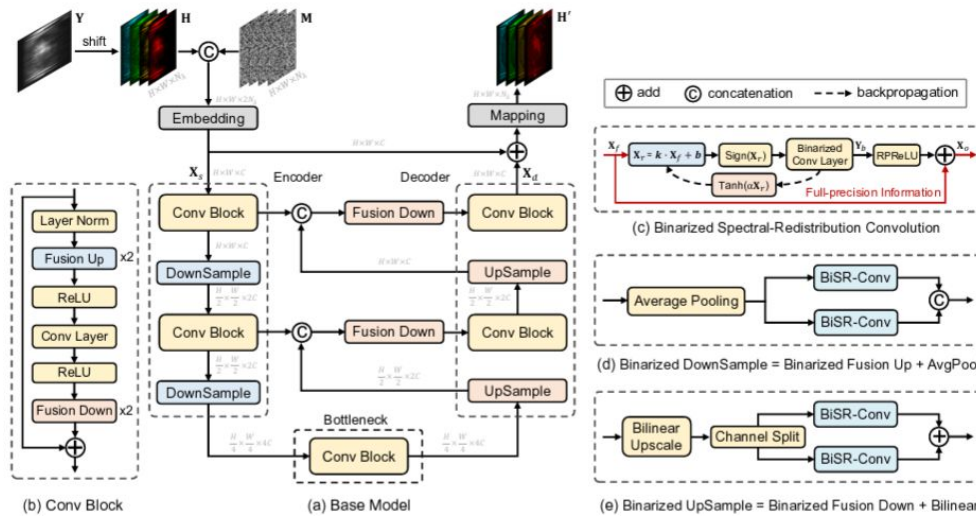
Recent developments in CNN-based hyperspectral reconstruction have increasingly focused on optimization strategies and computational efficiency. Figure 24 illustrates the HerosNet architecture, exemplifying sophisticated CNN-based designs that integrate multiple specialized modules (DGDM and HFIM) with encoding, decoding, and residual components to achieve efficient iterative reconstruction through optimized feature processing pathways. Zhang et al. addressed the fundamental challenge of data acquisition by integrating an optimal sampling strategy directly into the end-to-end learning framework<sup>86</sup>, thereby aligning the measurement process with the reconstruction objective. This co-design approach demonstrates significant improvements in reconstruction quality while maintaining interpretability of the learned sampling patterns.



**Fig. 24** The architecture of HerosNet, exemplifying advanced CNN-based hyperspectral reconstruction with integrated DGDM and HFIM modules<sup>86</sup>.

Hardware efficiency considerations have led to innovative quantization approaches, exemplified by Cai et al., who introduced binary CNN architectures for hyperspectral

reconstruction<sup>87</sup>. Figure 25 details the architecture of these Binarized Convolutional Modules, showing the encoder-decoder base model structure alongside specialized components including binarized spectral-redistribution convolution, binarized downsampling and upsampling operations that enable efficient computation while preserving spectral information processing capabilities. By employing binary weights and activations, this method achieves substantial reductions in memory usage and inference time—critical requirements for real-time or resource-constrained applications without significantly compromising reconstruction quality.



**Fig. 25** The architecture of Binarized Convolutional Modules, featuring encoder-decoder structure with binarized spectral-redistribution operations<sup>87</sup>.

Despite their proven effectiveness, CNN-based methods face inherent limitations in modeling long-range dependencies due to the localized nature of convolutional operations. The restricted receptive field of standard convolutions limits the network's ability to capture global spectral correlations and long-range spatial relationships, which are crucial for high-fidelity hyperspectral reconstruction. These limitations have motivated the recent exploration of alternative architectures, including Multi-Layer Perceptrons (MLPs) and Vision Transformers, which will be discussed in subsequent sections.

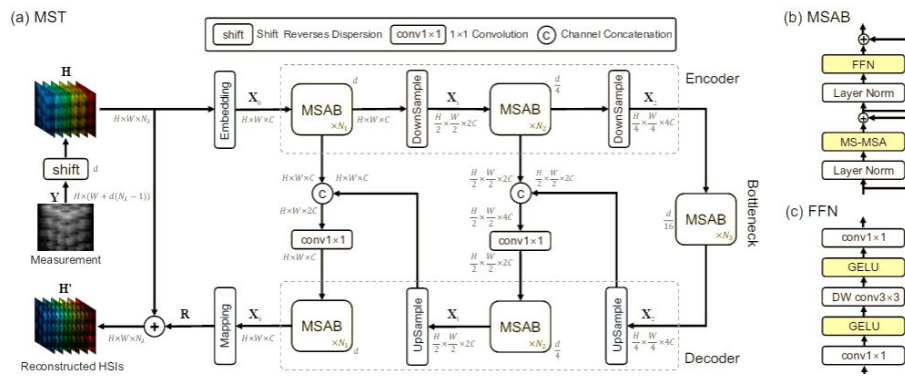
### 5.3 Transformer-based Reconstruction Algorithms

Transformer-based architectures have recently emerged as a transformative paradigm in end-to-end hyperspectral image reconstruction, addressing the fundamental limitations of CNN-based approaches in modeling long-range dependencies. Originally developed for natural language processing tasks and subsequently adapted for computer vision applications, Transformers excel at capturing complex correlations across both spatial and spectral dimensions—a critical capability for hyperspectral imaging where meaningful relationships often exist between distant regions in both domains. The self-attention mechanism inherent to Transformer architecture enables adaptive weighting of spatial-spectral features, allowing networks to focus on the most relevant information for reconstruction while maintaining global context awareness.

The adaptation of Transformer architectures to hyperspectral reconstruction has been driven by the need to model global spatial-spectral dependencies effectively. In snapshot compressive spectral imaging scenarios, these models typically employ window-based self-attention mechanisms to balance computational complexity with modeling capability, enabling efficient processing of high-dimensional hyperspectral data while preserving essential long-range correlations.

Cai et al. pioneered the application of Transformers to hyperspectral reconstruction with the introduction of the Mask-guided Spectral-wise Transformer (MST)<sup>18</sup>. Figure 26 illustrates the MST architecture, showing (a) the encoder-decoder framework with multiple MSAB (Multi-head Self-Attention Block) modules connected by skip connections, (b) the detailed MSAB structure containing spectral-wise multi-head self-attention (MS-MSA) and feed-forward networks, and (c) the FFN implementation with depthwise convolutions and GELU activations. This seminal work introduced two key innovations: spectral-wise multi-head self-attention (S-MSA) and a mask-

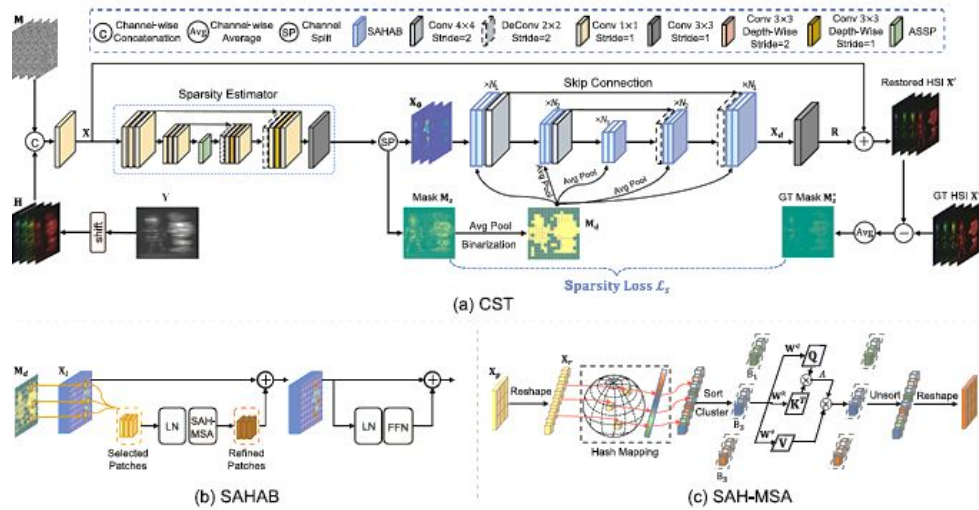
guided mechanism (MM) that incorporates prior physical knowledge from the sensing process. The architecture adopts a popular encoder-decoder design reminiscent of U-Net structures, progressively extracting and fusing hierarchical features while maintaining computational efficiency. The encoder-decoder arrangement facilitates both local feature extraction through convolutional operations and global dependency modeling via S-MSA, achieving state-of-the-art reconstruction quality with reduced memory and computational overhead compared to conventional approaches.



**Fig. 26** The architecture of MST, demonstrating encoder-decoder framework with Multi-head Self-Attention Blocks (MSAB), spectral-wise self-attention (S-MSA), and mask-guided mechanism<sup>18</sup>.

The success of MST motivated the development of more sophisticated multi-stage refinement strategies. Cai et al. extended this foundation with the Coarse-to-fine Sparse Transformer (CST)<sup>88</sup>, which integrates sparse attention mechanisms with hierarchical feature processing. Figure 27 illustrates the CST architecture, showing (a) the overall coarse-to-fine framework with sparsity estimator and skip connections for progressive refinement, (b) the SAHAB (Sparse Attention Hybrid Attention Block) structure with refined patch processing, and (c) the SAH-MSA mechanism that implements sparse attention through hash mapping and clustering operations to efficiently model long-range dependencies. This multi-stage approach effectively balances the modeling of global contextual information—critical for maintaining spectral fidelity—with the

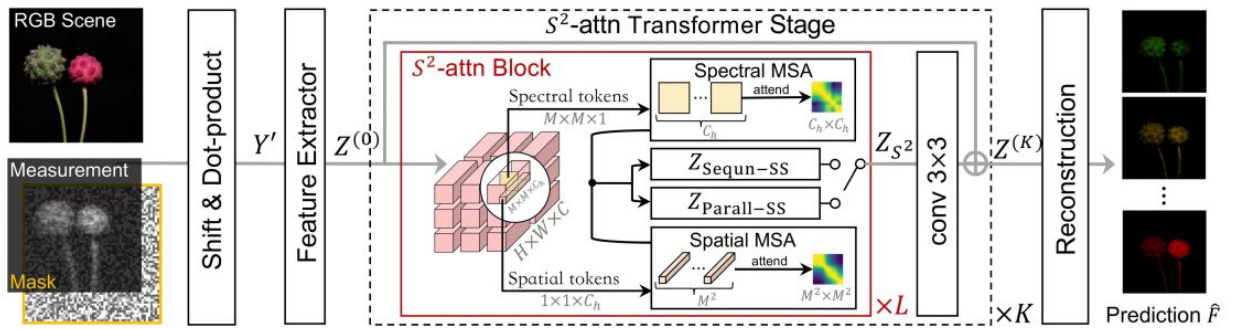
refinement of local spatial details, resulting in superior reconstruction performance across diverse spectral imaging scenarios. The coarse-to-fine paradigm addresses a fundamental challenge in hyperspectral reconstruction: the need to simultaneously preserve broad spectral characteristics while recovering fine spatial details. By decomposing the reconstruction process into multiple stages, CST enables progressive refinement that enhances both spatial resolution and spectral accuracy through iterative attention-based feature aggregation.



**Fig. 27** The architecture of CST, demonstrating coarse-to-fine framework with sparsity estimator, SAHAB (Sparse Attention Hybrid Attention Block), and SAH-MSA mechanism<sup>88</sup>.

Recognizing the distinct characteristics of spatial and spectral information in hyperspectral data, several studies have explored dual-stream architectures that process these dimensions independently before fusion. Wang et al. developed the Spatial-Spectral Transformer (S<sup>2</sup>-Tran)<sup>89</sup>, which employs two interconnected processing streams with dedicated spatial and spectral self-attention modules. Figure 28 shows the S<sup>2</sup>-Tran architecture with dual-stream processing - separate spectral and spatial self-attention pathways that process different token types while maintaining cross-dimensional information exchange. This design enables specialized feature extraction for each domain while maintaining cross-dimensional information exchange. A particularly

innovative aspect of  $S^2$ -Tran is its direct integration of mask information into both the attention mechanism and loss function. This integration addresses the spatial non-uniform degradation introduced by the sensing mask, effectively disentangling the complex measurement patterns caused by spatial shifting and mask coding. This physics-informed approach demonstrates how domain-specific knowledge can be seamlessly incorporated into Transformer architectures to enhance reconstruction fidelity.



**Fig. 28** The architecture of the spatial-spectral transformer, dual-stream Transformer with separate spatial and spectral self-attention modules and integrated mask processing<sup>90</sup>.

Computational efficiency remains a critical consideration in Transformer-based hyperspectral reconstruction, particularly for high-resolution applications. Luo et al. addressed this challenge with the Dual-Window Multi-scale Transformer (DWMT)<sup>91</sup>, which captures both fine local details and broad contextual information through innovative dual-window processing strategies. Figure 29 illustrates the DWMT architecture featuring parallel dual-window processing with different window sizes that undergo separate Multi-head Self-Attention operations before concatenation and linear projection. By implementing attention mechanisms separately within each window, DWMT significantly reduces computational burden compared to global self-attention while maintaining sufficient receptive field coverage to capture complex spatial-spectral relationships. The multi-scale processing capability of DWMT enables hierarchical feature extraction that mirrors the

multi-resolution nature of hyperspectral data, where different spectral bands may exhibit varying spatial characteristics and correlation patterns.

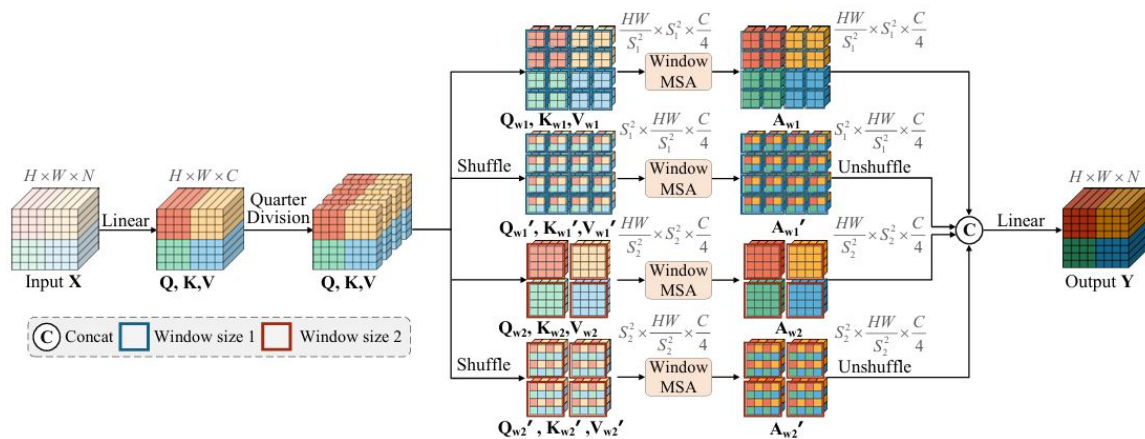
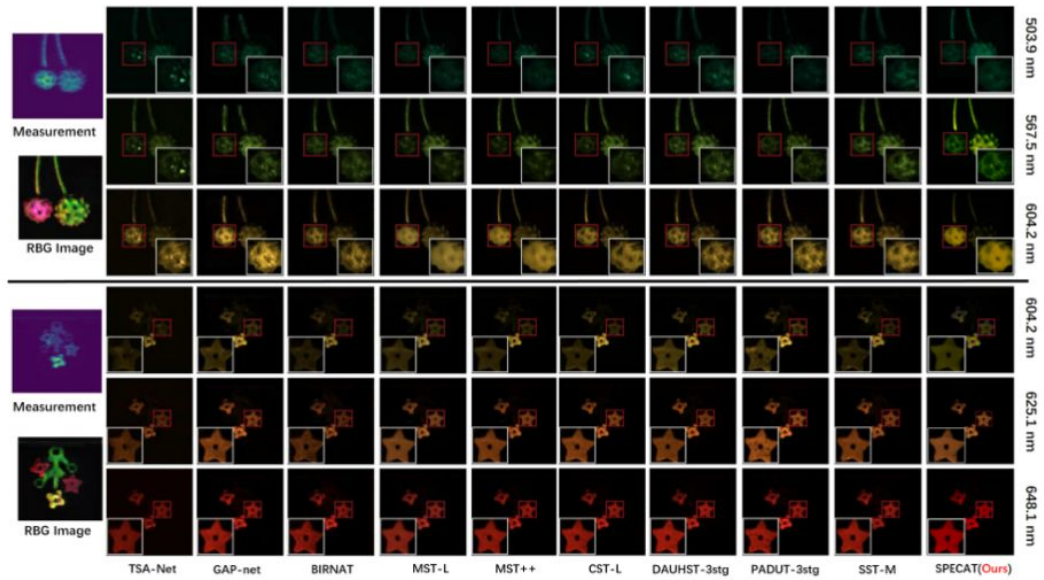


Fig. 29 The dual-window multiscale multi-head self-attention in DWMT<sup>91</sup>.

The evolution of attention mechanisms has led to increasingly sophisticated approaches for hyperspectral reconstruction. Yao et al. introduced the Spatial-Spectral Cumulative-Attention Transformer (SPECAT)<sup>90</sup>, specifically designed for high-resolution hyperspectral reconstruction in Fabry-Pérot Filter Array (FPFA) sensors. SPECAT employs a dual-branch architecture that independently captures spatial and spectral dependencies through cumulative attention modules, enabling effective cross-dimensional feature fusion. Figure 31 demonstrates SPECAT's superior reconstruction performance through visual comparison with state-of-the-art methods across multiple spectral bands, showing improved detail preservation and spectral fidelity.

The cumulative attention mechanism progressively integrates multi-scale context, allowing the network to build increasingly refined representations of the hyperspectral data. This approach attains superior reconstruction quality while ensuring computational efficiency, demonstrating the potential for specialized attention mechanisms tailored to specific sensor configurations.



**Fig. 30** Comparison of reconstruction algorithms of some algorithms<sup>90</sup>.

Transformer-based architectures have consistently demonstrated state-of-the-art performance in hyperspectral image reconstruction, offering a powerful framework for extracting complex spatial-spectral features under compressed sensing scenarios. The success of these methods stems from their ability to model global dependencies while incorporating domain-specific knowledge through innovative attention mechanisms and architectural designs.

The continued evolution of Transformer architectures for hyperspectral reconstruction suggests promising directions for future research, including the development of more efficient attention mechanisms, integration of physics-informed constraints, and adaptation to emerging sensor technologies and acquisition paradigms.

#### 5.4 Comparative Analysis of Resource Consumption and Requirements

Beyond reconstruction accuracy, the practical applicability of an algorithm is heavily dependent on its resource consumption. This includes hardware requirements, computational complexity, and data demands. A comprehensive comparison across these dimensions is essential

for selecting the appropriate algorithm for a given application, such as real-time medical imaging versus offline remote sensing analysis. This section provides a comparative analysis of the three main categories of algorithms discussed: model-based, hybrid, and learning-based methods. A summary of this comparison is presented in Table 6.

**Table 6** Comparative Analysis of Resource Consumption for Different Algorithm Categories.

Aspect	Model-based methods	Hybrid methods	Learning-based Methods
Hardware Requirements	Moderate: Primarily CPU-based. Memory scales with data size	Moderate to High: Similar to model-based, but can be more memory-intensive	High: GPU is essential for training. GPU preferred for fast inference. High storage for models
Computational Complexity	Training: None. Inference: High (iterative optimization)	Training: None. Inference: Very High (more complex optimization)	Training: Very High (days to weeks). Inference: Very Low (single forward pass)
Training Data Requirements	None. Relies on mathematical priors	None. Relies on mathematical priors	Very High. Requires large, diverse, and high-quality datasets of measurement-truth pairs
Robustness to SNR	Moderate. Performance degrades with noise unless the model is specifically designed for it (e.g., TV)	Good. Combining multiple priors often improves noise robustness	Dependent on training data. Can be very robust if trained with noise augmentation, but may fail on unseen noise types

Model-based and hybrid methods share fundamental characteristics regarding resource consumption. They do not require a training phase, which entirely bypasses the need for large datasets and powerful GPUs for model preparation. Their primary computational cost occurs during inference, where they solve an optimization problem for each new measurement. This iterative process can be time-consuming, making them less suitable for real-time applications. The complexity of hybrid methods is typically higher than that of single-prior model-based approaches, as the optimization problem involves balancing multiple competing constraints. Hardware requirements are generally moderate, with standard CPUs being sufficient, though memory usage scales with the dimensions of the hyperspectral datacube. Their performance in the presence of noise is directly tied to the chosen priors; while regularization terms like Total Variation (TV) are

inherently robust to noise, other priors might be more sensitive. Hybrid methods often exhibit better noise robustness by leveraging the complementary strengths of different priors.

Learning-based methods, in stark contrast, shift the computational burden almost entirely to the training phase. Training deep neural networks like CNNs and Transformers is a resource-intensive process, demanding large-scale, high-quality datasets and prolonged access to high-performance GPUs. The trained models themselves can also require significant storage space. However, the major advantage of this paradigm is the extremely fast inference speed. Once trained, a reconstruction is achieved through a single forward pass of the network, making these methods ideal for real-time or near-real-time applications. Their hardware requirement for deployment can be flexible; while a GPU accelerates inference, models can also be run on CPUs or specialized edge devices, especially if techniques like model quantization are employed. The reliance on data is both a strength and a weakness: these models can learn complex, non-linear priors that outperform hand-crafted ones, but their performance is fundamentally limited by the diversity and quality of the training data. A model trained only on high-SNR data may not generalize well to noisy real-world measurements unless specific data augmentation strategies are used.

In conclusion, the choice of a reconstruction algorithm involves a critical trade-off. Model-based and hybrid approaches offer data-independence and high interpretability at the cost of slow, iterative inference. Learning-based methods provide unparalleled inference speed suitable for real-time systems, but this comes at the cost of a demanding training process that requires extensive data and computational resources.

## 6 Conclusions and Future Perspectives

### 6.1 Conclusion

This review has systematically examined the evolution and current state of reconstruction algorithms for broadband spectral modulation-based hyperspectral imaging systems. Through comprehensive analysis of model-based methods, hybrid frameworks, and learning-based approaches, several key insights emerge regarding the fundamental trade-offs and opportunities in this field.

Model-based reconstruction methods, encompassing sparsity-based, regularization-based, low-rank, and tensor decomposition approaches, provide strong theoretical foundations and interpretable results. These methods effectively exploit the intrinsic properties of hyperspectral data, including spectral correlations, spatial smoothness, and low-dimensional structures. However, their performance remains constrained by the accuracy of underlying mathematical assumptions and the computational burden of iterative optimization. The requirement for manual parameter tuning and limited adaptability to diverse imaging conditions further restrict their practical deployment in dynamic scenarios.

The emergence of hybrid constraint methods represents a natural progression toward more comprehensive modeling of hyperspectral data characteristics. By integrating complementary priors within unified optimization frameworks, these approaches achieve superior reconstruction quality compared to single-constraint methods. Nevertheless, the increased algorithmic complexity and expanded parameter space introduce new challenges in optimization convergence and computational efficiency. The delicate balance between different constraint terms requires careful calibration, often demanding domain expertise and extensive empirical validation.

Learning-based methods, particularly deep unfolding networks and transformer architectures, have demonstrated remarkable capabilities in capturing complex spatial-spectral relationships directly from data. These approaches offer substantial advantages in terms of reconstruction speed and adaptability to varying imaging conditions. The integration of physical knowledge through network architecture design, as exemplified by mask-guided transformers and physics-informed neural networks, suggests a promising direction for combining the strengths of model-based and data-driven paradigms. However, the generalization capability of these methods across different sensor configurations and scene types remains an open challenge, particularly when training data is limited or domain shifts are significant.

## *6.2 Future Perspectives*

Many learning-based reconstruction methods require GPU platforms to achieve high inference speeds. However, on embedded devices, these methods often experience significantly slower inference, limiting their application to post-acquisition processing. Model quantization techniques<sup>92</sup> offer a promising solution to enhance inference efficiency on resource constrained platforms. Advanced quantization strategies, such as learned step size quantization<sup>93</sup> and lattice vector quantization<sup>94, 95</sup>, have shown potential in maintaining accuracy while significantly reducing computational complexity. By enabling more compact and efficient network architectures, these techniques pave the way for real-time, on-device reconstruction during data acquisition, thereby expanding the range of practical and time-sensitive applications.

Beyond spectral domain reconstruction, recent studies have started to investigate the joint modeling of spectral characteristics and 3D geometric structures, aiming to achieve more comprehensive scene understanding in hyperspectral imaging. One representative direction is hyperspectral novel view synthesis (HNVS), which extends hyperspectral imaging beyond single-

view spectral recovery to the reconstruction of consistent, material-aware hyperspectral images from arbitrary viewpoints. This capability is crucial for applications that require both geometric reasoning and spectral fidelity, such as remote sensing, robotics, and medical diagnostics. By enabling spatial-spectral modeling across views, HNVS facilitates more accurate material discrimination, enhances 3D scene comprehension, and enables the reconstruction of hyperspectral images from unobservable viewpoints. One recent effort in this direction is HyperGS<sup>96</sup>, which integrates 3D Gaussian Splatting (3DGS)<sup>97</sup> with a latent hyperspectral representation to improve both rendering speed and spectral fidelity. While HyperGS represents a promising step toward efficient and accurate hyperspectral novel view synthesis, it inherits the high memory footprint of the original 3D Gaussian Splatting framework. To address this limitation, recent advances in 3DGS compression could be incorporated into HyperGS to reduce storage costs without compromising rendering quality<sup>98-100</sup>. In addition, techniques from point cloud compression may offer valuable insights for developing more compact and scalable hyperspectral 3DGS systems<sup>101-103</sup>.

### *Disclosures*

The authors declare no conflicts of interest.

### *Acknowledgments*

This work was supported by the National Natural Science Foundation of China (12274262, 62305144), the National Key Research and Development Program of China(2022YFC2807702) and the Jinan Government Project (202228042).

## References

1. W. Yang, K. Gao, J. Yuan, X. Li, W. Jiang, D. Li, Y. Liu, and B. Sun, Single-pixel computational spectrometer based on nonlinear spectrum modulation. *Optics express*, 33(3), 4269-4279 (2025).
2. J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar, Multispectral imaging using multiplexed illumination. in *2007 IEEE 11th International Conference on Computer Vision*, IEEE, (2007), 1-8.
3. J. Fang, K. Huang, R. Qin, Y. Liang, E. Wu, M. Yan, and H. Zeng, Wide-field mid-infrared hyperspectral imaging beyond video rate. *Nature Communications*, 15(1), 1811 (2024).
4. H. Meng, Y. Gao, X. Wang, X. Li, L. Wang, X. Zhao, and B. Sun, Quantum dot-enabled infrared hyperspectral imaging with single-pixel detection. *Light: Science & Applications*, 13(1), 121 (2024).
5. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1), 105-115 (2013).
6. J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, 1(2), 6-36 (2013).
7. H. H. Radamson, H. Zhu, Z. Wu, X. He, H. Lin, J. Liu, J. Xiang, Z. Kong, W. Xiong, and J. Li, State of the art and future perspectives in advanced CMOS technology. *Nanomaterials*, 10(8), 1555 (2020).
8. H. Li, F. Si, L. Xi, F. Lin, Y. Jiang, F. Liu, Y. Zeng, Y. Han, and K. Wu, Calibration of Short-Wave Infrared Spectrometer for Atmosphere Methane Monitoring. *Remote Sensing*, 17(5), 851 (2025).
9. G. Xia, Q. Liu, H. Zhou, and F. Yu, A non-linearity correction method of charge-coupled device array spectrometer. in *AOPC 2015: Optical Test, Measurement, and Equipment*, SPIE, (2015), 116-121.
10. H.-x. Zhou, H.-l. Qin, L.-p. Bai, Q.-c. Liu, X. Geng, and B.-j. Wang, Nonuniformity correction algorithm with nonlinear model for infrared focal plane arrays. *Infrared physics & technology*, 53(1), 10-16 (2010).
11. E. J. Candès, and M. B. Wakin, An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2), 21-30 (2008).

12. D. L. Donoho, Compressed sensing. *IEEE Transactions on information theory*, 52(4), 1289-1306 (2006).
13. A. Chakrabarti, and T. Zickler, Statistics of real-world hyperspectral images. in *CVPR 2011*, IEEE, (2011), 193-200.
14. A. Wagadarikar, R. John, R. Willett, and D. Brady, Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics*, 47(10), B44-B51 (2008).
15. M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics express*, 15(21), 14013-14027 (2007).
16. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600-612 (2004).
17. Z. Meng, J. Ma, and X. Yuan, End-to-end low cost compressive spectral imaging with spatial-spectral self-attention. in *European conference on computer vision*, Springer, (2020), 187-204.
18. Y. Cai, J. Lin, X. Hu, H. Wang, X. Yuan, Y. Zhang, R. Timofte, and L. Van Gool, Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (2022), 17502-17511.
19. L. Wang, C. Sun, Y. Fu, M. H. Kim, and H. Huang, Hyperspectral image reconstruction using a deep spatial-spectral prior. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2019), 8032-8041.
20. Z. Xiong, Z. Shi, H. Li, L. Wang, D. Liu, and F. Wu, Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections. in *Proceedings of the IEEE international conference on computer vision workshops*, (2017), 518-525.
21. T. Ranchin, and L. Wald, Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric engineering and remote sensing*, 66(1), 49-61 (2000).
22. R. M. Nguyen, D. K. Prasad, and M. S. Brown, Training-based spectral reconstruction from a single RGB image. in *European Conference on Computer Vision*, Springer, (2014), 186-201.
23. I. Choi, M. H. Kim, D. Gutierrez, D. S. Jeon, and G. Nam, High-quality hyperspectral reconstruction

- using a spectral prior. (2017).
24. J. Ma, X.-Y. Liu, Z. Shou, and X. Yuan, Deep tensor admm-net for snapshot compressive imaging. in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, (2019), 10223-10232.
  25. X. Miao, X. Yuan, Y. Pu, and V. Athitsos, l-net: Reconstruct hyperspectral images from a snapshot measurement. in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, (2019), 4059-4069.
  26. H. W. Engl, and C. W. Groetsch, *Inverse and ill-posed problems*, ed., Elsevier, (2014).
  27. X. Lin, Y. Liu, J. Wu, and Q. Dai, Spatial-spectral encoded compressive hyperspectral imaging. *ACM Transactions on Graphics (TOG)*, 33(6), 1-11 (2014).
  28. J. Oliver, W.-B. Lee, and H.-N. Lee, Filters with random transmittance for improving resolution in filter-array-based spectrometers. *Optics express*, 21(4), 3969-3989 (2013).
  29. A. S. Charles, B. A. Olshausen, and C. J. Rozell, Learning sparse codes for hyperspectral imagery. *IEEE Journal of Selected Topics in Signal Processing*, 5(5), 963-978 (2011).
  30. L. Zhang, W. Wei, Y. Zhang, C. Shen, A. Van Den Hengel, and Q. Shi, Dictionary learning for promoting structured sparsity in hyperspectral compressive sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12), 7223-7235 (2016).
  31. S. Yang, M. Wang, P. Li, L. Jin, B. Wu, and L. Jiao, Compressive hyperspectral imaging via sparse tensor and nonlinear compressed sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11), 5943-5957 (2015).
  32. C. V. Correa, H. Arguello, and G. R. Arce, Snapshot colored compressive spectral imager. *Journal of the Optical Society of America A*, 32(10), 1754-1763 (2015).
  33. S. Zhang, Y. Dong, H. Fu, S.-L. Huang, and L. Zhang, A spectral reconstruction algorithm of miniature spectrometer based on sparse optimization and dictionary learning. *Sensors*, 18(2), 644 (2018).
  34. Q. Wang, L. Ma, L. Tang, C. Li, and Y. Zhou, Hyperspectral compressive sensing reconstruction based on spectral sparse model. *Journal of Infrared and Millimeter Waves*, 35(6), 723-730 (2016).

35. D. Needell, and R. Ward, Near-optimal compressed sensing guarantees for total variation minimization. *IEEE Transactions on Image Processing*, 22(10), 3941-3949 (2013).
36. J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng, Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing. *IEEE Transactions on Image Processing*, 29(7889-7903 (2020).
37. H. Xiao, Z. Wang, X. Cui, L. Wang, and H. Yang, Hyperspectral Compressed Sensing Reconstruction Applying Multi-TV Collaboration. *Journal of Sensors*, 2023(1), 5186977 (2023).
38. Y. Hu, and M. Jacob, Higher degree total variation (HDTV) regularization for image recovery. *IEEE Transactions on Image Processing*, 21(5), 2559-2571 (2012).
39. P. Meza, I. Ortiz, E. Vera, and J. Martinez, Compressive hyperspectral imaging recovery by spatial-spectral non-local means regularization. *Optics express*, 26(6), 7043-7055 (2018).
40. J. Li, Q. Yuan, H. Shen, and L. Zhang, Hyperspectral image recovery employing a multidimensional nonlocal total variation model. *Signal Processing*, 111(230-248 (2015).
41. Y. Wang, L. Lin, Q. Zhao, T. Yue, D. Meng, and Y. Leung, Compressive sensing of hyperspectral images via joint tensor tucker decomposition and weighted total variation regularization. *IEEE Geoscience and Remote Sensing Letters*, 14(12), 2457-2461 (2017).
42. C. Chen, W. Li, E. W. Tramel, and J. E. Fowler, Reconstruction of hyperspectral imagery from random projections using multihypothesis prediction. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1), 365-374 (2013).
43. Y.-B. Sun, H. Li, M. Wu, Z.-B. Wu, J.-P. He, and Q.-S. Liu, Compressed sensing reconstruction of hyperspectral image using the graph sparsity regularized multiple measurement vector model. *Journal of Electronics & Information*, 36(12), 2942-2948 (2014).
44. Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, and D. 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).

45. A. L. Mur, P. Leclerc, F. Peyrin, and N. Ducros, Single-pixel image reconstruction from experimental data using neural networks. *Optics express*, 29(11), 17097-17110 (2021).
46. S. Takemoto, K. Naganuma, and S. Ono, Graph spatio-spectral total variation model for hyperspectral image denoising. *IEEE Geoscience and Remote Sensing Letters*, 19(1-5) (2022).
47. G. Martín, J. M. Bioucas-Dias, and A. Plaza, HYCA: A new technique for hyperspectral compressive sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5), 2819-2831 (2014).
48. M. Golbabaee, and P. Vandergheynst, Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery. in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Ieee, (2012), 2741-2744.
49. A. Gogna, A. Shukla, H. Agarwal, and A. Majumdar, Split Bregman algorithms for sparse/joint-sparse and low-rank signal recovery: Application in compressive hyperspectral imaging. in *2014 IEEE International Conference on Image Processing (ICIP)*, IEEE, (2014), 1302-1306.
50. K. Gunasheela, and H. Prasantha, Sparse Reconstruction of Hyperspectral Image using Bregman Iterations. in *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, IEEE, (2018), 1-7.
51. T. Xie, X. Kang, R. Dian, T. Wang, and L. Liu, Multishot Compressive Hyperspectral Imaging Based on Tensor Fibered Rank Minimization and Its Primal-Dual Algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17(4466-4477) (2024).
52. P. Xu, L. Liu, Y. Jia, H. Zheng, C. Xu, and L. Xue, A refinement boosted and attention guided deep FISTA reconstruction framework for compressive spectral imaging. *IEEE Transactions on Geoscience and Remote Sensing*, 61(1-12) (2023).
53. C. Zhou, J. Cao, H. Yao, H. Cui, H. Zhang, and Q. Hao, Enhanced ghost imaging reconstruction via a Chambolle-Pock inspired deep unfolding network. *Optics & Laser Technology*, 191(113323) (2025).
54. A. Mousavi, A. B. Patel, and R. G. Baraniuk, A deep learning approach to structured signal recovery. in *2015 53rd annual allerton conference on communication, control, and computing (Allerton)*, IEEE,

- (2015), 1336-1343.
55. X. Yuan, Generalized alternating projection based total variation minimization for compressive sensing. in *2016 IEEE International conference on image processing (ICIP)*, IEEE, (2016), 2539-2543.
  56. D. T. Eason, and M. Andrews, Total variation regularization via continuation to recover compressed hyperspectral images. *IEEE Transactions on Image Processing*, 24(1), 284-293 (2014).
  57. Y. Fu, Y. Zheng, I. Sato, and Y. Sato, Exploiting spectral-spatial correlation for coded hyperspectral image restoration. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016), 3727-3736.
  58. X. Yin, L. Su, X. Chen, H. Liu, Q. Yan, and Y. Yuan, Hyperspectral Image Reconstruction of SD-CASSI Based on Nonlocal Low-Rank Tensor Prior. *IEEE Transactions on Geoscience and Remote Sensing*, 62(1-15) (2024).
  59. C. Cao, J. Li, P. Wang, W. Jin, R. Zou, and C. Qi, Hyperspectral Reconstruction Method Based on Global Gradient Information and Local Low-Rank Priors. *Remote Sensing*, 16(24), 4759 (2024).
  60. X. Huang, B. Du, D. Tao, and L. Zhang, Spatial-spectral weighted nuclear norm minimization for hyperspectral image denoising. *Neurocomputing*, 399(271-284) (2020).
  61. C. He, Q. Cao, Y. Xu, L. Sun, Z. Wu, and Z. Wei, Weighted order-p tensor nuclear norm minimization and its application to hyperspectral image mixed denoising. *IEEE Geoscience and Remote Sensing Letters*, 20(1-5) (2023).
  62. Y. Liu, X. Yuan, J. Suo, D. J. Brady, and Q. Dai, Rank minimization for snapshot compressive imaging. *IEEE transactions on pattern analysis and machine intelligence*, 41(12), 2990-3006 (2018).
  63. H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8), 4729-4743 (2013).
  64. T. G. Kolda, and B. W. Bader, Tensor decompositions and applications. *SIAM review*, 51(3), 455-500 (2009).
  65. H. Xiang, B. Li, L. Sun, Y. Zheng, Z. Wu, J. Zhang, and B. Jeon, Hyperspectral compressive image

- reconstruction with deep tucker decomposition and spatial–spectral learning network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16(725-737) (2022).
66. J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, Nonlocal tensor sparse representation and low-rank regularization for hyperspectral image compressive sensing reconstruction. *Remote Sensing*, 11(2), 193 (2019).
67. M. Wang, D. Hong, Z. Han, J. Li, J. Yao, L. Gao, B. Zhang, and J. Chanussot, Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review. *IEEE Geoscience and Remote Sensing Magazine*, 11(1), 26-72 (2023).
68. J. Zhang, J. Erway, X. Hu, Q. Zhang, and R. Plemmons, Randomized SVD methods in hyperspectral imaging. *Journal of Electrical and Computer Engineering*, 2012(1), 409357 (2012).
69. S. Zhang, L. Wang, L. Zhang, and H. Huang, Learning tensor low-rank prior for hyperspectral image reconstruction. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (2021), 12006-12015.
70. Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, Hyperspectral computational imaging via collaborative Tucker3 tensor decomposition. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(1), 98-111 (2020).
71. Y. Chen, T.-Z. Huang, W. He, N. Yokoya, and X.-L. Zhao, Hyperspectral image compressive sensing reconstruction using subspace-based nonlocal tensor ring decomposition. *IEEE Transactions on Image Processing*, 29(6813-6828) (2020).
72. H. Fan, Y. Chen, Y. Guo, H. Zhang, and G. Kuang, Hyperspectral image restoration using low-rank tensor recovery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(10), 4589-4604 (2017).
73. B. Du, M. Zhang, L. Zhang, R. Hu, and D. Tao, PLTD: Patch-based low-rank tensor decomposition for hyperspectral images. *IEEE Transactions on Multimedia*, 19(1), 67-79 (2016).
74. M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, An augmented Lagrangian approach to the

- constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3), 681-695 (2010).
75. Y.-Q. Zhao, and J. Yang, Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1), 296-308 (2014).
76. Z. Wu, X. Chen, W. Shi, L. Chen, and S. Hu, Hyperspectral image reconstruction via block low-rank and three-dimension weighted total variation constraint. *IEEE Access*, 7(47698-47713 (2019).
77. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. J. I. J. o. S. A. i. I. T. Willett, Deep learning techniques for inverse problems in imaging. *1(1)*, 39-56 (2020).
78. H. Jiang, C. Xu, and L. Liu, Hyperspectral Image Restoration via Tensor Multimode Low-Rank Prior and Spatial-Spectral Smoothness Regularization. *Computers and Electrical Engineering*, 115(109133 (2024).
79. M. Wang, Q. Wang, J. Chanussot, and D. Hong,  $L_0$ - $L_1$  hybrid total variation regularization and its applications on hyperspectral image mixed noise removal and compressed sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9), 7695-7710 (2021).
80. H. Zeng, S. Huang, Y. Chen, H. Luong, and W. Philips, All of low-rank and sparse: A recast total variation approach to hyperspectral denoising. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16(7357-7373 (2023).
81. V. Monga, Y. Li, and Y. C. Eldar, Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2), 18-44 (2021).
82. J. Zhang, and B. Ghanem, ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2018), 1828-1837.
83. X. Han, B. Wu, Z. Shou, X.-Y. Liu, Y. Zhang, and L. Kong, Tensor FISTA-Net for real-time snapshot compressive imaging. in *Proceedings of the AAAI Conference on Artificial Intelligence*, (2020), 10933-10940.
84. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed

- machine learning. *Nature Reviews Physics*, 3(6), 422-440 (2021).
85. X. Hu, Y. Cai, J. Lin, H. Wang, X. Yuan, Y. Zhang, R. Timofte, and L. Van Gool, Hdnet: High-resolution dual-domain learning for spectral compressive imaging. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2022), 17542-17551.
  86. X. Zhang, Y. Zhang, R. Xiong, Q. Sun, and J. Zhang, Herosnet: Hyperspectral explicable reconstruction and optimal sampling deep network for snapshot compressive imaging. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2022), 17532-17541.
  87. Y. Cai, Y. Zheng, J. Lin, X. Yuan, Y. Zhang, and H. Wang, Binarized spectral compressive imaging. *Advances in Neural Information Processing Systems*, 36(38335-38346) (2023).
  88. Y. Cai, J. Lin, X. Hu, H. Wang, X. Yuan, Y. Zhang, R. Timofte, and L. Van Gool, Coarse-to-fine sparse transformer for hyperspectral image reconstruction. in *European conference on computer vision*, Springer, (2022), 686-704.
  89. J. Wang, K. Li, Y. Zhang, X. Yuan, and Z. Tao,  $S^2$ -transformer for mask-aware hyperspectral image reconstruction. *IEEE transactions on pattern analysis and machine intelligence*, (2025).
  90. Z. Yao, S. Liu, X. Yuan, and L. Fang, Specat: Spatial-spectral cumulative-attention transformer for high-resolution hyperspectral image reconstruction. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2024), 25368-25377.
  91. F. Luo, X. Chen, X. Gong, W. Wu, and T. Guo, Dual-window multiscale transformer for hyperspectral snapshot compressive imaging. in *Proceedings of the AAAI Conference on Artificial Intelligence*, (2024), 3972-3980.
  92. M. Nagel, M. Fournarakis, R. A. Amjad, Y. Bondarenko, M. Van Baalen, and T. Blankevoort, A white paper on neural network quantization. *arXiv preprint arXiv:2106.08295*, (2021).
  93. S. K. Esser, J. L. McKinstry, D. Bablani, R. Appuswamy, and D. S. Modha, Learned step size quantization. *arXiv preprint arXiv:1902.08153*, (2019).
  94. A. Tseng, J. Chee, Q. Sun, V. Kuleshov, and C. De Sa, Quip\#: Even better llm quantization with

- hadamard incoherence and lattice codebooks. *arXiv preprint arXiv:2402.04396*, (2024).
95. H. Xu, X. Wu, and X. Zhang, Multirate Neural Image Compression with Adaptive Lattice Vector Quantization. in *Proceedings of the Computer Vision and Pattern Recognition Conference*, (2025), 7633-7642.
96. C. Thirgood, O. Mendez, E. Ling, J. Storey, and S. Hadfield, Hypergs: Hyperspectral 3d gaussian splatting. *Proceedings of the Computer Vision and Pattern Recognition Conference*, 5970--5979 (2025).
97. B. Kerbl, G. Kopanas, T. Leimkhler, and G. Drettakis, 3D Gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4), 139--131 (2023).
98. Y. Chen, Q. Wu, W. Lin, M. Harandi, and J. Cai, Hac: Hash-grid assisted context for 3d gaussian splatting compression. in *European Conference on Computer Vision*, Springer, (2024), 422-438.
99. Y. Wang, Z. Li, L. Guo, W. Yang, A. Kot, and B. Wen, Contextgs: Compact 3d gaussian splatting with anchor level context model. *Advances in Neural Information Processing Systems*, 37(51532--51551 (2024).
100. H. Xu, X. Wu, and X. Zhang, 3DGS Compression with Sparsity-guided Hierarchical Transform Coding. *arXiv preprint arXiv:2505.22908*, (2025).
101. Y. He, X. Ren, D. Tang, Y. Zhang, X. Xue, and Y. Fu, Density-preserving deep point cloud compression. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2333--2342 (2022).
102. R. Song, C. Fu, S. Liu, and G. Li, Efficient hierarchical entropy model for learned point cloud compression. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14368--14377 (2023).
103. H. Xu, X. Zhang, and X. Wu, Fast Point Cloud Geometry Compression with Context-based Residual Coding and INR-based Refinement. *European Conference on Computer Vision*, 270--288 (2024).

## **Caption List**

**Fig. 1** Development Roadmap of Broadband Spectral Modulation-based HSI.

**Fig. 2** (a) An active spectral modulation system using tunable laser illumination. (b) A passive spectral modulation system employing optical filters.

**Fig. 3** Hyperspectral data process based on spectral modulation.

**Fig. 4** Hyperspectral reconstruction algorithm classification.

**Fig. 5** Comparison of hyperspectral compressive imaging reconstruction algorithms based on sparse representation. (a) Comparison of the spectral signature for the highlighted spatial points, measured with a commercial spectrometer and the SCCSI reconstructions using two, four, and six colors<sup>32</sup>. (b) Spectral Sparsity-Based Spectral Reconstruction and Curve Removal for Corn and Alfalfa<sup>34</sup>. (c) Comparison of Ground Truth and Reconstructed LED Spectra Using Dictionary Learning with L1-Norm Minimization<sup>33</sup>.

**Fig. 6** Comparison of hyperspectral compressive imaging reconstruction algorithms based on regularization. (a) Reconstruction results in Ref. 18. (b) Reconstruction results in Ref. 15. (c) Reconstruction results in Ref. 23.

**Fig. 7** Low-rank matrix formation from hyperspectral images via patch extraction, showing both spectral and spatial non-local low-rank characteristics utilized in reconstruction algorithms<sup>57</sup>.

**Fig. 8** Dual low-rank approximation framework demonstrating spectral band correlation and spatial non-local patch similarity exploitation in hyperspectral image reconstruction<sup>62</sup>.

**Fig. 9** The flowchart of dual low-rank reconstruction algorithm for coded hyperspectral image restoration by using spectral low-rank and spatial non-local lowrank regularizations<sup>57</sup>.

**Fig. 10** Tensor Decomposition of a Third-Order Tensor<sup>67</sup>. (a) CP decomposition. (b) Tucker decomposition. (c) t-SVD.

**Fig. 11** The flowchart of the NTSRLR algorithms<sup>66</sup>, which consists of two steps: sensing and reconstruction. First, it acquires the compressive measurement  $y$  by a random sampling matrix  $H$ . Second, NTSRLR recovers an HSI from the measurements  $y=Hx$ .

**Fig. 12** The flowchart of the PLTD algorithm<sup>73</sup>, illustrating patch-based tensor representation, nonlocal similarity clustering, and low-rank tensor decomposition.

**Fig. 13** Comparison of hyperspectral compressive imaging reconstruction algorithms based on tensor decomposition<sup>71</sup>.

**Fig. 14** (a) False color image of the original hyperspectral image<sup>76</sup>. (b) The sum of gradients at each pixel position across all spectral bands. (c) The top 1% pixels with maximum gradients. (d) Blocks using 3DWTV regularization.

**Fig. 15** The flowchart of the MLRTAS Algorithm for HSI compressive sensing reconstruction<sup>78</sup>, 4D tensor construction from similar cubic patches with nonlocal tensor multi-mode low-rank properties and spatial-spectral pixel-wise smoothness constraints for HSI compressive sensing reconstruction.

**Fig. 16** Illustration of the l0-l1 hybrid total variation regularization<sup>79</sup>, alternating BDTV and SSTV operations for balancing edge preservation and noise suppression.

**Fig. 17** Sparsity and low-rank characteristics of hyperspectral image gradient domains<sup>80</sup>: evidence for joint regularization approaches combining  $\ell_1$  and tensor nuclear norm constraints.

**Fig. 18** Architectures of partial reconstruction network. (a) E2E network. (b) Deep unfolding network. (c) Transformer-based network.

**Fig. 19** The architecture of ISTA-Net ISTA-Net architecture for hyperspectral compressive imaging: deep unfolding of iterative ISTA algorithm with learnable nonlinear transforms and soft-thresholding operations across multiple phases<sup>82</sup>.

**Fig. 20** The architecture of Tensor FISTA-Net, maintaining 3D tensor format across iterative phases to preserve spatial-spectral correlations and reduce information loss in high-dimensional hyperspectral reconstruction<sup>83</sup>.

**Fig. 21** The architecture of ReAttFISTA-Net with attention-guided fusion and multi-component reconstruction design<sup>52</sup>.

**Fig. 22** The architecture of HDNet<sup>85</sup>, a dual-domain learning network for spectral compressive imaging.

**Fig. 23** (a) U-net architecture used in the reconstruction stage of  $\lambda$ -Net. (b) The attention module of  $\lambda$ -Net<sup>25</sup>.

**Fig. 24** The architecture of HerosNet, exemplifying advanced CNN-based hyperspectral reconstruction with integrated DGDM and HFIM modules<sup>86</sup>.

**Fig. 25** The architecture of Binarized Convolutional Modules, featuring encoder-decoder structure with binarized spectral-redistribution operations<sup>87</sup>.

**Fig. 26** The architecture of MST, demonstrating encoder-decoder framework with Multi-head Self-Attention Blocks (MSAB), spectral-wise self-attention (S-MSA), and mask-guided mechanism<sup>18</sup>.

**Fig. 27** The architecture of CST, demonstrating coarse-to-fine framework with sparsity estimator, SAHAB (Sparse Attention Hybrid Attention Block), and SAH-MSA mechanism<sup>88</sup>.

**Fig. 28** The architecture of the spatial-spectral transformer, dual-stream Transformer with separate spatial and spectral self-attention modules and integrated mask processing<sup>90</sup>.

**Fig. 29** The dual-window multiscale multi-head self-attention in DWMT<sup>91</sup>.

**Fig. 30** Comparison of reconstruction algorithms of some algorithms<sup>90</sup>.

**Table 1** Comparison of Regularization Functionals.

**Table 2** Optimization Algorithms for Regularization.

**Table 3** Comparison of Low-rank Method Variants.

**Table 4** Computational complexity comparison of tensor decomposition methods.

**Table 5** Comparison of hybrid reconstruction algorithms for spectral modulation-based hyperspectral imaging.

**Table 6** Comparative Analysis of Resource Consumption for Different Algorithm Categories.