

# 编码孔径快照光谱成像重构算法综述

马祥天<sup>1)</sup> 王立志<sup>1)</sup> 黄华<sup>2)</sup>

<sup>1)</sup>(北京理工大学计算机学院 北京 100081)

<sup>2)</sup>(北京师范大学人工智能学院 北京 100875)

**摘要** 光谱图像含有丰富的空间和光谱信息,能够反映目标的组成、结构和材料特性,在航天遥感、医疗诊断和机器视觉等领域具有重要的应用价值.近年来,光谱成像技术作为热点研究领域受到广泛关注.传统光谱成像技术采用沿空间维度或光谱维度扫描的方式,依次获取待测物体表面的光谱信息.由于曝光时间较长,传统光谱成像技术不适用于拍摄动态场景.编码孔径快照光谱成像(Coded Aperture Snapshot Spectral Imaging, CASSI)是光谱成像的前沿技术方案,能够从单次曝光中快速获取动态场景的光谱图像,其包括两个阶段:对高维光谱图像的“编码降维采集”和对低维观测图像的“解码升维重构”.CASSI的早期研究聚焦于“编码降维采集”,通过物理系统设计提升图像编码的有效性,包括编码模板设计和双相机系统设计.目前,CASSI的“编码降维采集”物理系统趋于稳定,其“解码升维重构”决定了光谱成像的质量和效率.本文综述CASSI的重构算法,首先介绍CASSI的物理系统和前向模型,详细描述物理系统的组成元件和硬件参数,推导CASSI前向模型的数学表达;其次梳理CASSI重构的特点和挑战,其挑战主要存在于系统前向模型、先验表示模型、算法灵活性、算法复杂度、实物数据集等方面;之后重点归纳重构算法的研究现状,包括基于优化模型的重构算法和基于深度学习的重构算法.基于优化模型的重构算法利用凸优化模型求解线性逆问题,结合平滑、稀疏、低秩等手工设计的先验表示模型降低逆问题的欠定性;基于深度学习的重构算法利用数据驱动的方式建立先验表示模型,结合端到端全网络、深度展开、即插即用等框架求解重构图像.接着比较主流算法的重构质量和计算效率,以峰值信噪比、结构相似度、光谱角制图为重构质量的评价指标,以模型参数量、浮点计算量为计算效率的评价指标.最后讨论现有工作的不足和未来研究趋势,指出当前仍未解决的领域痛点,展望进一步的研究方向,为本领域开拓创新提供参考.

**关键词** 快照光谱成像;编码孔径;图像重构;优化模型;深度学习

**中图法分类号** TP391 **DOI号** 10.11897/SP.J.1016.2024.00190

## A Survey of Reconstruction Algorithms for Coded Aperture Snapshot Spectral Imaging

MA Xiang-Tian<sup>1)</sup> WANG Li-Zhi<sup>1)</sup> HUANG Hua<sup>2)</sup>

<sup>1)</sup>(School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081)

<sup>2)</sup>(School of Artificial Intelligence, Beijing Normal University, Beijing 100875)

**Abstract** Spectral images contain a wealth of spatial and spectral information, enabling them to effectively reflect an object's composition, structure, and material properties. They have significant application value in aerospace remote sensing, medical diagnosis, and machine vision. In recent years, spectral imaging technology has got significant attention and emerged as a prominent research area. Conventional spectral imaging techniques scan the spatial or spectral dimensions, enabling the sequential acquisition of spectral information from the object's surface. Due to the extended exposure time, these techniques are unsuitable for capturing dynamic scenes. Coded Aperture Snapshot Spectral Imaging (CASSI) is a cutting-edge technique for spectral imaging that allows for the rapid acquisition of spectral images of dynamic scenes from a single exposure. It

consists of “encoding dimension-reduction collection” of high-dimensional spectral images and “decoding dimension-increase reconstruction” of low-dimensional measurements. Early research on CASSI primarily focused on the “encoding dimension-reduction collection” stage, aiming to enhance the effectiveness of image encoding through physical system design, including the design of coded aperture and dual-camera system. At present, the physical system for the “encoding dimension-reduction collection” stage has become fixed. The “decoding dimension-increase reconstruction” stage plays a crucial role in determining the quality and efficiency of spectral imaging. This paper presents a comprehensive overview of CASSI reconstruction algorithms, which aims to provide readers with a detailed understanding of the inner workings and intricacies of the various algorithms. First, we introduce the physical system and forward model of CASSI, providing a detailed description of the components and hardware parameters of the physical system and deriving the mathematical expression of the CASSI forward model. Second, we outline the characteristics and challenges of CASSI reconstruction, which mainly contain the forward model, prior representation model, algorithm flexibility, algorithm complexity, and real-world datasets. Next, we summarize the current research status of reconstruction algorithms, including optimization-based and learning-based reconstruction algorithms. The optimization-based reconstruction algorithms employ convex optimization models to address the challenging linear inverse problems effectively. These algorithms apply crafted prior representation models, including but not limited to smoothness, sparsity, and low-rank, to tackle the inverse problem’s inherent ill-posedness. The learning-based reconstruction algorithms take a different data-driven approach to establishing prior representation models. These algorithms leverage the power of deep learning frameworks, such as end-to-end networks, deep unfolding, and plug-and-play, to effectively solve the reconstruction problem. With the capabilities of deep learning, these algorithms can learn and fit the underlying patterns and structures within the data, leading to enhanced reconstruction performance. By contrasting the optimization-based and learning-based methods, we comprehensively understand the diverse methodologies of CASSI reconstruction to explore the inner workings and potential benefits and limitations. Furthermore, a thorough comparison is conducted utilizing various evaluation metrics to assess mainstream algorithms’ reconstruction quality and computational efficiency. These metrics include peak signal-to-noise ratio, structural similarity, and spectral angle mapping for evaluating the reconstruction quality. In addition, model parameter count and floating-point operations are utilized to measure the computational efficiency. Finally, the shortcomings of existing work and future research trends are discussed. The unresolved pain points in the current field are identified, and potential research directions are highlighted, providing valuable insights for further innovation and advancement.

**Keywords** snapshot spectral imaging; coded aperture; image reconstruction; optimization model; deep learning

## 1 引言

光谱图像能够同时反映物体的空间结构和光学特性,具有“空谱一体”的特点,在航天遥感<sup>[1]</sup>、医疗诊断<sup>[2]</sup>、机器视觉<sup>[3]</sup>等领域具有重要应用价值。光谱图像具有高维数据结构,传统的采集方案通常采用基于扫描的方式,包括掸扫<sup>[4]</sup>、推扫<sup>[5]</sup>和光谱扫描<sup>[6-7]</sup>

等,通过多次曝光获取完整的数据立方体。这种方式牺牲了时间分辨率,无法捕获动态场景的光谱信息。

编码孔径快照光谱成像<sup>[8-10]</sup>(Coded Aperture Snapshot Spectral Imaging, CASSI)结合光学信息调制和数字图像处理,大幅提升了信息捕获效率。该项技术将光谱获取过程划分为两个阶段:首先将场景信息“降维”采集得到观测图像,之后通过计算、校正等步骤将观测图像“升维”重构得到原始光谱图

像。基于上述过程, CASSI 能够在单次曝光中快速捕获场景光谱信息, 可应用于动态场景的光谱成像。在“降维”采集方式确定的情况下, CASSI 最终获取光谱图像的质量和效率取决于“升维”重构算法。如何高质量、高效率地将观测图像“升维”重构至原始图像, 是 CASSI 的重要科学问题, 也是 CASSI 的瓶颈问题。

CASSI 重构是指从 CASSI 观测图像中预测原始光谱图像, 这是经典的压缩感知<sup>[11-13]</sup>问题, 具有高度欠定性。求解欠定问题的主流方案是刻画准确的图像先验表示模型, 构建欠定问题求解的约束条件。根据不同的理论框架, 一般将其归纳为基于优化模型的重构算法和基于深度学习的重构算法, 基于优化模型的重构算法通常以手工设计的先验表示模型作为约束条件, 建立凸优化问题并迭代求解。基于深度学习的重构算法通常以数据驱动的方式建立图像的先验表示模型, 约束欠定问题的解空间。

目前尚无针对 CASSI 重构算法的综述文献。为帮助学者快速且有针对性地了解本领域研究现状, 并进一步开拓创新, 本文结合最新研究成果, 对 CASSI 重构算法进行梳理和评价。具体地说, 本文归纳了现有 CASSI 重构算法, 包括基于优化模型的重构算法和基于深度学习的重构算法, 同时列举了常用数据集和评价指标, 汇总了主流算法的源代码, 并分析了主流算法的重构质量和计算效率, 此外讨论了现有工作的不足, 展望了未来研究趋势。图 1 展示了本文脉络和 CASSI 重构算法的发展历程。

## 2 编码孔径快照光谱成像系统

CASSI 由杜克大学 David Brady 教授团队提出<sup>[8]</sup>, 是一种线性成像系统, 其硬件组成包括物镜、编码模板、中继透镜、滤光片、色散棱镜和全色相机。图 2 是作者团队搭建的 CASSI 实物系统。其中, 编码模板由随机二进制图案构成, 共包含  $300 \times 300$  个基本单元, 基本单元尺寸为  $10 \mu\text{m} \times 10 \mu\text{m}$ ; 滤光片允许通过的波长范围是  $450 \text{ nm} \sim 700 \text{ nm}$ ; 色散棱镜为双阿米西棱镜; 全色相机型号为 PointGrey FL3-U3-13Y3M, 分辨率为  $1280 \times 1024$ , 像素间距为  $4.8 \mu\text{m} \times 4.8 \mu\text{m}$ , 编码模板中各基本单元被分别映射到全色相机中的  $2 \times 2$  个像素。



图 2 CASSI 实物系统示意图

图 3 展示了 CASSI 的成像过程: 场景光信号首先通过物镜, 被一次投影至编码模板平面, 由编码模板沿空间维度进行 0-1 二值调制; 接着, 经过调制的光信号通过中继透镜, 被二次投影至相机的焦平面阵列 (Focal Plane Array, FPA), 在投影过程中, 由色散棱镜沿光谱维度进行色散调制; 最后, 光信号被传感器接收, 形成观测图像。由成像过程可知, CASSI 能够通过单次曝光获取完整的光谱图像, 具有快速捕获动态场景光谱信息的显著优势。

针对 CASSI 的早期研究聚焦于硬件系统设计, 通过改变系统前向模型, 提升“降维”压缩过程的效能, 以减小欠定性对重构质量的影响。硬件系统设计主要包括编码模板设计<sup>[14-23]</sup>和双相机系统设计<sup>[24-28]</sup>, 编码模板设计致力于优化观测图像的混叠方式, 双相机系统设计致力于增加观测图像的采样数量。

编码模板设计是指在标准前向模型的基础上, 对传统随机编码方式进行优化, 以减少调制过程中的信息损失。Arguello 等人<sup>[14-16]</sup>将编码模板优化转化为最佳滤波系数求解问题, 通过使观测图像更好地满足有限等距性质 (Restricted Isometry Property,

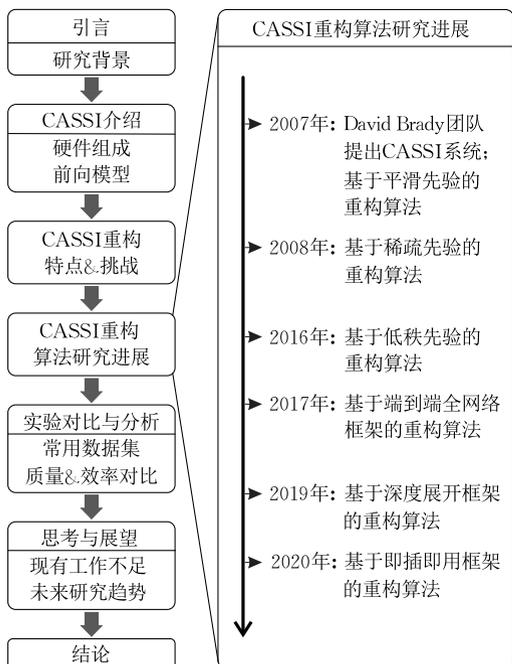


图 1 本文组织结构图

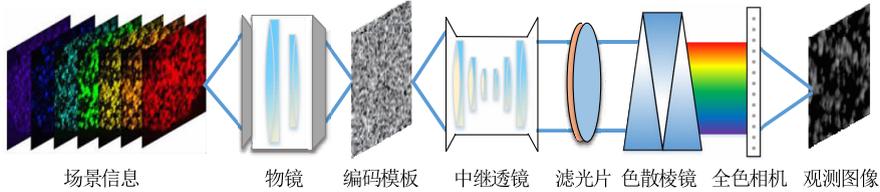


图 3 CASSI 成像过程示意图

RIP), 提升重构精度. Rueda 等人<sup>[17]</sup> 使用彩色编码模板替代传统模板, 同时从空间维度和光谱维度对场景信息进行调制. Gelvez 等人<sup>[18]</sup> 通过最大化编码模板中半透明元素之间的距离求得最优编码模板, 再利用矩阵补全算法重构光谱图像. Pinilla 等人<sup>[19]</sup> 基于蒙特卡洛模拟确定编码模板的最佳透射率, 并以此改进模板设计. Ma 和 Zhang 等人<sup>[20-21]</sup> 引入侧光采集目标场景的灰度图像以提供先验信息, 并使用随机阈值方法自适应生成编码模板. Shen 等人<sup>[22]</sup> 以电子光罩替代传统的编码模板, 通过控制光罩上的二元频闪实现可变编码. Tao 等人<sup>[23]</sup> 基于迭代优化策略, 利用遗传算法和梯度下降算法联合优化编码模板和字典.

双相机系统设计是指在标准前向模型的基础上, 对传感矩阵进行扩充, 增加额外的采样模型, 以弥补原传感矩阵造成的信息损失. Wang 等人<sup>[24-25]</sup> 增加一个与 CASSI 同型号的全色相机作为分支, 对原始图像的光谱维度直接积分, 解决了因三维体素位移导致的空间模糊. Yuan 等人<sup>[26]</sup> 和 Wang 等人<sup>[27]</sup> 提出了包含 CASSI 和 RGB 相机的混合双相机系统, 通过融合 CASSI 分支获取的场景编码信息和 RGB 分支获取的颜色信息, 实现高保真度的图像重构. Bacca 等人<sup>[28]</sup> 利用单像素相机提供的光谱维度信息, 和三维编码孔径快照光谱成像系统提供的空间维度信息, 重构高分辨率的光谱图像.

近年来, CASSI 的“降维”压缩方式已基本固定, 研究重点从硬件系统设计逐渐转移到“升维”重构算法设计. 重构算法设计的重点是对逆成像问题的求解. 在求解过程中, 需要建立准确的系统前向模型, 其建模过程如下:

给定原始图像  $\mathbf{x}$  的张量形式  $\mathcal{X} \in \mathbb{R}^{H \times W \times C}$  和编码模板  $\phi \in \mathbb{R}^{H \times W}$ , 其中  $H \times W$  表示空间尺寸,  $C$  表示谱带数, 光谱图像的空间调制过程如式(1)所示:

$$\mathcal{X}'_{h,w,c} = \phi_{h,w} \cdot \mathcal{X}_{h,w,c} \quad (1)$$

其中,  $h \in [1, H]$  和  $w \in [1, W]$  表示像素点的空间坐标,  $c \in [1, C]$  表示像素点对应的谱带序号.

在第二次投影中, 经过空间调制的光谱图像

$\mathcal{X}'_{h,w,c}$  通过色散棱镜, 沿空间竖直方向(或水平方向)分散, 并最终积分至相机传感器平面. 该过程中, 相邻谱带图像在传感器平面的投影相差 1 个像素,  $c$  谱带上空间位置为  $(h+c, w)$  的点被积分至传感器平面  $(h, w)$  像素位置. 如式(2)所示:

$$\mathcal{Y}_{h,w} = \sum_{c=1}^C \mathcal{X}'_{h+c,w,c} = \sum_{c=1}^C \phi_{h+c,w} \cdot \mathcal{X}_{h+c,w,c} \quad (2)$$

其中  $\mathcal{Y} \in \mathbb{R}^{(H+C-1) \times W}$  表示观测图像  $\mathbf{y}$  的张量形式. 为简化数学描述, 通常将上述模型表示为矩阵向量形式, 如式(3)所示:

$$\mathbf{y} = \Phi \mathbf{x} \quad (3)$$

其中,  $\mathbf{x}$  和  $\mathbf{y}$  分别表示原始光谱图像和观测图像,  $\Phi$  表示传感矩阵.

基于式(3)所示的前向模型, 给定传感矩阵  $\Phi$ , 从观测图像  $\mathbf{y}$  中估计原始光谱图像  $\mathbf{x}$  是 CASSI 重构的基本问题, 其解决方案一般基于两种思想, 即通过构建约束条件求解凸优化问题, 或利用神经网络实现数据拟合. 本文将结合最新研究进展, 对重构算法进行详细介绍和归纳.

### 3 CASSI 图像重构的特点和挑战

图像重构是图像复原领域(如图像去噪、去模糊、超分辨率等)的经典问题, 针对自然图像的图像复原算法已有多年工作积累. 然而, 针对自然图像的图像复原算法不适用于 CASSI 重构, 其原因在于 CASSI 重构需充分考虑光学成像模型, 且成像过程具有空间随机调制、信息混叠程度高等特点. 近年来, 研究人员结合 CASSI 重构任务的特点提出了诸多挑战, 归纳如下:

(1) 仿真实验难以精确模拟光学成像过程

CASSI 的光学元件多、噪声引入点多, 成像过程具有不确定性, 其前向模型在仿真实验中难以准确描述. 例如, 仿真实验中, 通常忽略衍射效应、像差、畸变等, 将点扩散函数理想化为冲激函数, 在此前提下建立的 CASSI 前向模型, 与真实光学系统存在差异, 进而限制了仿真算法在实际系统中的应用.

### (2) 先验表示模型难以适配复杂场景

先验表示模型是对图像一般性规律的数学描述,通过手工设计或数据驱动的方式构建,用作优化问题的约束条件.由于手工设计方式调参难度较高,数据驱动方式过度依赖数据分布,现有先验表示模型通常难以适配复杂多样的实际场景.以谱间相关性为例,现实世界的光源种类多样,在不同光源下拍摄的场景图像,其光谱维度的相关性不同,易导致先验信息失效.

### (3) 重构算法对前向模型的灵活性不足

CASSI 采用随机编码方式对空间信息进行调制,不同光学系统之间可能采用不同的编码模板,使得 CASSI 前向模型具有多样性.同时,对系统的轻微扰动可能导致 CASSI 前向模型发生变化.上述原因易导致针对特定前向模型建立的映射关系发生错误.

### (4) 重构质量和计算效率需要权衡

CASSI 提升了光谱图像采集效率,同时也引入了计算重构过程,该过程比较耗时.基于优化模型的重构算法因迭代次数多导致重构速度慢,无法适用计算资源受限或有时效性要求的应用场景.相比之下,深度学习方法在重构速度上更具优势,但对硬件计算性能要求较高,在移动端设备上的应用存在限制.

### (5) 缺少实物数据集和统一评价标准

目前,领域内缺少同时包含 CASSI 观测图像和光谱图像的实物数据集.由于光学系统模拟困难,利用仿真方法从光谱图像中合成的观测图像,无法等同于真实光学系统中的观测图像,使得训练数据不够准确.此外,对模型性能的评价标准不够统一.例如,数据集的划分方式不同,用于测试的图像区域不同,峰值信噪比的计算方式不同等.

## 4 基于优化模型的重构算法研究进展

基于优化模型重构光谱图像,其核心思想是利用凸优化模型求解线性逆问题.由于该问题是欠定的,理论上无穷多个解满足系统前向模型的约束,使得原始图像求解困难.为缩小解空间,传统方法是结合图像的一般性规律,通过手工设计的方式建立先验表示模型,作为优化问题的正则项.

给定如式(3)所示的系统前向模型,以求解最优光谱图像  $\mathbf{x}$  为目标建立的数学优化模型如式(4)所示:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \gamma \cdot P(\mathbf{x}) \quad (4)$$

其中,  $P(\cdot)$  为优化模型的正则项,即先验表示模型,  $\gamma$  表示正则项系数.

在光谱图像重构领域,主流的先验信息类型包括平滑先验、稀疏先验和低秩先验,三者从不同角度建模图像结构,指导重构过程.

### 4.1 基于平滑先验的重构算法

图像的平滑先验,是指一般图像中边缘和噪声对应区域是高频非平滑的,其余图像区域是低频平滑的.基于平滑先验, Rudin 等人<sup>[29]</sup>最早提出了全变差(Total Variation, TV)正则化,用于在保持图像边缘锐度的前提下去除噪声.具体地说,TV 是由水平方向和垂直方向的有限差分离散表达,利用局部差分算子  $\nabla$ , 计算相邻像素点之间的梯度.在光谱重构领域,通常以图像的 TV 表征空间非平滑特性,其数学表达式如式(5)所示:

$$TV(\mathbf{x}) = \sum_c \sum_{h,w} \|\nabla \mathcal{X}_{h,w,c}\|_p \quad (5)$$

其中,  $p$  的取值为 1 时表示空间各向异性 TV,  $p$  的取值为 2 时表示空间各向同性 TV,  $\nabla \mathcal{X}_{h,w,c}$  表示像素点在水平方向和垂直方向的差分运算,如式(6)所示:

$$\nabla \mathcal{X}_{h,w,c} = \begin{bmatrix} \mathcal{X}_{h+1,w,c} - \mathcal{X}_{h,w,c} \\ \mathcal{X}_{h,w+1,c} - \mathcal{X}_{h,w,c} \end{bmatrix} \quad (6)$$

在满足约束条件  $\mathbf{y} = \Phi \mathbf{x}$  的前提下,最小化光谱图像的 TV, 能够得到平滑度较高的重构结果.其优化模型如式(7)所示:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \gamma \cdot TV(\mathbf{x}) \quad (7)$$

针对式(7)所示的优化模型,其主流求解方法包括二步迭代收缩阈值(Two-step Iterative Shrinkage/Thresholding, TwIST)算法<sup>[30]</sup>、交替方向乘子法(Alternating Direction Method of Multipliers, ADMM)和广义交替投影(Generalized Alternating Projection, GAP)算法<sup>[31]</sup>等.

TwIST 算法由 Bioucas-Dias 等人<sup>[30]</sup>提出,基于梯度实现.该算法结合迭代加权收缩(Iterative Reweighted Shrinkage, IRS)算法<sup>[32]</sup>处理病态问题的能力和迭代收缩阈值(Iterative Shrinkage/Thresholding, IST)算法<sup>[33]</sup>的去噪性能,将图像重构作为广义非线性无约束优化问题求解.基于 TwIST 算法, Liu 等人<sup>[34]</sup>从残差数据中重构残差图像并将其叠加在模型输出上,解决 CASSI 双相机重构问题. Kittle 等人<sup>[35]</sup>针对单帧 CASSI 中复杂场景重构困难的问题,首次提出多帧 CASSI 并采用 TwIST 算法求解优化模

型. TwIST 算法具有比 IST 算法更快的收敛速度.

ADMM 算法是增广拉格朗日乘子法的改进, 基于分裂原理实现. Zhang 等人<sup>[36]</sup>利用此算法求解 TV 正则化约束下的 CASSI 双相机重构问题, 同时提出一套基于 GPU 加速的并行实现策略. Zhang 等人<sup>[37]</sup>提出一种非零均值噪声的建模方案并在 TV 正则化约束下重构光谱图像, 在重构过程中利用 ADMM 算法交替优化光谱保真项和噪声均值至收敛. Liang 等人<sup>[38]</sup>提出各向异性空间-光谱 TV 正则化描述图像的分段平滑性, 并提出一个张量分解模型判定 CASSI 双相机系统中光谱图像和全色图像的空间结构一致性, 最终采用 ADMM 算法对优化问题进行求解. Vargas 等人<sup>[39]</sup>提出以图像融合的方式重构光谱图像, 其中, 融合的对象是高空间分辨率低光谱分辨率的光谱图像, 和低空间分辨率高光谱分辨率的光谱图像. 该方法结合 TV 正则化和稀疏先验建立优化问题, 并利用 ADMM 算法求解. Wang 等人<sup>[40]</sup>结合光谱图像的全局相关性和非局部自相似性, 以非局部低秩先验和三维加权全变差作为正则项建立优化模型, 并使用 ADMM 算法求解. 其中, 低秩先验用于表征空间维度、光谱维度和非局部区域之间的相关性, 全变差用于表征光谱图像的局部非平滑特性. ADMM 算法具有比 TwIST 算法更快的收敛速度和重构效率.

GAP 算法是 ADMM 算法的特例, 由 Yuan<sup>[31]</sup>等提出, 在广义交替投影框架下解决 TV 约束的 CASSI 重构问题. 该算法通过引入欧式投影简化参数更新, 计算复杂度低于 ADMM 算法.

Split-Bregman 迭代算法基于分裂原理实现, 按照 Bregman 距离展开求解, 是 ADMM 算法的缩放形式. Greer 等人<sup>[41]</sup>结合空间维度和光谱维度的平滑先验建立正则项, 构建优化模型解决 CASSI 重构问题, 并采用 Split-Bregman 迭代求解. 该算法具有比 Bregman 迭代算法更低的计算量.

快速全变差去卷积算法 (Fast Total Variation deconvolution, FTVd)<sup>[42]</sup>基于分裂原理实现, 将有限元法<sup>[43]</sup>拓展为半二次型模型并连续化处理惩罚系数以加速收敛. 该算法能够处理各向同性 TV 约束下的 CASSI 重构问题, 一次迭代需使用三次快速傅里叶变换, 具有比半二次型模型更低的计算量.

基于 TV 正则化的重构算法能够保持图像在空间维度上的分段平滑性, 参数设置简单且计算效率高. 然而, 此类方法易在光滑区域产生阶梯效应, 易造成高频细节损失, 导致图像边缘过于平滑.

## 4.2 基于稀疏先验的重构算法

根据压缩感知原理<sup>[11-13]</sup>, 当信号足够稀疏时, 可通过求解稀疏优化问题从信号的稀疏表示中重构原始信号. 在 CASSI 重构中, 光谱图像具有三维数据结构, 能够由一组系数和过完备字典稀疏表示, 以系数的 1 范数作为正则项构建稀疏优化模型, 是 CASSI 重构领域的常用方法.

在最优化理论中, 正则化稀疏模型包含多个类别, 包括经典 LASSO (Least Absolute Shrinkage and Selection Operator) 模型、近似无偏稀疏模型、自动组效应稀疏模型、预设组效应稀疏模型等<sup>[44]</sup>. 其中, 经典 LASSO 模型由 Tibshirani 于 1996 年提出, 是一种凸优化模型, 具备同时处理变量选择和参数估计的优势, 常用于 CASSI 重构.

假设原始图像  $x \in \mathbb{R}^u$  ( $u = H \times W \times C$ ) 在特定变换域内是稀疏的, 给定一组表示基  $\Psi \in \mathbb{R}^{u \times u}$ , 则图像  $x$  的稀疏表示如式(8)所示:

$$x = \Psi t \quad (8)$$

其中,  $t$  是图像  $x$  在表示基  $\Psi$  下的系数.

针对 CASSI 的标准前向模型  $y = \Phi x$ , 给定传感矩阵  $\Phi \in \mathbb{R}^{v \times u}$  ( $v = (H + C - 1) \times W$ ) 满足  $v \ll u$ , 观测图像  $y \in \mathbb{R}^v$  可通过传感矩阵  $\Phi$ 、表示基  $\Psi$  和稀疏表示  $t$  计算得到, 如式(9)所示:

$$y = \Phi \Psi t \quad (9)$$

其中, 稀疏表示  $t$  的稀疏程度越高, 图像重构效果越好. 因此, 从观测图像  $y$  中重构原始图像  $x$ , 等价于对稀疏表示  $t$  的优化问题. 基于稀疏先验的优化模型如式(10)所示:

$$\hat{x} = \arg \min_x \|y - \Phi \Psi t\|_2^2 + \gamma \cdot \|t\|_1 \quad (10)$$

针对式(10)所示的优化模型, 其主流求解方法包括梯度投影稀疏重构 (Gradient Projection for Sparse Reconstruction, GPSR) 算法<sup>[45]</sup>、近似消息传递 (Approximate Message Passing, AMP) 算法、迭代阈值算法<sup>[46]</sup>和正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 算法<sup>[47]</sup>等.

GPSR 算法基于梯度实现, 通过负梯度方向上的投影计算搜索方向, 解决非光滑无约束凸优化问题. GPSR 算法由 Figueiredo 等人<sup>[45]</sup>提出, 用于求解从二维小波变换和一维离散余弦变换的乘积中建立的稀疏表示模型. Wagadarikar 等人<sup>[9]</sup>提出了 CASSI 的设计方案, 并使用 GPSR 算法解决图像重构问题. Correa 和 Arguello 等人<sup>[48-49]</sup>将图像划分为小图像块, 并通过 GPSR 算法分别重构, 其中, 相邻图像块之间存在重叠区域且在拼接时对重叠区域求

均值. Arguello 等人<sup>[50]</sup>提出单个体素在相机焦平面上的实际投影对应多个相邻像素,导致 CASSI 成像模糊,因此利用高阶离散化模型建模投影过程,并使用 GPSR 算法求解重构问题. Galvis 等人<sup>[51]</sup>提出数字微镜器件(Digital Micromirror Devices, DMD)与相机焦平面阵列的像素间距不具备对应关系,因此,在像素不匹配的前提下建立新的前向模型,并使用 GPSR 算法求解 CASSI 重构问题. Mejia 等人<sup>[52]</sup>从图像结构和稀疏性两方面考虑,在梯度下降过程中增加滤波步骤以提升模型收敛速度,该方法是对 GPSR 算法的改进. Gelvez 等人<sup>[53]</sup>提出以稀疏先验和低秩先验作为正则项建立优化模型并使用 GPSR 算法求解,该方法在 CASSI 系统中得到了验证. GPSR 算法在精度上具有优势,但比较耗时.

AMP 算法基于概率图思想,通过状态演化预测后续迭代,具有较快的收敛速度. Tan 等人<sup>[54]</sup>提出利用维纳滤波器消除 AMP 迭代过程中的图像噪声,并使用“阻尼”方法增强算法在稀疏矩阵上的收敛性,最终实现稳定的 CASSI 重构. 相比于 TwIST 算法和 GPSR 算法,AMP 算法无需手动调优参数,使用便捷,缺点是对字典的独立同分布性质要求较高,容易发散.

迭代阈值算法<sup>[46]</sup>基于梯度实现,以负梯度方向作为搜索方向,在实现方式上近似于 GPSR 算法. Wang 等人<sup>[55]</sup>从 CASSI 双相机系统获取的全色图像中学习自适应字典,并通过立体匹配估计光谱图像和全色图像之间的视差图,同时恢复场景的深度信息和光谱信息. 该方法利用迭代阈值算法<sup>[46]</sup>求解优化问题,并在实物系统上验证了性能. 此外, Wang 等人<sup>[56]</sup>结合空间和光谱维度的非局部相似性,提出了自适应非局部稀疏表示(Adaptive Nonlocal Sparse Representation, ANSR)方法,利用三维数据的非局部相似性提高 CASSI 双相机系统的光谱重构质量,同时建立一个联合相似度量,自适应地组合光谱图像相似度和全色图像相似度,最终利用迭代阈值算法<sup>[46]</sup>进行求解. 迭代阈值算法<sup>[46]</sup>具有解决高维稀疏优化问题的优势,但计算效率较低.

OMP 算法是对传统匹配追踪(Matching Pursuit, MP)算法的改进,基于贪婪迭代思想实现,在思想上不同于 GPSR 算法的凸优化逼近思想和贝叶斯框架. Wang 等人<sup>[25]</sup>针对 CASSI 双相机系统提出了一套光谱视频采集方案,通过全色相机采集高帧率全色视频,通过 CASSI 采集低帧率光谱压缩图像. 该

方法利用光谱维度的结构相似性,从全色视频中学习过完备字典,用于重构高帧率的光谱视频. Zhang 等人<sup>[57-59]</sup>提出了 OMP 算法的改进方案,用于求解光谱重构和图像分割的联合优化模型,通过对解的有限性和同质性假设,有效削减了重构过程中未知量的数目,进而提升重构效率和精度,该方法在 CASSI 系统上进行了验证. OMP 算法在分解过完备字典时,对所选择的原子进行正交化处理,收敛速度更快.

除上述主流求解方法外, Yuan 等人<sup>[26]</sup>提出一种贝叶斯字典学习模型和优化算法,从 CASSI 观测图像和 RGB 图像中联合学习耦合字典,解决 CASSI 双相机重构问题. Rajwade 等人<sup>[60]</sup>使用  $\beta$  过程因素分析(Beta Process Factor Analysis, BPPFA)<sup>[61]</sup>解决压缩感知中的优化问题,其中,字典从小尺寸的空间-光谱立方体中学习,各波段间的相关性由一个高斯过程表述. Correa 等人<sup>[62]</sup>提出一种基于多分辨率图像的 CASSI 重构算法,首先重构低分辨率光谱图像,再利用低分辨率图像的梯度强度图重构更高分辨率的光谱图像,该方法能够获得不同空间分辨率的重构结果且在噪声环境下表现良好. Zhang 等人<sup>[63]</sup>结合 RGB 相机的响应曲线实现自适应字典选择,从 CASSI 双相机观测图像中重构光谱图像.

基于稀疏先验的重构算法以图像在特定变换域内的稀疏性为先验信息,具有良好的鲁棒性. 然而,在优化过程中,此类方法通常将高维数据表示为向量或矩阵,导致高维结构信息损失. 此外,该类方法复杂度较高,计算资源消耗大.

### 4.3 基于低秩先验的重构算法

图像的秩能够表征图像信息的丰富程度和噪声含量. 图像噪声含量越高,图像包含的相关内容越少、相关性越弱,图像的秩越大. 反之,清晰图像因其具有自相似性和高信噪比,通常是低秩的. 光谱图像也满足这一性质. 在光谱重构领域,以最小化图像的秩为目标建立约束条件,是求解欠定问题的主流方法之一. 针对 CASSI 的标准前向模型,基于低秩先验建立的优化问题如式(11)所示:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \gamma \cdot R(\mathbf{x}) \quad (11)$$

其中,  $R(\cdot)$  表示图像的秩.

光谱图像  $\mathcal{X} \in \mathbb{R}^{H \times W \times C}$  可使用三阶张量表示,其秩有多种定义方式. 在光谱重构领域,其主流定义方式包括基于 Tucker 分解<sup>[64]</sup>定义的 Tucker 秩和基于典范/平行因子(Candecomp/Parafac, CP)分解<sup>[65]</sup>

的 CP 秩. 具体地说, Tucker 分解将高维张量分解为一个核心张量和多个展开矩阵的乘积, Tucker 秩即所有二维展开矩阵的秩之和; CP 分解将高阶张量分解为多个秩 1 张量之和, CP 秩即秩 1 张量的最小个数.

给定光谱图像  $\mathcal{X} \in \mathbb{R}^{H \times W \times C}$ , Tucker 分解的数学描述如式(12)所示:

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \times_3 \mathbf{A}_3 + \mathcal{Z} \quad (12)$$

其中,  $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  表示核心张量,  $\mathbf{A}_1 \in \mathbb{R}^{H \times I_1}$ 、 $\mathbf{A}_2 \in \mathbb{R}^{W \times I_2}$ 、 $\mathbf{A}_3 \in \mathbb{R}^{C \times I_3}$  分别表示模式- $n$  矩阵, 且各个矩阵的列相互正交,  $\mathcal{Z}$  表示加性误差张量.

如式(12)所示, 对分解得到的模式- $n$  矩阵, 可直接使用奇异值分解(Singular Value Decomposition, SVD)等矩阵分析方法进行研究. 然而, 极小化矩阵的秩是一个非凸优化问题, 是 NP 难的. 为简化此问题的求解, 通常将极小化矩阵的秩转化为极小化矩阵的奇异值之和, 即将优化目标替换为模式- $n$  矩阵的核范数之和, 进而将原问题转化为一个凸优化问题. 简化后的优化模型如式(13)所示:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \min_{\mathbf{x}} \sum_{n=1}^3 \|\mathbf{A}_n\|_* \\ &= \arg \min_{\mathbf{x}} \sum_{n=1}^3 \sum_r \sigma_r(\mathbf{A}_n) \quad \text{s. t. } \mathbf{y} = \Phi \mathbf{x} \end{aligned} \quad (13)$$

其中,  $\|\cdot\|_*$  表示核范数,  $\sigma(\cdot)$  表示矩阵的奇异值,  $r$  表示模式- $n$  矩阵奇异值的序数.

相较于 Tucker 分解, CP 分解保持了张量的高维特性, 能够降低分解过程中的结构信息损失, 其数学描述如式(14)所示:

$$\mathcal{X} = \sum_{k=1}^K \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{e}_k + \mathcal{Z} \quad (14)$$

其中,  $k \in [1, K]$ ,  $\mathbf{a}_k \in \mathbb{R}^{H \times 1}$ 、 $\mathbf{b}_k \in \mathbb{R}^{W \times 1}$ 、 $\mathbf{e}_k \in \mathbb{R}^{C \times 1}$  分别表示不同维度下的基向量, “ $\circ$ ”表示克罗内克积(Kronecker Product), 且  $\mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{e}_k$  构成一个秩 1 张量,  $K$  表示表示秩 1 张量的总数. 式(14)表示的分解过程如图 4 所示.



图 4 CP 分解过程示意图

针对式(11)所示的优化模型, 其主流求解方法包括共轭梯度(Conjugate Gradient, CG)算法、期望最大值(Expectation Maximization, EM)算法等.

CG 算法基于梯度实现, 通过梯度获取搜索方向, 在实现方式上近似于 TwIST 算法和 GPSR 算法. Fu

等人<sup>[66]</sup>在空间维度和光谱维度引入低秩先验建立 CASSI 重构的优化模型, 利用加权核范数自适应地调整奇异值权重, 同时引入空间维度的非局部相似性作为额外的先验信息. Cheng 等人<sup>[67]</sup>考虑到一维向量或二维矩阵无法充分利用空间自相似性、谱间相关性和空间-光谱联合相关性, 提出一种加权高阶奇异值正则化, 用于约束低秩张量恢复模型, 并通过 CG 算法求解 CASSI 重构问题. Zhang 等人<sup>[68-69]</sup>提出维度判别低秩张量恢复(Dimension-discriminative Low-rank Tensor Recovery, DLTR)模型判别不同维度的先验信息, 该方法首先利用光谱图像的非局部相似性建立三维高阶张量, 再通过加权核范数正则化约束 DTLR 模型, 使其自适应地表征光谱图像的空间自相似性、谱间相关性和联合相关性, 最后使用 CG 算法求解 CASSI 重构问题. CG 算法具有二次中止性, 空间复杂度低, 收敛稳定性强, 且无需手工设置参数.

EM 算法通过迭代进行极大似然估计, 是一类算法的总称. Yang 等人<sup>[70]</sup>利用边缘极大似然估计, 从线性观测模型的测量值中学习高斯混合模型(Gaussian Mixture Model, GMM), 同时在高斯协方差矩阵具有近似低秩表示的前提下, 利用低秩约束解决 CASSI 重构问题, EM 算法被用于学习混合因子分解(Mixture of Factor Analysis, MFA)的参数. Wei 等人<sup>[71]</sup>为保持光谱数据的三维特性, 将光谱图像分解为秩 1 张量作为输入, 并结合交替最小二乘(Alternating Least Squares, ALS)算法和 EM 算法估计三维张量的最大秩 1 张量, 求解 CASSI 重构问题. EM 算法计算复杂度低, 容易收敛, 但依赖较好的初始化, 且在求解高维问题时容易陷入局部最优.

除上述主流求解方法外, He 等人<sup>[72]</sup>结合空间非局部相似性和低秩特性建立优化问题, 并使用 FTVd 算法求解, 通过联合优化正交基矩阵和输出结果, 以较低的复杂度实现图像复原. 在后续工作中, He 等人<sup>[73]</sup>针对 CASSI 双相机系统, 提出将低维光谱图像的子空间特性划分为光谱基和空间系数, 并通过 CASSI 观测图像和 RGB 图像分别估计. 该模型在优化过程中无需迭代或输入传感矩阵, 能够灵活、快速地重构光谱图像. Liu 等人<sup>[74]</sup>结合光谱图像的非局部相似性和低秩特性构建约束条件, 针对 CASSI 提出一个联合非凸优化模型, 并使用 ADMM 算法求解.

基于低秩先验的重构算法能够对高维数据的结

构信息进行建模,且在捕获上下文信息方面具有优势.然而,秩的求解或近似需消耗大量计算资源,并且图像的秩对图像内容、噪声等较敏感,导致参数设置困难.

## 5 基于深度学习的重构算法研究进展

先验信息表示是一个复杂的科学问题,基于经验手工建立数学模型,通常不能准确拟合目标函数,且泛化能力有限.深度学习为此问题提供了新的解决方案.具体地说,基于深度学习的重构算法采用数据驱动的方式,从大规模图像数据中学习一般性规律,并建立先验表示模型.与基于优化模型的重构算法相比,基于深度学习的重构算法具备三点优势:

(1) 实时性强. 基于传统数学方法求解优化问题时,通常需要多次迭代逼近最优解,时间效率低.基于端到端深度学习的方法将模型训练和模型推理过程分隔开,通过预训练保存模型参数,在推理阶段快速加载模型执行推理.该过程无需迭代,适用于高实时性任务.

(2) 泛化性强. 传统的光谱重构方案通常基于经验知识手工建模,对先验表示等复杂问题拟合困难,当算法被迁移到其他数据或任务上时,易产生严重的性能退化.深度学习方法基于数据驱动,在数据量足够的前提下,利用数据的统计学规律建立鲁棒模型.实验证明,基于深度学习方法建立的先验表示模型,通常具有更好的泛化性能.

(3) 拟合能力强. 深度学习方法基于神经网络实现,其模型具有网络层数深、权重参数多的特点,因此,在理论上神经网络模型能够逼近任意函数.在光谱重构任务中,神经网络能够依靠其强大的非线性拟合能力,学习准确的先验表示模型,更好地约束欠定问题.

基于深度学习的优势,利用神经网络解决光谱重构问题已成为近年来的研究热点.根据设计思想的不同,基于深度学习的重构算法可依托于以下三类框架实现:端到端全网络框架、即插即用框架和深度展开框架.

### 5.1 基于端到端全网络框架的重构算法

神经网络是一种内部一体化模型,通过反向传播实现参数的整体优化,依靠单步执行得到网络模型的输出结果,此类框架称为端到端(End-to-end).基于端到端全网络框架的重构算法,利用神经网络同时实现先验信息建模和光谱数据拟合,在输入观

测图像后,通过单步执行,直接输出重构图像,如图5所示.给定原始图像 $x$ 、观测图像 $y$ ,光谱重构任务的目标是建立映射关系 $g: y \rightarrow x$ .

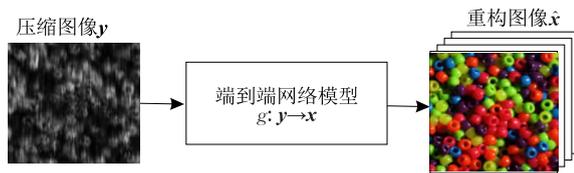


图5 端到端全网络框架示意图

针对光谱重构任务,基于端到端全网络框架的网络模型设计可分为:无注意力机制的网络模型、基于传统软注意力机制的网络模型和基于自注意力机制的网络模型.

无注意力机制的网络模型不考虑特征权重的动态变化.Yorimoto等人<sup>[75]</sup>提出了一种网络模块解决CASSI重构问题,该模块用于捕获空间域和光谱域的联合相关性,并以堆叠的方式组成网络主体,同时使用MixConv层<sup>[76]</sup>替代传统卷积层以减少网络参数.Wang等人<sup>[77]</sup>以编解码联合优化为设计思想,提出HyperReconNet同时实现CASSI编码模板优化和光谱图像重构.该网络由空间子网络和光谱子网络组成,分别学习空间维度和光谱维度的映射关系.此外,具有二值化权重系数的BinaryNet<sup>[78]</sup>被用于模拟编码模板,伴随网络收敛过程逐步优化.HyperReconNet的网络结构如图6(a)所示,由2个子网络组成,分别实现空间维度和光谱维度的图像重构.无注意力机制的网络结构简单,但因缺乏权重调节机制,信息利用率较低.

传统软注意力机制旨在从外部信息中建模特征通道间的相关性.Fu和Zhang等人<sup>[79-80]</sup>提出一种通道注意力网络DEIL建模空间维度和光谱维度的相关性,从外部数据集中学习深度先验信息,同时在全变差正则化约束下学习输入CASSI观测图像的内部特性,最终结合外部和内部的先验信息重构光谱图像.DEIL的网络结构如图6(b)所示,图中相同样式虚线框内的网络结构相同,该网络采用堆叠式网络设计,通过分支网络结构学习通道注意力.Yorimoto等人<sup>[81]</sup>结合多阶段残差网络和注意力机制提出一种深度网络模型解决CASSI重构问题,其中,残差网络用于图像的细节恢复,注意力模块用于从空间维度和光谱维度中自适应地选择重要特征向量.Wang等人<sup>[82]</sup>提出一种堆叠残差网络(Stacked Residual Network, SRN)用于CASSI重构和视频重构.针对光谱图像重构,在SRN模型基础上,提出

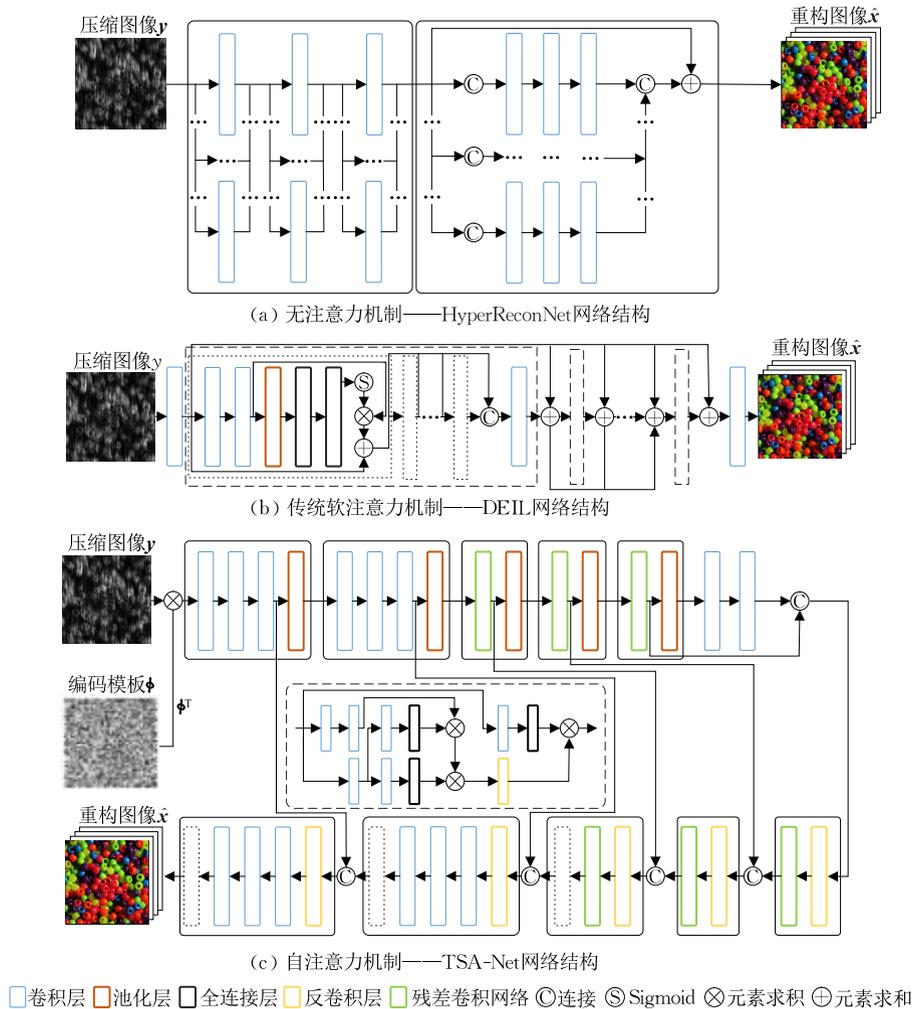


图 6 CASSI 重构典型网络结构示意图

了包含通道注意力增强模块的 CAE-SRN 框架,利用神经网络建模光谱维度的相关性并估计缺失的图像细节。传统注意力机制能够增加信息利用率,降低冗余计算量,但其难以捕获长距离依赖,在建模非局部相关性时存在局限。

自注意力机制旨在挖掘特征内部的相互关系。Miao 等人<sup>[83]</sup>针对 CASSI 提出了包含重构阶段和细化阶段的两阶段生成模型。在重构阶段,结合自注意力生成对抗网络(Generative Adversarial Network, GAN)和 U-Net<sup>[84]</sup>学习图像的非局部相似性,并分级重构出完整的光谱图像。在细化阶段,结合小型 U-Net 和残差网络分别提升各个光谱通道的图像重构质量。Meng 等人<sup>[85]</sup>使用低成本的色散棱镜复现了 CASSI 系统,并提出 TSA-Net 模型以顺序无关方式学习空间维度和光谱维度的相关性,利用自注意力机制提升重构精度。TSA-Net 的网络结构如图 6(c)所示,图中中央虚线框内的网络结构同其余虚线框。该网络以 U-Net 网络为基本框架,融合了

自注意机制,将经典 U-Net 网络中的 3 个卷积模块替换为残差卷积模块。在后续工作中,Meng 等人<sup>[86]</sup>考虑感知在图像恢复任务中的重要性,提出在保持网络结构不变的前提下,引入感知损失<sup>[87]</sup>以提升模型性能。Hu 等人<sup>[88]</sup>设计了一种自注意力模块学习图像空间维度和光谱维度的相关性,同时在目标函数中加入频域损失,提升高频细节的重构质量。Cai 等人<sup>[89]</sup>结合 Transformer 模块<sup>[90]</sup>提出 MST 网络,通过自注意力模块捕获光谱图像的谱间相似性和远程依赖,并提出一种模板引导机制,挖掘 CASSI 编码模板在重构过程中的指导作用。在后续工作中,Cai 等人<sup>[91]</sup>提出 CST 网络估计信息丰富的图像区域,并使用 Transformer 建模光谱图像的空间相关性。Wang 等人<sup>[92]</sup>结合推理注意力网络提出了 GST 模型,通过建模编码模板的不确定性,提升了算法在 CASSI 系统校准错误状态下的鲁棒性。在后续工作中,Wang 等人<sup>[93]</sup>利用 Transformer 建模空谱联合相关性,并从 CASSI 编码模板中预测像素的重构难度,

进而调整惩罚系数. Cheng 等人<sup>[94]</sup>设计 Transformer 模块预测单谱带图像,并通过 RNN 预测完整光谱图像. 自注意力机制能够捕获长距离依赖关系,具有更大的感受野,在建模非局部相关性时具有优势,但其计算量较大,对显存容量需求较高.

基于端到端框架的重构算法能够准确建立原始图像和观测图像之间的映射关系,参数设置便捷,模型推理速度快. 然而,此类方法对系统前向模型的灵活性存在局限,改变前向模型易导致算法性能退化.

## 5.2 基于深度展开框架的重构算法

深度展开(Unrolling/unfolding)框架<sup>[95]</sup>集成了优化模型和神经网络模型,将优化模型解耦为数据项和先验项,并利用数学方法和神经网络分别求解. 该框架将传统迭代优化方式展开到多阶段网络模型中,通过模型的阶段数定义迭代次数,并以端到端的方式进行训练. 给定观测图像  $y$ , 基于深度展开框架的算法流程如图 7 所示,其中,  $\hat{x}^{(0)}$  表示初始化的光谱图像,  $\beta$  表示总阶段数.

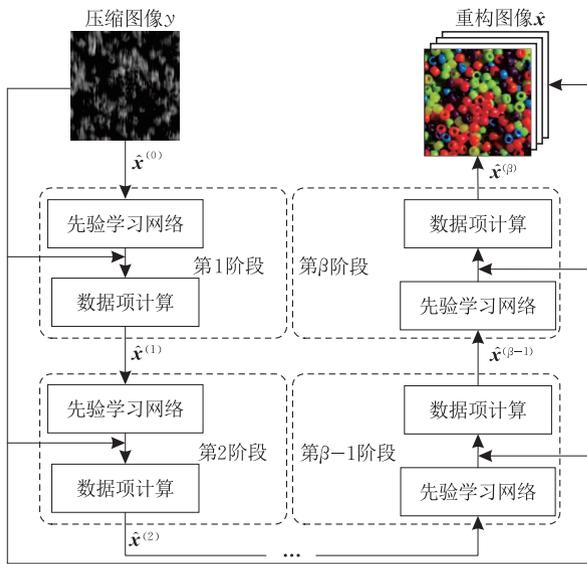


图 7 深度展开框架示意图

针对光谱重构任务,能够和深度展开框架结合的主流近端算法包括 IST 算法、半二次方分裂(Half Quadratic Splitting, HQS)算法、GAP 算法和 ADMM 算法.

基于 IST 算法, Zhang 等人<sup>[96]</sup>提出 ISTA-Net 用于自然图像的压缩感知任务. 具体地说,利用神经网络的非线性变换解决稀疏诱导正则化相关的近端映射问题,同时以有监督学习的方式更新网络参数. 该方法具备良好的重构速度和可解释性,为深度展开方法在图像重构领域的应用提供了参考. 在

CASSI 重构任务中, Zhang 等人<sup>[97]</sup>结合深度展开框架和 IST 算法解决光谱快照压缩成像任务,提出 HerosNet 用于二进制传感矩阵和光谱重构网络的联合优化. 此外,为增强模型的泛化性和灵活性,提出动态梯度下降模块模拟传感矩阵,并在近端映射中分级融合和交互先前阶段的隐藏状态,最终该框架在 CASSI 前向模型上验证了性能. 基于 IST 算法设计的深度展开框架能够避免大规模矩阵求逆运算,计算效率高,但收敛速度较慢.

基于 HQS 算法, Wang 等人<sup>[98]</sup>结合深度展开框架提出 DSSP 模型解决 CASSI 重构问题,这是深度展开框架在光谱重构领域的首次应用. 该模型以数据驱动的方式学习光谱图像的局部相关性和动态特性,并对图像的先验信息进行建模. 在后续工作中, Wang 等人<sup>[99]</sup>利用传感矩阵的分块对角性质,推导出新的递归公式,并在此基础上结合 HQS 算法提出了 DNU 模型. 该模型以加权的方式融合局部先验信息与非局部先验信息,在 CASSI 重构任务中取得了良好性能. Zhang 等人<sup>[100]</sup>结合深度展开框架和 HQS 算法提出了 TLPLN 模型,该模型基于 CP 分解理论,利用多个秩 1 张量的加权和构造光谱图像的低秩表示,在建模低秩先验的同时,保持图像空间维度和光谱维度的上下文信息. Cai 等人<sup>[101]</sup>从 CASSI 编码模板中预测计算数据项使用的惩罚系数,和输入去噪先验网络的噪声等级,实现正则参数自适应,同时使用 Transformer 建模光谱图像的非局部相关性. 基于 HQS 算法设计的深度展开框架在计算数据项时无需迭代逼近,计算简洁,容易收敛.

基于 GAP 算法, Meng 等人<sup>[102]</sup>结合深度展开框架提出网络模型 GAP-Net,该模型采用自编码神经网络作为去噪器,用于 CASSI 重构任务和快照视频成像任务. Wang 等人<sup>[103]</sup>结合 Transformer 提出 GAP-CCoT 模型,其主要模块是由传统卷积层和 Transformer 组成的分支网络,用于同时捕获光谱图像的全局结构信息和局部细节. 基于 GAP 算法设计的深度展开框架收敛速度较快.

基于 ADMM 算法, Sogabe 等人<sup>[104]</sup>分析了 HQS 算法和 ADMM 算法对优化问题的解耦原理,在深度展开框架下,引入可学习参数在两种算法之间折中,提升了模型收敛速度.

除上述近端算法外, Huang 等人<sup>[105]</sup>将 CASSI 重构问题描述为高斯混合尺度(Gaussian Scale

Mixture, GSM)模型下的最大后验概率估计问题,提出一种深度学习学习 GSM 先验信息,用于建模空间维度和光谱维度之间的相关性.此外,该方法将优化模型解耦为两个独立的子问题,利用神经网络和数学方法分别优化.

基于深度展开框架的重构算法对系统前向模型具有灵活性,重构速度快,可解释性强.然而,阶段间信息传递存在损失,且该框架下的迭代次数是一个常量,若需改变,则需要更改网络的阶段数并重新训练.

### 5.3 基于即插即用框架的重构算法

即插即用(Plug-and-Play, PnP)框架<sup>[106]</sup>是一种非凸框架,和深度展开框架具有相似性,二者均是数学优化模型和神经网络模型的结合. PnP 框架的不同之处在于,该框架基于迭代优化方式,具有数据无关的特点.在光谱重构任务中,即插即用框架能够和近端算法结合,将优化模型解耦为数据项和先验项,利用数学计算求解数据项,利用预训练网络模型求解先验项.给定观测图像  $y$ , 最大迭代次数  $\Pi$ , 基于即插即用框架的算法流程如图 8 所示. 其中,  $\hat{x}^{(0)}$  表示初始化的光谱图像,  $\pi \in [1, \Pi]$ ,  $\hat{x}^{(\pi)}$  表示第  $\pi$  次迭代结果.

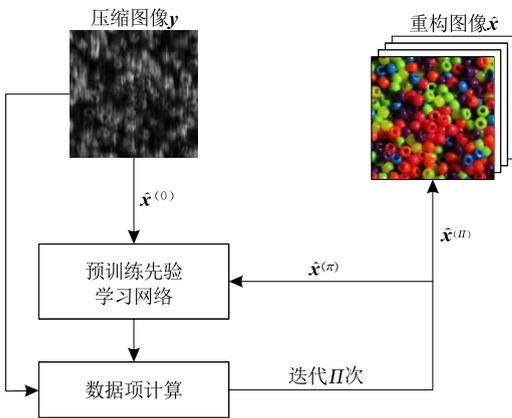


图 8 即插即用框架示意图

针对光谱重构任务,能够和即插即用框架结合的主流近端算法包括 ADMM 算法和 GAP 算法.

基于 ADMM 算法, Chan 等人<sup>[107]</sup>证明了对任意有界去噪器, PnP-ADMM 框架均可收敛至定点,为即插即用框架在图像复原任务的应用提供了重要参考. Lai 等人<sup>[108]</sup>提出了一种深度光谱去噪器作为统一框架,处理包括图像去噪、图像重构、超分辨率和图像修复在内的多种图像复原任务.该框架利用门控递归卷积单元、跳跃连接和噪声等级图对先验知识进行隐式建模,解决光谱图像去噪问题.此

外,在去噪任务下预训练的模型被直接插入到 PnP-ADMM 框架中,并在面向 CASSI 的图像重构任务上验证了有效性. Zheng 等人<sup>[109]</sup>提出了一种深度光谱去噪网络并将其插入到 PnP-ADMM 框架和 PnP-TwIST 框架中,该网络能够学习空间维度和光谱维度的先验信息,且在数据尺寸和噪声等级方面具备灵活性. Meng 等人<sup>[110]</sup>提出了一种自监督框架解决 CASSI 重构问题,该方法使用未经训练的神经网络直接从场景的压缩测量中学习先验信息,并将该网络集成到 PnP-ADMM 框架中.此外,该框架和预训练的光谱图像去噪网络<sup>[109]</sup>结合,证明了该方法对真实环境下的泊松噪声具有鲁棒性. PnP-ADMM 框架噪声鲁棒性强,但计算复杂度相对较高,原因是优化模型被解耦为 3 个子问题(通常为 2 个),计算开销更大.

基于 GAP 算法, Yuan 等人<sup>[111]</sup>将 ADMM 算法中的正则项替换为欧几里得投影,并提出 PnP-GAP 框架,用于从二维快照观测图像中恢复超高清视频,该方法在 CASSI 前向模型的约束下具有全局收敛性. Qiu 等人<sup>[112]</sup>提出不同类型的先验知识能够优势互补,因此以深度去噪网络 FFDNet<sup>[113]</sup>和空间平滑特性作为先验信息,将二者的加权融合集成在 PnP-GAP 框架中,实现 CASSI 图像重构.相比于 PnP-ADMM 框架, PnP-GAP 框架的计算复杂度更低,收敛速度更快.

基于即插即用框架的重构算法节约了模型训练产生的计算成本,有效解决了训练数据缺乏的问题,且具有较强的灵活性和可解释性.然而,迭代过程产生的高时间复杂度,导致此类方法不适用于实时性任务.此外,即插即用框架通常依赖良好的初始化,需要同传统优化算法配合使用.

## 6 实验对比与分析

算法的精度和复杂度决定图像重构质量和计算效率,在置信度要求和资源受限的任务场景中,对算法性能的评估尤为重要.本节首先介绍编码孔径快照光谱成像领域的常用数据集和评价指标,其次对比并分析算法的重构质量,最后收集已公开的源程序,并参考程序中的参数设置分析算法的计算效率.相关算法已汇总于网站<sup>[114]</sup>中.

### 6.1 数据集与评价指标

为公平对比算法性能,需要在相同数据和相同

仿真实验设置上,对模型测试得到的定量结果和定性结果进行考量.其中,数据集作为衡量基准,其制作方式通常依赖于高精度的采集设备和人工标注.在光谱重构领域,常用数据集包括 ICVL 数据集<sup>[115]</sup>、Harvard 数据集<sup>[116]</sup>、CAVE 数据集<sup>[117]</sup>和 KAIST 数据集<sup>[118]</sup>,如图 9 所示.

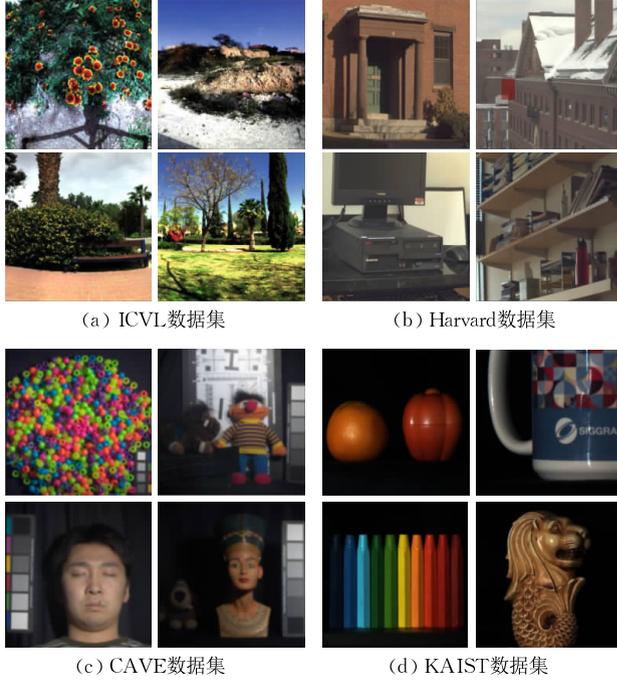


图 9 CASSI 重构任务常用数据集

ICVL 数据集使用 Specim PS Kappa DX4 光谱仪拍摄,共包含 201 幅光谱图像,单幅图像由 31 个谱带组成,谱带范围 400 nm~700 nm. Harvard 数据集使用 CRI Nuance FX 光谱仪拍摄,共包含 50 幅光谱图像,单幅图像由 31 个谱带组成,谱带范围 420 nm~720 nm. CAVE 数据集使用 Apogee Alta U260 相机和 VariSpec 可调谐滤波器拍摄,共包含 32 幅光谱图像,单幅图像由 31 个谱带组成,谱带范围 400 nm~700 nm. KAIST 数据集使用 PointGrey Grasshopper 9.1MP 灰度相机和 VariSpec 可调谐滤波器拍摄,共包含 30 幅光谱图像,单幅图像由 28 个谱带组成,谱带范围 400 nm~720 nm.

此外,在模型重构性能的定量评估中,峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性指标(Structure Similarity Index Measure, SSIM)<sup>[119]</sup>和光谱角制图(Spectral Angle Mapper, SAM)<sup>[120]</sup>是常用的质量评价指标,从多个角度客观量化重构效果.

PSNR 描述图像的最大像素值和背景噪声之间

的关系,PSNR 越大,图像保真度越高.给定无噪声图像  $\mathcal{F} \in \mathbb{R}^{H \times W \times C}$  和待评价图像  $\mathcal{Q} \in \mathbb{R}^{H \times W \times C}$ ,PSNR 计算方法如式(15)和式(16)所示:

$$MSE_c = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W [\mathcal{F}_{h,w,c} - \mathcal{Q}_{h,w,c}]^2 \quad (15)$$

$$PSNR = \frac{1}{C} \sum_{c=1}^C 10 \cdot \log_{10} \left( \frac{\max^2(\mathcal{F}_{\cdot,\cdot,c})}{MSE_c} \right) \quad (16)$$

其中, $MSE_c$  表示图像在第  $c$  维通道的均方误差, $\max(\mathcal{F}_{\cdot,\cdot,c})$  表示图像在第  $c$  维通道的最大像素值,“ $\cdot$ ”表示不对张量的相应维度进行操作.在计算过程中,对归一化后的图像,最大像素值有两种定义方式,即实际最大像素值或 1.0.不同的定义方式导致性能评价标准不统一.

SSIM 从亮度、对比度和结构性三个方面衡量图像之间的相似度,符合肉眼感知.具体地说,对单个谱带的二维图像分量,分别使用像素值的期望和标准差衡量亮度 and 对比度,使用图像矩阵的相关系数衡量结构差异.SSIM 越大,图像的视觉效果越好,其计算方法如式(17)所示:

$$SSIM = \sum_{c=1}^C \left( \frac{2E(\mathcal{F}_{\cdot,\cdot,c})E(\mathcal{Q}_{\cdot,\cdot,c}) + \tau_1}{E^2(\mathcal{F}_{\cdot,\cdot,c}) + E^2(\mathcal{Q}_{\cdot,\cdot,c}) + \tau_1} \cdot \frac{2cov(\mathcal{F}_{\cdot,\cdot,c}, \mathcal{Q}_{\cdot,\cdot,c}) + \tau_2}{D(\mathcal{F}_{\cdot,\cdot,c}) + D(\mathcal{Q}_{\cdot,\cdot,c}) + \tau_2} \right) \quad (17)$$

其中, $E(\cdot)$ 和  $D(\cdot)$ 分别表示均值函数和方差函数, $cov(\cdot)$ 表示协方差函数, $\mathcal{F}_{\cdot,\cdot,c}$ 、 $\mathcal{Q}_{\cdot,\cdot,c}$ 分别表示光谱图像  $\mathcal{F}$ 和  $\mathcal{Q}$ 的第  $c$  维通道.对归一化后的图像, $\tau_1$  的默认值为  $10^{-4}$ , $\tau_2$  的默认值为  $9 \times 10^{-4}$ .注意,协方差运算的时间复杂度满足  $O(W^2H)$ ,为节约计算资源,通常使用小尺寸的滑动窗口计算局部图像块的 SSIM,以其均值作为完整图像的量化结果.

SAM 衡量两幅图像中对应像素点之间的光谱相似度.具体地说,将任意两像素点的光谱信息分别拟作一维向量,并计算向量夹角,以此评估重构图像的光谱保真度.SAM 越小,表示重构得到的光谱信息越准确,其计算方法如式(18)<sup>[120]</sup>所示:

$$SAM = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \cos^{-1} \left( \frac{\mathbf{f}_{h,w,\cdot}^T \mathbf{q}_{h,w,\cdot}}{\|\mathbf{f}_{h,w,\cdot}\|_2 \|\mathbf{q}_{h,w,\cdot}\|_2} \right) \quad (18)$$

其中, $\|\cdot\|_2$ 表示向量的 2 范数, $\mathbf{f}_{h,w,\cdot}$ 和  $\mathbf{q}_{h,w,\cdot}$ 分别表示光谱图像  $\mathcal{F}$ 、 $\mathcal{Q}$ 在像素位置  $(h, w)$  上的光谱信息,且以向量形式参与运算.

## 6.2 重构质量对比与分析

结合常用评价指标,本文汇总了当前主流算法的仿真实验结果,并分别比较了模型在虚拟编码模板和真实编码模板下的重构性能,如表 1 和表 2 所示.

表 1 ICVL 数据集和 Harvard 数据集上的重构质量对比

年份	算法类型	算法名称	数据来源	测试图空间尺寸	ICVL 数据集			Harvard 数据集		
					PSNR	SSIM	SAM	PSNR	SSIM	SAM
2007	TV 优化	TwIST <sup>[30]</sup>	文献[100]	256×256	25.44	0.91	0.05	26.23	0.89	0.11
			文献[77,98-99]	512×512	26.16	0.94	0.05	27.16	0.92	0.12
2008	稀疏优化	GPSR <sup>[9]</sup>	文献[77,98-99]	512×512	24.56	0.91	0.09	24.96	0.91	0.20
2011	TV 优化	ADMM <sup>[121]</sup>	文献[77]	512×512	26.47	0.94	0.05	27.35	0.92	0.13
2016	稀疏优化	AMP <sup>[54]</sup>	文献[77,98]	512×512	26.77	0.95	0.05	26.67	0.94	0.16
2016	非局部 & 低秩优化	LRMA <sup>[66]</sup>	文献[100]	256×256	29.79	0.94	0.03	30.50	0.93	0.08
			文献[98-99]	512×512	29.16	0.96	0.05	29.68	0.95	0.10
			文献[80]	512×512	29.98	0.97	0.03	30.11	0.96	0.09
2017	全网络	HSCNN <sup>[122]</sup>	文献[100]	256×256	28.45	0.93	0.04	27.60	0.90	0.11
			文献[77,98-99]	512×512	29.48	0.97	0.04	28.55	0.94	0.12
			文献[75]	256×256	36.64*	0.96	0.08	35.09*	0.94	0.09
2017	非局部 & 稀疏优化	3DNSR <sup>[56]</sup>	文献[100]	256×256	27.47	0.91	0.05	28.64	0.90	0.11
			文献[77,98-99]	512×512	27.95	0.96	0.05	28.51	0.94	0.13
2017	全网络	Autoencoder <sup>[118]</sup>	文献[100]	256×256	28.07	0.92	0.04	29.21	0.91	0.09
			文献[98-99]	512×512	30.44	0.97	0.04	30.30	0.95	0.10
2018	深度展开	ISTA-Net <sup>[96]</sup>	文献[100]	256×256	30.50	0.95	0.04	29.87	0.91	0.11
			文献[98-99]	512×512	31.73	0.98	0.04	31.13	0.97	0.11
2019	低秩优化	DLTR <sup>[68]</sup>	文献[100]	256×256	30.40	0.94	0.03	31.14	0.93	0.08
2019	全网络	$\lambda$ -Net <sup>[83]</sup>	文献[100]	256×256	29.01	0.95	0.05	29.37	0.91	0.13
			文献[80]	512×512	28.85	0.97	0.06	29.28	0.95	0.21
2019	深度展开	DSSP <sup>[98]</sup>	原文	512×512	34.13	0.99	0.03	32.84	0.98	0.09
			文献[75]	256×256	39.67*	0.98	0.05	37.62*	0.96	0.13
2019	全网络	HyperReconNet <sup>[77]</sup>	原文	512×512	32.36	0.99	0.04	30.34	0.96	0.12
			文献[75]	256×256	38.43*	0.97	0.06	36.04*	0.94	0.17
2020	深度展开	DNU <sup>[99]</sup>	原文	512×512	34.27	0.99	0.03	32.71	0.98	0.09
			文献[80]	512×512	33.78	0.99	0.03	32.55	0.98	0.09
2021	深度展开	TLPLN <sup>[100]</sup>	原文	256×256	34.53	0.98	0.03	32.43	0.94	0.09
2021	全网络	DEIL <sup>[80]</sup>	原文	512×512	36.57	0.99	0.02	34.05	0.98	0.09
2021	全网络	DRAN <sup>[81]</sup>	原文	256×256	38.90*	0.98	0.06	35.69*	0.95	0.09

注：\* 表示计算 PSNR 时使用的最大像素值是 1.0，而非实际最大像素值。

表 2 KAIST 数据集上的重构质量和计算效率对比

年份	算法类型	算法名称	数据来源	模型性能		模型复杂度	
				PSNR	SSIM	参数量(M)	FLOPs(GMac)
2007	TV 优化	TwIST <sup>[30]</sup>	文献[89,105]	23.12	0.67	/	$\geq 1000$
			[85]	22.44	0.70		
2015	全网络	U-Net <sup>[84]</sup>	文献[85]	26.80	0.80	31.32	58.99
2016	TV 优化	GAP-TV <sup>[31]</sup>	文献[89,105]	24.36	0.70	/	$\geq 1000$
			文献[85,97]	23.73	0.68		
2019	全网络	$\lambda$ -Net <sup>[83]</sup>	文献[89,105]	28.53	0.84	58.25	44.59
			文献[85,97]	29.25	0.89		
2019	低秩优化	DeSCI <sup>[74]</sup>	文献[89,105]	25.27	0.72	/	$\geq 1000$
			文献[85,97]	25.86	0.79		
2019	深度展开	DSSP <sup>[98]</sup>	文献[89,105]	30.35	0.85	0.30	20.14
			文献[85,97]	28.93	0.83		
2020	深度展开	DNU <sup>[99]</sup>	文献[89,105]	30.74	0.86	4.47	293.90
2020	全网络	TSA-Net <sup>[85]</sup>	原文	30.15	0.89	44.25	135.12
			[89][105]	31.46	0.89		
2020	深度展开	GAP-Net <sup>[102]</sup>	原文	32.13	0.92	4.27	84.08
			[82]	32.47	0.93		
2021	深度展开	DGSM <sup>[105]</sup>	原文	32.63	0.92	3.76	647.8
2021	即插即用	PnP-DIP-HSI <sup>[110]</sup>	原文	31.30	0.90	33.85	$\geq 3000$
2021	深度展开	HerosNet <sup>[97]</sup>	原文	34.45	0.97	11.75	447.18

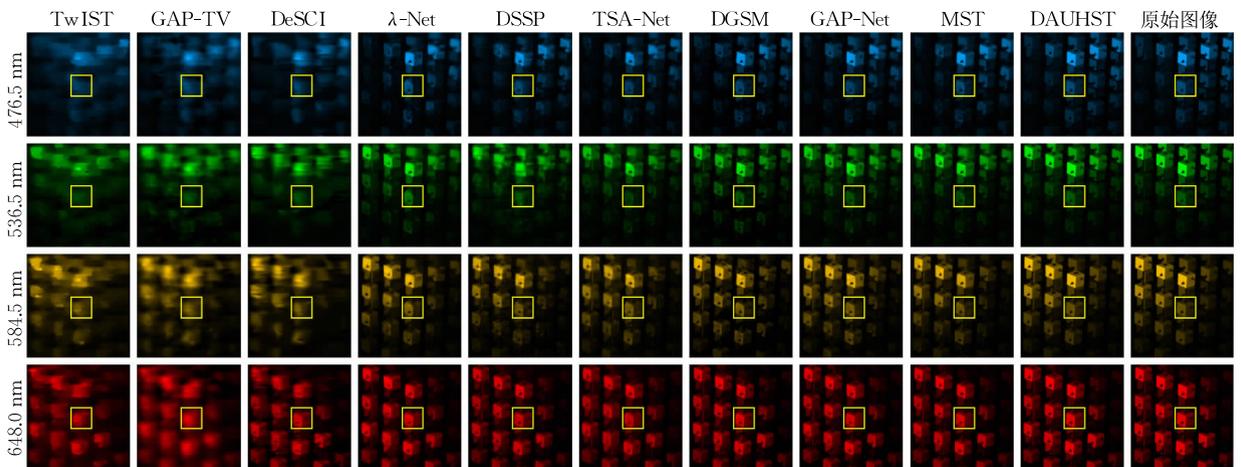
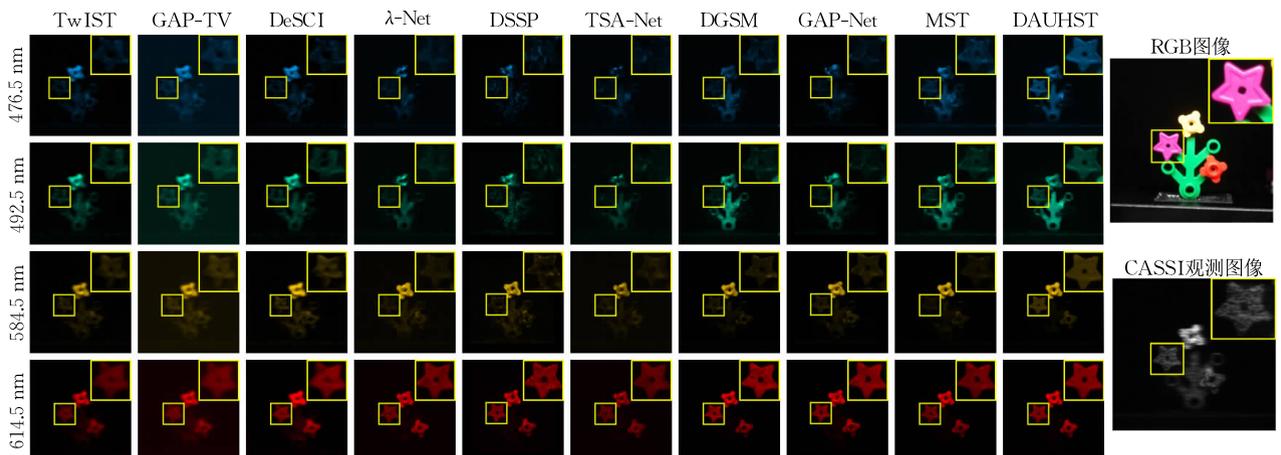
(续 表)

年份	算法类型	算法名称	数据来源	模型性能		模型复杂度	
				PSNR	SSIM	参数量(M)	FLOPs(GMac)
2021	即插即用	PnP-HSI <sup>[109]</sup>	文献[110]	25.67	0.70	1.96	≥3000
2022	全网络	CAE-SRN <sup>[82]</sup>	原文	33.26	0.93	1.25	83.06
2022	全网络	HDNet <sup>[88]</sup>	原文	34.97	0.94	2.37	154.76
2022	全网络	MST <sup>[89]</sup>	原文	35.18	0.95	2.46	31.40
2022	深度展开	GAP-CCoT <sup>[103]</sup>	原文	35.26	0.95	8.04	95.60
2022	全网络	CST <sup>[91]</sup>	原文	36.12	0.96	3.00	40.10
2022	全网络	S <sup>2</sup> -Transformer <sup>[93]</sup>	原文	36.48	0.96	1.80	27.21
2022	全网络	BIRNAT <sup>[94]</sup>	原文	36.14	0.97	4.40	2122.66
2022	深度展开	DAUHST <sup>[101]</sup>	原文	38.36	0.97	6.15	79.50
2023	深度展开	RDFNet <sup>[123]</sup>	原文	33.34	0.96	1.29	604.88

表 1 展示了虚拟编码模板下的重构质量对比。在实验过程中,编码模板随机生成且满足  $p=0.5$  的伯努利分布,数据集被划分为训练集和测试集,其中,训练集用于神经网络模型的参数更新,测试集用于评估算法性能。该组实验分别在 ICVL 数据集和 Harvard 数据集上进行。表 2 展示了真实编码模板下的重构质量和计算效率对比。在实验过程中,编码模板采用 Meng 等人<sup>[85]</sup>拍摄的真实编码模板,在

CAVE 数据集上训练模型,并在 KAIST 数据集中选取 10 个场景图用作测试。

在相同实验设置下,各文献汇报的实验结果不尽相同,其原因在于深度学习具有随机性和不可复制性,以及数据划分、软件库版本可能对实验结果造成影响。针对此问题,表 1 和表 2 展示了部分算法的多组实验结果,并指出其数据来源。此外,图 10、图 11 对比了各算法在仿真实验、实物实验中的定性结果。

图 10 仿真实验定性结果对比图<sup>[101]</sup>图 11 实物实验定性结果对比图<sup>[101]</sup>

结合表 1 和表 2 中的实验数据,可得到如下结论:

(1) 在仿真实验中,基于深度学习的重构算法性能总体高于基于优化模型的重构算法.其原因在于神经网络具有良好的非线性拟合能力,能够更好地学习输入、输出之间的映射关系.受上述优势影响,近年来深度学习成为光谱重构领域的主要研究方向.

(2) 由表 1 可知,对传统优化算法,引入非局部相似性作为辅助先验信息能够提升重构质量.其原因在于非局部相似性能够利用全局范围内的先验信息,与局部相似性结合能够更好地约束重构问题.

(3) 由表 1 可知,相比基于端到端全网络框架的算法,同时期基于深度展开框架的算法性能更优.其原因在于全网络方法难以结合系统的前向模型,限制了欠定问题的求解质量.

(4) 由表 2 可知,基于即插即用框架的算法性能高于传统优化算法.其原因在于依靠神经网络学习得到的先验表示模型优于手工设计的先验表示模型,能够更好地约束优化问题的解空间.

(5) 由表 2 可知,对真实拍摄的编码模板,同时期基于深度展开框架的算法性能更好.其原因在于深度展开框架利用数学方法,结合前向模型推导数据保真项,因此,对前向模型具有更好的灵活性,能够适应真实的编码模板.

### 6.3 计算效率对比与分析

计算效率决定算法在实时任务上的应用潜力.针对光谱重构问题,通常在图像尺寸固定的前提下,计算多幅图像的平均重构时间用于效率对比.然而,底层算子优化和程序并行度不同,易导致仿真实验中对硬件算力的利用率不同,因此,将运行时间等价于算法效率存在局限.

本文采用模型参数量和浮点运算量(Floating-point Operations, FLOPs)评估模型的计算复杂度,如表 2 所示,模型参数量表示神经网络模型包含的参数个数, FLOPs 表示模型推理过程中产生的浮点运算总量.其中,网络层的 FLOPs 计算方式参考了 Molchanov 等人<sup>[124]</sup>的工作.注意,对传统优化算法和基于即插即用框架的重构算法,其收敛时的迭代次数难以估算,因此本文不做精确评价;对基于深度展开框架的重构算法,本文采用和原文一致的阶段数计算.结合表 2 中的实验数据,可得到如下结论:

(1) 基于深度学习的模型参数量呈总体降低趋势,表明算法模型正在向轻量化发展,在计算资源受

限的系统中,深度学习算法的应用潜力逐渐提升.

(2) 基于即插即用框架的重构算法 FLOPs 较高,其原因在于即插即用框架采用传统优化方法迭代求解,且在单次迭代中执行至少一次模型推理,导致浮点运算量大.然而,此类算法无需预训练,具备数据无关特性,在模型的泛化性研究中具备重要意义.

(3) 基于深度展开框架的重构算法 FLOPs 浮动较大,其原因在于先验学习网络的复杂度不同,且各方法的阶段数设置不同.深度展开框架能够与任意复杂度的先验网络集成.

## 7 思考与展望

自 2007 年 CASSI 提出以来,国内外学者针对其重构算法提出了诸多改进方案,在仿真实验中取得了突破性进展.然而,现有工作仍存在不足,重点体现在实物场景的重构质量方面.本节将讨论现有工作在解决 CASSI 重构挑战中仍然存在的不足,并展望未来研究趋势.

### (1) 仿真实验难以精确模拟光学成像过程

低精度前向模型和 CASSI 物理系统存在差异,使重构算法性能受限.尽管高精度建模<sup>[125]</sup>、不确定性建模<sup>[92]</sup>、训练中引入噪声<sup>[85]</sup>改进了此类问题,但物理系统中由衍射、错位、畸变、噪声等引起的干扰仍然是现有重构算法难以应用于真实场景的主要原因.

为缓解低精度前向模型和 CASSI 物理系统之间差异对重构质量的影响,可参考以下研究方向:①从物理系统出发,优化系统的前向模型或光学元件设计,减小物理系统中的误差;②从数学建模出发,充分考虑物理系统中的错位、衍射和畸变,完善高精度建模方案,减小前向模型和物理系统的差距,研究高精度标定方案,降低系统误差对实验结果的影响;③从软件算法出发,针对 CASSI 的噪声产生过程精确建模<sup>[126]</sup>,设计噪声鲁棒的重构算法.

### (2) 先验表示模型难以适配复杂场景

现有先验表示模型的泛化能力有限,导致复杂场景重构困难.针对此问题,内外部联合学习<sup>[79-80]</sup>通过获取待测图像的先验信息完善重构效果,无监督学习<sup>[110,127]</sup>通过数据无关的方式获取先验信息,提升模型泛化能力.然而,上述方法产生了额外的计算时间,未完全解决此类问题.

为提升先验表示模型的泛化能力,可参考以下研究方向:①从数据分布出发,基于数据驱动方式

建立的先验表示模型,其泛化能力和样本数据分布有关,可研究数据增强技术<sup>[128]</sup>,在数据特性不变的前提下获取分布均匀的样本;②从目标函数出发,基于深度学习的重构算法,其目标函数决定模型收敛方向和特征拟合方式,可研究目标函数选择策略或融合策略<sup>[129]</sup>,使神经网络拟合更具泛化性的光谱特征;③从建模方法出发,基于手工设计的先验表示模型表达能力有限,基于深度学习建立的先验表示模型过度依赖样本,可研究集成学习策略<sup>[130]</sup>或早停方案<sup>[131]</sup>,在不过度依赖样本的前提下挖掘光谱数据的先验信息。

### (3) 重构算法对前向模型的灵活性不足

基于端到端全网络框架的重构算法,通常针对特定前向模型进行数据拟合,导致前向模型改变时,重构质量下降.深度展开框架<sup>[95,99]</sup>致力于解决灵活性问题,但其采用的端到端训练方式,易导致模型在训练时的编码模板上过拟合。

为提升重构算法对前向模型的灵活性,可参考以下研究方向:①从全局优化出发,可研究前向模型和重构算法的联合优化方案<sup>[132]</sup>,将二者作为整体,同时确定编码耦合方式和解码重构方式,降低灵活性问题对重构质量的影响;②从任务相关性出发,可研究任务无关的光谱先验建模方案,利用小样本学习<sup>[133]</sup>从数据、模型、算法三个维度增强监督信息,以自监督方式学习光谱先验知识,降低重构算法和前向模型的相关性;③从重构算法出发,可研究深度展开框架的改进方案,引入分阶段训练策略<sup>[134]</sup>替代传统的端到端训练策略,逐步将样本特征注入网络,降低网络对前向模型的过拟合。

### (4) 重构质量和计算效率需要权衡

对时间复杂度高于平方阶的重构算法,将大尺寸图像切割为图像块参与测试能有效降低计算复杂度<sup>[77,100]</sup>,但同时将引入图像块的边缘伪影.此外,为降低模型参数量,深度展开框架和即插即用框架利用数学方法计算数据保真项,削减了网络参数,然而,两种框架下的浮点运算量仍有待优化。

为降低模型复杂度,可参考以下研究方向:①从模型量化<sup>[135]</sup>出发,在重构质量无损失或微小损失的前提下,构建轻量化网络模型,节省计算资源以提升重构任务的实时性;②从模型剪枝<sup>[136]</sup>出发,利用网络剪枝技术,从预训练的 CASSI 重构模型中去除冗余参数,降低计算复杂度,提升重构速度;③从知识蒸馏<sup>[137]</sup>出发,利用知识蒸馏技术从复杂光谱重

构模型中获取监督信息,训练参数更少的小模型。

### (5) 缺少实物数据集和统一评价标准

为准确评估真实场景下的模型性能,更好地结合理论与实际,应建立包含 CASSI 观测图像的实物数据集.首先利用扫描光谱仪拍摄高质量光谱图像,之后利用 CASSI 拍摄同一场景得到观测图像,最后对光谱图像和观测图像进行校准、对齐,制作数据对.此外,鼓励研究人员共建开源社区,设立数据中心和相应共享机制,逐步实现性能评价方式的统一。

## 8 结 论

重构算法是编码孔径快照光谱成像的重要组成部分,致力于逆成像问题的求解.本文首先介绍了 CASSI 的硬件设计和前向模型;其次,重点归纳了基于优化模型和深度学习的重构算法;接着,在公开数据集上对比了主流算法的重构质量和计算效率;最后,结合当前研究进展总结并讨论了现存问题和未来研究趋势。

总而言之,CASSI 重构算法已在仿真实验上取得了显著成果,但在实际场景中的应用仍存在不足,具体表现为泛化能力差、计算效率低.在未来,对算法模型的泛化性研究和轻量化设计,将推动 CASSI 在实际生产生活中的应用。

## 参 考 文 献

- [1] Li Hong, Liu Fang, Yang Shu-Yuan, et al. Remote sensing image fusion based on deep support value learning networks. *Chinese Journal of Computers*, 2016, 39(8): 1583-1596 (in Chinese)  
(李红, 刘芳, 杨淑媛等. 基于深度支撑值学习网络的遥感图像融合. *计算机学报*, 2016, 39(8): 1583-1596)
- [2] Li Wei, Lyu Meng, Chen Tian-Hong, et al. Application of a hyperspectral image in medical field: A review. *Journal of Image and Graphics*, 2021, 26(8): 1764-1785 (in Chinese)  
(李伟, 吕蒙, 陈天虹等. 高光谱图像在生物医学中的应用. *中国图象图形学报*, 2021, 26(8): 22)
- [3] He Lin, Pan Quan, Di Wei, et al. Supervised detection for hyperspectral imagery based on high-dimensional multiscale autoregression. *Acta Automatica Sinica*, 2009, 35(5): 509-518 (in Chinese)  
(贺霖, 潘泉, 邸巍等. 高光谱图像高维多尺度自回归有监督检测. *自动化学报*, 2009, 35(5): 509-518)
- [4] Porter W M, Enmark H T. A system overview of the airborne visible/infrared imaging spectrometer (AVIRIS)//*Proceedings*

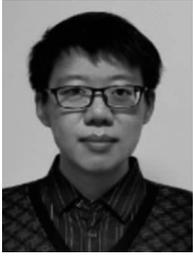
- of the 1987 SPIE Conference on Imaging Spectroscopy II. San Diego, USA, 1987: 22-31
- [5] Basedow R W, Carmer D C, Anderson M E. HYDICE system: Implementation and performance//Proceedings of the 1995 SPIE Conference on Imaging Spectrometry. Orlando, USA, 1995: 258-267
- [6] Gat N. Imaging spectroscopy using tunable filters: A review//Proceedings of the 2000 SPIE Conference on Wavelet Applications VII. Orlando, USA, 2000: 50-64
- [7] Schechner Y Y, Nayar S K. Generalized mosaicing: Wide field of view multispectral imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(10): 1334-1348
- [8] Gehm M E, John R, Brady D J, et al. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*, 2007, 15(21): 14013-14027
- [9] Wagadarikar A, John R, Willett R, et al. Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics*, 2008, 47(10): B44-B51
- [10] Arce G R, Brady D J, Carin L, et al. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 2014, 31(1): 105-115
- [11] Candes E J. Compressive sampling//Proceedings of the 2006 EMS International Congress of Mathematicians. Madrid, Spain, 2006: 1433-1452
- [12] Candes E J, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 2006, 52(2): 489-509
- [13] Gleichman S, Eldar Y C. Blind compressed sensing. *IEEE Transactions on Information Theory*, 2011, 57(10): 6958-6975
- [14] Arguello H, Arce G R. Code aperture optimization for spectrally agile compressive imaging. *Journal of Optical Society of America A*, 2011, 28(11): 2400-2413
- [15] Arguello H, Arce G R. Rank minimization code aperture design for spectrally selective compressive imaging. *IEEE Transactions on Image Processing*, 2013, 22(3): 941-954
- [16] Arguello H, Parada A, Arce G R. Optimization of pseudo-random coded apertures for compressive spectral imaging//Proceedings of the 2013 SPIE Conference on Compressive Sensing II. Baltimore, USA, 2013: 86-93
- [17] Rueda H F, Arguello H, Arce G R. Compressive spectral imaging based on colored coded apertures//Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Florence, Italy, 2014: 7799-7803
- [18] Gelvez T, Arguello H, Rueda H F. Coded aperture design for hyper-spectral image recovery via matrix completion//Proceedings of the 2015 20th IEEE Symposium on Signal Processing, Images and Computer Vision. Bogota, Colombia, 2015: 1-7
- [19] Pinilla S E, Vargas H M, Arguello H. Probability of correct reconstruction in compressive spectral imaging. *Ingenieria E Investigacion*, 2016, 36(2): 68-77
- [20] Ma X, Zhang H, Ma X, et al. Snapshot compressive spectral imaging based on adaptive coded apertures//Proceedings of the 2018 SPIE Conference on Compressive Sensing VII. Orlando, USA, 2018: 11-19
- [21] Zhang H, Ma X, Arce G R. Compressive spectral imaging approach using adaptive coded apertures. *Applied Optics*, 2020, 59(7): 1924-1938
- [22] Shen J, Chen Y, Ding Z, et al. Development of energy correction based multishot snapshot spectral imaging system. *Optics Communications*, 2019, 432: 116-122
- [23] Tao C, Zhu H, Sun P, et al. Simultaneous coded aperture and dictionary optimization in compressive spectral imaging via coherence minimization. *Optics Express*, 2020, 28(18): 26587-26600
- [24] Wang L, Xiong Z, Gao D, et al. Dual-camera design for coded aperture snapshot spectral imaging. *Applied Optics*, 2015, 54(4): 848-858
- [25] Wang L, Xiong Z, Huang H, et al. High-speed hyperspectral video acquisition by combining nyquist and compressive sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 857-870
- [26] Yuan X, Tsai T H, Zhu R, et al. Compressive hyperspectral imaging with side information. *IEEE Journal on Selected Topics in Signal Processing*, 2015, 9(6): 964-976
- [27] Wang L, Xiong Z, Shi G, et al. Compressive hyperspectral imaging with complementary RGB measurements//Proceedings of the 2016 IEEE Conference on Visual Communications and Image Processing (VCIP). Chengdu, China, 2016: 1-4
- [28] Bacca J, Correa C V, Arguello H. Noniterative hyperspectral image reconstruction from compressive fused measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(4): 1231-1239
- [29] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, 1992, 60(1-4): 259-268
- [30] Bioucas-Dias J M, Figueiredo M A T. A new twist: Twostep iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Process*, 2007, 16(12): 2992-3004
- [31] Yuan X. Generalized alternating projection based total variation minimization for compressive sensing//Proceedings of the 2016 IEEE International Conference on Image Processing. Phoenix, USA, 2016: 2539-2543
- [32] Figueiredo M A T, Bioucas-Dias J M, Nowak R D. Majorization-minimization algorithms for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 2007, 16(12): 2980-2991

- [33] Khoramian S. An iterative thresholding algorithm for linear inverse problems with multi-constraints and its applications. *Applied and Computational Harmonic Analysis*, 2012, 32(1): 109-130
- [34] Liu X, Tao C, Xie Q, et al. An improved method based on TwIST and dual-camera design for hyperspectral image reconstruction//*Proceedings of the 2021 SPIE Conference on Optoelectronic Imaging and Multimedia Technology VIII*. Nantong, China, 2021: 165-173
- [35] Kittle D, Choi K, Wagadarikar A, et al. Multiframe image estimation for coded aperture snapshot spectral imagers. *Applied Optics*, 2010, 49(36): 6824-6833
- [36] Zhang S, Huang H, Fu Y. Fast parallel implementation of dual-camera compressive hyperspectral imaging system. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(11): 3404-3414
- [37] Zhang M, Wang L, Zhang L, et al. Compressive hyperspectral imaging with non-zero mean noise. *Optics Express*, 2019, 27(13): 17449-17462
- [38] Liang Z, Xu Y, Xiao L, et al. Spatial-spectral total variation constrained collaborative tensor regularization for dual-camera compressive hyperspectral imaging//*Proceedings of the 2021 IEEE Conference on International Geoscience and Remote Sensing Symposium (IGARSS)*. 2021: 3873-3876
- [39] Vargas E, Espitia O, Arguello H, et al. Spectral image fusion from compressive measurements. *IEEE Transactions on Image Processing*, 2019, 28(5): 2271-2282
- [40] Wang Y, Han Y, Wang K, et al. Total variation regularized nonlocal low-rank tensor train for spectral compressive imaging. *Signal Processing*, 2022, 195: 108464:1-108464:10
- [41] Greer J B, Flake J C. Accurate reconstruction of hyperspectral images from compressive sensing measurements//*Proceedings of the 2013 SPIE Conference on Compressive Sensing II*. Baltimore, USA, 2013: 94-110
- [42] Wang Y, Yang J, Yin W, et al. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences*, 2008, 1(3): 248-272
- [43] Courant R. Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of the American Mathematical Society*, 1943, 49(1): 1-24
- [44] Liu Jian-Wei, Cui Li-Peng, Liu Ze-Yu, et al. Survey on the regularized sparse models. *Chinese Journal of Computers*, 2015, 38(7): 1307-1325(in Chinese)  
(刘建伟, 崔立鹏, 刘泽宇等. 正则化稀疏模型. *计算机学报*, 2015, 38(7): 1307-1325)
- [45] Figueiredo M A T, Nowak R D, Wright S J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal on Selected Topics in Signal Processing*, 2007, 1(4): 586-597
- [46] Daubechies I, Defrise M, Mol C D. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 2004, 57(11): 1413-1457
- [47] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666
- [48] Correa C V, Arguello H, Arce G R. Block-based reconstructions for compressive spectral imaging//*Proceedings of the 2013 SPIE Conference on Compressive Sensing II*. Baltimore, USA, 2013: 111-119
- [49] Arguello H, Correa C V, Arce G R. Fast lapped block reconstructions in compressive spectral imaging. *Applied Optics*, 2013, 52(10): D32-D45
- [50] Arguello H, Rueda H F, Wu Y, et al. Higher-order computational model for coded aperture spectral imaging. *Applied Optics*, 2013, 52(10): D12-D21
- [51] Galvis L, Arguello H, Arce G R. Synthetic coded apertures in compressive spectral imaging//*Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing*. Florence, Italy, 2014: 3181-3185
- [52] Mejia Y, Arguello H. Filtered gradient reconstruction algorithm for compressive spectral imaging. *Optical Engineering*, 2016, 56(4): 1-11
- [53] Gelvez T, Rueda H, Arguello H. Joint sparse and low rank recovery algorithm for compressive hyperspectral imaging. *Applied Optics*, 2017, 56(24): 6785-6795
- [54] Tan J, Ma Y, Rueda H F, et al. Compressive hyperspectral imaging via approximate message passing. *IEEE Journal on Selected Topics in Signal Processing*, 2016, 10(2): 389-401
- [55] Wang L, Xiong Z, Shi G, et al. Simultaneous depth and spectral imaging with a cross-modal stereo system. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(3): 812-817
- [56] Wang L, Xiong Z, Shi G, et al. Adaptive nonlocal sparse representation for dual-camera compressive hyperspectral imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(10): 2104-2111
- [57] Zhang Q, Plemmons R J, Kittle D, et al. Reconstructing and segmenting hyperspectral images from compressed measurements //*Proceedings of the 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. Lisbon, Portugal, 2011: 1-4
- [58] Zhang Q, Plemmons R J, Kittle D, et al. Joint segmentation and reconstruction of hyperspectral images from a single snapshot //*Proceedings of the 2011 SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*. Orlando, USA, 2011: 432-443
- [59] Zhang Q, Plemmons R J, Kittle D, et al. Joint segmentation and reconstruction of hyperspectral data with compressed measurements. *Applied Optics*, 2011, 50(22): 4417-4435

- [60] Rajwade A, Kittle D S, Tsai T H, et al. Coded hyperspectral imaging and blind compressive sensing. *SIAM Journal on Imaging Sciences*, 2013, 6(2): 782-812
- [61] Zhou M, Chen H, Paisley J, et al. Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images. *IEEE Transactions on Image Processing*, 2012, 21(1): 130-144
- [62] Correa C V, Arguello H, Arce G R. Multi-resolution reconstructions from compressive spectral coded projections// *Proceedings of the 2018 26th European Signal Processing Conference (EUSIPCO)*. Rome, Italy, 2018: 1995-1999
- [63] Zhang Shi-Peng, Wang Li-Zhi, Fu Ying, et al. Color adaptive dictionary based reconstruction algorithm for dual camera compressive hyperspectral imaging. *Chinese Journal of Computers*, 2020, 43(1): 151-164(in Chinese)  
(张仕鹏, 王立志, 付莹等. 基于颜色自适应字典的双相机光谱系统重建算法. *计算机学报*, 2020, 43(1): 151-164)
- [64] Tucker L R. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 1966, 31(3): 279-311
- [65] Kolda T G, Bader B W. Tensor decompositions and applications. *SIAM Review*, 2009, 51(3): 455-500
- [66] Fu Y, Zheng Y, Sato I, et al. Exploiting spectral-spatial correlation for coded hyperspectral image restoration// *Proceedings of the 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 2016: 3727-3736
- [67] Cheng N, Huang H, Zhang L, et al. Snapshot hyperspectral imaging based on weighted high-order singular value regularization// *Proceedings of the 2020 25th International Conference on Pattern Recognition*. Milan, Italy, 2020: 1267-1274
- [68] Zhang S, Wang L, Fu Y, et al. Computational hyperspectral imaging based on dimension-discriminative low-rank tensor recovery// *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, South Korea, 2019: 10182-10191
- [69] Wang L, Zhang S, Huang H. Adaptive dimension-discriminative low-rank tensor recovery for computational hyperspectral imaging. *International Journal of Computer Vision*, 2021, 129(10): 2907-2926
- [70] Yang J, Liao X, Yuan X, et al. Compressive sensing by learning a gaussian mixture model from measurements. *IEEE Transactions on Image Processing*, 2015, 24(1): 106-119
- [71] Wei C, Li Q, Zhang X, et al. A fast snapshot hyperspectral image reconstruction method based on three-dimensional low rank constraint. *Canadian Journal of Remote Sensing*, 2021, 47(4): 588-595
- [72] He W, Yao Q, Li C, et al. Non-local meets global: An integrated paradigm for hyperspectral image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 44(4): 2089-2107
- [73] He W, Yokoya N, Yuan X. Fast hyperspectral image recovery of dual-camera compressive hyperspectral imaging via non-iterative subspace-based fusion. *IEEE Transactions on Image Processing*, 2021, 30: 7170-7183
- [74] Liu Y, Yuan X, Suo J, et al. Rank minimization for snapshot compressive imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(12): 2990-3006
- [75] Yorimoto K, Han X. HyperMixNet: Hyperspectral image reconstruction with deep mixed network from a snapshot measurement// *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops*. British Columbia, Canada, 2021: 1184-1193
- [76] Tan M, Le Q V. MixConv: Mixed depthwise convolutional kernels// *Proceedings of the 2019 30th British Machine Vision Conference*. Cardiff, UK, 2019: 116.1-116.13
- [77] Wang L, Zhang T, Fu Y, et al. HyperReconNet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. *IEEE Transactions on Image Processing*, 2019, 28(5): 2257-2270
- [78] Courbariaux M, Bengio Y, David J P. BinaryConnect: Training deep neural networks with binary weights during propagations// *Proceedings of the 2015 Conference on Neural Information Processing Systems*. Quebec, Canada, 2015: 3123-3131
- [79] Zhang T, Fu Y, Wang L, et al. Hyperspectral image reconstruction using deep external and internal learning// *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, South Korea, 2019: 8559-8568
- [80] Fu Y, Zhang T, Wang L, et al. Coded hyperspectral image reconstruction using deep external and internal learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(7): 3404-3420
- [81] Yorimoto K, Han X. Deep residual attention network for hyperspectral image reconstruction// *Proceedings of the 2020 25th International Conference on Pattern Recognition*. Milan, Italy, 2021: 8547-8553
- [82] Wang J, Zhang Y, Yuan X, et al. A simple and efficient reconstruction backbone for snapshot compressive imaging. *arXiv preprint, arXiv: 2108.07739*, 2021
- [83] Miao X, Yuan X, Pu Y, et al. Lambda-Net: Reconstruct hyperspectral images from a snapshot measurement// *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, South Korea, 2019: 4058-4068
- [84] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation// *Proceedings of the 2015 18th Springer Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany, 2015: 234-241
- [85] Meng Z, Ma J, Yuan X. End-to-end low cost compressive spectral imaging with spatial-spectral self-attention// *Proceedings of the 2020 16th European Conference on Computer Vision*. Glasgow, UK, 2020: 187-204

- [86] Meng Z, Yuan X. Perception inspired deep neural networks for spectral snapshot compressive imaging//Proceedings of the 2021 International Conference on Image Processing. Anchorage, USA, 2021: 2813-2817
- [87] Justin J, Alexandre A, Li F F. Perceptual losses for realtime style transfer and super-resolution//Proceedings of the 2016 14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 694-711
- [88] Hu X, Cai Y, Lin J, et al. HDNet: High-resolution dual-domain learning for spectral compressive imaging//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 17521-17530
- [89] Cai Y, Lin J, Hu X, et al. Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 17481-17490
- [90] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need//Proceedings of the 2017 Conference on Neural Information Processing Systems. Long Beach, USA, 2017: 5998-6008
- [91] Lin J, Cai Y, Hu X, et al. Coarse-to-fine sparse transformer for hyperspectral image reconstruction//Proceedings of the 2022 17th European Conference on Computer Vision. Tel Aviv, Israel, 2022: 686-704
- [92] Wang J, Zhang Y, Yuan X, et al. Modeling mask uncertainty in hyperspectral image reconstruction//Proceedings of the 2022 17th European Conference on Computer Vision. Tel Aviv, Israel, 2022: 112-129
- [93] Wang J, Li K, Zhang Y, et al. S<sup>2</sup>-transformer for mask-aware hyperspectral image reconstruction. arXiv preprint, arXiv: 2209.12075, 2022
- [94] Cheng Z, Chen B, Lu R, et al. Recurrent neural networks for snapshot compressive imaging. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(2): 2264-2281
- [95] Gregor K, LeCun Y. Learning fast approximations of sparse coding//Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel, 2010: 399-406
- [96] Zhang J, Ghanem B. ISTA-Net: Interpretable optimization inspired deep network for image compressive sensing//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 1828-1837
- [97] Zhang X, Zhang Y, Xiong R, et al. HerosNet: Hyperspectral explicable reconstruction and optimal sampling deep network for snapshot compressive imaging//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 17511-17520
- [98] Wang L, Sun C, Fu Y, et al. Hyperspectral image reconstruction using a deep spatial-spectral prior//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 8032-8041
- [99] Wang L, Sun C, Zhang M, et al. DNU: Deep non-local unrolling for computational spectral imaging//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 1658-1668
- [100] Zhang S, Wang L, Zhang L, et al. Learning tensor low-rank prior for hyperspectral image reconstruction//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA, 2021: 12006-12015
- [101] Cai Y, Lin J, Wang H, et al. Degradation-aware unfolding half-shuffle transformer for spectral compressive imaging//Proceedings of the 2022 Conference on Neural Information Processing Systems. New Orleans, USA, 2022: 37749-37761
- [102] Meng Z, Jalali S, Yuan X. GAP-net for snapshot compressive imaging. arXiv preprint, arXiv: 2012.08364, 2020
- [103] Wang L, Wu Z, Zhong Y, et al. Spectral compressive imaging reconstruction using convolution and spectral contextual transformer. arXiv preprint, arXiv: 2201.05768, 2022
- [104] Sogabe Y, Sugimoto S, Kurozumi T, et al. Adminspired reconstruction network for compressive spectral imaging//Proceedings of the 2020 IEEE International Conference on Image Processing. Abu Dhabi, United Arab Emirates, 2020: 2865-2869
- [105] Huang T, Dong W, Yuan X, et al. Deep gaussian scale mixture prior for spectral compressive imaging//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA, 2021: 16216-16225
- [106] Venkatakrisnan S V, Bouman C A, Wohlberg B. Plug-and-play priors for model based reconstruction//Proceedings of the 2013 IEEE Global Conference on Signal and Information Processing. Austin, USA, 2013: 945-948
- [107] Chan S, Wang X, Elgendy O A. Plug-and-play ADMM for image restoration: Fixed-point convergence and applications. IEEE Transactions on Computational Imaging, 2017, 3(1): 84-98
- [108] Lai Z, Wei K, Fu Y. Deep plug-and-play prior for hyperspectral image restoration. Neurocomputing, 2022, 481: 281-293
- [109] Zheng S, Liu Y, Meng Z, et al. Deep plug-and-play priors for spectral snapshot compressive imaging. Photonics Research, 2021, 9(2): B18-B29
- [110] Meng Z, Yu Z, Xu K, et al. Self-supervised neural networks for spectral snapshot compressive imaging//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Quebec, Canada, 2021: 2622-2631
- [111] Yuan X, Liu Y, Suo J, et al. Plug-and-play algorithms for large-scale snapshot compressive imaging//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 1444-1454

- [112] Qiu H, Wang Y, Meng D. Effective snapshot compressive-spectral imaging via deep denoising and total variation priors//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA, 2021: 9127-9136
- [113] Zhang K, Zuo W, Zhang L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, 2018, 27(9): 4608-4622
- [114] <https://github.com/MaxtBIT/Summary-of-Codes-for-HSI-Reconstruction>
- [115] Arad B, Ben-Shahar O. Sparse recovery of hyperspectral signal from natural RGB images//Proceedings of the 2016 14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 19-34
- [116] Chakrabarti A, Zickler T E. Statistics of real-world hyperspectral images//Proceedings of the 2011 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA, 2011: 193-200
- [117] Yasuma F, Mitsunaga T, Iso D, et al. Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, 2010, 19(9): 2241-2253
- [118] Choi I, Jeon D S, Nam G, et al. High-quality hyperspectral reconstruction using a spectral prior. *ACM Transactions on Graphics*, 2017, 36(6): 218;1-218:13
- [119] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [120] Kruse F A, Lefkoff A B, Boardman J W, et al. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 1993, 44(2-3): 145-163
- [121] Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2010, 3(1): 1-122
- [122] Xiong Z, Shi Z, Li H, et al. HSCNN: CNN-based hyperspectral image recovery from spectrally undersampled projections//Proceedings of the 2017 IEEE/CVF International Conference on Computer Vision Workshops. Venice, Italy, 2017: 518-525
- [123] Zhou S, Xu T, Dong S. RDFNet: Regional dynamic FISTA-Net for spectral snapshot compressive imaging. *IEEE Transactions on Computational Imaging*, 2023, 9: 490-501
- [124] Molchanov P, Tyree S, Karras T, et al. Pruning convolutional neural networks for resource efficient inference//Proceedings of the 2017 5th International Conference on Learning Representations. Toulon, France, 2017: 1-17
- [125] Song L, Wang L, Kim M, et al. High-accuracy image formation model for coded aperture snapshot spectral imaging. *IEEE Transactions on Computational Imaging*, 2022, 8: 188-200
- [126] Han B, Yao Q, Liu T, et al. A survey of label-noise representation learning: Past, present and future. *arXiv preprint, arXiv: 2011.04406*, 2020
- [127] Sun Y, Yang Y, Liu Q, et al. Unsupervised spatial-spectral network learning for hyperspectral compressive snapshot reconstruction. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5514314;1-5514314:14
- [128] Cubuk E D, Zoph B, Mane D, et al. AutoAugment: Learning augmentation strategies from data//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 113-123
- [129] Kendall A, Gal Y, Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics //Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 7482-7491
- [130] Celikkanat A, Malliaros F D. Exponential family graph embeddings//Proceedings of the 2020 Conference on Artificial Intelligence. New York, USA, 2020: 3357-3364
- [131] Li J, Nguyen T V, Hegde C, et al. Implicit sparse regularization: The impact of depth and early stopping//Proceedings of the 2021 Conference on Neural Information Processing Systems. 2021: 28298-28309
- [132] Li L, Wang L, Song W, et al. Quantization-aware deep optics for diffractive snapshot hyperspectral imaging//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 19748-19757
- [133] Keshari R, Vatsa M, Singh R, et al. Learning structure and strength of CNN filters for small sample size training//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 9349-9358
- [134] Barshan E, Fieguth P W. Stage-wise training: An improved feature learning strategy for deep models//Proceedings of the 2015 Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 49-59
- [135] Gholami A, Kim S, Dong Z, et al. A survey of quantization methods for efficient neural network inference. *arXiv preprint, arXiv: 2103.13630*, 2021
- [136] Luo J H, Zhang H, Zhou H Y, et al. ThiNet: Pruning CNN filters for a thinner net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(10): 2525-2538
- [137] Hinton G, Vinyals O, Dean J, et al. Distilling the knowledge in a neural network. *Computer Science*, 2015, 14(7): 38-39



**MA Xiang-Tian**, Ph.D. candidate.

His research interests focus on computational photography.

**WANG Li-Zhi**, Ph.D., associate professor. His research interests include computational photography, visual information acquisition and reconstruction.

**HUANG Hua**, Ph.D., professor. His research interests include image and video processing.

## Background

Spectral images can provide both spatial and spectral information about objects, making them highly valuable in various fields such as remote sensing and medical diagnosis. The conventional technology of spectral imaging relies on scanning, which leads to low time resolution. This limitation poses a significant challenge, rendering conventional spectral imaging unsuitable for capturing dynamic scene shooting.

Coded aperture snapshot spectral imaging (CASSI) is a cutting-edge technology that derives complete spectral images from a single exposure and reconstruction algorithms. The ability utilized in a dynamic scene is the primary advantage of CASSI. Early research on CASSI focused on physical system design, including coded aperture design and dual-camera design. In recent years, the physical system of CASSI has become generally fixed. The reconstruction algorithms of CASSI have become the bottleneck of imaging quality. Researchers have developed several algorithm frameworks for CASSI, but a related survey has yet to be published. Conducting a comprehensive survey of existing works in the field is essential, as it provides a valuable reference for other scholars and serves as a guiding tool for researchers interested in exploring this particular domain. With this objective, we have collected and classified the existing works of CASSI. Furthermore, we briefly evaluate these works, offering an overview of the advancements and contributions in this area.

In this paper, we review and evaluate the reconstruction algorithms of CASSI. Specifically, we first introduce the physical system and forward model of CASSI. The physical system has become fixed. We describe the hardware structure and provide a set of design parameters. The system is modeled as a typical linear imaging system. Afterward, we list the characteristics and challenges of CASSI reconstruction. Unlike natural image restoration, CASSI reconstruction must consider the forward model fully. Then, we classify and summarize the existing reconstruction algorithms, including model-based and learning-based reconstruction algorithms. The main difference between model and learning is how the prior representation model is modeled. The former uses manually designed prior representation models such as smoothing, sparsity, and low rank, while the latter models prior knowledge in a data-driven manner. After that, we introduce commonly utilized datasets and evaluation indicators. Meanwhile, we summarize the source codes and analyze mainstream algorithms' reconstruction quality and computational efficiency. By comparison, each method's computational power requirements and time consumption can be estimated. Finally, we discuss the weakness of existing works and the trend of future research.

This work was supported in part by the National Natural Science Foundation of China under Grant No. 62131003.