

引用格式: ZHU Junjie, ZHAO Jufeng, TIAN Haijun, et al. Dual-camera Compressed Spectral Image Reconstruction Algorithm Based on Non-local Self-similarity[J]. Acta Photonica Sinica, 2023, 52(1):0111003

朱骏捷, 赵巨峰, 田海军, 等. 基于非局部自相似性的双相机压缩光谱图像重建算法[J]. 光子学报, 2023, 52(1):0111003

基于非局部自相似性的双相机压缩光谱 图像重建算法

朱骏捷^{1,2}, 赵巨峰^{1,2}, 田海军^{1,2}, 崔光茫^{1,2}, 石振^{1,2}

(1 杭州电子科技大学 碳中和新能源研究院, 杭州 310018)

(2 杭州电子科技大学 电子信息学院, 杭州 310018)

摘要: 针对压缩光谱成像的图像重建问题, 提出了一种基于非局部稀疏表示与双相机系统的压缩光谱重建方法。首先, 利用 RGB 观测来构建一种三维图像块, 使用 K 均值聚类对图像块进行分类, 并以聚类结果来指导目标高光谱图像的光谱块分类, 通过主成分分析获取每个簇的特征用来稀疏表示其他光谱块。然后用构建的三维图像块估计目标光谱图像非局部相似性, 并构建目标函数。最后, 通过迭代收缩算法与共轭梯度下降法来交替优化目标函数完成重建。仿真和实拍结果表明, 所提方法能大幅提升重建质量与精度, 在空间和光谱维度上重建误差更小, RGB 观测辅助字典学习与相似块估计的方法能有效提升双相机系统的计算效率。

关键词: 光谱成像; 压缩感知; 编码孔径; 非局部自相似性; 稀疏性; 双相机

中图分类号: O433.1

文献标识码: A

doi: 10.3788/gzxb20235201.0111003

0 引言

高光谱成像相比常用的 RGB(红、绿、蓝)三色成像, 能够拍摄更多的光谱通道, 获得更多被测目标的特征和细节, 从而广泛应用于遥感^[1]、医学成像^[2]、环境监测^[3]等领域。高质量、高精度的光谱成像与重建方法是本领域追求的目标之一。

高空间、光谱、时间分辨率的光谱成像技术是未来的发展方向。为了快速获取目标光谱信息, BRADY D J 等提出了编码孔径快照式光谱成像系统(Coded Aperture Snapshot Spectral Imager, CASSI)^[4-6]。这是基于压缩感知理论^[7]发展起来的一种快照式的光谱成像技术, 它对三维光谱数据进行稀疏低维采样, 在二维探测器上获得混叠的光谱信息, 然后通过计算重建出三维光谱数据, 只需要一次曝光就能获得目标的光谱信息, 具有获取光谱视频的能力。由于从二维采样中重建三维光谱数据是一个欠定的问题, 重建误差不可避免, 如何减小重建的误差是该成像方式的关键问题之一。

为了提高重建质量和精度, 研究者在重建框架中引入了不同的正则化方法。使用全变差正则化^[8-9]作为约束的算法是一种常用的重建算法, 它假设图像具有分段光滑性, 但是这种正则化方法会使图像过度平滑, 损失一些细节。针对采样数量的限制, KITTLE D 等^[10]通过拍摄多帧编码图像来进行重建, 即通过增加采样数量提升重建质量, 但同时也破坏了 CASSI 获取光谱视频的能力。WANG Lizhi 等^[11-12]在原来的 CASSI 基础上增加一路相机进行互补采样, 在提升重建质量的同时保持了快照特性。TAN Jin 等^[13]将小波变换作为稀疏基, 提出了一种基于 3D 维纳滤波的近似消息传递法, 但是这种预先定义的正交基不能很好地表示自然场景, 因此限制了重建质量。一些研究者通过研究高光谱图像的内在特性, 例如非局部性自相似性^[14-15]、低秩特性^[14,16]等, 提出了不同的重建模型, 但通常重建时间过长或者重建质量有限。近几年, 深度学习被应用于 CASSI 重建, 研究者们设计了不同的网

基金项目: 浙江省自然科学基金(Nos. LY22F050002, LGF20F050003, LQ20F030011), 浙江省科协“育才工程”项目(No. SKX201901)

第一作者: 朱骏捷, 836808746@qq.com

通讯作者: 赵巨峰, dabaozjf@hdu.edu.cn

收稿日期: 2022-06-15; 录用日期: 2022-07-19

<http://www.photon.ac.cn>

络结构用来重建编码图像。MIAO Xin等^[17]联合注意力机制与U-net设计了λ-net来实现端到端的高光谱图像重建。文献[18]中提出了一种具有空间光谱自注意力机制的模块来利用空间光谱相关性。然而,基于学习的方法通常需要数据来训练模型,并且重建效果容易受到训练数据的影响,限制了模型的泛化能力。

非局部自相似性在图像重建、去噪、超分辨率^[19-21]中有广泛的应用。在过去的一些方法中,由于只有退化的图像而没有更多的信息,相似性估计等操作都是通过重建的中间结果进行的,并且在重建过程中需要不断进行估计,对于尺寸较大的光谱图像非常费时。为了获得高质量的光谱图像重建,本文提出了一种基于非局部自相似性与双相机系统的重建算法,用双相机系统来获取目标信息,然后用RGB观测获取的额外信息来辅助重建。首先,利用高光谱图像存在的空间局部、非局部相似性以及谱间相关性,提出一种压缩光谱图像的重建模型,将恢复问题转化为变分优化模型;然后,采用RGB观测来辅助字典学习,并对目标的光谱图像存在的非局部相似性进行估计;最后,采用一种交替优化的方案来完成高光谱图像的重建。

1 双相机系统模型

双相机系统如图1所示,入射光进入系统后被平均分成两份,一份进入CASSI系统被灰度相机捕获,一份被RGB相机捕获。定义目标场景为 $f_0(i, j, \lambda) \in \mathbb{R}^{M \times N \times \Omega}$,其中 i, j 表示空间坐标, λ 表示光谱坐标, M, N 表示空间尺寸, Ω 表示通道数量。CASSI分支 (i, j) 的压缩测量可以表示为

$$y_c(i, j) = 0.5 \sum_{\lambda=1}^{\Omega} \omega_c(\lambda) T(i - \varphi(\lambda), j) f(i - \varphi(\lambda), j, \lambda) \quad (1)$$

式中, $\omega_c(\lambda)$ 表示灰度相机的响应曲线, $T(i, j)$ 表示编码模板, $\varphi(\lambda)$ 表示棱镜的色散函数。CASSI分支的采样可以写成线性矩阵形式,即

$$y_c = H_c F + V_c \quad (2)$$

式中, $H_c \in \mathbb{R}^{N(M+\Omega-1) \times MN\Omega}$,表示感知矩阵, $y_c \in \mathbb{R}^{N(M+\Omega-1)}$ 和 $V_c \in \mathbb{R}^{N(M+\Omega-1)}$ 分别表示观测图像和噪声的向量形式, $F \in \mathbb{R}^{MN\Omega}$ 表示目标高光谱图像向量。对于RGB相机分支,3个通道的观测图像分别是目标高光谱图像有效光谱范围内的积分,RGB分支的测量可以表示为

$$y_o(i, j) = 0.5 \sum_{\lambda=1}^{\Omega} \omega_o(\lambda) f_0(i, j, \lambda) \quad (3)$$

相应的线性矩阵形式为

$$Y_o = H_o F + V_o \quad (4)$$

式中, $H_o \in \mathbb{R}^{3MN \times MN\Omega}$ 表示RGB相机的前向响应矩阵, $V_o \in \mathbb{R}^{3MN}$ 是观测噪声,则

$$Y_o = [(y^b)^T, (y^g)^T, (y^r)^T]^T \quad (5)$$

式中, y^r, y^g, y^b 分别为3个通道的观测图像的向量表示。联合两个分支的测量,整个双相机系统的前向模型可以表示为

$$Y = HF + V \quad (6)$$

其中

$$Y = \begin{bmatrix} Y_c \\ Y_o \end{bmatrix}, H = \begin{bmatrix} H_c \\ H_o \end{bmatrix}, V = \begin{bmatrix} V_c \\ V_o \end{bmatrix} \quad (7)$$

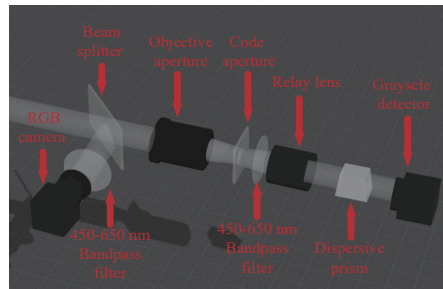


图1 双相机系统

Fig. 1 The system of dual camera

2 高光谱图像重建算法

高光谱图像是一个高度稀疏的信号,在空间和谱间都存在着丰富的信息冗余。具体表现为:光谱维度上,自然场景的光谱可以由6~8个主成分来表示^[22],并且同一场景中存在很多光谱曲线相似的区域。空间维度上,同一张图上可以找到很多相似结构的图像块。在充分考虑这些先验信息的基础上,构建了一种三维非局部稀疏表示模型来提升重建性能。

首先,使用RGB观测进行辅助字典学习。将初始化的高光谱图像 F 按步长为1分割成空间尺寸为 $b \times b \times \Omega$ 的跨全光谱波段的重叠立方体块。然后,将RGB相机分支获得的RGB图像的3个通道图像也分别按步长为1分割成尺寸为 $b \times b$ 的二维图像块。假设 $\mathbf{x}_{i,j}^r, \mathbf{x}_{i,j}^g, \mathbf{x}_{i,j}^b$ 分别是RGB三通道位于坐标 (i, j) 的图像块的向量表示,定义

$$\mathbf{X}_{i,j} = [(\mathbf{x}_{i,j}^r)^T, (\mathbf{x}_{i,j}^g)^T, (\mathbf{x}_{i,j}^b)^T]^T \quad (8)$$

利用K均值聚类对不同坐标的 $\mathbf{X}_{i,j}$ 进行分类,分成 K 个簇,然后以该聚类结果为指导将高光谱图像对应的立方体块进行相同的分类。定义 $\mathbf{M}_k \in \mathbb{R}^{n \times l_k}$,表示第 k 个簇的立方体块向量组成的矩阵。其中 $n = b \times b \times \Omega$ 表示立方体块向量的长度, l_k 表示第 k 个簇的立方体块个数。对每个簇的立方体块矩阵使用主成分分析降维,获取该簇的特征作为这个簇的字典,定义第 k 个簇的字典为 $\Phi_k \in \mathbb{R}^{n \times n}$ 。

定义向量 $\mathbf{x}_{i,j} \in \mathbb{R}^n$ 表示从高光谱图像 F 中取出的一个三维光谱块向量表示,处于坐标 (i, j) 的位置,即

$$\mathbf{x}_{i,j} = \mathbf{R}_{i,j} \mathbf{F} \quad (9)$$

式中, $\mathbf{R}_{i,j} \in \mathbb{R}^{n \times M\Omega}$ 是一个块提取矩阵,从 F 中以步长为 s 提取三维光谱块 $\mathbf{x}_{i,j}$ 。根据聚类结果找到 $\mathbf{x}_{i,j}$ 所属的簇,利用获取的字典 Φ_k 给当前簇中的块 $\mathbf{x}_{i,j}$ 进行稀疏编码,通过一个簇中共有的特征对当前簇中的其它块进行约束。对于给定的字典 Φ_k ,在充分考虑高光谱图像的谱间相似性、空间局部相似性与非局部相似性的基础上,构建重建模型为

$$(\hat{F}, \hat{\alpha}) = \arg \min_{F, \alpha} \frac{1}{2} \|Y - HF\|_2^2 + \sum_{k=1}^K \sum_{(i,j) \in C_k} \left(\frac{\gamma}{2} \|R_{i,j}F - \Phi_k \alpha_{i,j}\|_2^2 + \eta \|\alpha_{i,j} - \beta_{i,j}\|_1 \right) \quad (10)$$

式中, $\hat{F}, \hat{\alpha}$ 分别是重建的高光谱图像和其稀疏表示系数; α 为所有光谱块稀疏表示系数 $\alpha_{i,j}$ 的级联, C_k 表示第 k 个簇的光谱块的集合, γ 和 η 是正则化参数; $\|\alpha_{i,j} - \beta_{i,j}\|_1$ 项是利用高光谱图像的非局部冗余来提升稀疏表示的准确度;第二个求和项表示通过自身训练得到的字典来稀疏表示其他三维光谱块误差能最小,即通过利用高光谱图像的非局部冗余来提升重建效果; $\beta_{i,j}$ 为当前光谱块的邻域内 J 个相似块的加权和的稀疏表示系数,它是对真实稀疏表示系数的估计,即

$$\beta_{i,j} = (\Phi_k)^T \sum_{q \in \Omega_{i,j}} \omega_{i,j,q} \mathbf{x}_{i,j,q} \quad (11)$$

式中, $\mathbf{x}_{i,j,q}$ 表示在图像块 $\mathbf{x}_{i,j}$ 为中心的一个窗口 $\Omega_{i,j}$ 内第 q 个相似块, $\omega_{i,j,q}$ 是相应的加权系数。

由于除了退化的图像外没有更多的信息, $\mathbf{x}_{i,j}$ 的相似块 $\mathbf{x}_{i,j,q}$ 通常是根据重建的中间结果来估计^[19-21],但是对于尺寸比较大的三维光谱块非常费时,并且在重建过程中需要不断进行相似块的估计,进一步增加了时间花费。另一方面,由于重建误差的存在,从中间结果来估计相似块不一定准确。针对这一问题,用双相机系统的RGB观测获取的额外信息来进行相似块的估计,由于RGB图像每一个通道是目标高光谱图像一小段光谱范围的积分,如果两个图像块在RGB三个通道的距离都非常近,那么目标高光谱图像的两个光谱块之间距离也非常近。

RGB图像估计相似块具体方法为:假设 $\mathbf{x}_{i,j}^r, \mathbf{x}_{i,j}^g, \mathbf{x}_{i,j}^b$ 分别是RGB三通道位于坐标 (i, j) 的图像块的向量表示, $\mathbf{x}_{m,n}^r, \mathbf{x}_{m,n}^g, \mathbf{x}_{m,n}^b$ 分别是三通道位于坐标 (m, n) 的图像块的向量表示,定义距离 $D_{m,i}$ 为

$$D_{m,i} = \left\| \left[(\mathbf{x}_{i,j}^r)^T, (\mathbf{x}_{i,j}^g)^T, (\mathbf{x}_{i,j}^b)^T \right]^T - \left[(\mathbf{x}_{m,n}^r)^T, (\mathbf{x}_{m,n}^g)^T, (\mathbf{x}_{m,n}^b)^T \right]^T \right\|_2^2 \quad (12)$$

以 $D_{m,i}$ 的大小来估计目标高光谱图像的两个图像块之间的相似度。因此,可以定义

$$\omega_{i,j,q} = \frac{\exp(-D_{m,i}/h)}{\sum_{q \in \Omega_{i,j}} \exp(-D_{m,i}/h)} \quad (13)$$

通过 $D_{m,i}$ 的大小来为每个三维图像块 $x_{i,j}$ 找到与它距离最近的 J 个块。

在实验过程中,正则化参数 η 的大小对重建结果影响明显,如果 η 比较大,能比较好地抑制稀疏编码噪声,但同时也会抑制一些有用的成分。当 η 比较小时,会导致重建时受稀疏编码噪声影响较大,出现收敛速度过慢或者重建结果变差的问题。加权 l_p 范数能提升稀疏性,得到更好的结果^[23-24],因此设置了自适应参数

$$\eta_{i,j,l} = \frac{c}{|\alpha_{i,j,l} - \beta_{i,j,l}| + \epsilon} \quad (14)$$

式中, l 表示向量 $\alpha_{i,j}$ 中第 l 个元素, c 是一个预先设置的常数, ϵ 是一个比较小的常数。因此,将重建模型扩展为

$$(\hat{F}, \hat{\alpha}) = \arg \min_{F, \alpha} \frac{1}{2} \|Y - HF\|_2^2 + \sum_{k=1}^K \sum_{(i,j) \in C_k} \left(\frac{\gamma}{2} \|R_{i,j}F - \Phi_k \alpha_{i,j}\|_2^2 + \|\eta_{i,j}(\alpha_{i,j} - \beta_{i,j})\|_1 \right) \quad (15)$$

式中, $\eta_{i,j}$ 是一个对角矩阵, $\eta_{i,j,l}$ 是对角线上第 l 个元素的值。

为了求解式(15),将优化问题分解成两个子优化问题,然后交替优化来求解。在 $(t+1)$ 次迭代中,对于固定的 F ,稀疏编码系数 α 的优化为

$$\alpha^{(t+1)} = \arg \min_{\alpha} \sum_{k=1}^K \sum_{(i,j) \in C_k} \left(\frac{\gamma}{2} \|R_{i,j}F^{(t)} - \Phi_k \alpha_{i,j}^{(t)}\|_2^2 + \|\eta_{i,j}(\alpha_{i,j}^{(t)} - \beta_{i,j}^{(t)})\|_1 \right) \quad (16)$$

用迭代收缩算法^[25]来求解式(16),经过计算, α 中的每个元素可以由式(17)得到。

$$\alpha_{ijl} = \max(|v_{ijl} - \beta_{ijl}| - \frac{\eta_{i,j,l}}{\gamma}, 0) \odot \text{sign}(v_{ijl} - \beta_{ijl}) + \beta_{ijl} \quad (17)$$

式中, l 是一个向量中的第 l 个元素, α_{ijl} 表示向量 $\alpha_{i,j}$ 的第 l 个元素。

$$v_{ij} = \Phi_k^T R_{i,j} F^{(t)} \quad (18)$$

对于固定的 α, F 可以通过式(19)的子问题进行更新。

$$F^{(t+1)} = \arg \min_F \frac{1}{2} \|Y - HF^{(t)}\|_2^2 + \sum_{k=1}^K \sum_{(i,j) \in C_k} \left(\frac{\gamma}{2} \|R_{i,j}F^{(t)} - \Phi_k \alpha_{i,j}^{(t+1)}\|_2^2 \right) \quad (19)$$

这是一个二次优化问题,可以求出闭合解为

$$F^{(t+1)} = \left(H^T H + \gamma \sum_{k=1}^K \sum_{(i,j) \in C_k} R_{i,j}^T R_{i,j} \right)^{-1} \left(H^T Y + \gamma \sum_{k=1}^K \sum_{(i,j) \in C_k} R_{i,j}^T \Phi_k \alpha_{i,j}^{(t+1)} \right) \quad (20)$$

由于矩阵 H 尺寸很大,无法直接求解,用共轭梯度下降法来求近似解。通过式(17)和式(20)交替更新 α 和 F ,然后每 T 次迭代更新一次字典 Φ_k ,直到收敛,就能获得重建的目标光谱图像。

3 实验结果

3.1 实验细节

为了证明所提方法的有效性,用CAVE数据集^[26]进行测试。CAVE数据集包含了31幅400~700 nm波段的光谱图像,分辨率为10 nm。实验中,测试图片裁剪出大小 $256 \times 256 \times 31$ 的中心区域作为参考图像,并归一化到 $[0, 1]$ 。实验中光谱块尺寸大小为 $4 \times 4 \times 31$,采样步长 s 为 2×2 ,簇的个数 K 为50,对于每个图像块都在 25×25 的窗口中找到距离最近的15个块,迭代次数为400,编码模板服从 $p=0.5$ 的伯努利分布。实验中的对比方法有两种,分别是基于秩最小化的算法(DeSCI)^[14,27],基于自监督神经网络的算法(PnP-DIP)^[28-29]。本文方法与DeSCI的实验是在CPU为i7-10750H,内存为16 G的笔记本电脑上使用MATLAB 2016a完成,PnP-DIP实验是在GPU 1080Ti的服务器上完成。

为了评估重建图像的质量,用峰值信噪比(Peak Signal to Noise Ratio, PSNR)、光谱角制图(Spectral Angle Mapping, SAM)^[30]、结构相似度(Structural Similarity, SSIM)^[31]来衡量重建质量。其中PSNR与SSIM衡量重建的空间质量,值越大表示空间信息重建质量越好;SAM衡量重建的光谱曲线准确度,值越小表示重建的曲线与真实曲线越相似。

3.2 仿真数据测试

3.2.1 重建质量与精度

在CAVE数据集上的实验结果如表1所示,每个图像的最佳结果以粗体突出显示。可以看出本文方法

在 PSNR、SSIM、SAM 上都有显著的提高,在空间质量和光谱精度上优于其它方法。观察 PSNR 的变化,DeSCI 的平均 PSNR 是 25.80 dB, PnP-DIP 是 30.45 dB, 本文方法是 42.68 dB, 相比 DeSCI 平均提升了 16.88 dB, 约 65%, 比 PnP-DIP 平均提升 12.23 dB, 约 40%。从 SSIM 和 SAM 中可以看出, 本文方法重建的光谱图像在空间结构和光谱曲线的精度上也是最高的。

表 1 CAVE 数据集重建结果对比
Table 1 Comparison of reconstruction results of CAVE dataset

Image	DeSCI			PnP-DIP			Proposed method		
	PSNR	SSIM	SAM	PSNR	SSIM	SAM	PSNR	SSIM	SAM
Beads	20.43	0.698 1	0.310 5	26.00	0.799 4	0.196 7	36.42	0.978 6	0.067 7
Chart	23.06	0.834 5	0.220 1	23.15	0.846 5	0.205 8	43.49	0.992 9	0.046 0
Balloon	35.04	0.977 2	0.061 1	34.37	0.912 8	0.115 3	51.08	0.996 9	0.016 3
Oil_painting	22.68	0.651 9	0.266 8	29.48	0.822 4	0.116 6	40.45	0.983 6	0.043 5
Cloth	21.15	0.794 7	0.217 8	28.18	0.853 8	0.106 2	36.28	0.964 6	0.044 3
Glass	21.15	0.794 7	0.217 8	29.15	0.880 7	0.127 8	37.02	0.982 5	0.045 1
Paint	26.29	0.927 5	0.123 8	31.81	0.925 7	0.155 1	43.15	0.993 4	0.043 1
Lemons	35.59	0.959 4	0.088 5	37.94	0.927 2	0.108 4	51.87	0.997 3	0.023 7
Thread	26.38	0.839 1	0.185 7	32.25	0.857 2	0.143 4	43.31	0.991 4	0.036 8
Lemon_slice	26.23	0.868 7	0.201 0	32.14	0.886 6	0.175 5	43.19	0.989 8	0.058 7
Average	25.80	0.834 6	0.189 3	30.45	0.871 2	0.145 1	42.63	0.987 1	0.042 5

3.2.2 重建视觉效果对比

为了进一步比较重建质量,图 2 给出了不同方法在 670 nm 波段的重建图像,对重建图像的视觉质量进行比较。可以看出本文方法在视觉效果上好于其他方法。DeSCI 利用了高光谱图像的组稀疏性,通过加权核范数最小化的方式进行重建,这可能会导致重建的图像比较平滑。PnP-DIP 通过从观测图像中学习先验,并在重建过程中不断优化网络,获得了更好的重建结果,但仍然存在明显的伪影。从两个场景的重建图片和局部放大图片中,可以看出本文方法对于一些细小的纹理仍然能比较好的地重建出来,而其他方法在细节上都存在不同程度的损失。本文所提方法在视觉效果上也明显好于其他三种方法,不管在定量还是定性的比较上都获得了更好的效果。

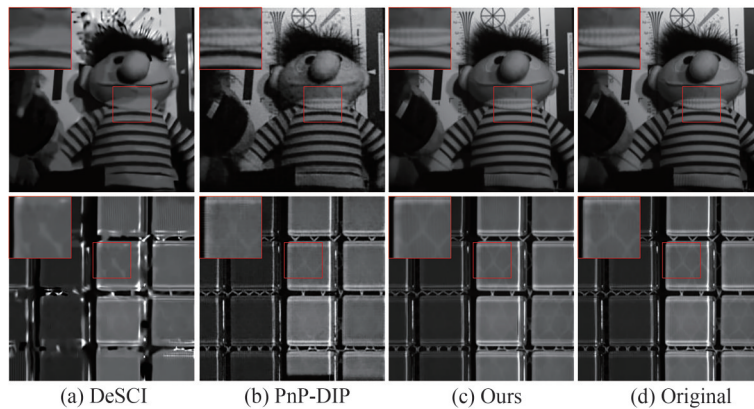


图 2 chart 和 glass 在 670 nm 的重建结果对比。第一行为 chart 场景,第二行为 glass 场景

Fig. 2 Comparison of reconstruction results of chart and glass at 670 nm. The first line is the chart scene, the second line is the glass scene

3.2.3 光谱维度重建误差分析

为了直观比较光谱域的重建误差,取重建的高光谱图像 chart 中两点进行比较,图 3 显示了 A、B 两点在不同方法下各个波段的绝对误差。在所有结果中,本文方法在所有波段都有最小的重建误差,且较为稳定,其他方法在不同波段的重建误差波动较大。表 2 显示了 chart 场景重建结果所有点光谱曲线平均绝对误差的均值,可以看出本文方法在整个场景光谱曲线的重建上也有明显更小的误差。

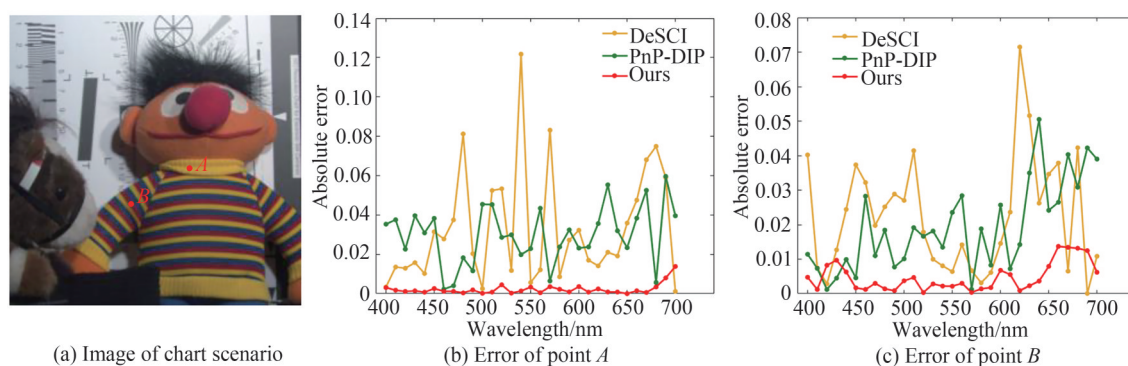


图 3 光谱曲线重建误差对比

Fig.3 Comparison of spectral curve reconstruction errors

表 2 平均绝对误差对比

Table 2 Comparison of mean absolute error

Method	DeSCI	PnP-DIP	Proposed method
Mean absolute error	0.033 4	0.026 6	0.003 9

3.2.4 兼顾质量的重建效率分析

首先对比了本文方法与 DeSCI, PnP-DIP 的计算时间。针对 chart 场景,三种方法重建时间如表 3 所示。DeSCI 需要进行相似块的匹配需要花费较多的时间,但是由于算法较快的收敛速度所以所需的时间比本文方法要少。PnP-DIP 从观测图像中学习先验,需要花费较长的时间,但是重建结果要好于 DeSCI。本文方法重建模型比较复杂,达到收敛所需的迭代次数比较多,所以花费时间较多,但是仍比 PnP-DIP 要快,并且重建质量在三种方法中是最高的。

表 3 chart 场景中不同方法重建时间对比

Table 3 Comparison of reconstruction time of different methods in chart scenarios

Method	DeSCI	PnP-DIP	Proposed method
Reconstruction time/s	2 608	7 024	3 355

事实上,本文方法在兼顾质量的重建效率提升方面也做了改进工作,主要体现在针对双相机体系,提出的 RGB 辅助字典学习与非局部相似性估计有效降低了重建时间。为了对比提出的 RGB 辅助字典学习与非局部相似性估计的策略和过去常用的中间结果字典学习与相似性估计^[23-25]策略的重建结果与重建效率的区别,用本文提出的重建模型在不同场景下测试,同时在 chart 场景测试了两种方法的重建时间。重建的 PSNR 对比结果如图 4 所示,重建效率对比如表 4 所示。从图中可以看出两种方法重建结果非常接近,但是在重建时间上 RGB 观测辅助重建的方法节约了约 47%。

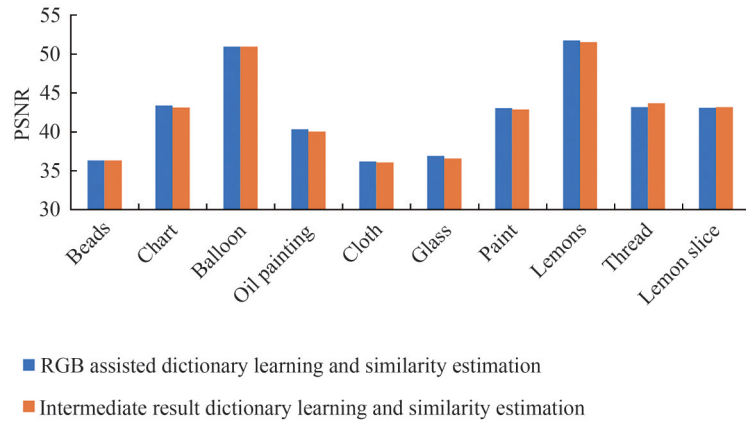


图4 RGB辅助字典学习与相似性估计方法以及中间结果字典学习与相似性估计方法的结果对比
Fig. 4 Comparison between RGB assisted dictionary learning and similarity estimation method and intermediate result dictionary learning similarity estimation method

表4 chart场景中RGB辅助字典学习与相似性估计方法以及中间结果字典学习相似性估计方法的重建时间对比
Table 4 Comparison of reconstruction time between RGB assisted dictionary learning and similarity estimation method and intermediate result dictionary learning similarity estimation method in the chart scene

Method	RGB auxiliary	Intermediate result
Reconstruction time/s	3 355	6 332

3.3 实拍数据测试

搭建了如图5所示的系统来进一步测试所提方法。该系统使用的物镜为 azure-0918M3M(焦距9 mm), 中继镜为 Edmund C45762, 灰度相机为 PointGray FL2-20S4M, RGB相机为 PointGray FL2-20S4C, 两个相机的像素尺寸都为 $4.4 \mu\text{m} \times 4.4 \mu\text{m}$ 。编码模板是一个 300×300 的随机二值矩阵, 每个最小单元尺寸为 $8.8 \mu\text{m} \times 8.8 \mu\text{m}$ 。棱镜为定制的双阿米西棱镜, 该棱镜使得 550 nm 的光能笔直穿过棱镜。

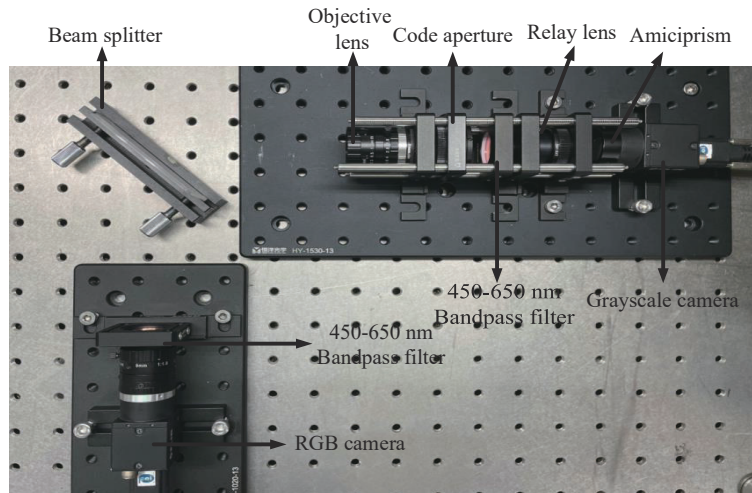


图5 双相机成像系统
Fig. 5 Dual camera imaging system

系统拍摄的图像如图6,重建的结果如图7所示。展示了 543 nm、606 nm、650 nm 三个波段的重建结果, 并使用滤光片拍摄图像作为参考。可以看出本文方法重建的图像视觉效果要更好, 保留了更多的细节, 与参考图像相似度最高。

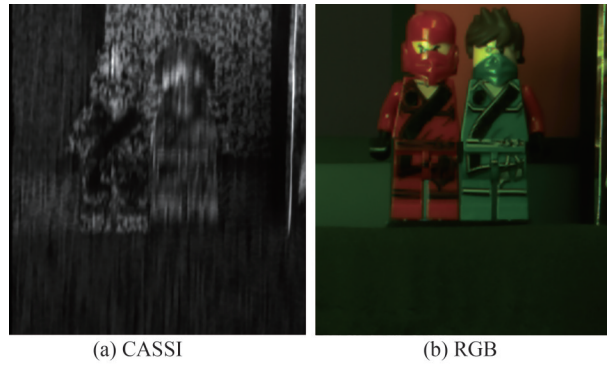


图6 两个分支的观测图像
Fig. 6 Observation images of two branches



图7 实拍数据重建结果
Fig. 7 The reconstruction results of real data

3.4 参数的讨论

重建过程中,图像块大小 b 、采样步长 s 和参数 c 对重建结果的影响比较明显,结果如图8所示。图像块太小时会破坏图像块的结构特征导致结果变差;图像块太大会使得恢复信号的难度增大,并且增加重建的耗时。采样步长比较小时会放大重建的误差使得结果变差,并且重建时间也会变长;而步长太大时块效应会变得明显。将图像块大小设置为 $4 \times 4 \times 31$,采样步长 s 设置为2。

参数 c 的影响如图8(c)所示,可以看出 c 比较大时迭代前期PSNR提升会比较快,收敛更快, c 比较小时收敛比较慢,但最后PSNR值更高,重建的空间误差更小。为了平衡这两个因素选择 $c = 0.0015$ 。

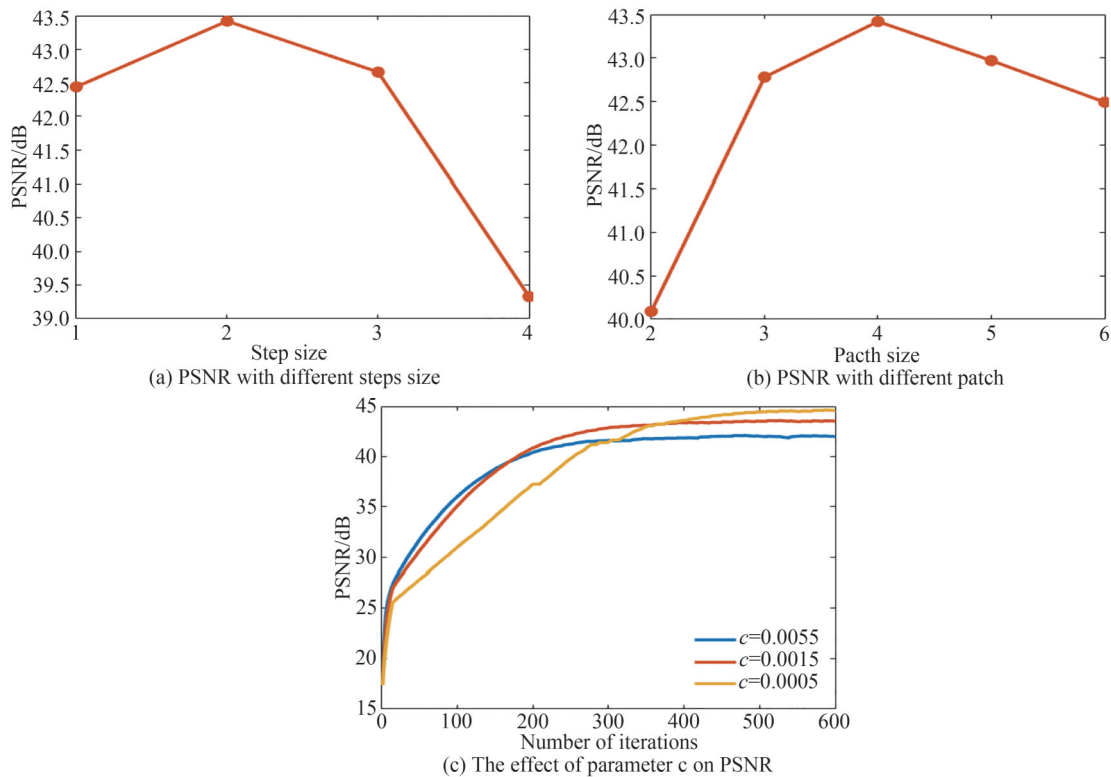


图8 不同参数对重建结果的影响

Fig. 8 Influence of different parameters on reconstruction results

4 结论

本文提出了一种基于非局部自相似性与双相机的高光谱图像重建方法。利用高光谱图像的内在特性,包括空间自相似性、光谱相关性和联合相关性,构建了模型来提高重建性能,并使用RGB观测来辅助字典学习和相似块估计,最后通过迭代重建出高光谱图像。实验表明本文提出的重建模型能大幅提升光谱图像的空间和光谱维度的重建质量,并且使用RGB观测辅助重建有效减少重建时间。虽然本文方法能获得高的重建质量,且在RGB观测的辅助下降低了重建时间,但是重建时间仍然比较长,如何进一步减少重建时间是未来的研究方向之一。

参考文献

- [1] STEIN D W J, BEAVEN S G, HOFF L E, et al. Anomaly detection from hyperspectral imagery [J]. IEEE Signal Processing Magazine, 2002, 19(1): 58-69.
- [2] PANASYUK S V, YANG Shi, FALLER D V, et al. Medical hyperspectral imaging to facilitate residual tumor identification during surgery[J]. Cancer Biology & Therapy, 2007, 6(3): 439-446.
- [3] STUART M B, MCGONIGLE A J S, WILLMOTT J R. Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems[J]. Sensors, 2019, 19(14): 3071.
- [4] GEHM M E, JOHN R, BRADY D J, et al. Single-shot compressive spectral imaging with a dual-disperser architecture [J]. Optics Express, 2007, 15(21): 14013-14027.
- [5] WAGADARIKAR A, JOHN R, WILLETT R, et al. Single disperser design for coded aperture snapshot spectral imaging [J]. Applied Optics, 2008, 47(10): B44-B51.
- [6] WAGADARIKAR A A, PITSIANIS N P, SUN Xiaobai, et al. Video rate spectral imaging using a coded aperture snapshot spectral imager[J]. Optics Express, 2009, 17(8): 6368-6388.
- [7] BARANIUK R G. Compressive sensing [lecture notes][J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.
- [8] YUAN Xin. Generalized alternating projection based total variation minimization for compressive sensing[C]. 2016 IEEE International Conference on Image Processing (ICIP), 2016: 2539-2543.
- [9] ZHANG Shipeng, HUANG Hua, FU Ying. Fast parallel implementation of dual-camera compressive hyperspectral imaging system[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(11): 3404-3414.
- [10] KITTLE D, CHOI K, WAGADARIKAR A, et al. Multiframe image estimation for coded aperture snapshot spectral

- imagers[J]. *Applied Optics*, 2010, 49(36): 6824–6833.
- [11] WANG Lizhi, XIONG Zhiwei, GAO Dahua, et al. Dual-camera design for coded aperture snapshot spectral imaging[J]. *Applied Optics*, 2015, 54(4): 848–858.
- [12] WANG Lizhi, XIONG Z, SHI Guangming, et al. Compressive hyperspectral imaging with complementary RGB measurements[C]. *2016 Visual Communications and Image Processing (VCIP)*, IEEE, 2016: 1–4.
- [13] TAN Jin, MA Yanting, RUEDA H, et al. Compressive hyperspectral imaging via approximate message passing[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2015, 10(2): 389–401.
- [14] LIU Yang, YUAN Xin, SUO Jinli, et al. Rank minimization for snapshot compressive imaging[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 41(12): 2990–3006.
- [15] WANG Lizhi, XIONG Zhiwei, SHI Guangming, et al. Adaptive nonlocal sparse representation for dual-camera compressive hyperspectral imaging[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(10): 2104–2111.
- [16] FU Ying, ZHENG Yinqiang, SATO I, et al. Exploiting spectral-spatial correlation for coded hyperspectral image restoration[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 3727–3736.
- [17] MIAO Xin, YUAN Xin, PU Yunchen, et al. l-net: reconstruct hyperspectral images from a snapshot measurement[C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019: 4059–4069.
- [18] MENG Ziyi, MA Jiawei, YUAN Xin. End-to-end low cost compressive spectral imaging with spatial-spectral self-attention[C]. *European Conference on Computer Vision*. Springer, Cham, 2020: 187–204.
- [19] DONG Weisheng, SHI Guangming, LI Xin. Nonlocal image restoration with bilateral variance estimation: a low-rank approach[J]. *IEEE Transactions on Image Processing*, 2012, 22(2): 700–711.
- [20] DONG Weisheng, ZHANG Lei, SHI Guangming, et al. Nonlocally centralized sparse representation for image restoration [J]. *IEEE transactions on Image Processing*, 2012, 22(4): 1620–1630.
- [21] DONG Weisheng, SHI Guangming, LI Xin, et al. Compressive sensing via nonlocal low-rank regularization[J]. *IEEE Transactions on Image Processing*, 2014, 23(8): 3618–3632.
- [22] VRHEL M J, GERSHON R, IWAN L S. Measurement and analysis of object reflectance spectra[J]. *Color Research & Application*, 1994, 19(1): 4–9.
- [23] CANDES E J, WAKIN M B, BOYD S P. Enhancing sparsity by reweighted ℓ_1 minimization[J]. *Journal of Fourier Analysis and Applications*, 2008, 14(5): 877–905.
- [24] DONG Weisheng, ZHANG Lei, LUKAC R, et al. Sparse representation based image interpolation with nonlocal autoregressive modeling[J]. *IEEE Transactions on Image Processing*, 2013, 22(4): 1382–1394.
- [25] DAUBECHIES I, DEFRISE M, DE MOL C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint[J]. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 2004, 57(11): 1413–1457.
- [26] YASUMA F, MITSUNAGA T, ISO D, et al. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum[J]. *IEEE Transactions on Image Processing*, 2010, 19(9): 2241–2253.
- [27] Decompress snapshot compressive imaging[EB/OL].[2022-06-15].<https://github.com/liuyang12/DeSCI>.
- [28] MENG Ziyi, YU Zhenming, XU Kun, et al. Self-supervised neural networks for spectral snapshot compressive imaging [C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 2622–2631.
- [29] CASSI-self-supervised[EB/OL].[2022-06-15].<https://github.com/mengziyi64/CASSI-Self-Supervised>.
- [30] 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[J]. *Remote Sensing of Environment*, 1993, 44(2–3): 145–163.
- [31] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612.

Dual-camera Compressed Spectral Image Reconstruction Algorithm Based on Non-local Self-similarity

ZHU Junjie^{1,2}, ZHAO Jufeng^{1,2}, TIAN Haijun^{1,2}, CUI Guangmang^{1,2}, SHI Zhen^{1,2}

(1 *Institute of Carbon Neutrality and New Energy, Hangzhou Dianzi University, Hangzhou 310018, China*)

(2 *School of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018, China*)

Abstract: Coded aperture spectral imaging is a snapshot spectral imaging method, but it usually has the

problems of large reconstruction error and high reconstruction computational complexity. To solve this problem, this paper proposes a compressed spectral reconstruction method based on non-local sparse representation and dual-camera system. First, a dual camera system is used to obtain the spectral and spatial data of the target. This dual camera system has two branches, the light is divided into two paths through a spectroscope, half enters the coded aperture spectral imaging system to obtain encoded images, and the other half is received by an RGB camera to obtain RGB images. The RGB observation image is used to construct 3D image patches, and k-means clustering is used to classify these 3D image patches. Then we propose a method to estimate the non-local similarity of target spectral image by RGB observation. The clustering and similarity estimation results of 3D image patches are used to guide the classification and similarity estimation of target spectral images. Divide the initialized target spectral image into a series of three-dimensional spectral patches, and classify the spectral patches based on the previous clustering results. Perform principal component analysis on each cluster, obtain the common features between different patches of the target spectral image, and use them to sparsely represent other spectral patches. For each patch, the sparse representation coefficients of the current patches are estimated by the weighted sum of sparse representation coefficients of nonlocal similar patches, and the weighted coefficients are calculated from the 3D image patches constructed by RGB observation. In order to improve the reconstruction quality, we set adaptive regularization parameters for sparse representation coefficients. We transform these operations into a variational optimization model, and then adopt an alternative optimization scheme to solve the objective function. We use conjugate gradient descent method and iterative threshold shrinkage method to optimize alternately. After every fifteen iterations, perform a principal component analysis on the classified three-dimensional spectral patches to obtain a new dictionary, and continue to repeat the iterative process. Through multiple repetitions, the final objective function converges, and the reconstructed spectral image can be obtained. We have done simulation experiments on the public spectral image dataset, and the experimental results show that our method has smaller spatial and spectral dimensions errors than other methods. We conducted simulation experiments on public spectral datasets, and the results show that our method has smaller errors in both spatial and spectral dimensions. From the perspective of spatial dimension, the proposed method can retain more details. From the spectral dimension, the method has smaller error and smaller error fluctuation in almost all wavebands than other methods. In addition, we compare the RGB auxiliary dictionary learning and similarity estimation method proposed in this paper with the common intermediate result dictionary learning and similarity estimation methods, the RGB auxiliary reconstruction method saves nearly half of the time while maintaining the same reconstruction quality. Finally, we set up an imaging system to do experiments on real data, and took images with filters for reference. The experiments show that our method can also obtain the best reconstruction quality on real data, which is most similar to the images obtained with filters. We also analyzed the influence of some factors, such as sampling step size and patch size, and selected the most appropriate parameter settings through a large number of experiments. Experiments on simulation data and real data show that our reconstruction model can greatly improve the reconstruction quality of spectral images in spatial and spectral dimensions, and the RGB observation assisted reconstruction method can effectively reduce the reconstruction time.

Key words: Spectral imaging; Compressed sensing; Coded aperture; Non-local similarity; Sparsity; Dual camera

OCIS Codes: 110.4234; 300.6380; 110.1220; 200.4560