

# 基于非线性光谱字典学习的单幅 RGB 图像光谱重建方法研究

左楚<sup>1</sup>, 谢德红<sup>2\*</sup>, 万晓霞<sup>3</sup>

1. 南京林业大学轻工与食品学院, 江苏 南京 210037
2. 南京林业大学信息科学技术学院, 江苏 南京 210037
3. 武汉大学, 湖北省文物颜色信息数字化与虚拟再现工程研究中心, 湖北 武汉 430079

**摘要** 针对单幅 RGB 图像重建光谱图像中的病态问题, 提出一种基于非线性光谱字典学习的非线性重建方法。为了适应线性和非线性数据, 该方法首先改进了基于自联想神经网络模型的非线性主成分分析算法, 并利用其从训练光谱集中学习低维光谱字典, 用于光谱重建的求逆方程中, 以缓解病态状况。再在此光谱字典基础上, 利用阻尼高斯牛顿法结合截断奇异值分解的正则化方法, 进一步缓解该非线性反演的病态问题, 实现单幅 RGB 图像重建光谱图像。在实验中, 采用 Munsell 以及 Munsell+Pantone 两个光谱训练集学习光谱字典, 同时利用 CAVE 和 UEA 光谱图像库进行光谱重建测试。该方法测试结果与现有方法比较发现, 该方法在不同光谱训练集下重建 CAVE 和 UEA 两库光谱图像的均方根差的平均值最低, 分别为 0.212 4, 0.255 4, 0.229 4 和 0.294 9, 均方根差的标准偏差接近最好方法的效果, 分别为 0.068 5, 0.084 7, 0.066 8 和 0.087 0。此结果表明该方法针对单幅 RGB 图像重建光谱图像在重建精度和稳定性上均存在优势。

**关键词** 光谱重建; RGB 图像; 非线性; 光谱字典; 学习

**中图分类号:** O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)07-2092-09

## 引言

光谱反射率因维度高、与光源、设备无关, 并能较全面、真实、客观地描述颜色、物质内部化学构成等信息<sup>[1]</sup>, 被广泛地应用于颜色和物质分析<sup>[2-3]</sup>、病症分析与判断<sup>[4]</sup>、地貌地貌的识别<sup>[5-6]</sup>。目前, 光谱图像常见获取方法有: 光谱相机<sup>[7]</sup>、装备宽带滤光片或窄带光源的普通相机<sup>[8-9]</sup>。光谱相机装备简单, 但价格昂贵、且所获光谱图像空间分辨率低; 装备宽带滤光片或窄带光源的普通相机较复杂、难携带且不易推广使用。因此, 近些年利用单幅 RGB 图像直接重建光谱图像成为一个重要研究课题。

通常, 光谱图像中光谱反射率维数远大于 RGB 图像中 RGB 值维数, 直接利用伪逆方法重建光谱反射率病态问题严重, 易产生较大重建误差<sup>[10]</sup>。针对此问题, 现有方法分为两类: 一类归为间接重建, 一类归为直接重建。间接重建方法认为像素的光谱反射率与 RGB 值在各自数据空间中的权重具有某种特定关系(如相等)。基于此, 间接重建方法从理论上回避了上述病态问题。如, Chou 等<sup>[11]</sup>在自然邻域插值

算法<sup>[12]</sup>的思想, 构建了光谱反射率与 RGB 值在各自数据空间插值权重相等的假设前提, 并在此前提下利用 Voronoi 网络重建了光谱反射率。此方法有效地回避了三对多重重建的病态问题, 但须测试样本与训练样本的光谱数据特征相似、且光谱反射率随波长均变化平缓<sup>[13]</sup>。现实生活中, 许多物质(如果荧光颜料)颜色鲜艳、且光谱反射率随波长变化剧烈、甚至存在窄波带波峰或波谷。Arad 等<sup>[14]</sup>提出了基于稀疏表示的重建方法(即 SR\_SR 方法), 该方法先利用 K-SVD 分别训练样本的光谱反射率数据与 RGB 数据, 得光谱和 RGB 两稀疏字典、以及两个字典之间的映射矩阵, 再依据两字典及两字典之间的关系重建了光谱反射率。此方法虽缓解了对光谱反射率曲线的要求, 但重建效果仍依赖于字典学习方法、以及测试样本集与训练样本集的数据特征相似程度。

依据色度学原理, 光谱反射率数据与 RGB 数据之间存在一个固有映射方程, 如式(1)。直接重建方法为缓解反演此方程病态问题的方法。基于主成分分析法(principal component analysis, PCA)重建方法<sup>[15-19]</sup>为典型的直接重建方法, 它假设任意颜色的光谱反射率均可由少数固有光谱征向量(即低维光谱字典)线性组成。预先利用 PCA 获得低维光

收稿日期: 2021-04-27, 修订日期: 2021-11-11

基金项目: 国家自然科学基金项目(61275172, 61575147), 南京林业大学青年科技创新基金项目(CX2018024)资助

作者简介: 左楚, 女, 1997年生, 南京林业大学硕士研究生 e-mail: zuochu@njfu.edu.cn

\* 通讯作者 e-mail: dehong.xie@gmail.com

谱字典后, RGB 数据与光谱反射率数据之间的反演问题则转变为 RGB 数据与光谱特征向量系数之间的反演问题。因此, 相对伪逆方法, 该方法降低了求解的维数、改善了病态状况, 提高了光谱反射率的重建精度。实际上, 该方法能否有效降低求解维数在很大程度上依赖于光谱字典学习方法。PCA 只适用于线性数据, 所获光谱字典难以低维表示非线性数据<sup>[20-23]</sup>。Marjan<sup>[23]</sup>针对非线性光谱反射率数据, 提出利用非线性主成分分析法 (nonlinear principal component analysis, NLPCA) 训练光谱字典。此 NLPCA<sup>[24]</sup>是一个基于优化方程的自联想神经网络模型。在 NLPCA 的神经网络框架中, 非线性传递函数可使获得地特征向量 (或低维字典) 切合数据的非线性特征, 但也易使优化方程在求解过程中局部收敛、而难获得全局最优解, 最终影响光谱重建精度。

鉴于此, 本文提出一种基于非线性光谱字典学习的非线性重建方法。为了适应光谱反射率数据的线性或非线性特征, 该方法首先通过改进传统 NLPCA 的优化方程, 以适用于光谱反射率数据的线性或非线性特征, 从而获得精度高、稳定性好的低维光谱字典, 再在此光谱字典基础上利用阻尼高斯牛顿法结合截断奇异值分解的正则化方法缓解该非线性反演的病态问题, 以实现精度高、稳定性好的光谱图像重建。

## 1 光谱直接重建框架

图像 RGB 数据与光谱反射率数据之间的映射方程如式 (1)

$$\begin{aligned} p_r(m) &= \sum_{n=1}^N E(\lambda_n) s_r(\lambda_n) r(\lambda_n, m) \\ p_g(m) &= \sum_{n=1}^N E(\lambda_n) s_g(\lambda_n) r(\lambda_n, m) \\ p_b(m) &= \sum_{n=1}^N E(\lambda_n) s_b(\lambda_n) r(\lambda_n, m) \end{aligned} \quad (1)$$

式(1)中,  $p_r(m)$ ,  $p_g(m)$  和  $p_b(m)$  分别为第  $m$  个像素点的 RGB 数据、且  $m=1, \dots, M$ ,  $M$  为图像总的像素数,  $E(\lambda_n)$  为光源的光谱功率分布,  $s_r(\lambda_n)$ ,  $s_g(\lambda_n)$  和  $s_b(\lambda_n)$  分别为 RGB 三个光谱响应曲线,  $r(\lambda_n, m)$  为第  $m$  像素、第  $n$  波段数的波长  $\lambda_n$  处的光谱反射率,  $N$  为可见光光谱范围内光谱反射率波段数。由此, 映射方程可进一步表示为

$$\mathbf{p} = \mathbf{A}\mathbf{r} \quad (2)$$

式(2)中,  $\mathbf{p} = [p_r(m) \ p_g(m) \ p_b(m)]^T \in R^{3 \times 1}$  为 RGB 数据构成的列向量,  $\mathbf{A} \in R^{3 \times N}$  为由  $E(\lambda_n) s_r(\lambda_n)$ ,  $E(\lambda_n) s_g(\lambda_n)$  和  $E(\lambda_n) s_b(\lambda_n)$  构成的光谱响应矩阵,  $\mathbf{r} = [r(\lambda_1, m) \ \dots \ r(\lambda_n, m) \ \dots \ r(\lambda_N, m)] \in R^{N \times 1}$  表示光谱反射率数据构成的列向量。由式(2), 重建光谱反射率的关键在于矩阵  $\mathbf{A}$  的求逆。  $\mathbf{A}$  为  $3 \times N$  矩阵且  $N \gg 3$ , 则矩阵  $\mathbf{A}$  的秩  $\text{rank}(\mathbf{A}) \leq \min(3, N) = 3$ , 但所求光谱反射率  $\mathbf{r}$  的维数为  $N$ , 由此  $\mathbf{r}$  的解不唯一。依上文所述, 假设已知光谱反射率数据的低维光谱字典, 式(2)可重写为式(3)

$$\mathbf{p} = \mathbf{A}\mathbf{f}(\mathbf{D}, \mathbf{u}) \quad (3)$$

式(3)中,  $\mathbf{D} \in R^{N \times K}$  表示低维光谱字典,  $\mathbf{u} \in R^{K \times 1}$  为系数向

量,  $K$  为系数向量的维数, 且  $K \ll N$ 。  $\mathbf{r} = \mathbf{f}(\mathbf{D}, \mathbf{u})$  为光谱低维表达,  $\mathbf{f}(\cdot)$  表示线性或非线性光谱表达函数, 若  $\mathbf{f}(\cdot)$  是线性的, 光谱表达可为  $\mathbf{r} = \mathbf{D}\mathbf{u}$ , 且式(3)可写作式(4)

$$\mathbf{p} = \mathbf{f}(\mathbf{A}\mathbf{D}, \mathbf{u}) \quad (4)$$

依据定理<sup>[25]</sup>, 式(4)所求系数向量  $\mathbf{u}$  的维数为  $K$ 。式(4)中,  $\mathbf{A}\mathbf{D} \in R^{N \times K}$  为  $N \times K$  矩阵, 其秩为  $\text{rank}(\mathbf{A}\mathbf{D}) \leq \min(N, K) = K$ 。若矩阵  $\mathbf{A}\mathbf{D}$  的秩等于  $K$ , 则式(4)存在  $\mathbf{u}$  的唯一解。在此情况下,  $K$  越接近 3, 越有利于缓解  $\mathbf{u}$  反演的病态, 但光谱表达精度越低。在本文中, 为了平衡光谱字典  $\mathbf{D}$  准确表达原光谱反射率数据与降低式(4)反演的病态问题, 除了  $\mathbf{f}(\cdot)$  线性或非线性自适应于训练光谱反射率数据外,  $K$  的取值一般远小于  $N$  且大于等于 3, 本文  $K$  取 6。此外, 式中  $\mathbf{A} \in R^{3 \times N}$  非方阵, 式(2)方程求解一般需乘以  $\mathbf{A}^T$ , 转换为方阵  $\mathbf{A}^T \mathbf{A} \in R^{N \times N}$ , 即

$$\mathbf{A}^T \mathbf{p} = \mathbf{A}^T \mathbf{A}\mathbf{r} \quad (5)$$

式(5)中, 方阵  $\mathbf{A}^T \mathbf{A}$  的秩  $\text{rank}(\mathbf{A}^T \mathbf{A}) = \min(\text{rank}(\mathbf{A}^T), \text{rank}(\mathbf{A})) \leq N$ 。唯一解的充分必要条件是  $\mathbf{A}^T \mathbf{A}$  满秩, 即  $\text{rank}(\mathbf{A}^T \mathbf{A}) = N$ 。若  $\mathbf{A}^T \mathbf{A}$  满秩, 其  $N$  个特征值  $\sigma_1 > \sigma_2 > \dots > \sigma_{N-1} > \sigma_N$  均不为 0。依据衡量解稳定性指标, 病态条件数  $\text{cond}(\mathbf{A}^T \mathbf{A}) = |\sigma_1|/|\sigma_N| \geq 1$  越大则解的稳定性越差。若采用光谱字典策略、且  $K > 3$ , 虽其病态条件数  $\text{cond}([\mathbf{A}\mathbf{D}]^T [\mathbf{A}\mathbf{D}]) = |\tilde{\sigma}_1|/|\tilde{\sigma}_K|$  可能小于  $\text{cond}(\mathbf{A}^T \mathbf{A})$ , 但仍不能保证其获得一个临近 1 的病态条件数。因此, 本文方法在使用字典策略的同时, 还利用阻尼高斯牛顿法结合截断奇异值分解的正则化方法, 降低病态状况, 改善解的稳定性。

鉴于此, 本文光谱图像重建框架如图 1 所示: 首先, 选取光谱反射率数据作为光谱训练数据集, 并利用本文所提光谱字典学习方法获得低维光谱字典与光谱表达函数, 再利用阻尼高斯牛顿法结合截断奇异值分解的正则化法, 获得单幅 RGB 图像的光谱图像。

## 2 光谱字典学习

### 2.1 基于 NLPCA 的光谱字典学习方法

鉴于上文所述, NLPCA<sup>[24]</sup>作为 PCA 的非线性扩展, 被用以训练低维光谱字典 (即固有光谱特征向量组), 获得了一定效果。NLPCA 是基于五个层结构的自联想人工神经网络的一种特征向量提取方法, 其优化求解方程如式(6)和式(7)

$$\hat{\mathbf{D}} = \arg \min_{\mathbf{D}} (J(\mathbf{D})) \quad (6)$$

$$J(\mathbf{D}) = \langle \|\mathbf{r} - \tilde{\mathbf{r}}\|_2 \rangle \quad (7)$$

式(7)中,  $\langle \cdot \rangle$  表示平均操作  $\frac{1}{M} \sum_{m=1}^M \{\|\mathbf{r} - \tilde{\mathbf{r}}\|_2\}_m$ ,  $\mathbf{D} = [d_1 \ \dots \ d_k \ \dots \ d_K] \in R^{N \times K}$  表示获得光谱字典。  $\mathbf{r}$  和  $\tilde{\mathbf{r}}$  均为光谱反射率数据,  $\mathbf{r}$  为测量值,  $\tilde{\mathbf{r}}$  为此 NLPCA 利用前馈神经网络方法重构值, 其重建公式如式(8)

$$\tilde{\mathbf{r}} = f_4(\mathbf{W}^{(D)} \mathbf{h}^{(D)} + \bar{\mathbf{b}}^{(D)}) \quad (8)$$

式(8)中, 传递函数  $f_4(\cdot)$  表示在神经网络框架下将  $\mathbf{h}^{(D)}$  映射成  $\tilde{\mathbf{r}}$  的过程。  $\mathbf{h}^{(D)}$  表示神经网络框架下隐含层神经元,  $\mathbf{h}^{(D)}$  的公式为式(9)

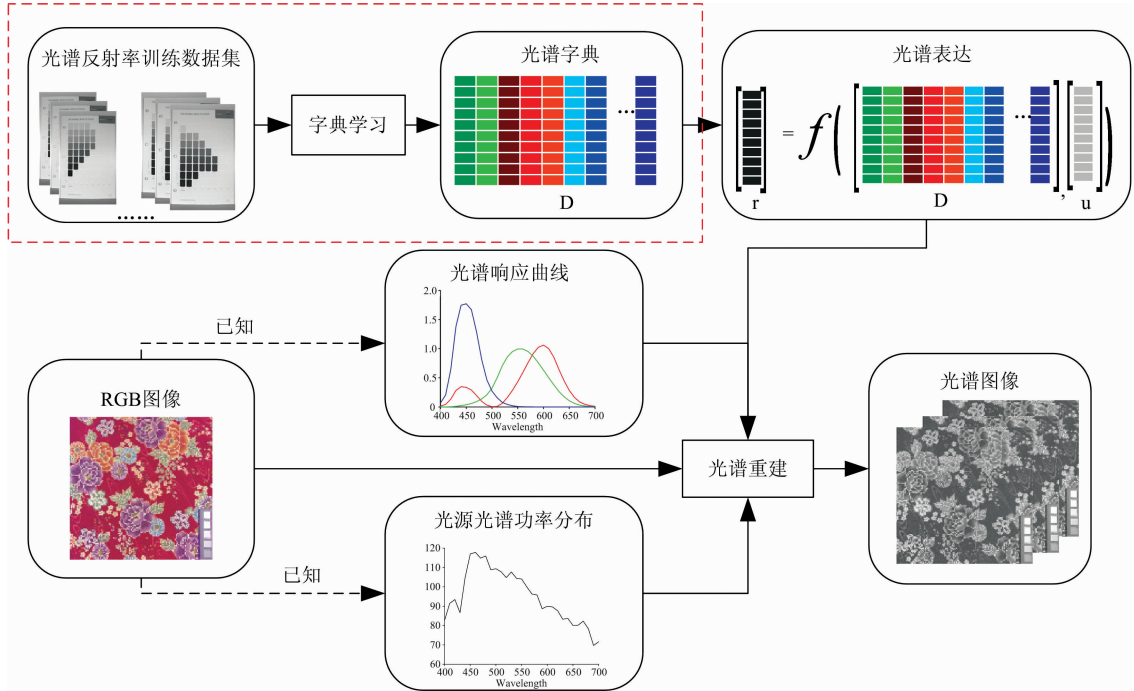


图 1 本文光谱重建框架

Fig. 1 Overview of our proposed spectral reconstruction method

$$\mathbf{h}^{(D)} = f_3(\mathbf{w}^{(D)}\mathbf{D} + \mathbf{b}^{(D)}) \quad (9)$$

式(9)中, 传递函数  $f_3(\cdot)$  表示在神经网络框架下将光谱字典  $\mathbf{D}$  映射成  $\mathbf{h}^{(D)}$  的过程。光谱字典的公式为式(10)

$$\mathbf{D} = f_2(\mathbf{w}^{(r)}\mathbf{h}^{(r)} + \bar{\mathbf{b}}^{(r)}) \quad (10)$$

式(10)中, 传递函数  $f_2(\cdot)$  表示在神经网络框架下将隐含层神经元  $\mathbf{h}^{(r)}$  映射成光谱字典  $\mathbf{D}$  的过程。隐含层神经元  $\mathbf{h}^{(r)}$  的公式为式(11)

$$\mathbf{h}^{(r)} = f_1(\mathbf{W}^{(r)}\mathbf{r} + \mathbf{b}^{(r)}) \quad (11)$$

式(11)中, 传递函数  $f_1(\cdot)$  表示在神经网络框架下将反射率光谱  $\mathbf{r}$  映射成  $\mathbf{h}^{(r)}$  的过程。 $\mathbf{W}^{(r)}$ ,  $\mathbf{b}^{(r)}$ ,  $\mathbf{w}^{(r)}$ ,  $\bar{\mathbf{b}}^{(r)}$ ,  $\mathbf{w}^{(D)}$ ,  $\mathbf{b}^{(D)}$ ,  $\bar{\mathbf{b}}^{(D)}$  为各层权重和偏置参数。 $f_1(\cdot)$  和  $f_3(\cdot)$  一般采用双曲正切函数,  $f_2(\cdot)$  和  $f_4(\cdot)$  一般采用恒等函数<sup>[24]</sup>。

## 2.2 基于改进 NLPCA 的光谱字典学习

PCA 是 NLPCA 的特殊线性形式, 即当上述优化方程中传递函数为恒定函数时, NLPCA 则等同于 PCA。因此, 在 NLPCA 的神经网络框架中, 传递函数的非线性化可使获得的特征向量更切合非线性化特征的数据, 但也易使优化方程在局部最小收敛, 而难以获得全局最优。本文将在 NLPAC 的基础上, 在优化方程中添加正则化项, 用以约束非线性传递函数, 避免导致过度非线性化, 以获得最优光谱字典。

如上所述,  $f_1(\cdot)$  和  $f_3(\cdot)$  采用是双曲正切函数, 即  $f_1(x) = f_3(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ 。当其函数自变量  $x$  在某个范围内 (如  $x \in (-L, L)$ ) 时, 若权重值  $w$  太小, 传递函数趋于线性 (即  $f_1(wx) = f_3(wx) \approx wx$ ), 获得的特征向量则难以切合数据非线性特征; 若权重值  $w$  太大, 传递函数趋于一个阶梯函数, 优化方程则可能在局部最小收敛。因此, 本文在式(7)优化方程的基础上, 约束权重  $\mathbf{W}^{(r)}$  和  $\mathbf{W}^{(D)}$ , 避免产

生较大权重值, 以有效控制非线性程度, 使所得特征向量更切合训练数据特征、且解趋于全局最优。改进地优化方程如式(12)

$$J_M(\mathbf{D}) = \langle \|\mathbf{r} - \tilde{\mathbf{r}}\|_2 \rangle + \mu(\|\mathbf{W}^{(r)}\|_2 + \|\mathbf{W}^{(D)}\|_2) \quad (12)$$

式(12)中,  $\|\cdot\|_2$  表示  $l_2$  范数,  $\mu > 0$  为正则化参数。依据优化方法分析<sup>[26]</sup>,  $\mu$  值越大, 权重矩阵  $\mathbf{W}^{(r)}$  和  $\mathbf{W}^{(D)}$  元素值则越小, 非线性程度则越低; 但若  $\mu$  值过大, 权重矩阵的元素值则均趋于 0, 式(12)优化方程则等同于式(7)。

为了获取最佳正则化参数  $\mu$  和权重与偏置参数  $\mathbf{W}^{(r)}$ ,  $\mathbf{b}^{(r)}$ ,  $\mathbf{w}^{(r)}$ ,  $\bar{\mathbf{b}}^{(r)}$ ,  $\mathbf{w}^{(D)}$ ,  $\mathbf{b}^{(D)}$ ,  $\mathbf{W}^{(D)}$ ,  $\bar{\mathbf{b}}^{(D)}$ , 本文利用独立验证集, 并以此数据集的色度属性随机分成训练集和验证集。其中, 训练集和验证集中样本均尽量均匀分布在整个色域中。 $\mu \in [0.001 \ 0.01 \ 0.1 \ 0.2 \ \dots \ 0.9 \ 1]$  在 0.001~1 范围内依次进行测试, 并通过均方差最小准则获得最优值为 0.1。此外, 本文利用反向传播的共轭梯度下降法求解上述优化方程, 获取最优参数。

## 3 非线性光谱重建

由上所述, 通过一组训练光谱数据集获得低维光谱字典  $\mathbf{D}$ , 并依据式(4)的光谱表达函数, 建立 RGB 数据  $\mathbf{p}$  与其光谱反射率数据  $\mathbf{r}$  在光谱字典  $\mathbf{D}$  下的系数向量  $\mathbf{u}$  的求解优化方程, 如式(13)

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \|\mathbf{p} - f(\mathbf{AD}, \mathbf{u})\|^2 \quad (13)$$

式(13)中,  $\mathbf{u}$  的维数为  $K$ , 远小于光谱  $\mathbf{r}$  维数 (即  $K \ll N$ ), 但仍高于维数  $\mathbf{p}$  (即  $K > 3$ )。因此, 式(13)的求解仍然面临着病

态非线性数据反演问题。本文将采用阻尼高斯-牛顿法<sup>[27]</sup>对相关系数矩阵  $\mathbf{u}$  进行求解,并结合截断奇异值分解<sup>[28]</sup>的正则化方法以缓解该非线性反演的病态问题<sup>[29]</sup>。

令  $h(\mathbf{u}) = \mathbf{p} - f(\mathbf{A}\mathbf{D}, \mathbf{u})$ , 式(13)则为

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \|h(\mathbf{u})\|^2 \quad (14)$$

利用高斯牛顿法解决此最小二乘问题,则迭代公式为

$$h(\mathbf{u}_{k+1}) \simeq h(\mathbf{u}_k) + \mathbf{J}(\mathbf{u}_k)\mathbf{s}_k \quad (15)$$

式(15)中,下标  $k$  表示迭代次数,  $\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{s}_k$ ,  $\mathbf{J}(\mathbf{u})$  为  $h(\mathbf{u})$  的雅可比行列式,即

$$\mathbf{J}(\mathbf{u}) = \frac{\partial h(\mathbf{u})}{\partial \mathbf{u}} \quad (16)$$

在每次迭代  $k$  中,  $\mathbf{s}_k$  求解优化方程如式(17)

$$\min_{\mathbf{s}} \|h(\mathbf{u}_k) + \mathbf{J}_k \mathbf{s}\|^2 \quad (17)$$

式(17)中,  $\mathbf{J}_k$  为  $\mathbf{J}(\mathbf{u}_k)$  (雅可比行列式)或有限差分近似计算获得。鉴于  $\mathbf{J}(\mathbf{u}_k)$  计算过于复杂性,本文采用有限差分近似表示如式(18)

$$\mathbf{J}(\mathbf{u}_k) = \frac{\partial h(\mathbf{u}_k)}{\partial \mathbf{u}_k} = \frac{h(\mathbf{u}_k + \delta) - h(\mathbf{u}_k)}{\delta} \quad (18)$$

式(18)优化方程的最小二乘法求解可得如式(19)表示

$$(\mathbf{J}_k)^T (\mathbf{J}_k \mathbf{s} + h(\mathbf{u}_k)) = 0 \quad (19)$$

式(19)中,上标 T 表示矩阵转置。由式(19)可得迭代公式如式(20)

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{s}_k = \mathbf{u}_k - (\mathbf{J}_k)^T h(\mathbf{u}_k) \quad (20)$$

为了保证收敛性且避免陷入局部最小,引入一个阻尼因子  $\alpha$ , 获得阻尼高斯-牛顿法的迭代公式如式(21)

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \alpha \mathbf{s}_k = \mathbf{u}_k - \alpha (\mathbf{J}_k)^T h(\mathbf{u}_k) \quad (21)$$

式(21)中,  $(\mathbf{J}_k)^T$  为  $\mathbf{J}_k$  的 Moore-Penrose 伪逆。阻尼因子  $\alpha$  是通过结合 Armijo-Goldstein 原理<sup>[30]</sup>确定的。

如上所析,式(17)求解方程仍旧是病态的,本文将 TSVD 方法进行正则化处理,式(17)的优化方程更新如式(22)

$$\min_{\mathbf{s} \in S} \|\mathbf{L}\mathbf{s}\|^2, \mathbf{s} = \{s \in R^K : (\mathbf{J}_k)^T \mathbf{J}_k \mathbf{s} + (\mathbf{J}_k)^T h(\mathbf{u}_k) = 0\} \quad (22)$$

式(22)中,  $\mathbf{L} \in R^{t \times K}$  ( $t < K$ ) 表示正则化矩阵,通常为单位矩阵,且满足  $N(\mathbf{J}) \cap N(\mathbf{L}) = \{0\}$  假设前提。 $\mathbf{J}_k$  和  $\mathbf{L}$  的广义奇

异值分解为式(23)

$$\mathbf{J}_k = \mathbf{C}\boldsymbol{\Sigma}_J \mathbf{Z}^{-1}, \mathbf{L} = \mathbf{V}\boldsymbol{\Sigma}_L \mathbf{Z}^{-1} \quad (23)$$

式(23)中,  $\mathbf{C}$  和  $\mathbf{V}$  为正交矩阵,  $\mathbf{Z}$  为非奇异矩阵,  $\boldsymbol{\Sigma}_J$  和  $\boldsymbol{\Sigma}_L$  为对角矩阵,表示如式(24)

$$\boldsymbol{\Sigma}_J = \begin{bmatrix} 0 & \mathbf{B} & 0 \\ 0 & 0 & \mathbf{I}_{K-t} \end{bmatrix}, \boldsymbol{\Sigma}_L = \begin{bmatrix} \mathbf{I}_{K-3} & 0 & 0 \\ 0 & \mathbf{X} & 0 \end{bmatrix} \quad (24)$$

式(24)中,  $\mathbf{B} = \text{diag}(b_1, \dots, b_t)$ ,  $\mathbf{X} = \text{diag}(x_1, \dots, x_t)$ 。

由此,优化方程式(22)求解如式(25)

$$\mathbf{s}^{(l)} = - \sum_{i=p-t+1}^{\bar{p}} \frac{\mathbf{c}_i^T h(\mathbf{u}_k)}{b_i} \mathbf{z}_{K-3+i} - \sum_{i=p+1}^3 [\mathbf{c}_i^T h(\mathbf{u}_k)] \mathbf{z}_{K-3+i} \quad (25)$$

式(25)中,  $l = 0, 1, \dots, \bar{p}$ , 且  $\bar{p} = 3 - K + t$ 。其中  $\kappa = \text{rank}(\mathbf{J}_k)$ ,  $l$  为正则化参数。

结合式(25),更新式(21)为如式(26)正则化阻尼高斯-牛顿法公式

$$\mathbf{u}_{k+1}^{(l)} = \mathbf{u}_k^{(l)} + \alpha_k \mathbf{s}_k^{(l)} \quad (26)$$

## 4 结果与讨论

为了验证方法的有效性,本文将分两个步骤进行评价:光谱字典评价和光谱反射率重建评价。

### 4.1 光谱字典评价

如上所析,为了验证本文光谱字典学习方法对线性数据集和非线性数据集的适用性,本文选用了两个数据集: Munsell 数据集和 Munsell+Pantone 数据集。Munsell 数据集的样本来源于 Munsell 颜色书<sup>[21]</sup>,包含 1269 个样本,其样本的颜色均匀分布在可见颜色空间[如图 2(a)],且样本的光谱反射率随波长变化均比较平缓; Pantone 数据集来源于 Pantone 荧光色卡 FFN100<sup>[31]</sup>,包含 21 个样本,其样本均为荧光颜色、非常鲜艳、可扩大 Munsell 颜色书的色域[如图 2(b)],且样本的光谱反射率随波长变化较为剧烈。由此,依据参考文献<sup>[23]</sup>,Munsell 构成的光谱数据集具有线性特征, Munsell+Pantone 构成光谱数据集具有非线性特征。数据集中,光谱反射率采集仪器为美国 X-Rite Ci64 型号的分光光度计,测量范围为 400~700 nm、采样步长为 10 nm。与此同

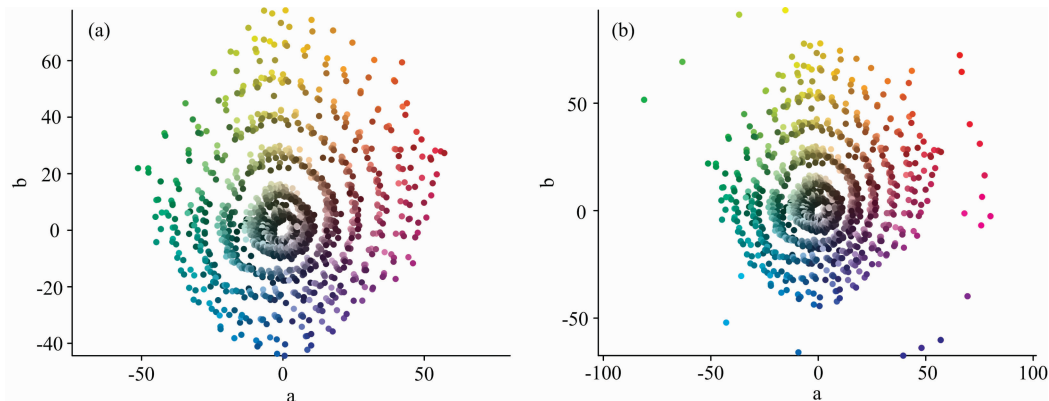


图 2 训练光谱字典的光谱数据集在颜色空间的分布

(a): Munsell; (b): Pantone+Munsell

Fig. 2 The distribution of two data sets in color space from the spectral dataset for training the spectral dictionary

(a): Munsell; (b): Pantone+Munsell

时, 还将其与现有线性方法 PCA<sup>[15]</sup> 和非线性方法 NLP-CA<sup>[23]</sup> 进行对比, 并采用重建结果均方根差 (root mean square error, RMSE) 的平均值和标准偏差值验证方法的准确性和稳定性。

表 1 中的数据为 PCA, NLP-CA 和本文改进 NLP-CA 所得光谱字典重建光谱反射率与原测量光谱反射率之间均方根差的平均值和标准偏差值。其中, RMSE 平均值表征光谱字典重建光谱反射率的精度, RMSE 标准偏差值表征光谱字典重建光谱反射率的稳定性。平均值越小, 重建精度越高; 标准偏差值越小, 重建地稳定性越好。由此, 从 RMSE 平均值看, 本文改进 NLP-CA 方法所得光谱字典重建的光谱反射率精度最高; 从 RMSE 标准偏差值看, 本文改进 NLP-CA 方法所得光谱字典重建的光谱反射率稳定性也好于 PCA 方法和 NLP-CA 方法。此外, 表 1 给出了 Munsell 和 Munsell+Pantone 两个数据集的重建结果, 且 Munsell 数据集是线性的、Munsell+Pantone 数据集是非线性的。从这两个数据集的重建结果看, 本文改进 NLP-CA 方法所得光谱字典的重建效果仍旧优于其他两种方法。由此, 说明本文改进 NLP-CA 方法对线性和非线性数据集均具有优势。

表 1 不同训练集的重建光谱反射率 RMSE 结果

Table 1 RMSE results of reconstructed spectral reflectance for different train sets

方法	Munsell		Munsell+Pantone	
	平均值	标准偏差	平均值	标准偏差
PCA	0.036 4	0.021 1	0.037 3	0.021 8
NLP-CA	0.017 4	0.010 7	0.016 9	0.010 0
改进 NLP-CA	0.015 6	0.009 5	0.014 5	0.009 5

表 2 不同光源下 RGB 图像重建光谱反射率图像的 RMSE 结果

Table 2 RMSE results of the reconstructed spectral images from their RGB images under different illuminants

光源	重建方法	Munsell+Pantone(训练)		Munsell(训练)	
		(平均值±标准偏差值)		(平均值±标准偏差值)	
		CAVE	UEA	CAVE	UEA
A	SR_SR <sup>[14]</sup>	0.254 9±0.053 0	0.354 3±0.100 0	0.304 9±0.052 0	0.343 7±0.073 6
	NLP-CA_SR <sup>[23]</sup>	0.411 6±0.253 1	0.422 0±0.231 2	0.378 1±0.131 9	0.385 1±0.097 2
	本文方法	0.202 2±0.067 0	0.253 4±0.087 8	0.207 6±0.061 9	0.271 1±0.066 4
D65	SR_SR <sup>[14]</sup>	0.254 9±0.053 0	0.354 3±0.100 0	0.304 9±0.052 0	0.343 7±0.073 6
	NLP-CA_SR <sup>[23]</sup>	0.411 5±0.103 0	0.414 6±0.114 7	0.368 9±0.095 6	0.375 8±0.092 8
	本文方法	0.202 2±0.069 7	0.237 4±0.087 6	0.217 6±0.060 7	0.293 6±0.088 8
F2	SR_SR <sup>[14]</sup>	0.254 9±0.053 0	0.354 3±0.100 0	0.304 9±0.052 0	0.343 7±0.073 6
	NLP-CA_SR <sup>[23]</sup>	0.235 6±0.076 6	0.281 3±0.104 8	0.440 4±0.236 7	0.360 7±0.168 1
	本文方法	0.232 7±0.068 8	0.275 5±0.078 8	0.263 1±0.077 9	0.320 1±0.105 9
Avg.	SR_SR <sup>[14]</sup>	0.254 9±0.053 0	0.354 3±0.100 0	0.304 9±0.052 0	0.343 7±0.073 6
	NLP-CA_SR <sup>[23]</sup>	0.352 9±0.144 2	0.372 6±0.150 2	0.395 8±0.154 7	0.373 9±0.119 4
	本文方法	0.212 4±0.068 5	0.255 4±0.084 7	0.229 4±0.066 8	0.294 9±0.087 0

表 2 中, 第一列与第三列(或第二列与第四列)的训练数据集不同、测试数据集相同。从一、三列(或二、四列)对比看, 本文方法重建的 RMSE 平均值和标准偏差值均明显低于 NLP-CA\_SR 方法, 且未随训练数据集的变化而出现明显变

## 4.2 光谱重建评价

为了进一步评价所得光谱字典以及验证本文基于光谱字典的光谱重建方法, 本文在获得上述光谱字典的基础上, 从 CAVE<sup>[32]</sup> 和 UEA<sup>[33]</sup> 两个图像库中选择图像作为重建的测试数据, 并将其重建结果与基于 NLP-CA 重建方法(NLP-CA\_SR)<sup>[23]</sup> 以及基于稀疏表示的重建方法(SR\_SR)<sup>[14]</sup> 进行比较。CAVE<sup>[32]</sup> 和 UEA<sup>[33]</sup> 图像库中分别包含 32 和 23 幅图, 每幅图的光谱图像的波长范围为 400~700 nm、间隔 10 nm。为了对比 NLP-CA\_SR 重建方法与本文重建方法对光源的鲁棒性能, 其测试 RGB 图像分别由其光谱图像在光源 A、D65 和 F2 下通过式(1)所得。其中, RGB 三个光谱响应曲线为 CIE1931 2° 视场的标准观察者。

表 2 中, 训练数据集分别为上述的 Munsell 数据集和 Munsell+Pantone 数据集, 光谱重建的测试数据集分别采用了 CAVE 和 UEA 图像库中的图像; 比较方法选取了 NLP-CA\_SR 方法<sup>[23]</sup> 和 SR\_SR 方法<sup>[14]</sup>。其中, NLP-CA\_SR 方法为利用 NLP-CA<sup>[23]</sup> 获得低维光谱字典后再利用本文光谱重建步骤所得, SR\_SR 方法<sup>[14]</sup> 为一个优秀的单幅 RGB 图像重建光谱图像间接方法。表 2 中数据为 RGB 图像的重建光谱反射率与原光谱图像的光谱反射率之间 RMSE 的平均值和标准偏差值, 且每格数据中: ± 号前数据表示平均值, ± 号后的数据表示标准偏差值。从整体看, 本文方法重建的 RMSE 平均值均最小, 表明本文方法具有较好光谱重建精度; RMSE 标准偏差值虽不全为最小值(有些略高于 SR\_SR 方法的结果), 但仍明显低于同类 NLP-CA\_SR 方法的结果, 此表明本文方法在稳定性上与 SR\_SR 方法不分伯仲、较之 NLP-CA\_SR 方法具有一定优势。

化。本文方法与 NLP-CA\_SR 方法均属于直接类重建方法。如上文所述, 直接类重建方法的重建精度依赖于光谱字典表达测试数据集中光谱数据的准确性, 而此准确性又依赖于光谱字典学习方法。由此, 对比结果表明: 本文改进 NLP-CA



方法更优，获得的光谱字典表达训练数据集的准确率更高、稳定性更好、更有益于获得高精度重建光谱图像。

依据方程病态问题反演定义可知，病态状况越严重时，即使数据变化(如光源变化导致 RGB 数值变化)很小，反演结果(如重建的光谱反射率)差异却很大。表 2 中，不同行表示不同光源(如 A, D65 和 F2 光源)下 RGB 图像重建光谱图像的结果。理论上，不同光源下重建光谱图像差异性越大，重建方法的病态问题则越严重。从行数据对比看，较之 NLPCA\_SR 方法，本文方法重建光谱图像的 RMSE 平均值和标准偏差值受 RGB 值变化(即光源变化)的影响均较小，此现象验证了本文方法也可很好地缓解重建方程反演的病态问题；SR\_SR 方法基本不受光源变化的影响，验证了 SR\_SR 方法在解决病态问题上有明显地优越性。但是，在不同训练数据集与测试数据集组合下，SR\_SR 方法重建效果差异也同时验证了此方法对训练数据集与测试数据集相似性要求的局限性。

图 3 和图 4 显示了 A、D65 和 F2 光源下 RGB 测试图像以及三种方法重建光谱图像与原获取光谱图像之间的光谱差

图。其中，测试图像来源于 CAVE 图像库，图 3 的训练数据集为 Munsell，图 4 的训练数据集为 Munsell+Pantone。在光谱差图中，越接近深蓝色表示光谱差值越小、越接近白色表示光谱差值越大。在 SR\_SR 方法所得光谱差图中，图中整个区域基本呈现蓝色，且此现象基本未随着光源变化而变化，表明该方法重建的光谱图像精度较高，且缓解病态问题效果较好，因而重建光谱图像不随测试图像光源的变化而变化；在 NLPCA\_SR 方法所得光谱差图中，图中大部分区域呈浅蓝色、小部区域呈现黄色，且此现象随光源变化较大，表明 NLPCA\_SR 方法重建的精度和稳定性均略差，且缓解病态问题效果也较差，因而重建光谱图像随测试图像光源变化较大；在本文方法所得光谱差图中，图中整个区域基本呈深蓝色、且该现象未随光源变化而变化，表明本文方法重建光谱图像的精度与稳定性均较优，且缓解病态问题效果较好，因而对测试图像的光源呈现了较好的鲁棒性。此外，由图 3 与图 4 对比可见，本文方法在两组不同训练集下均可重建较好效果的光谱图像。

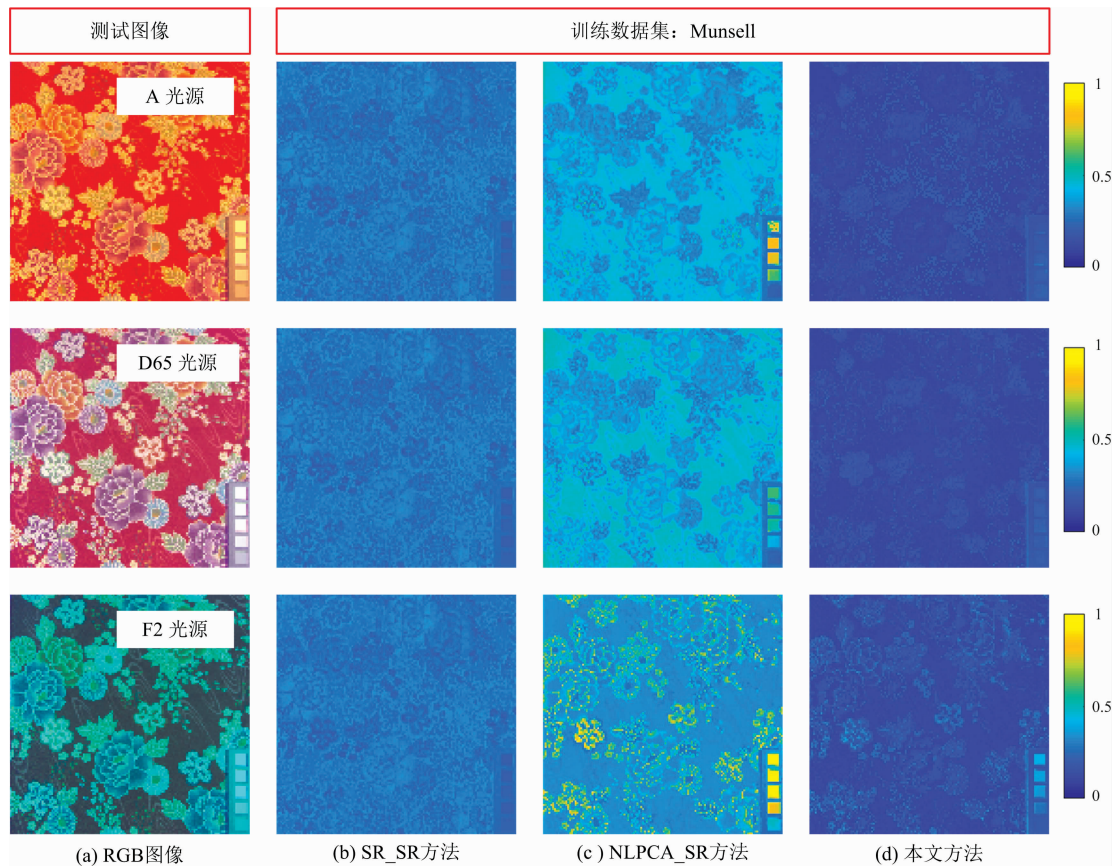


图 3 不同光源下的 RGB 图像与重建光谱图像的光谱差图，重建光谱图像的训练数据集为 Munsell

Fig. 3 RGB images under different illuminats and residual error map between their reconstructed spectral images and their original captured spectral images, the training dataset of the reconstructed spectral images is Munsell

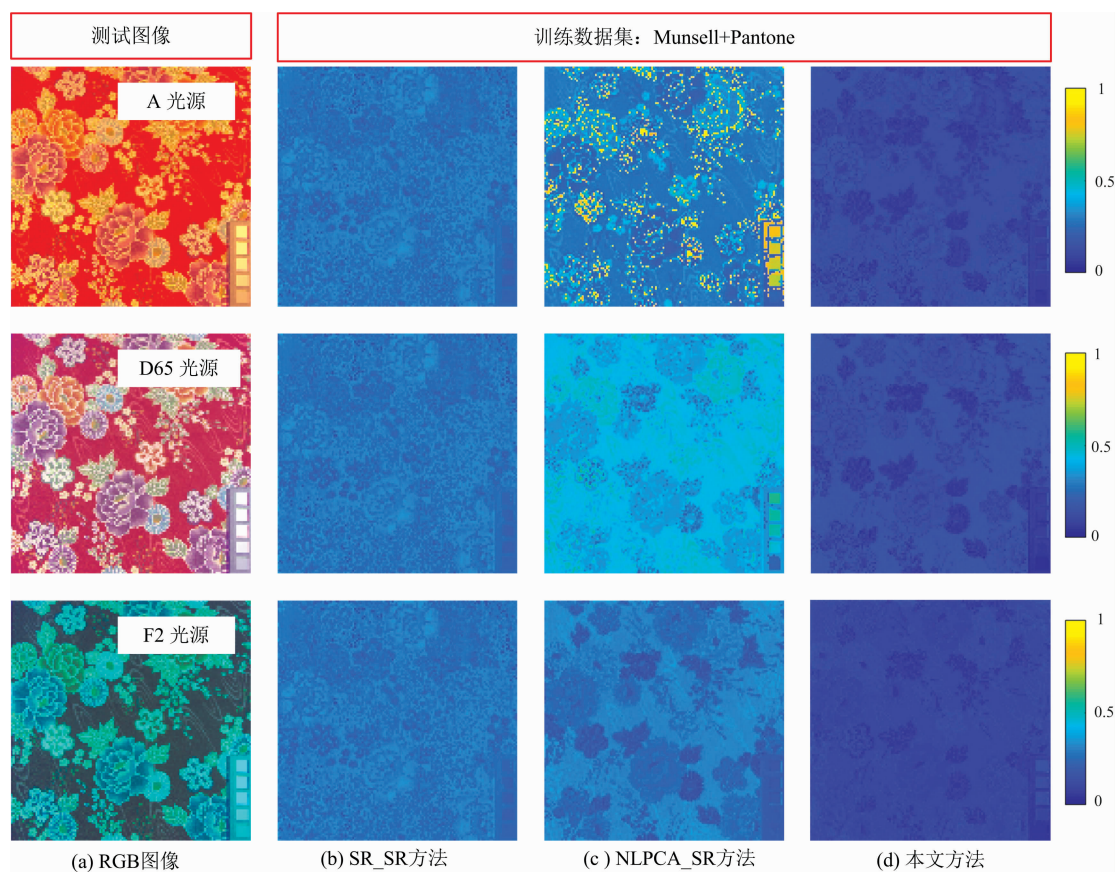


图 4 不同光源下的 RGB 图像与重建光谱图像的光谱差图, 重建光谱图像的训练数据集为 Munsell+Pantone

Fig. 4 RGB images under different illuminants and residual error map between their reconstructed spectral images and their original captured spectral images, the training dataset of the reconstructed spectral images is Munsell+Pantone

## 5 结 论

提出了一种基于非线性光谱字典学习的非线性重建方法实现单幅 RGB 图像重建光谱图像。在此方法中, 首先利用了改进的 NLPCA 方法从训练数据集中学习出一个光谱重建精度高、稳定性好的低维光谱字典, 并用此光谱字典缓解 RGB 图像重建光谱图像方程反演的病态问题, 然后又利用了阻尼

高斯牛顿法结合截断奇异值分解的正则化方法, 进一步缓解基于光谱字典重建方程反演的病态问题, 最终获得重建精度和稳定性较好的光谱图像。通过不同训练光谱数据集以及不同测试图像库, 并通过重建光谱图像与原光谱图像之间 RMSE 的平均值和标准偏差值验证了: 在重建精度和稳定性方面, 本文方法较之现有两种典型方法, 具有一定的优势; 通过不同光源下 RGB 图像重建光谱图像效果还验证了: 本文方法对 RGB 图像的光源具有较好鲁棒性能。

## References

- [ 1 ] Khan M J, Khan H S, Yousaf A, et al. IEEE Access, 2018, 6: 14118.
- [ 2 ] Kuhns T, Messinger D W. Spectral Phenomenology of Historical Parchments and Inks to Aid Cultural Heritage Imaging System Development. VelezReyes M, Messinger D W. Proceedings of SPIE, 2018.
- [ 3 ] Del Pozo S, Rodriguez-Gonzalvez P, Sanchez-Aparicio L J, et al. Multispectral Imaging in Cultural Heritage Conservation. Hayes J, Quiimet C, Quintero M S, et al. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Göttingen: Copernicus Gesellschaft MBH, 2017. 155.
- [ 4 ] He Y, Jiao W, Shi Y, et al. IEEE Journal of Biomedical and Health Informatics, 2020, 24(2): 493.
- [ 5 ] Brossard M, Marion R, Carrere V. Remote Sensing Letters, 2016, 7(6): 581.
- [ 6 ] Zazi L, Boutaleb A, Guettouche M S. Arabian Journal of Geosciences, 2017, 10(11): 252.
- [ 7 ] Thomas J, Lapray P, Gouton P, et al. Sensors, 2016, 16(7): 993.
- [ 8 ] Ji X, Miao C, Zhang Y, et al. Optics Communications, 2017, 404(SI): 11.
- [ 9 ] Wu R, Li Y, Xie X, et al. Sensors, 2019, 19(13): 2905.

- [10] Neumaier A. *SIAM Rev.*, 1998, 40: 636.
- [11] Chou T, Hsieh C, Chen E. *Color Research and Application*, 2019, 44(4): 508.
- [12] Sibson R. *A Brief Description of Natural Neighbor Interpolation*, John Wiley & Sons, 1981.
- [13] Cao B, Liao N, Li Y, et al. *Optical Engineering*, 2017, 56(5): 053104.
- [14] Arad B, Ben-Shahar O. *Sparse Recovery of Hyperspectral Signal from Natural RGB Images*. Leibe B, Matas J, Sebe N, et al. *Lecture Notes in Computer Science*, 2016, 9911: 19.
- [15] Oh S W, Brown M S, Pollefeys M, et al. *Do It Yourself Hyperspectral Imaging with Everyday Digital Cameras*. *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2016: 2461.
- [16] Cao Q, Wan X, Li J, et al. *Optical Review*, 2016, 23(5): 753.
- [17] Wu D, Tian J, Tang Y. *Optical Review*, 2014, 21(2): 117.
- [18] Flinkman M, Laamanen H, Tuomela J, et al. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 2013, 30(9): 1806.
- [19] Chou Y, Cheung V, Luo M R, et al. *Spectral Representation of Object Colours*. *Color and Imaging Conference Final Program and Proceedings*, 2012: 57.
- [20] Rayat A, Amirshahi S H, Agahian F. *Color Research and Application*, 2014, 39(2): 136.
- [21] Usui S, Nakauchi S, Nakano M. *J. Opt. Soc. Am. A*, 1992, 9(4): 516.
- [22] LI Chan, WAN Xiao-xia, LIU Qiang, et al(李 婵, 万晓霞, 刘 强, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2016, 36(5): 1400.
- [23] Barakzehi M, Amirshahi S H, Peyvandi S, et al. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 2013, 30(9): 1862.
- [24] Licciardi G, Chanussot J. *European Journal of Remote Sensing*, 2018, 51(1): 375.
- [25] Snee R D, Marquardt D W. *The American Statistician*, 1984, 38(2): 83.
- [26] Hsieh W W. *Tellus Series A-Dynamic Meteorology and Oceanography*, 2001, 53(5): 599.
- [27] Björck Å. *Numerical Methods for Least Squares Problems*, 1996. 339
- [28] Hansen P C. *BIT Numerical Mathematics*, 1987, 27(4): 534.
- [29] Deidda G P, Fenu C, Rodriguez G. *Inverse Problems*, 2014, 30(12): 125014.
- [30] Armijo L. *Pacific Journal of Mathematics*, 1966, 16(1): 1.
- [31] <https://www.pantone.com/nylon-brights-set>.
- [32] Yasuma F, Mitsunaga T, Iso D, et al. *IEEE Transactions on Image Processing*, 2010, 19(9): 2241.
- [33] Finlayson G D, Hordley S D, Morovic P. *CGIV 2004; 2nd European Conference on Color in Graphics, Imaging, and Vision*, 2004. 268.

## Research on Spectral Image Reconstruction Based on Nonlinear Spectral Dictionary Learning From Single RGB Image

ZUO Chu<sup>1</sup>, XIE De-hong<sup>2\*</sup>, WAN Xiao-xia<sup>3</sup>

1. School of Light Industry and Food, Nanjing Forestry University, Nanjing 210037, China

2. College of Information Science and Technology, Nanjing Forestry University, Nanjing 210037, China

3. Hubei Province Engineering Technical Center for Digitization and Virtual Roproduction of Color Information of Culture Relics, Wuhan University, Wuhan 430079, China

**Abstract** A nonlinear reconstruction method based on nonlinear spectral dictionary learning was proposed to solve the ill-posed problem of spectral image reconstruction from a single RGB image. In order to adapt to the linear and nonlinear data, the method firstly improves the nonlinear principal component analysis algorithm based on a modified self-association neural network model. It uses to learn the low-dimensional spectral dictionary from the training spectrum set, which is used in the inverse equation of spectral reconstruction to alleviate the ill condition. In addition, based on the spectral dictionary, the damped Gaussian Newton method combined with the truncated singular value decomposition regularization method is used further to alleviate the ill-posed problem of the nonlinear inversion, and the spectral image can be reconstructed from a single RGB image. In the experiment, two different spectral training sets, Munsell and Munsell+Pantone, were used to learn the spectral dictionary. Meanwhile, CAVE and UEA spectral image libraries were used for the spectral reconstruction tests. Compared with the existing methods, it is found that the average root means square error of CAVE and UEA spectral images reconstructed by this method under different spectral training sets were the lowest, which were 0.212 4, 0.255 4, 0.229 4 and 0.294 9 respectively. The standard



deviations of root mean square error was close to the effect of the best method, which was 0.068 5, 0.084 7, 0.066 8 and 0.087 0 respectively. The results show that the method for reconstructing the spectral image from a single RGB image has advantages in accuracy and stability.

**Keywords** Spectral reconstruction; RGB image; Nonlinear; Spectral dictionary; Learning

(Received Apr. 27, 2021; accepted Nov. 11, 2021)

\* Corresponding author

---

## 《光谱学与光谱分析》期刊社决定采用 ScholarOne Manuscripts 在线投稿审稿系统

《光谱学与光谱分析》期刊社与汤森路透集团签约,自 2010 年 12 月 1 日起《光谱学与光谱分析》决定采用 Thomson Reuters 旗下的 ScholarOne Manuscripts 在线投稿审稿系统。

- ScholarOne Manuscripts, 该系统不仅能轻松处理稿件,而且能提速科技交流。
- 全球已有 360 多家学会和出版社的 3 800 多种期刊选用了 ScholarOne Manuscripts 系统作为在线投稿、审稿平台,全球拥有超过 1 350 万的注册用户,代表着全球学术期刊在线投审稿的一流水平。
- ScholarOne Manuscripts 与 EndNote, Web of Science 无缝链接和整合;使科研探索、论文评阅和信息传播效率大为提高。
- ScholarOne Manuscripts 是汤森路透科技集团的一个业务部门,拥有丰富的学术期刊业务经验,为学术期刊提供综合管理工作流程系统,使期刊更有效管理投稿、同行评审、加工和发表过程,提高作者心中的专业形象,缩短论文发表时间,削减管理成本,帮助期刊提高科研绩效和实现学术创新。

《光谱学与光谱分析》采用“全球学术期刊首选的在线投稿审稿系统—ScholarOne Manuscripts”,势必对 2010 年 11 月 30 日以前向本刊投稿的作者在查阅稿件信息时,会带来某些不便,在此深表歉意!为了推进本刊的网络化、数字化、国际化进程,以实现与国际先进出版系统对接;为了不断提高期刊质量,加快网络化、数字化建设,加快与国际接轨的进程,希望能得到广大作者、读者们的支持与理解,对您的理解和配合深表感激。这是一件新事物,肯定有不周全、不完善的地方,让我们共同努力,不断改进和完善起来。

《光谱学与光谱分析》期刊社

2010 年 12 月 1 日