基于分离字典的图像超分辨率重建

张凤珍^{1,2,3}, 岑翼刚^{2,3*}, 赵瑞珍^{2,3}, 王艳红^{2,3}, 张琳娜⁴, 胡绍海^{2,3}

1. 中国科学院国家天文台太阳活动重点实验室, 北京 100101

2. 北京交通大学信息科学研究所, 北京 100044

3. 现代信息科学与网络技术北京市重点实验室, 北京 100044

4. 贵州大学机械工程学院, 贵阳 550025

* 通信作者. E-mail: ygcn@bjtu.edu.cn

收稿日期: 2018-09-14; 接受日期: 2019-01-17; 网络出版日期: 2020-02-12

国家自然科学基金 (批准号: 61872034, 61572067, 61572063, 61572461, 11790305)、贵州省自然科学基金 (批准号: [2019]1064) 和中央高校基本科研业务费 (批准号: 2017JBZ108) 资助项目

摘要 传统的基于稀疏表示的图像超分辨率重建算法, 需要将图像进行分块并列为向量, 这样就破坏了图像块内邻域像素间的相关性. 为了更好地利用图像邻域内的结构信息, 本文结合分离字典能从不同方向对图像块进行稀疏表示的特性, 提出了基于分离字典的图像超分辨率重建算法. 实验结果表明, 与传统基于稀疏表示的图像超分辨率重建算法相比, 本文算法不仅提高了图像重建的速度, 而且在 PSNR 和 SSIM 两个衡量指标上都优于传统基于稀疏表示的超分辨率重建算法 (PSNR 提高约 0.2 dB, SSIM 提高约 0.01).

关键词 图像超分辨率重建, 2D 稀疏编码, 黎曼流形, 稀疏表示, 分离字典

1 引言

随着计算机显示技术的不断发展, 生活中人们对图像的分辨率要求越来越高, 但由于受到硬件设备等因素的限制, 得到的图像分辨率往往不能满足人们的实际需要. 因此通过相应的算法或软件技术实现图像的超分辨率重建技术就受到了广泛的关注. 目前, 图像超分辨率重建技术的应用领域主要有医学图像重建、视频监控和人脸识别等.

图像的超分辨率重建 SR^[1~3] (image super-resolution reconstruction) 技术是指由一帧或多帧低分辨率 LR 图像 (low resolution image) 经过图像处理技术来重建一帧高分辨率 HR 图像 (high resolution image) 的技术.

常见的图像超分辨率重建算法可以分为 3 类^[4,5]: 基于插值的算法、基于重建的算法和基于学习的算法. 基于插值的图像超分辨率重建算法因简单快速而受到了广泛的关注, 但该方法对纹理复杂的

引用格式: 张凤珍, 岑翼刚, 赵瑞珍, 等. 基于分离字典的图像超分辨率重建. 中国科学: 信息科学, 2020, 50: 275–288, doi: 10.1360/N112018-00192
Zhang F Z, Cen Y G, Zhao R Z, et al. The algorithm of image super-resolution reconstruction via separable dictionaries (in Chinese). Sci Sin Inform, 2020, 50: 275–288, doi: 10.1360/N112018-00192

图像重建效果并不理想. 基于重建的图像超分辨率重建算法由于受到计算复杂度和存储空间的限制, 使其应用也受到了一定的影响. 近年来, 随着稀疏表示理论的深入研究, 使得基于学习的图像超分辨率重建算法受到了越来越多的研究. Yang 等^[6,7]和 Zedye 等^[8]学者提出将稀疏表示应用于图像超分辨率重建过程, 该类算法利用图像特征信息在不同分辨率字典下具有相同稀疏表示的原理, 来重建高分辨率图像. 基于稀疏表示的图像超分辨率重建算法不仅提高了重建速度, 同时也提高了重建质量. 但是, 传统的基于稀疏表示的图像超分辨率重建算法在对图像进行分块后, 还需要将所得的图像块列化为向量的形式, 这样不仅增加了训练与重建过程中的计算复杂度, 同时也破坏了图像块内的结构信息. 为了充分利用图像邻域内的结构信息, 本文提出一种新的基于分离字典的图像超分辨率重建算法.

2 稀疏表示理论及其在图像超分辨率重建中的应用

2.1 稀疏表示理论基础

稀疏表示, 其目的是在给定的过完备字典中用尽可能少的原子来表示信号, 从而简化信号表达形式. 设信号为 $\mathbf{y} \in \mathbb{R}^m$, 过完备字典为 $\mathbf{D} \in \mathbb{R}^{m \times K}$ ($K \geq m$), 稀疏表示的优化问题可以表示如下:

$$\begin{aligned} \arg \min_{\mathbf{D}, \mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t. } \mathbf{y} = \mathbf{D}\mathbf{x}, \\ \text{or } \arg \min_{\mathbf{D}, \mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t. } \|\mathbf{y} - \mathbf{D}\mathbf{x}\| \leq \varepsilon, \end{aligned} \quad (1)$$

其中 $\varepsilon \in \mathbb{R}^+$ 是正极小值, 用来约束重建误差. $\|\cdot\|_0$ 表示 ℓ_0 范数, 用来统计 $\cdot$ 中非零元素的个数. 将 ℓ_0 范数用 ℓ_1 范数代替, 将问题 (1) 转化为无约束的最优化目标函数, 即

$$\arg \min_{\mathbf{D}, \mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (2)$$

λ 是一个权重系数, 用来平衡重建误差和 ℓ_1 范数之间的关系.

稀疏表示的核心在于字典的选择, 因为字典的性能决定了是否能对自然图像/自然信号形成简洁的表达形式^[9]. 经典的字典学习算法主要有 MOD (method of optimal directions)^[10,11], 和 K-SVD (K-singular value decomposition)^[12], GOAL (geometric analysis operator learning)^[13] 等. 因为经过 K-SVD 算法学习所得的字典在信号表示方面具有较强的适应性, 所以该字典学习算法被广泛地应用于图像处理中, 如图像超分辨率重建^[6~8]过程. K-SVD 是一种交替迭代算法, 其迭代过程包括: 稀疏表示和字典优化, 其字典学习过程如图 1 所示.

2.2 基于稀疏表示的图像超分辨率重建

基于稀疏表示的图像超分辨率重建算法, 其原理是从高低分辨率图像块/样本提取特征, 在对应的过完备字典下具有相同的稀疏表示, 整个过程分为: 离线高低分辨率字典的学习和图像超分辨率重建过程. 基于稀疏表示的图像超分辨率重建如图 2 所示. 设 $\mathbf{X}^h = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ 是从高分辨率图像块所取得的特征, $\mathbf{x}_i \in \mathbb{R}^n$ 是图像块特征经过拉伸后生成的向量, $\mathbf{Y}^l = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ 是从与 $\mathbf{X}^h$ 对应的低分辨率图像块提取的特征, 其中 $\mathbf{y}_i \in \mathbb{R}^m$ ($m \leq n$), 高低分辨率字典的学习可以表示为^[6]

$$\begin{aligned} \mathbf{D}_h = \arg \min_{\mathbf{D}_h, \mathbf{Z}} \|\mathbf{X}^h - \mathbf{D}_h \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad \text{s.t. } \|\mathbf{D}_h(:, i)\|_2^2 \leq 1, \\ \mathbf{D}_l = \arg \min_{\mathbf{D}_l, \mathbf{Z}} \|\mathbf{Y}^l - \mathbf{D}_l \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad \text{s.t. } \|\mathbf{D}_l(:, i)\|_2^2 \leq 1, \end{aligned} \quad (3)$$

图 1 K-SVD 算法流程图

Figure 1 The flow chart of K-SVD algorithm

其中, $D_h(:, i)$ 表示字典 D_h 中的 i -th 列. 将式 (3) 中的两个问题进行联合, 重新将该问题表述为

$$\arg \min_{D_h, D_l, Z} \frac{1}{n} \|X^h - D_h Z\|_2^2 + \frac{1}{m} \|Y^l - D_l Z\|_2^2 + \lambda \left(\frac{1}{n} + \frac{1}{m} \right) \|Z\|_1. \quad (4)$$

设

$$X_c = \begin{bmatrix} \left(\frac{1}{\sqrt{n}} \right) X^h \\ \left(\frac{1}{\sqrt{m}} \right) Y^l \end{bmatrix}, \quad D_c = \begin{bmatrix} \left(\frac{1}{\sqrt{n}} \right) D_h \\ \left(\frac{1}{\sqrt{m}} \right) D_l \end{bmatrix},$$

$\hat{\lambda} = \lambda \left(\frac{1}{n} + \frac{1}{m} \right)$, 则问题 (3) 等价于

$$\arg \min_{D_c, Z} \|X_c - D_c Z\|_2^2 + \hat{\lambda} \|Z\|_1. \quad (5)$$

在上述问题中, 高低分辨率字典 D_h 和 D_l 都是经过学习得到的. 在学习过程中, 需要不断强调高低分辨率特征在对应的字典下具有相同的稀疏表示. 所以问题 (5) 通过将高低分辨率特征和字典级联来保证二者稀疏表示系数的一致性. Zedye 等^[8] 提出分别优化高低分辨率字典, 他们使用 K-SVD 算法先得到低分辨率字典 D_l , 然后将高分辨率字典 D_h 学习的目标函数转化为

$$D_h = \arg \min_{D_h} \|X^h - D_h Z\|_2^2. \quad (6)$$

经过分析发现, 问题 (6) 是关于 D_h 的二次优化问题, 采用最小二乘法对该问题进行求解, 其解如下:

$$D_h = X^h Z^\dagger = X^h Z^T (Z Z^T)^{-1}, \quad (7)$$

其中, $Z^\dagger$ 表示 Z 的伪逆, 且 $Z^\dagger = Z^T (Z Z^T)^{-1}$.

如图 2 所示, 在完成字典学习过程得到 D_h 和 D_l 之后, 对于输入的低分辨率图像块先在低分辨率字典 D_l 下进行稀疏表示, 然后通过得到的稀疏系数和高分辨率字典 D_h 重建对应的高分辨率图像块.

图 2 (网络版彩图) 基于稀疏表示的图像超分辨率重建

Figure 2 (Color online) The image super-resolution reconstruction based on sparse representation

3 分离字典学习及其在图像超分辨率重建中的应用

虽然上述基于稀疏表示的图像超分辨率重建算法无论是在运算速度, 还是在重建质量上都有一定的提升, 但它们都有一个共同的缺点: 需要将图像块重置为向量的形式. 对图像块的矢量化操作, 不仅增加了学习和重建过程的计算量, 同时也破坏了图像块邻域内的结构信息. 所以我们利用分离字典可以捕捉图像块在不同方向上的结构特征, 提出基于分离字典的图像超分辨率重建算法.

3.1 分离字典

定义1 设矩阵 $A \in \mathbb{R}^{m \times \sqrt{K}}$ 和 $B \in \mathbb{R}^{m \times \sqrt{K}}$, 二者的 Kronecker 内积 $A \otimes B$ 则是大小为 $m^2 \times K$ 的矩阵, 其表示如下:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1\sqrt{K}}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{m\sqrt{K}}B \end{bmatrix}. \quad (8)$$

字典的分离性是指一个尺寸较大的字典可以通过两个尺寸较小的字典的 Kronecker 内积得到, 例如 $D = B \otimes A$, 而且对于信号 $y \in \mathbb{R}^{m^2}$ 有如下的关系:

$$y = D\alpha = (B \otimes A)\alpha = \text{vec}(\text{Avec}^{-1}(\alpha)B^T), \quad (9)$$

其中, $\text{vec}(\cdot)$ 表示将矩阵按列排列成向量, $\text{vec}^{-1}(\cdot)$ 为 $\text{vec}(\cdot)$ 的逆操作. 我们称式 (9) 中的矩阵 A 和 B 为分离字典. 与传统稀疏表示算法相比, 通过这样的结构分离, 不仅能从不同方向上捕捉信号的结构信息, 而且还降低了算法的计算复杂度和存储空间. 如式 (9) 中字典 D 与稀疏系数 α 相乘的算法复杂度是 $O(m^2K)$, 需要存储的字典 D 中的参数个数为 m^2K , 而 3 个矩阵 A , B 和 $\text{vec}^{-1}(\alpha)$ 相乘的计算复杂度是 $O(mK) + O(m^2\sqrt{K})$, 需要存储的参数也减少为 $2 \times (m\sqrt{K})$.

3.2 分离字典在图像超分辨率重建中的应用

已有的基于稀疏表示的图像超分辨率重建算法虽然在重建质量和重建速度上已经有一定的提高, 但是因为在稀疏表示时需要将图像进行分块并列化为向量的形式, 会破坏图像中邻域内的相关性和结构信息, 同时因为将图像块列化为向量的操作也会增加在稀疏表示过程中的计算复杂度. 所以为了克

服这个不足,我们将分离字典应用于图像超分辨率重建过程中,提出基于分离字典的图像超分辨率重建算法,这样就可以利用分离字典可以从不同方向上提取图像块特征的特点,达到提高重建质量的目的. 设高低分辨率训练样本块 $\{\mathbf{X}_j^h\}$ 和 $\{\mathbf{Y}_j^l\}$, 且 $\mathbf{X}_j^h \in \mathbb{R}^{n \times n}$ 和 $\mathbf{Y}_j^l \in \mathbb{R}^{n \times n}$, $\forall j = 1, 2, \dots, N$, 高低分辨率分离字典大小为 $n \times K$ ($K \geq n$), 所以基于分离字典的图像超分辨率重建字典学习的目标函数可以表示为

$$\begin{aligned} \arg \min_{\{\mathbf{A}, \mathbf{B}\}, \{\mathbf{C}, \mathbf{F}\}, \{\mathbf{Z}_j\}_{j=1}^N} & \frac{1}{2N} \sum_{j=1}^N \|\mathbf{Y}_j^l - \mathbf{A} \mathbf{Z}_j \mathbf{B}^T\|_F^2 + \frac{1}{2N} \sum_{j=1}^N \|\mathbf{X}_j^h - \mathbf{C} \mathbf{Z}_j \mathbf{F}^T\|_F^2 + \lambda \sum_{j=1}^N g(\mathbf{Z}_j) \\ \text{s.t. } & \text{ddiag}(\mathbf{A}^T \cdot \mathbf{A}) = \mathbf{I}_K, \text{ddiag}(\mathbf{B}^T \cdot \mathbf{B}) = \mathbf{I}_K, \\ & \text{ddiag}(\mathbf{C}^T \cdot \mathbf{C}) = \mathbf{I}_K, \text{ddiag}(\mathbf{F}^T \cdot \mathbf{F}) = \mathbf{I}_K, \end{aligned} \quad (10)$$

其中, $\|\cdot\|_F$ 是矩阵的 Frobenius 范数: $\|\mathbf{A}\|_F^2 = \sum \sum a_{i,j}^2$, $g(\mathbf{Z}_j)$ 是矩阵 $\mathbf{Z}_j$ 的稀疏度量函数, 可以使用 $\log$ 、幂指数函数等, $\text{ddiag}(\cdot)$ 表示由 $\cdot$ 的主对角线元素组成的对角矩阵, 如 $\mathbf{D}$ 是正交矩阵, 则 $\text{ddiag}(\mathbf{D}^T \mathbf{D})$ 是单位矩阵. $\mathbf{I}_K$ 是大小为 $K \times K$ 的单位矩阵. $\{\mathbf{Z}_j\}_{j=1}^N$ 是由稀疏系数矩阵组成的矩阵序列, 而且 $\forall \mathbf{Z}_j \in \mathbb{R}^{K \times K}$. $\{\mathbf{A}, \mathbf{B}\}$ 和 $\{\mathbf{C}, \mathbf{F}\}$ 分别是与高低分辨率训练样本对应的分离字典对, N 是训练样本个数.

因为高低分辨率字典之间是相互独立的, 所以将问题 (10) 拆分为低分辨率字典的优化问题 P_1 和高分辨率字典的优化问题 P_2 :

$$\begin{aligned} P_1 = & \arg \min_{\{\mathbf{A}, \mathbf{B}\}, \{\mathbf{Z}_j\}_{j=1}^N} \frac{1}{2N} \sum_{j=1}^N \|\mathbf{Y}_j^l - \mathbf{A} \mathbf{Z}_j \mathbf{B}^T\|_F^2 + \lambda \sum_{j=1}^N g(\mathbf{Z}_j) \\ \text{s.t. } & \text{ddiag}(\mathbf{A}^T * \mathbf{A}) = \mathbf{I}_K, \text{ddiag}(\mathbf{B}^T * \mathbf{B}) = \mathbf{I}_K, \\ P_2 = & \arg \min_{\{\mathbf{C}, \mathbf{F}\}} \frac{1}{2N} \sum_{j=1}^N \|\mathbf{X}_j^h - \mathbf{C} \mathbf{Z}_j \mathbf{F}^T\|_F^2 \\ \text{s.t. } & \text{ddiag}(\mathbf{C}^T * \mathbf{C}) = \mathbf{I}_K, \text{ddiag}(\mathbf{F}^T * \mathbf{F}) = \mathbf{I}_K. \end{aligned} \quad (11)$$

接下来, 我们将通过不同的算法求解问题 P_1 和 P_2 .

3.2.1 低分辨率分离字典的学习

对于问题 P_1 是通过分离字典 $\{\mathbf{A}, \mathbf{B}\}$ 提取低分辨率训练样本 $\{\mathbf{Y}_j^l\}_{j=1}^N$ 中的本质特征, 并求解训练样本在对应分离字典下的稀疏表示系数 $\{\mathbf{Z}_j\}_{j=1}^N$, $g(\mathbf{Z}_j)$ 是稀疏度量函数. Hawe 等^[14] 提出使用 $\log$ 函数拟合 ℓ_0 范数和对矩阵的约束条件, 并将字典投影到黎曼流形 (Riemannian manifold) 上, 采用共轭梯度法对字典和稀疏系数进行优化求解. 其求解过程如图 3 所示.

从图 3 中可以看出, 分离字典和稀疏系数是一起进行更新的, 而且对于未知量 $\{\mathbf{A}, \mathbf{B}\}$ 和 $\{\mathbf{Z}_j\}_{j=1}^N$ 在更新过程中采用相同的移动步长. 但实际上, 如图 4 所示, 分离字典和稀疏系数无论是在矩阵性质、约束条件等方面都存在着很大的差异. 例如字典 $\mathbf{A}$ 要求满足^[13~15]: 字典中原子的 ℓ_2 范数为 1, $\|\mathbf{A}(:, i)\|_2 = 1$; 若 $\mathbf{A} \in \mathbb{R}^{n \times K}$ ($K \geq n$), $\mathbf{A}$ 的秩需要满足 $\text{rank}(\mathbf{A}) = n$; 字典的原子之间是线性不相关的, 要求若 $i \neq j$ 则 $\mathbf{A}(:, i) \neq \mathbf{A}(:, j)$. 而稀疏系数矩阵 $\mathbf{Z}_j$ 只要求矩阵中大部分元素值为 0, 或者元素的绝对值比较小, 如图 4(b) 所示.

考虑到字典和稀疏系数之间的不同, 我们在文献 [15, 16] 中提出将分离字典学习过程分为两个过程: 2D 稀疏编码和基于黎曼流形的分离字典优化. 文献 [15, 16] 中分离字典的优化过程如图 5 所示.

图 3 基于流形的分离字典学习流程^[14]

Figure 3 The flow chart of separable dictionary learning based on oblique manifold^[14]

图 4 (网络版彩图) (a) 字典与 (b) 稀疏系数图示

Figure 4 (Color online) The graphics of (a) dictionary and (b) sparse represent coefficients

图 5 基于流形和 2D 稀疏编码的分离字典学习流程^[15, 16]

Figure 5 The flow chart of separable dictionary learning based on 2D sparse coding and oblique manifold^[15, 16]

在文献 [15] 中, 提出一种基于黎曼流形的分离字典学习算法 ASeDiL (analytic separable dictionary learning), 在该算法中, 用 ℓ_0 范数作为稀疏度量函数, 将字典学习过程在 2D-OMP 稀疏编码算法和分离字典优化两个过程中迭代进行, 其目标函数如下:

$$\arg \min_{\mathbf{A}, \mathbf{B}, \{\mathbf{Z}_j\}_{j=1}^N} \frac{1}{N} \sum_{j=1}^N \|\mathbf{A}\mathbf{Z}_j\mathbf{B}^T - \mathbf{Y}_j^I\|_F^2 + \rho \sum_{j=1}^N \|\mathbf{Z}_j\|_0 + \lambda [r(\mathbf{A}) + r(\mathbf{B})] + \beta [h(\mathbf{A}) + h(\mathbf{B})], \quad (12)$$

其中, $h(\cdot)$ 和 $r(\cdot)$ 分别表示对字典的秩和列不相关性的约束条件, 且

$$\begin{aligned} h(\mathbf{A}) &= -\frac{1}{m \log(m)} \log \det \left(\frac{1}{n} \mathbf{A} \mathbf{A}^T \right), \\ r(\mathbf{A}) &= -\sum_{k < l} \log \left(1 - ((\mathbf{A}(:, k))^T (\mathbf{A}(:, l)))^2 \right), \\ h(\mathbf{B}) &= -\frac{1}{m \log(m)} \log \det \left(\frac{1}{n} \mathbf{B} \mathbf{B}^T \right), \\ r(\mathbf{B}) &= -\sum_{k < l} \log \left(1 - ((\mathbf{B}(:, k))^T (\mathbf{B}(:, l)))^2 \right), \end{aligned}$$

其中, ρ , λ 和 β 分别是稀疏系数, 满秩约束条件和列不相关性约束条件的权重系数. 在 ASeDiL 算法中, 因为使用 ℓ_0 范数作为稀疏度量函数, 所以 2D 稀疏编码过程中使用的是 2D-OMP 算法^[17] (2-dimensional orthogonal matching pursuit) (2D-OMP 算法是 OMP 算法在 2 维空间的拓展, 在相同条件下与 OMP 算法重建结果相同, 但重建速度远胜 OMP 算法) 求解稀疏系数, 然后再将分离字典 $\{\mathbf{A}, \mathbf{B}\}$ 投影到黎曼流形上作为一个整体进行更新. 因为传统的字典学习算法如 K-SVD, 更新过程中针对的是字典中的原子, 并不能保证字典是全局最优解^[18]. 所以在 ASeDiL 算法中, 将字典投影到黎曼流形上作为整体进行更新, 能够提高字典对信号的表示能力.

凸优化算法为信号的稀疏重建提供了一个有力的途径, 因为该方法已经有很多可用的准确的数值求解程序. 因此在文献 [16] 中, 提出对 ℓ_0 范数进行松弛, 用 ℓ_1 范数作为稀疏度量函数, 其目标函数可以重新表述为

$$\arg \min_{\mathbf{A}, \mathbf{B}, \{\mathbf{Z}_j\}_{j=1}^N} \frac{1}{N} \sum_{j=1}^N \|\mathbf{A} \mathbf{Z}_j \mathbf{B}^T - \mathbf{Y}_j\|_F^2 + \rho \sum_{j=1}^N \|\mathbf{Z}_j\|_1 + \lambda [r(\mathbf{A}) + r(\mathbf{B})] + \beta [h(\mathbf{A}) + h(\mathbf{B})]. \quad (13)$$

因此, 在 2D 稀疏编码过程中, 问题 (13) 可以表示关于稀疏系数矩阵 $\{\mathbf{Z}_j\}_{j=1}^N$ 的 Frobenius 范数和 ℓ_1 范数的加权和. 所以在文献 [16] 中提出用分离的快速迭代收缩阈值 SFISTA (separable fast iterative shrinkage-thresholding algorithm) 算法进行 2D 稀疏编码, 然后再将分离字典投影到黎曼流形上进行更新.

3.2.2 高分辨率分离字典的学习

在基于分离字典的图像超分辨率重建优化问题 (11) 中, 通过 3.2.1 小节内容对问题 P_1 进行求解, 获得低分辨率分离字典 $\{\mathbf{A}, \mathbf{B}\}$ 和稀疏系数 $\{\mathbf{Z}_j\}_{j=1}^N$ 之后, 再对问题 P_2 进行求解, 以获得高分辨率分离字典 $\{\mathbf{C}, \mathbf{F}\}$.

对于问题 P_2 中的两个变量 $\{\mathbf{C}, \mathbf{F}\}$, 将其拆分为

$$\arg \min_{\mathbf{C}} \left\| [\mathbf{X}_1^h, \dots, \mathbf{X}_N^h] - \mathbf{C} [\mathbf{Z}_1 \mathbf{F}^T, \dots, \mathbf{Z}_N \mathbf{F}^T] \right\|_F^2, \quad \text{s.t. } \text{ddiag}(\mathbf{C}^T \mathbf{C}) = \mathbf{I}_K, \quad (14)$$

和

$$\arg \min_{\mathbf{F}} \left\| [\mathbf{X}_1^T, \dots, \mathbf{X}_N^T] - \mathbf{F} [\mathbf{Z}_1^T \mathbf{C}^T, \dots, \mathbf{Z}_N^T \mathbf{C}^T] \right\|_F^2, \quad \text{s.t. } \text{ddiag}(\mathbf{F}^T \mathbf{F}) = \mathbf{I}_K. \quad (15)$$

经分析发现, 式 (14) 和 (15) 是同一类问题, 下面内容以式 (14) 为例对这类问题进行求解.

令 $\mathbf{X}_c = [\mathbf{X}_1^h, \mathbf{X}_2^h, \dots, \mathbf{X}_N^h]$, $\mathbf{U} = [\mathbf{Z}_1 \mathbf{F}^T, \dots, \mathbf{Z}_N \mathbf{F}^T]$, 将式 (14) 简化为

$$\arg \min_{\mathbf{C}} \|\mathbf{X}_c - \mathbf{C} \mathbf{U}\|_F^2, \quad \text{s.t. } \|\mathbf{C}(:, i)\|_2 = 1. \quad (16)$$

式 (16) 是含有约束条件的最优化问题, 所以在问题中引入 Lagrange 乘子 ϑ , 将其转化为无约束问题, 其表达形式如下:

$$\begin{aligned} L(\mathbf{C}, \{\vartheta_1, \dots, \vartheta_K\}) &= \|\mathbf{X}_c - \mathbf{C}\mathbf{U}\|_F^2 + \sum_{j=1}^K \vartheta_j \left(\sum_{i=1}^n \mathbf{C}(i, j)^2 - 1 \right) \\ &= \text{tr}((\mathbf{X}_c - \mathbf{C}\mathbf{U})^T (\mathbf{X}_c - \mathbf{C}\mathbf{U})) + \sum_{j=1}^K \vartheta_j \left(\sum_{i=1}^n \mathbf{C}(i, j)^2 - 1 \right), \end{aligned} \quad (17)$$

其中 $\{\vartheta_1, \dots, \vartheta_K\}$ 是引入的 Lagrange 乘子, $\vartheta_i \geq 0$. 令 $\mathbf{J} = \text{diag}(\{\vartheta_1, \dots, \vartheta_K\})$ ($\text{diag}(\cdot)$ 表示由 $\cdot$ 构成的对角矩阵), $\mathbf{J}^T = \mathbf{J}$. 式 (17) 是关于 $\mathbf{C}$ 和 Lagrange 乘子 $\{\vartheta_1, \dots, \vartheta_K\}$ 的最优化问题, 因为这个问题比较难解, 所以引入 Lagrange 对偶理论, 将式 (17) 转化为对偶问题

$$\begin{aligned} \text{dual}(\mathbf{J}) &\triangleq \arg \max_{\mathbf{J}} \arg \min_{\mathbf{C}} L(\mathbf{C}, \mathbf{J}) \\ &= \arg \max_{\mathbf{J}} \arg \min_{\mathbf{C}} \text{tr}((\mathbf{X}_c - \mathbf{C}\mathbf{U})^T (\mathbf{X}_c - \mathbf{C}\mathbf{U})) + \mathbf{J}(\mathbf{C}^T \mathbf{C} - \mathbf{I}_K). \end{aligned} \quad (18)$$

对偶问题 $\text{dual}(\mathbf{J})$ 是关于 $\mathbf{J}$ 的凸二次规划问题, 通过 $\mathbf{C}$ 的解析解, 将式 (18) 表示为只含有 $\mathbf{J}$ 的问题. 通过 $L(\mathbf{C}, \mathbf{J})$ 对 $\mathbf{C}$ 求偏导得到 $\mathbf{C}$ 的解析解

$$\mathbf{C} = (\mathbf{X}_c \mathbf{U}^T) (\mathbf{U} \mathbf{U}^T + \mathbf{J})^{-1}. \quad (19)$$

将 $\mathbf{C}$ 的解析解带入对偶问题 $\text{dual}(\mathbf{C})$ 中, 得到

$$\text{dual}(\mathbf{J}) = \text{tr}(\mathbf{X}_c^T \mathbf{X}_c - (\mathbf{X}_c \mathbf{U}^T) (\mathbf{U} \mathbf{U}^T + \mathbf{J})^{-1} (\mathbf{X}_c \mathbf{U}^T)^T - \mathbf{J}). \quad (20)$$

对偶问题 $\text{dual}(\mathbf{J})$ 是关于 $\mathbf{J}$ 的连续可微函数, 这类问题可以通过梯度下降法求解, 而且 $\text{dual}(\mathbf{J})$ 对 $\mathbf{J}$ 的导数为

$$\frac{d(\text{dual}(\mathbf{J}))}{d\mathbf{J}} = \|\mathbf{X}_c \mathbf{U}^T (\mathbf{U} \mathbf{U}^T + \mathbf{J})^{-1} \mathbf{I}\|^2 - \mathbf{I}. \quad (21)$$

因此在求解问题 (21) 得到 $\mathbf{J}$ 的最优解 $\mathbf{J}^*$ 之后, 将 $\mathbf{J}^*$ 代入式 (19), 得到 $\mathbf{C}$ 的最优解

$$\mathbf{C}^* = (\mathbf{X}_c \mathbf{U}^T) (\mathbf{U} \mathbf{U}^T + \mathbf{J}^*)^{-1}. \quad (22)$$

问题 (15) 的求解过程类似, 在本文内容中将不再赘述. 通过引入 Lagrange 对偶理论, 可以降低问题中求解参数的个数, 如 $\mathbf{C} \in \mathbb{R}^{m \times K}$, 原问题中需要求解的参数个数是 $m \times K$, 而采用 Lagrange 对偶问题后, 求解参数个数减少为 K [19].

4 实验结果与分析

本文中的训练样本来自文献 [8] 中提到的 91 幅自然图像, 对训练图像进行分块, 然后随机选择其中的图像块作为训练样本. 测试图像不仅包括图像超分辨率重建中经常用到的 Set5 和 Set14 两个数据集, 还包括图像分割中经常使用的 Berkeley 数据集, 选择该数据集中的 100 幅图像作为测试数据集, 该数据集简称为 B100. 除了这 3 个自然图像数据集外, 我们还对 LFW (labeled face in the wild crop dataset), ORL 和 Yale 这 3 个人脸图像数据集进行超分辨率重建.

本文所有实验都是在 Matlab2017b 环境下完成, 实验仿真所使用的硬件设备为 Intel(R) Core(TM) I5-4590CPU, 频率 3.30 GHz, 内存 8.0 G.

问题 P_1 中稀疏度量函数 $g(\mathbf{Z}_j)$ 选择不同, 就可以使用不同的分离字典学习方法优化低分辨率分离字典. 当 $g(\mathbf{Z}_j)$ 选择 $\log$ 函数时, 可以用 SeDiL 算法优化低分辨率分离字典; 当 $g(\mathbf{Z}_j) = \|\mathbf{Z}_j\|_0$ 时, 可以用 ASeDiL 算法优化低分辨率分离字典; 当 $g(\mathbf{Z}_j) = \|\mathbf{Z}_j\|_1$ 时, 低分辨率分离字典可以通过 ISeDiL 算法优化. 在实验中, 将通过这 3 种算法学习低分辨率分离字典, 然后进行超分辨率重建的算法分别称为 SeDiLSR 算法 (image super-resolution via separable dictionary learning), ASeDiLSR 算法 (image super-resolution via analytic separable dictionary learning) 和 ISeDiLSR 算法 (image super-resolution via improved separable dictionary learning).

在低分辨率字典学习算法中, SeDiL 算法的最大迭代次数为 500, 参数 $\alpha = 1.35 \times 10^{-6}$, $\beta = 1.29 \times 10^{-6}$. ASeDiL 算法的最大迭代次数为 200, 参数 $\alpha = \beta = 10^{-3}$. ISeDiL 算法的最大迭代次数为 200, 参数 $\rho = 0.07$. 在字典学习和超分辨率重建过程中, 若图像块的大小为 $m \times m$, 则高低分辨率分离字典 $\{\mathbf{A}, \mathbf{B}\}$ 和 $\{\mathbf{C}, \mathbf{F}\}$ 的大小为 $m \times K$, 且 $K = 2m$.

4.1 不同算法的图像超分辨率重建结果与分析

本文比较的图像超分辨率算法包括: Yang 等^[6], Zedye 等^[8], ANR (anchored neighborhood regression)^[20], SRCNN (image super-resolution using deep convolutional networks)^[21]. 其中, Yang 等, Zedye 等和 ANR 算法所使用的字典原子数为 1024 ($K = 1024$). SRCNN 的训练参数为 $f_1 = 9$, $f_2 = 1$, $f_3 = 5$, $n_1 = 64$, $n_2 = 32$ (f 表示卷积核的大小, n 表示对应的卷积个数).

Set14 数据集中的 flower 图像 4 倍超分辨率重建结果如图 6 所示. 从图中可以看出, 对图像超分辨率重建效果最好的是 SRCNN 算法, 但是该算法的训练过程对样本和设备的要求很高. 与 SRCNN 算法相比, 本文提出的 3 种基于分离字典的图像超分辨率重建算法, 虽然 PSNR 与 SRCNN 算法相差 0.5 dB 左右, 但是 SSIM 与 SRCNN 算法非常接近, 即与原图的整体相似性相近, 这也说明了分离字典可以很好地保存图像中的结构信息.

对 Set5, Set14 和 B100 这 3 个数据集中共计 119 幅图像分别进行 2, 3 和 4 倍超分辨率重建, 结果如表 1 所示. 从表中可以看出, SRCNN 算法的重建结果最好, 但重建过程所需时间较长. ANR 算法作为一种快速的超分辨率重建算法, 与其他传统超分辨率重建算法相比, PSNR 提高了 0.3~0.5 dB, 重建时长是 Zedye 算法的 1/5 左右. 而本文提出的 3 种基于分离字典的图像超分辨率重建算法 SeDiLSR, ISeDiLSR 和 ASeDiLSR 算法, 重建质量 PSNR 比 ANR 算法提高了 0.2~0.3 dB, 虽然与 SRCNN 算法相比中间质量略差, 但重建时间远远小于 SRCNN 算法, 约为 SRCNN 算法的 1/14 左右.

4.2 人脸图像的超分辨率结果与分析

为了进一步检验分离字典在图像超分辨率重建中的应用, 除了对 Set5, Set14 和 B100 这 3 个数据集中的图像进行超分辨率重建之外, 本文还对 LFW, ORL 和 Yale 这 3 个人脸图像数据集进行超分辨率重建.

LFW 数据集一共有 13233 张大小为 64×64 的图像, 随机选择其中的 6×10^3 张图像作为训练样本, 剩余图像作为测试样本. Yale 图像数据集含有 100 人的共计 700 张图像, 其中每个人的图像表情和姿势都比较单一, 每张图像的大小为 55×40 , 选择每个人的前 5 张图像作为训练样本, 剩余两张作为测试样本. ORL 数据集含有 40 人共计 400 张图像, 每张图像的大小为 112×92 , 其中人脸图像的表情、姿势、是否佩戴眼镜等情况各不相同, 选择每个人的前 9 张图像作为训练样本, 剩余 1 张作

图 6 (网络版彩图) flower 图像的 4 倍超分辨率重建结果

Figure 6 (Color online) The super-resolution results of flower with magnification 4. (a) The original image; (b) Bicubic; (c) Zedye [8]; (d) ANR [20]; (e) SRCNN [21]; (f) SeDiLSR (ours); (g) ISeDiLSR (ours); (h) ASeDiLSR (ours)

为测试样本. 设输入的人脸图像大小为 $m \times n$, 字典 A 和 C 的大小 $m \times 2m$, B 和 F 的大小为 $n \times 2n$. 在人脸图像的超分辨率重建过程中, 没有对人脸图像进行分块, Yang 等、Zedye 等图像超分辨率重建算法在个人电脑上无法实现, 所以本组实验所比较的算法只有 SeDiLSR, ISeDiLSR 和 ASeDiLSR 这 3 种算法.

用 SeDiLSR, ASeDiLSR 和 ISeDiLSR 3 种基于分离字典的图像超分辨率重建算法对 LFW, Yale 和 ORL 人脸图像数据集进行 2, 3 和 4 倍超分辨率重建, 其重建结果如表 2 所示, 重建过程所耗费的时间如表 3 所示.

从表中 2 可以看出, LFW 数据库中因为用足够的人脸图像作为训练样本, 所以超分辨率重建质量提升比较明显, 在 2 倍超分辨率重建时, PSNR 比三线性插值 Bicubic 算法提高了约 0.8 dB. Yale 数据库虽然训练样本比较少, 只有 500 张人脸图像, 但是整个数据库的人脸图像姿势和表情都比较单一, 所以超分辨率重建效果提升也很明显. ORL 数据库中一共只有 400 张图像, 选择其中的 360 张作为训练样本, 不仅训练样本比较少, 而且该数据库中的人脸图像表情和姿态变化都非常丰富, 其中部分人脸还戴有眼镜等遮挡物, 所以 ORL 数据库的超分辨率重建结果相对比较差, 与 Bicubic 三线性插值相比, PSNR 仅提升了 0.3~0.4 dB, SSIM 提高了约 0.01. 从表 3 可以看出, SeDiLSR, ASeDiLSR 和 ISeDiLSR 这 3 种基于分离字典的超分辨率重建算法除了 4 倍超分辨率重建时后两种算法所需时间较长外, 其余超分辨率重建所需时间都比较相近. 部分人脸图像的 3 倍超分辨率重建结果如图 7 所示.

5 总结

本文主要对基于分离字典的图像超分辨率重建问题进行建模与求解, 将整个过程分为: 高低分辨率分离字典的学习过程和图像超分辨率重建过程.

在高低分辨率分离字典的学习过程中, 提出使用已有的 3 种分离字典学习算法优化低分辨率分离字典 $\{A, B\}$, 然后再采用 Lagrange 对偶理论优化高分辨率分离字典 $\{C, F\}$. 在图像的超分辨率重建

表 1 Set5, Set14 和 B100 的超分辨率重建结果

Table 1 Average PSNR, SSIM and time (s) on Set5, Set14 and B100

		Set5			Set14			B100		
		2	3	4	2	3	4	2	3	4
Bicubic	PSNR	33.66	30.39	28.42	30.23	27.54	26	29.32	27.15	25.92
	SSIM	0.93	0.87	0.81	0.87	0.77	0.69	0.832	0.733	0.663
Yang et al.	PSNR	–	31.41	–	–	28.31	–	–	27.72	–
	SSIM	–	0.88	–	–	0.79	–	–	0.759	–
	Time	–	42.14	–	–	96.82	–	–	30.04	–
Zedye et al.	PSNR	35.78	31.9	29.69	31.81	28.66	26.89	30.39	27.87	26.51
	SSIM	0.95	0.89	0.85	0.898	0.81	0.73	0.866	0.767	0.693
	Time	4.63	2.08	1.25	3.92	4.14	1.22	2.49	1.25	0.79
ANR	PSNR	35.79	31.9	26.69	31.76	28.64	26.86	30.42	27.88	26.51
	SSIM	0.949	0.898	0.844	0.899	0.808	0.734	0.869	0.77	0.696
	Time	0.78	0.45	0.34	0.97	0.9	0.46	0.59	0.4	0.31
SRCNN	PSNR	36.66	32.389	30.48	32.454	29	27.5	31.36	28.21	26.9
	SSIM	0.954	0.905	0.865	0.906	0.813	0.749	0.887	0.778	0.7
	Time	3.854	3.546	3.863	8.345	7.81	8.396	5.766	5.51	5.804
SeDiLSR (ours)	PSNR	35.52	31.945	29.66	32.028	28.93	27.05	31.32	28.46	26.95
	SSIM	0.95	0.9	0.85	0.917	0.831	0.76	0.9	0.8	0.73
	Time	3.135	0.8016	0.38	5.65	0.84	0.6	3.69	0.595	0.44
ASeDiLSR (ours)	PSNR	35.34	32.0	29.7	32.059	28.97	27.09	31.11	25.5	26.8
	SSIM	0.953	0.902	0.85	0.918	0.833	0.76	0.904	0.806	0.733
	Time	2.879	0.437	0.34	5.6	0.816	0.57	3.7	0.55	0.36
ISeDiLSR (ours)	PSNR	35.56	32.05	29.72	32.062	29.01	27.1	31.35	28.53	26.99
	SSIM	0.954	0.903	0.849	0.918	0.83	0.76	0.9	0.81	0.734
	Time	2.829	0.444	0.315	5.596	0.82	0.576	3.69	0.619	0.39

表 2 人脸数据库的超分辨率重建结果

Table 2 The super-resolution results of the datasets of LFW, ORL and Yale

		Bicubic		SeDiLSR (ours)		ASeDiLSR (ours)		ISeDiLSR (ours)	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
2	LFW	35.122	0.9616	35.5396	0.968	36.3721	0.9698	35.9968	0.9661
	ORL	30.3001	0.9123	30.5224	0.9195	30.7368	0.9209	30.7202	0.9201
	Yale	26.2624	0.8817	26.1103	0.8817	26.5393	0.8854	26.8538	0.8873
3	LFW	29.9856	0.8849	30.2636	0.8961	30.6429	0.8984	30.7381	0.8986
	ORL	27.6624	0.8345	27.8313	0.8433	27.9484	0.8444	27.9499	0.8435
	Yale	23.0754	0.7635	23.0021	0.7624	23.502	0.7735	23.8688	0.7892
4	LFW	27.0311	0.7886	27.2905	0.8043	27.4714	0.806	27.6124	0.8102
	ORL	26.0258	0.7593	26.2167	0.7691	26.2728	0.7698	26.3004	0.7704
	Yale	20.9202	0.4181	20.8746	0.4188	21.3545	0.5681	21.7321	0.5265

过程中, 其目标是将待处理样本/图像块在对应字典下进行稀疏表示, 然后再通过高分辨率分离字典和稀疏系数矩阵重建对应的高分辨率样本/图像块.

最后, 我们对不同的数据集 (Set5, Set14, B100 共 3 个自然图像数据集和 LFW, Yale, ORL 共 3 个

表 3 LFW, ORL 和 Yale 人脸图像超分辨率重建所需时间 (s)
Table 3 The time consumption for the facial images in LFW, ORL and Yale (s)

Algorithm	2			3			4		
	LFW	ORL	Yale	LFW	ORL	Yale	LFW	ORL	Yale
SeDiLSR (ours)	0.095	0.0841	0.138	0.0984	0.0703	0.07	0.0949	0.091	0.0581
ASeDiLSR (ours)	0.0522	0.0827	0.0168	0.028	0.0848	0.052	0.0287	0.1292	0.027
ISeDiLSR (ours)	0.0906	0.0826	0.1826	0.0964	0.0843	0.0796	0.0844	0.1268	0.1805

图 7 人脸图像 3 倍超分辨率重建结果
Figure 7 The super-resolution results of facial images with magnification 3

人脸图像数据集) 进行超分辨率重建. 实验结果表明, 基于分离字典的图像超分辨率重建算法SeDiLSR, ASeDiLSR 和 ISeDiLSR 算法, 无论是在重建质量 (PSNR 和 SSIM) 上, 还是在重建速度上, 与传统的超分辨率重建算法相比都有较强的优势. 与最新的基于深度学习的图像超分辨率重建算法 (SRCNN)

相比, 在重建质量相似的情况下, 重建速度的优势也非常明显.

参考文献

- 1 Tian Y S, Yap K H. Joint image registration and super-resolution from low-resolution images with zooming motion. *IEEE Trans Circ Syst Video Technol*, 2013, 23: 1224–1234
- 2 Yang M C, Wang Y C F. A self-learning approach to single image super-resolution. *IEEE Trans Multimedia*, 2013, 15: 498–508
- 3 Yang S F, Zhao R Z. Image super-resolution reconstruction based on low-rank matrix and dictionary learning. *J Comput Res Dev*, 2016, 53: 884–891 [杨帅锋, 赵瑞珍. 基于低秩矩阵和字典学习的图像超分辨率重建, *计算机研究与发展*, 2016, 53: 884–891]
- 4 Liu C X, Zhang F L, Li M. Image super-resolution based on extreme learning machine. *J Hebei Univ Technol*, 2016, 46: 11–16 [刘翠响, 张凤林, 李敏. 基于极限学习机的图像超分辨率重建算法. *河北工业大学学报*, 2017, 46: 11–16]
- 5 Jiang J, Tang J M, Pan Z Y. Super-resolution reconstruction based on structure tensor's eigenvalue and classification dictionary. In: *Proceedings of IEEE International Conference on Audio, Language and Image Processing (ICALIP)*, Shanghai, 2016. 452–456
- 6 Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation. *IEEE Trans Image Process*, 2010, 19: 2861–2873
- 7 Yang J C, Wang Z W, Lin Z, et al. Coupled dictionary training for image super-resolution. *IEEE Trans Image Process*, 2012, 21: 3467–3478
- 8 Zeyde R, Elad M, Protter M. On single image scale-up using sparse representation. In: *Proceeding of the 7th International Conference on Curves and Surfaces*, Avignon, 2010. 711–730
- 9 Xie H Y, Xie Z G, Hu W, et al. Optimization of dictionary atoms in image sparse representations and its application. *J Southeast Univ (Nat Sci Edit)*, 2014, 44: 116–122 [谢洪均, 谢正光, 胡伟, 等. 字典原子优化的图像稀疏表示及其应用. *东南大学学报 (自然科学版)*, 2014, 44: 116–122]
- 10 Engan K, Aase S O, Husøy J H. Multi-frame compression: theory and design. *Signal Process*, 2000, 80: 2121–2140
- 11 Engan K, Aase S O, Hakon Husoy J. Method of optimal directions for frame design. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, 1999
- 12 Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*, 2006, 54: 4311–4322
- 13 Hawe S, Kleinsteuber M, Diepold K. Analysis operator learning and its application to image reconstruction. *IEEE Trans Image Process*, 2013, 22: 2138–2150
- 14 Hawe S, Seibert M, Kleinsteuber M. Separable dictionary learning. In: *Proceedings of Conference on Computer Vision and Pattern Recognition*, Portland, 2013. 438–445
- 15 Zhang F Z, Cen Y G, Zhao R Z, et al. Analytic separable dictionary learning based on oblique manifold. *Neurocomputing*, 2016, 236: 32–38
- 16 Zhang F Z, Cen Y G, Zhao R Z, et al. Improved separable dictionary learning. In: *Proceedings of IEEE International Conference on Signal Processing*, Chengdu, 2016. 884–889
- 17 Fang Y, Wu J J, Huang B M. 2D sparse signal recovery via 2D orthogonal matching pursuit. *Sci China Inf Sci*, 2012, 55: 889–897
- 18 Rubinstein R, Bruckstein A M, Elad M. Dictionaries for sparse representation modeling. *Proc IEEE*, 2010, 98: 1045–1057
- 19 Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms. In: *Proceedings of International Conference on Neural Information Processing Systems*, Vancouver, 2006. 801–808
- 20 Timofte R, De V, Gool L V. Anchored neighborhood regression for fast example-based super-resolution. In: *Proceeding of IEEE Conference on Computer Vision*, Washington, 2013. 1920–1927
- 21 Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell*, 2016, 38: 295–307

The algorithm of image super-resolution reconstruction via separable dictionaries

Fengzhen ZHANG^{1,2,3}, Yigang CEN^{2,3*}, Ruizhen ZHAO^{2,3}, Yanhong WANG^{2,3},
Linna ZHANG⁴ & Shaohai HU^{2,3}

1. *Key Laboratory of Solar Activity, National Astronomical Observatory, Chinese Academy of Science, Beijing 100101, China;*

2. *Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China;*

3. *Key Laboratory of Advanced Information Science and Network Technology of Beijing, Beijing 100044, China;*

4. *College of Mechanical Engineering, Guizhou University, Guiyang 550025, China*

* Corresponding author. E-mail: ygcen@bjtu.edu.cn

Abstract Traditional sparse representation-based super-resolution algorithms need to divide images into patches and then stack them into columns. This operation ignores the intrinsic 2D structure and spatial correlation inherent in patches. In order to fully exploit 2D spatial correlation in image patches, we combine the sparse representation ability of the separable dictionary in both the horizontal and vertical directions, and propose an algorithm for image super-resolution based on a separable dictionary. The experimental results show that our proposed algorithm not only improves the efficiency of image super-resolution, but also improves the PSNR and SSIM (i.e., about 0.2-dB PSNR better than traditional methods, and 0.01 SSIM better than existing methods).

Keywords image super-resolution reconstruction, 2D sparse coding, Riemannian manifold, sparse representation, separable dictionary

Fengzhen ZHANG was born in 1983. She received a Ph.D. from the Institute of Information Science, Beijing Jiaotong University, in September 2018. In October 2018, she joined the National Astronomical Observatory, Chinese Academy of Science, as a postdoctoral fellow. Her research interests include compressed sensing, sparse representation, and astronomical-image processing.

Yigang CEN was born in 1978. He received a Ph.D. degree in control science and engineering in 2006 from HuaZhong University of Science and Technology. In September 2006, he joined Nanyang Technology University in Singapore as a research fellow IEEE signal processing. From January 2014 to January 2015, he was a visiting scholar at the Department of Computer Science, University of Missouri in the US. He is currently a professor at Beijing Jiaotong University. His research interests include compressed sensing, sparse representation, low-rank-matrix reconstruction, and wavelet-construction theory.

Ruizhen ZHAO was born in 1975. He received his Ph.D. from Xidian University. After that, he was with a postdoctoral fellow at the Institute of Automation, Chinese Academy of Sciences. He is currently a professor and supervisor of doctoral students. His research interests include wavelet transformation and its applications, algorithms for image and signal processing compression, and sparse representation.

Yanhong WANG was born in 1979. She received her B.E. degree from Qilu University of Technology, Jinan, China, in 2003, and her master's degree from Shandong University in 2009. Currently, she is pursuing her Ph.D. at Beijing Jiaotong University Institute of Information Science. Her research interests include multimedia-information retrieval, sparse representation, and computer vision.