

基于多尺度结构自相似性的单幅图像超分辨率算法

潘宗序¹ 禹晶¹ 胡少兴² 孙卫东¹

摘要 多尺度结构自相似性是指同一幅图像中存在相同尺度或不同尺度的相似结构, 这种多尺度图像结构自相似性广泛存在于遥感图像中. 本文提出了一种基于多尺度结构自相似性的单幅图像超分辨率 (Super resolution, SR) 算法, 该算法结合了压缩感知框架与图像结构自相似性, 利用非局部方法和基于图像金字塔的 K-SVD 字典学习方法, 将蕴含在相同尺度和不同尺度相似图像块中的附加信息在压缩感知的框架下加入到重构图像中. 本文算法的优势在于, 它仅借助于单幅低分辨率图像自身所蕴含的信息, 实现了空间分辨率的提升. 实验表明, 与 CSSS 算法和 ASDSAR 算法相比, 本文算法更有效地提升了遥感图像的空间分辨率.

关键词 超分辨率, 结构自相似性, 多尺度, 压缩感知, 非局部方法

引用格式 潘宗序, 禹晶, 胡少兴, 孙卫东. 基于多尺度结构自相似性的单幅图像超分辨率算法. 自动化学报, 2014, 40(4): 594–603

DOI 10.3724/SP.J.1004.2014.00594

Single Image Super Resolution Based on Multi-scale Structural Self-similarity

PAN Zong-Xu¹ YU Jing¹ HU Shao-Xing² SUN Wei-Dong¹

Abstract Multi-scale structural self-similarity refers to those similar structures either within the same scale or across different scales coming from the same image, which widely occur in remote sensing images. In this paper, we propose a single image super resolution (SR) method based on multi-scale structural self-similarity, which combines compressive sensing framework and structural self-similarity. In our method, the nonlocal and the pyramid-based K-SVD methods are used to add the extra information hidden in multi-scale structural self-similarity into the reconstructed image in the compressive sensing framework. The advantage of our method is that it only uses a single low-resolution image to promote spatial resolution by fully exploiting the extra information hidden in the image itself. Experimental results demonstrate that our method can improve spatial resolution more effectively compared with the CSSS and the ASDSAR methods.

Key words Super resolution (SR), structural self-similarity, multi-scale, compressive sensing, nonlocal

Citation Pan Zong-Xu, Yu Jing, Hu Shao-Xing, Sun Wei-Dong. Single image super resolution based on multi-scale structural self-similarity. *Acta Automatica Sinica*, 2014, 40(4): 594–603

在众多领域中, 高分辨率图像的获取具有重要意义. 对于遥感图像, 高分辨率图像有利于图像中目标的检测、识别以及进一步处理, 然而成像平台和成像设备的制造工艺、成本等因素制约了遥感图像的空间分辨率. 图像超分辨率 (Super resolution, SR) 方法是利用信号处理方法提升图像空间分辨率的一

项技术. SR 方法将低分辨率成像设备获取图像的过程看作是由高分辨率图像到低分辨率图像的降质过程, 并利用单幅或多幅低分辨率图像重构高分辨率图像. SR 方法的关键是在图像重构过程中加入一定的附加信息来弥补降质过程中损失的细节信息. 根据所获取的图像数量, 可将 SR 方法分为多幅图像 SR 方法和单幅图像 SR 方法. 多幅图像 SR 方法利用多幅低分辨率图像所提供的互补信息重构高分辨率图像, 而单幅图像 SR 方法利用学习获取的附加信息重构高分辨率图像. 在遥感领域, 由于同一时相、同一区域多幅遥感图像的获取十分困难, 因此单幅图像 SR 方法更具实用性.

目前广泛认可的单幅图像 SR 方法包括基于正则化约束的 SR 方法和基于结构自相似性 (Structural self-similarity) 的 SR 方法. 基于正则化约束的 SR 方法将附加信息作为正则化约束项加入到图像重构过程中, 将图像重构问题转换为寻求满足特定限制条件解的最优化问题. 基于全变分 (Total variation, TV) 模型的 SR 方法将图像梯度的范数作

收稿日期 2012-10-18 录用日期 2013-02-18
Manuscript received October 18, 2012; accepted February 18, 2013

国家自然科学基金 (61171117), 国家科技支撑计划项目 (2012BAH31B01), 北京市教育委员会科技计划重点项目 (KZ201310028035) 资助
Supported by National Natural Science Foundation of China (61171117), National Science and Technology Pillar Program of China (2012BAH31B01), and Key Project of the Science and Technology Development Program of Beijing Education Committee of China (KZ201310028035)

本文责任编辑 刘一军

Recommended by Associate Editor LIU Yi-Jun

1. 清华大学电子工程系 北京 100084 2. 北京航空航天大学机械工
程与自动化学院 北京 100083

1. Department of Electronic Engineering, Tsinghua University, Beijing 100084 2. School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083

为正则化约束项^[1-2]. 自然图像在某种字典下具有稀疏性, 可以将这种稀疏性作为正则化约束项, 并在压缩感知 (Compressive sensing) 框架下加入到图像重构过程中. 随着压缩感知技术的不断发展, 相继提出了多种基于压缩感知的 SR 方法^[3-5]. 基于压缩感知的 SR 方法的关键是如何选取用于表示图像的字典, 使得图像块在字典下具有稀疏表示形式. 文献 [3] 利用自然图像在小波基下具有稀疏性这一特点将小波基作为字典, 但由于在该方法中字典是固定的, 因而缺乏自适应性, 这也制约了这种方法的性能. 文献 [4-5] 利用由大量高分辨率图像组成的图像库作为训练样本来构建字典, 这种方法在重构过程中能够获得很多附加信息, 但无法保证这些附加信息的准确性和可靠性. 因此, 尽管这种方法获得的重构图像具有好的视觉效果, 然而均方误差较大. 在这种全局字典的构建过程中, 为了使各种不同类型的图像块在字典下均能具有稀疏表示, 必须利用大量的样本进行字典学习, 这使得字典的学习效率降低; 另外, 当用于字典学习的样本不能提供与待处理低分辨率图像相似的准确信息时, 这种方法的图像重构效果难以保证. 造成这种问题的根本原因在于不存在能够稀疏表示所有图像块的全局字典, SR 方法需要的是准确而不是大量不准确的附加信息, 从这一点上来讲, 尽管使用图像库可以利用大量的图像作为字典学习的样本, 然而这里的“多”并不一定代表“好”.

基于结构自相似性的 SR 方法利用了图像广泛具有的结构自相似性, 并将这种结构自相似性作为图像重构过程中的附加信息^[6-9]. 近年来相继出现了一些利用非局部 (Nonlocal, NL) 自相似结构的 SR 方法^[6-7], 这种方法的基本思路是在整幅图像中搜索相同尺度的相似图像块, 利用这些相似图像块所提供的互补信息重构高分辨率图像. 由于图像自身中相同尺度相似图像块所提供的附加信息具有一定的局限性, 因而制约了这种方法图像重构效果的进一步提升. 文献 [8] 构建与图像高、低频子图所对应的高、低频码本, 并通过高频码本所提供的附加信息实现图像重构, 这种方法实际上是利用了图像自身不同尺度相似图像块所提供的附加信息. 文献 [9] 将相同尺度和不同尺度相似图像块所提供的附加信息均加入到图像重构过程中, 利用相同尺度相似图像块就如同利用多幅低分辨率图像, 而不同尺度相似图像块又提供了不同分辨率图像块间的对应关系. 文献 [8-9] 在整幅图像中搜索具有相似结构的图像块, 因此存在运算复杂度较高的问题.

文献 [10] 将基于压缩感知的 SR 方法与基于结构自相似性的 SR 方法结合起来实现图像分辨率的提升, 提出了自适应字典选取与正则化算法 (Adaptive sparse domain selection and adaptive

regularization method, ASDSAR), ASDSAR 算法对图像库中的样本进行聚类, 并对每一类分别训练字典, 将 NL 方法作为正则化约束项加入到图像重构过程中, 但这种算法没有考虑不同尺度相似图像块所提供的附加信息. 在文献 [11] 中, 我们提出了一种结合压缩感知框架和结构自相似性的单幅图像超分辨率算法 (Compressive sensing and structural self-similarity based method, CSSS). CSSS 算法利用图像结构自相似性所提供的附加信息, 通过压缩感知框架实现图像重构, 并在图像重构过程中仅使用了待处理低分辨率图像的插值图像作为 K-SVD 字典学习的样本, 而没有采用其他任何高分辨率图像. 本文在对 CSSS 算法补充和完善的基础上, 提出了一种基于多尺度结构自相似性的单幅图像超分辨率算法. CSSS 算法仅利用了相同尺度的相似图像块, 而没有利用不同尺度的相似图像块, 本文算法则结合相同尺度和不同尺度的相似图像块, 利用非局部方法和基于图像金字塔的 K-SVD 字典学习方法, 将蕴含在相同尺度和不同尺度相似图像块中的附加信息在压缩感知的框架下加入到重构图像中. 此外, CSSS 算法采用了正交匹配追踪 (Orthogonal matching pursuit, OMP) 方法构建稀疏表示系数, 但 OMP 方法固有的局限性使得算法在 K-SVD 字典学习时必须严格限制图像块的稀疏性, 从而降低了字典学习的精度, 本文算法则利用迭代收缩算法 (Iterative shrinkage algorithm)^[12] 取代 OMP 方法构建稀疏表示系数, 从而克服了这一问题.

本文后续结构组织如下, 第 1 节介绍图像的稀疏表示与基于压缩感知的 SR 方法数学模型, 第 2 节阐述本文如何将蕴含在图像结构自相似性中的附加信息加入到重构图像中, 第 3 节通过实验验证本文算法的有效性, 第 4 节为全文的结论.

1 图像的稀疏表示与基于压缩感知的 SR 方法数学模型

设 $\sqrt{n} \times \sqrt{n}$ 大小的图像块组成向量 \mathbf{x} , $\mathbf{x} \in \mathbf{R}^n$, 若存在矩阵 Ψ 以及向量 $\boldsymbol{\alpha}$ 使得

$$\mathbf{x} = \Psi\boldsymbol{\alpha}, \quad \|\boldsymbol{\alpha}\|_0 = k \ll n \quad (1)$$

则称向量 \mathbf{x} 在 Ψ 下具有稀疏性, 其中, $\Psi \in \mathbf{R}^{n \times t}$ 称为字典, 字典 Ψ 中的每一列称为字典中的一个元素; $\boldsymbol{\alpha} \in \mathbf{R}^t$ 为稀疏表示系数且满足 $\|\boldsymbol{\alpha}\|_0 = k \ll n$, 这里 $\|\boldsymbol{\alpha}\|_0$ 表示 $\boldsymbol{\alpha}$ 中非零元的个数. 图像块所对应的向量通常在某种字典下具有稀疏性, 即图像块可以表示为字典中少数几个元素的线性组合. 通常通过学习的方式来构建字典, 学习中用于构建字典的向量称为样本. K-SVD 字典学习方法^[13] 是一种有效的字典构建方法, 这种方法通过对样本的学习构建字典, 设样本 \mathbf{x}_i , $i = 1, 2, \dots, p$ 组成样本矩阵 X .

$= [\mathbf{x}_1, \dots, \mathbf{x}_p]$, $A = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_p]$ 为稀疏表示矩阵, $\boldsymbol{\alpha}_i$ 为 \mathbf{x}_i 对应的稀疏表示系数. K-SVD 字典学习方法在 Ψ 与 A 均未知的情况下通过求解如下的优化问题构建字典 Ψ :

$$\min_{\Psi, A} \left\{ \|X_s - \Psi A\|_F^2 \right\} \text{ s.t. } \forall i, \|\boldsymbol{\alpha}_i\|_0 \leq T \quad (2)$$

其中, T 为控制表示系数稀疏程度的常数.

设 $\mathbf{X} \in \mathbf{R}^N$ 表示高分辨率图像组成的向量, $\mathbf{Y} \in \mathbf{R}^M$ 表示低分辨率图像组成的向量, $\hat{\mathbf{X}} \in \mathbf{R}^N$ 表示重构图像组成的向量, 则高分辨率图像与低分辨率图像之间的关系可以表示为

$$\mathbf{Y} = D\mathbf{H}\mathbf{X} + \mathbf{v} \quad (3)$$

式中, D 表示降采样矩阵, H 表示模糊矩阵, \mathbf{v} 为加性噪声. 式 (3) 所示的观测模型说明低分辨率图像 \mathbf{Y} 是由高分辨率图像 \mathbf{X} 经过模糊、降采样以及加入噪声等过程获取的, 而 SR 方法要解决的问题则是如何由 \mathbf{Y} 重构 \mathbf{X} , 即求解如下的最小二乘问题:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{Y} - D\mathbf{H}\mathbf{X}\|_2^2 \quad (4)$$

由于式 (4) 的解不惟一, 为了获得惟一解或最优解, 需要在式 (4) 的目标函数中添加约束项. 为了将图像的稀疏性作为约束项加入式 (4) 所示的 SR 模型中, 通常需要将图像进行分块, 图像块之间可以重叠. 设 $\mathbf{x}_i \in \mathbf{R}^n$ 表示高分辨率图像块组成的向量, $\hat{\mathbf{x}}_i \in \mathbf{R}^n$ 表示重构图像块组成的向量, 则图像块与图像之间的关系可以表示为 $\mathbf{x}_i = R_i \mathbf{X}$, $i = 1, 2, \dots, p$, 其中 $R_i \in \mathbf{R}^{n \times N}$ 为抽取矩阵, 其作用是将图像块从图像中抽取出来. 重构图像块的稀疏表示为 $\hat{\mathbf{x}}_i = \Psi \boldsymbol{\alpha}_i$, $i = 1, 2, \dots, p$, 其中, Ψ 为字典, $\boldsymbol{\alpha}_i$ 为稀疏表示系数. 重构图像 $\hat{\mathbf{X}}$ 可以通过对所有的重构图像块 $\hat{\mathbf{x}}_i$ 进行平均处理获得^[14], 为了便于书写, 引入运算符 “ \circ ”, 则有:

$$\hat{\mathbf{X}} = \Psi \circ \boldsymbol{\alpha} = \left(\sum_{i=1}^p R_i^T R_i \right)^{-1} \sum_{i=1}^p (R_i^T \Psi \boldsymbol{\alpha}_i) \quad (5)$$

将式 (5) 代入式 (4), 并加入对稀疏表示系数的稀疏性约束, 基于压缩感知的 SR 方法的数学模型则可表示为

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\mathbf{Y} - D\mathbf{H}\Psi \circ \boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_0 \right\} \quad (6)$$

式中, 第 1 项表示观测模型对重构图像的约束, 第 2 项表示图像稀疏性对重构图像的约束. 式 (6) 中最小化 l_0 范数的优化问题是一个 NP 难问题, 通常需要使用追踪算法求解. 当向量稀疏时, l_1 范数可以代替 l_0 范数, 且 l_1 范数最小化问题是一个凸优化问题, 因此通常将最小化 l_1 范数的优化问题用于基于压缩感知的 SR 方法的数学模型中, 此时式 (6) 转换

为如下的凸优化问题:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\mathbf{Y} - D\mathbf{H}\Psi \circ \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^p \sum_{j=1}^t \lambda |\alpha_{i,j}| \right\} \quad (7)$$

通过求解式 (7) 可以得到图像的稀疏表示系数 $\boldsymbol{\alpha}$, 再将 $\boldsymbol{\alpha}$ 带入式 (5) 即可得到重构图像 $\hat{\mathbf{X}}$.

2 基于多尺度结构自相似性的 SR 方法

2.1 图像多尺度结构自相似性

图像的多尺度结构自相似性是指在同一幅图像中存在相同尺度或不同尺度的相似结构, 其具体表现为图像中相同尺度和不同尺度的相似图像块. 图 1 给出了两幅具有显在结构自相似性的图像例子, 图中存在大量结构相似的车辆, 位于左图同一行中的车辆具有相同尺度的相似结构, 而位于不同行中的车辆具有不同尺度的相似结构. 这种多尺度结构自相似性在遥感图像中是非常常见的, 它们通常以显在或潜在的形式普遍存在于道路、房屋以及自然景观中, 并为实现空间分辨率提升提供了必要的附加信息.

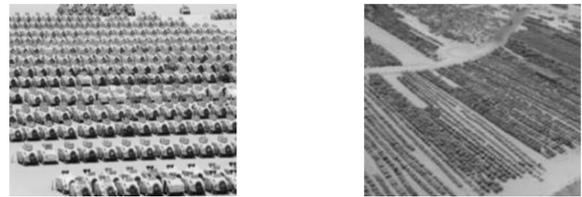


图 1 具有显在多尺度结构自相似的图像例子

Fig. 1 Typical images with explicit multi-scale structural self-similarity

利用相同尺度相似图像块如同利用多幅低分辨率图像, 图 2 直观说明了利用相同尺度相似图像块的 SR 方法原理. 在图 2(a) 中, 多幅图像 SR 方法利用多幅低分辨率图像中具有亚像素位移的图像块实现图像重构. 图 2(b) 显示了在单幅低分辨率图像中存在着诸多相同尺度的相似图像块, 利用这些相似图像块进行图像重构就如同利用多幅低分辨率图像中具有亚像素位移的图像块进行图像重构一样. 从这个意义上讲, 图 2(a) 与图 2(b) 在实现原理上是一致的, 当图像中存在相同尺度相似图像块时两者是近似等价的, 不同之处在于图 2(b) 仅利用了单幅低分辨率图像.

不同尺度相似图像块提供了高、低分辨率图像之间的对应关系, 图 3 直观说明了利用不同尺度相似图像块的 SR 方法原理. 在图 3(a) 中, HR 表示高分辨率图像, LR 表示与之对应的低分辨率图像, HR 图像尺寸为 LR 图像的 s 倍. 设 Ω_1^{HR} 和 Ω_2^{HR} 为 HR 图像中具有不同尺度的相似图像块, 且 Ω_2^{HR} 尺

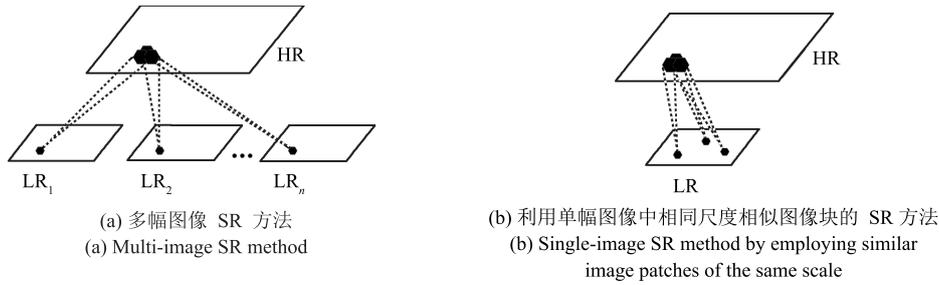


图 2 利用相同尺度相似图像块的 SR 方法原理

Fig. 2 Principle of applying similar image patches of the same scale to SR reconstruction

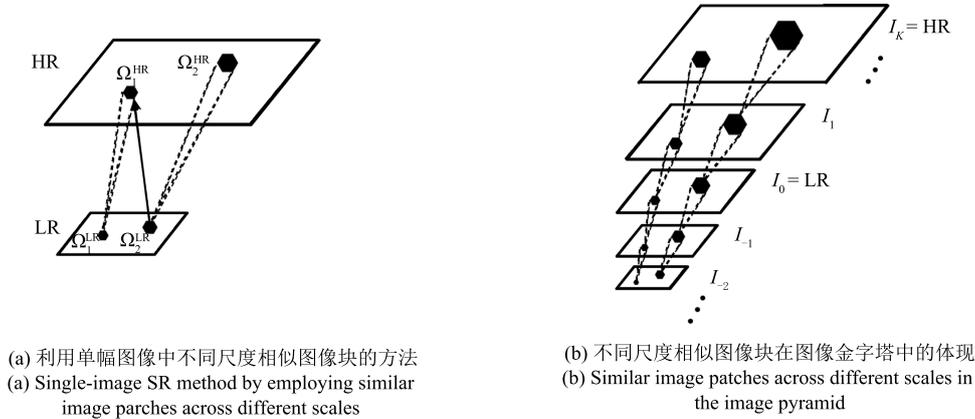


图 3 利用不同尺度相似图像块的 SR 方法原理

Fig. 3 Principle of applying similar image patches across different scales to SR reconstruction

寸为 Ω_1^{HR} 的 s 倍, 它们在 LR 图像中对应的图像块分别为 Ω_1^{LR} 和 Ω_2^{LR} . 此时, 低分辨率图像中的 Ω_1^{LR} 和 Ω_2^{LR} 就构成了一对不同尺度的相似图像块. 由于 HR 图像和 LR 图像之间的缩放系数与 Ω_2^{HR} 和 Ω_1^{HR} 之间的缩放系数相同, Ω_1^{HR} 与 Ω_2^{LR} 完全相同, 因此在对 LR 图像中的 Ω_1^{LR} 进行重构时, Ω_2^{LR} 就能提供准确的附加信息. 虽然这是一个极为理想的情况, 但是说明了图像中不同尺度相似图像块可以提供图像重构所需要的附加信息. 当对图像进行金字塔分解时, 多级金字塔结构可以提供不同尺度和相同尺度的相似图像块. 图 3(b) 直观地说明了不同尺度相似图像块在图像金字塔中的体现, 其中第 0 层 I_0 表示低分辨率图像, 第 K 层 I_K 表示高分辨率图像. 在实际应用中, 低分辨率图像下层的图像可以通过降采样获得, 而低分辨率图像上层的图像可以通过插值方法获得.

2.2 图像结构自相似性附加信息的获取

相似图像块分为相同尺度和不同尺度相似图像块两类, 本文通过 NL 方法将蕴含在相同尺度相似图像块中的附加信息加入到重构图像中, 通过基于图像金字塔的 K-SVD 字典学习方法同时将蕴含在相同尺度和不同尺度相似图像块中的附加信息加入

到重构图像中.

NL 方法通过迭代的方式将蕴含在相同尺度相似图像块中的附加信息以正则化约束项的形式加入到 SR 的数学模型中. 在 NL 方法中, 首先产生一幅初始重构图像, 本文将低分辨率图像的插值图像作为初始重构图像, 然后通过迭代逐次更新重构图像. 对于当前重构图像块 $\hat{\mathbf{x}}_i$, 在当前重构图像 $\hat{\mathbf{X}}$ 中遍历搜索与 $\hat{\mathbf{x}}_i$ 相似的图像块. 计算搜索图像块 $\hat{\mathbf{x}}_i^l$ 与当前重构图像块 $\hat{\mathbf{x}}_i$ 之间的差异 $e_i^l = \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_i^l\|_2^2$, 取 L 个与 $\hat{\mathbf{x}}_i$ 最为相近的图像块 $\hat{\mathbf{x}}_i^l$, $l = 1, 2, \dots, L$, 作为 \mathbf{x}_i 的相同尺度相似图像块. 由于在整幅图像中遍历搜索会消耗大量的运算时间, 实际只取 $\hat{\mathbf{x}}_i$ 的较大邻域即可. 对于遥感图像而言, 相同尺度相似图像块普遍出现在临近的范围内, 比如分段的道路、临近的建筑物和成片的自然景物等, 因此这种限制搜索范围的方法是行之有效的.

设 χ_i 和 χ_i^l 分别为图像块 \mathbf{x}_i 和 \mathbf{x}_i^l 中心像素的灰度值, 用 χ_i^l , $l = 1, 2, \dots, L$ 的线性表示对 χ_i 进行估计, 即 $\hat{\chi}_i = \sum_{l=1}^L \omega_i^l \chi_i^l$, 则 $\sum_{\chi_i \in \mathbf{X}} \|\chi_i - \sum_{l=1}^L \omega_i^l \chi_i^l\|_2^2$ 应该较小^[10], 其中, $\omega_i^l = \exp(-e_i^l/h) / \sum_{l=1}^L \exp(-e_i^l/h)$ 为权重, h 为控制常数. 令 $\boldsymbol{\omega}_i$ 表示 ω_i^l , $l = 1, 2, \dots, L$ 所组成的向量, $\boldsymbol{\chi}_i$ 表示

χ_i^l , $l = 1, 2, \dots, L$ 所组成的向量, 上式可写成 $\sum_{\chi_i \in \mathbf{X}} \|\chi_i - \omega_i^T \boldsymbol{\alpha}\|_2^2$, 将该式作为约束项加入到式 (7) 中, 则有:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\mathbf{Y} - DH\Psi \circ \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^p \sum_{j=1}^t \lambda |\alpha_{i,j}| + \eta \sum_{\chi_i \in \mathbf{X}} \|\chi_i - \omega_i^T \boldsymbol{\alpha}\|_2^2 \right\} \quad (8)$$

其中, 参数 η 控制约束项 $\sum_{\chi_i \in \mathbf{X}} \|\chi_i - \omega_i^T \boldsymbol{\alpha}\|_2^2$ 在代价函数中的权重. 式 (8) 可以用矩阵形式表示为

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\mathbf{Y} - DH\Psi \circ \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^p \sum_{j=1}^t \lambda |\alpha_{i,j}| + \eta \|(I - B)\Psi \circ \boldsymbol{\alpha}\|_2^2 \right\} \quad (9)$$

其中, I 为单位矩阵, B 为权值矩阵, 满足:

$$B(i, l) = \begin{cases} \omega_i^l, & \text{若 } \chi_i^l \text{ 是 } \boldsymbol{\chi}_i \text{ 中的一个元素} \\ 0, & \text{否则} \end{cases}$$

式 (9) 即为通过 NL 方法将蕴含在相同尺度相似图像块中的附加信息加入到重构图像的 SR 数学模型.

图像金字塔是将低分辨率图像进行不同尺度的缩放, 从而得到一系列具有不同分辨率的图像. 利用图像金字塔的 K-SVD 字典学习方法可以将蕴含在相同尺度和不同尺度相似图像块中的附加信息加入到重构图像中. 考虑式 (9) 所示的 SR 数学模型, 由于在图像金字塔中体现了相同尺度和不同尺度相似图像块, 本文将利用图像金字塔作为样本进行字典学习, 从而建立 SR 数学模型中的字典 Ψ . 与利用图像库的 SR 方法相比, 本文用于字典学习的样本取自于待处理低分辨率图像自身的图像金字塔, 这种样本选取方法充分利用了图像自身所蕴含的信息, 从而解决了采用图像库而带来的问题. 对于字典学习样本的选取, CSSS 算法采用了图像金字塔的第 K 层, 即低分辨率图像的插值图像, 这实际上是仅利用了相同尺度相似图像块所提供的附加信息, 因此属于本文算法的一种特殊情况. 与 CSSS 算法的不同之处在于, 本文利用 NL 方法获取相同尺度相似图像块所蕴含的附加信息, 与此同时借助图像金字塔将不同尺度相似图像块所蕴含的附加信息也加入到图像重构过程中.

2.3 本文算法的整体流程

将式 (9) 中的第 1 项与第 3 项合并, 则式 (9) 所示的 SR 数学模型可简化表示为如下形式:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\tilde{\mathbf{Y}} - K\Psi \circ \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^p \sum_{j=1}^t \lambda |\alpha_{i,j}| \right\} \quad (10)$$

式中,

$$\tilde{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix}, \quad K = \begin{bmatrix} DH \\ \eta(I - B) \end{bmatrix}$$

本文利用迭代收缩算法 (Iterative shrinkage algorithm)^[12] 求解式 (10) 所示的凸优化问题. 本文算法的基本流程描述如下:

步骤 1. 设置图像的初始估计 $\hat{\mathbf{X}}^{(0)}$ 、迭代中止误差 ϵ 以及最大迭代次数 K_{\max} ;

步骤 2. 根据图像的降质过程确定降采样矩阵 D 和模糊矩阵 H ;

步骤 3. 建立图像金子塔并将其作为 K-SVD 字典学习方法的样本, 通过字典学习构建字典 Ψ ;

步骤 4. 通过搜索图像中相同尺度的相似图像块, 获得 NL 方法中的权值矩阵 B ;

步骤 5. 更新图像的当前估计:

$$\hat{\mathbf{X}}^{(k+\frac{1}{2})} = \hat{\mathbf{X}}^{(k)} + K^T \left(\tilde{\mathbf{Y}} - K\hat{\mathbf{X}}^{(k)} \right) = \hat{\mathbf{X}}^{(k)} + \left((DH)^T \mathbf{Y} - U\hat{\mathbf{X}}^{(k)} - V\hat{\mathbf{X}}^{(k)} \right) \quad (11)$$

其中,

$$U = (DH)^T DH, \quad V = \eta^2 (I - B)^T (I - B)$$

步骤 6. 更新稀疏表示系数:

$$\boldsymbol{\alpha}_i^{(k+\frac{1}{2})} = \Psi^T R_i \hat{\mathbf{X}}^{(k+\frac{1}{2})}, \quad i = 1, 2, \dots, p \quad (12)$$

$$\boldsymbol{\alpha}_i^{(k+1)} = \text{soft} \left(\boldsymbol{\alpha}_i^{(k+\frac{1}{2})}, \tau \right) \quad (13)$$

其中, p 为图像块的个数, $\text{soft}(x, \tau)$ 为阈值 τ 的软阈值函数,

$$\text{soft}(x, \tau) = \text{sgn}(x) \max(|x| - \tau, 0) \quad (14)$$

其中, $\text{sgn}(x)$ 表示符号函数;

步骤 7. 更新图像的当前估计 $\hat{\mathbf{X}}^{(k+1)} = \Psi \circ \boldsymbol{\alpha}^{(k+1)}$;

步骤 8. 重复步骤 4~7, 进行下一次迭代, 直至连续两次的迭代结果满足 $\|\hat{\mathbf{X}}^{(k)} - \hat{\mathbf{X}}^{(k-1)}\|_2^2/p \leq \epsilon$ 或 $k \geq K_{\max}$ 迭代终止.

3 实验结果与分析

本节将通过两组实验从定性评价和定量分析两个方面来验证本文算法的有效性, 第 1 组实验在无高分辨率参考图像的条件下, 对图像中具有显在结构自相似性时本文算法的重构效果进行定量评价; 第 2 组实验利用全参考和无参考图像质量评价方法, 对本文算法应用于具有潜在结构自相似性的图像上的重构效果进行定量评价. 本文采用两种方案选

取 K-SVD 字典学习所用样本: 一种是仅使用低分辨率图像的插值图像, 称之为基于单尺度结构自相似性的单幅图像超分辨率 (Single-scale structural self-similarity, sSSIM) 算法. 与 CSSS 算法相比, sSSIM 算法将 NL 方法加入到了 CSSS 算法中, 并用迭代收缩算法代替 OMP 方法构建图像块的稀疏表示系数. 另一种是使用图像金字塔, 称之为基于多尺度结构自相似性的单幅图像超分辨率 (Multi-scale structural self-similarity, mSSIM) 算法. 实验中使用图像金字塔中的第 0 层和第 K 层, 即低分辨率图像和低分辨率图像的插值图像. 与 CSSS 算法相比, mSSIM 算法不仅将 NL 方法加入到了图像重构过程中, 同时借助图像金字塔的 K-SVD 字典学习方法将蕴含在不同尺度相似图像块中的附加信息也加入到重构图像中.

3.1 显在结构自相似性图像的无参考质量评价

第 1 组实验在无高分辨率参考图像的条件下, 对图像中具有显在结构自相似性时本文算法的重构效果进行定量评价, 实验中本文算法相关参数的选取如下: 图像块的尺寸为 7×7 ; 在 K-SVD 字典学习方法中, 字典 Ψ 中的元素个数为 49; 在 NL 方法

中相似图像块个数 L 取值为 10; 迭代收缩算法的最大迭代次数 K_{\max} 设置为 120, 软阈值函数的参数 τ 取值为 0.08, 迭代中止的误差 ϵ 设置为 2×10^{-6} . 将图 1 所示的两幅图像进行 2 倍分辨率提升, 图 4 给出了这两幅图像的双三次插值、ASDSAR 算法¹以及 mSSIM 算法的重构图像. 通过比较可以看出, 本文提出的 mSSIM 算法的重构图像在边缘以及其他细节处均有更好的重构效果. 实验结果表明在图像具有显在多尺度结构自相似性时, 本文算法具有很好的重构效果.

由于图 1 中的两幅图像没有高分辨率参考图像, 为了定量分析 SR 方法的重构效果, 本文采用了盲图像质量评价指标 (Blind image quality indices, BIQI)^{[15]2}, 对无高分辨率参考图像时的重构效果进行定量评价, BIQI 指标的值域为 $[0, 100]$, 该值越小表示图像质量越高. 首先, 利用 BIQI 指标对图 4 中的重构图像质量进行定量评价, 对于第 1 幅图像, 双三次插值、ASDSAR 算法以及 mSSIM 算法重构图像的 BIQI 指标分别为 77.4131、60.4345 和 54.4032; 对于第 2 幅图像, 这三种算法重构图像的 BIQI 指标分别为 75.5511、64.7425 和 58.6860. 由此可见, 在 BIQI 指标下, 本文算法重构图像质量优

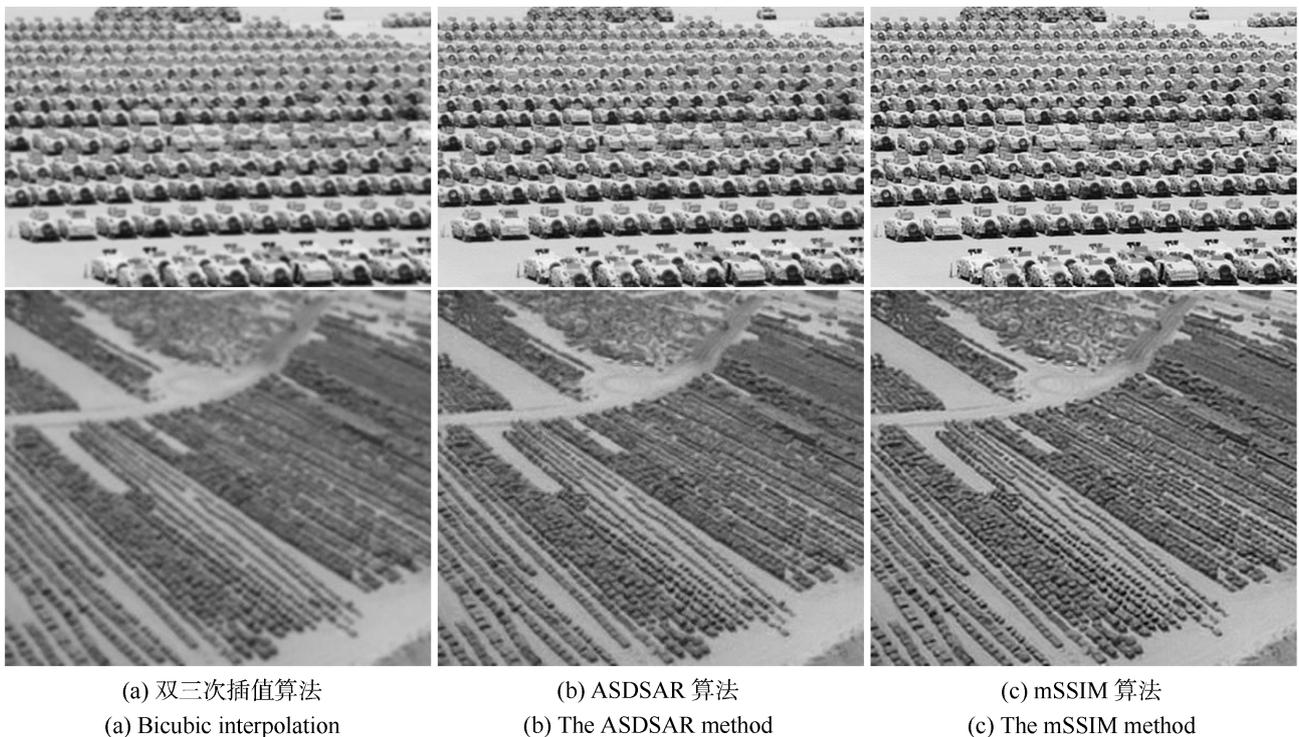


图 4 图 1 中两幅具有显在结构自相似性图像的不同算法重构效果比较

Fig. 4 Reconstructed images of Fig. 1 with explicit structural self-similarity using various SR methods

¹ASDSAR 算法的实现代码可参见 http://www4.comp.polyu.edu.hk/~cslzhang/ASDS_AReg.htm.

²BIQI 指标的实现代码可参见 <http://live.ece.utexas.edu/research/Quality/index.htm>.

于 ASDSAR 算法。同样采用 BIQI 指标, 本文进一步评价了本文算法在较大分辨率提升倍率下的重构效果。这里, 对多幅遥感图像分别进行 2、3、4、6、8、9 倍率的分辨率提升, 计算低分辨率图像与不同倍率分辨率提升后重构图像的 BIQI 指标, 并求取平均值。图 5 给出了平均 BIQI 指标与分辨率提升倍率之间的关系, 图中标记“+”、“×”和“*”的三条折线分别对应双三次插值、ASDSAR 算法以及 mSSIM 算法。从图中可以看出, 随着分辨率提升倍率的增加, 这三种算法重构图像的质量都有所下降, 但当分辨率提升倍率较大时双三次插值重构图像的质量下降十分明显, 而另两种算法在较大分辨率提升倍率的情况下重构图像仍然保持较好的质量; 与 ASDSAR 算法相比, mSSIM 算法重构图像质量下降速度更慢。实验表明本文算法优于 ASDSAR 算法, 并且在分辨率提升倍率较大时更为明显。

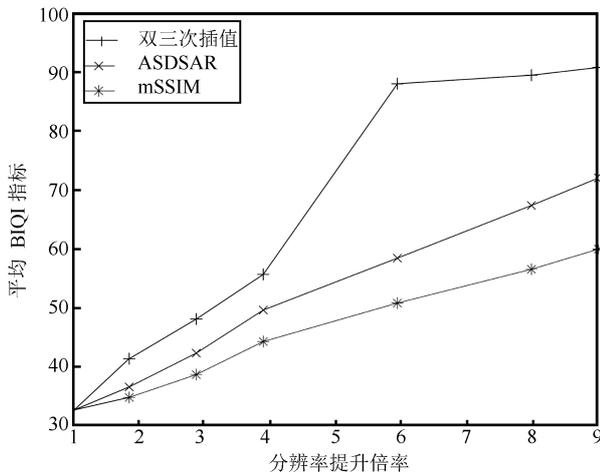


图 5 不同 SR 算法在不同分辨率提升倍率下的平均 BIQI 指标对比分析

Fig. 5 Average BIQI values of reconstructed images using different SR methods with various upscaling factors

3.2 潜在结构自相似性图像的全参考和无参考质量评价

第 2 组实验利用全参考和无参考图像质量评价方法, 对本文算法应用于具有潜在结构自相似性的图像上的重构效果进行定量评价。这里选用 SIPI 图像数据库中 10 幅尺寸为 512×512 的真实遥感图像作为高分辨率参考图像³。实验中将高分辨率参考图像进行高斯模糊以及降采样从而获得相应的低分辨率图像, 高斯模糊核的尺寸为 7×7 , 行和列的降采样因子均为 2, 分别利用双三次插值、CSSS 算法、ASDSAR 算法以及本文算法获得重构图像。实

验中相关参数的选取与第 3.1 节完全一致。

图 6 比较了其中一幅图像在各种 SR 算法下的重构效果。图 6(a) 给出了高分辨率参考图像。如图 6(b) 所示, 双三次插值图像边缘模糊且很多细节无法分辨, 尤其细节丰富区域, 例如下方的机场建筑、农田和右上角的小型建筑物。如图 6(c) 所示, CSSS 算法在一定程度上提高了边缘的锐度, 并增强了图像中的细节, 但是 CSSS 算法固有的问题限制了重构效果的改善。从图 6(d)~6(f) 中的重构图像可以看出, ASDSAR、sSSIM、mSSIM 算法均取得了较好的重构效果。仔细观察位于右上角的小型建筑物可以看出, 对于这样具有小尺度特征的目标, sSSIM 和 mSSIM 算法比 ASDSAR 算法恢复了更多的细节。本文利用峰值信噪比 (Peak signal to noise ratio, PSNR) 和结构相似性 (Structural similarity, SSIM) 对不同 SR 算法的重构图像进行了定量比较, 表 1 列出了不同 SR 算法的这两种图像质量评价指标的结果, 其中粗体标出的数值对应于相应图像质量评价指标下具有最优重构效果的算法。表 1 说明, 在全参考图像质量评价指标 PSNR 和 SSIM 下, sSSIM 与 mSSIM 算法均优于 ASDSAR 算法, 并且 mSSIM 算法是最优的。实验结果表明: 1) 对于具有潜在结构自相似性的遥感图像, 本文提出的 sSSIM 和 mSSIM 算法也取得了更好的重构效果; 2) mSSIM 算法借助图像金字塔的 K-SVD 字典学习方法能够将蕴含在多尺度结构自相似性中的附加信息加入到重构图像中。

本文利用模糊度量指标对各种 SR 算法的边缘保持性进行定量评价。无参考边缘模糊度 (No-reference edge-based blur measurement, NEBM) 通过估计边缘的宽度来度量图像的模糊程度^[16]。NEBM 指标越大, 边缘越模糊, 表明所对应算法具有较差的边缘保持能力。表 2 列出了双三次插值、CSSS 算法、ASDSAR 算法以及本文提出的 sSSIM 和 mSSIM 算法重构图像的 NEBM 指标, 表中粗体标出的数值表示所对应算法在 NEBM 指标下具有最优的重构效果。由表 2 中可见, 本文提出的 sSSIM 和 mSSIM 算法与 CSSS 算法和 ASDSAR 算法相比具有更强的边缘保持能力。

4 结论

本文提出了一种基于多尺度结构自相似性的单幅图像超分辨率算法, 该算法结合了压缩感知框架与图像结构自相似性, 利用了图像金字塔所提供的多尺度自相似结构, 并在压缩感知框架下通过 NL 方法和 K-SVD 字典学习方法将蕴含在多尺度结构自相似性中的附加信息加入到重构图像中。该算法充

³<http://sipi.usc.edu/database/database.php?volume=aerials>.

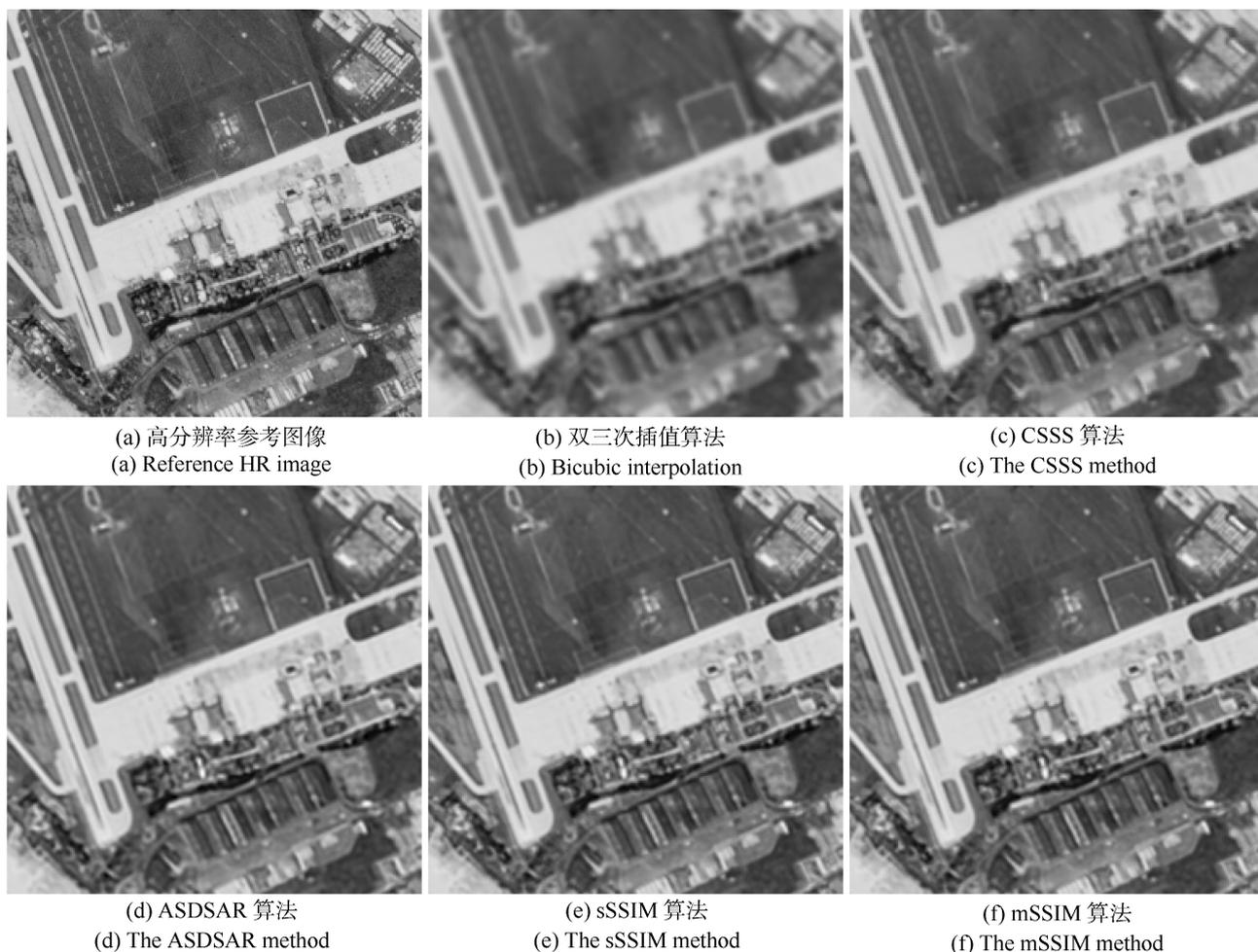


图 6 本文提出的 sSSIM 和 mSSIM 算法与双三次插值、CSSS 算法、ASDSAR 算法的重构效果比较
Fig. 6 Comparison of the proposed sSSIM and mSSIM methods with bicubic interpolation, CSSS, and ASDSAR

表 1 全参考图像质量评价指标 PSNR 和 SSIM 下不同 SR 算法重构效果的比较 (PSNR/SSIM)
Table 1 Comparison of PSNR and SSIM of reconstructed images with various SR methods (PSNR/SSIM)

图像序号	双三次插值	CSSS	ASDSAR	sSSIM	mSSIM
2.1.01	23.72/0.5404	24.59/0.6212	26.61/0.7583	26.63/0.7605	26.65/0.7616
2.1.02	20.90/0.4751	22.03/0.5373	23.83/0.7455	23.84/0.7480	23.86/0.7491
2.1.03	28.91/0.8097	30.04/0.8403	32.19/0.8929	32.22/0.8940	32.23/0.8942
2.1.04	24.25/0.4680	24.93/0.5247	26.91/0.7262	26.92/0.7286	26.94/0.7298
2.1.05	24.92/0.6937	26.19/0.7248	30.06/0.8742	30.07/0.8755	30.10/0.8763
2.1.06	24.44/0.5664	25.77/0.6149	28.19/0.8052	28.20/0.8070	28.23/0.8082
2.1.07	27.99/0.6982	28.34/0.7236	29.97/0.8039	30.01/0.8060	30.03/0.8067
2.1.08	31.33/0.7795	32.68/0.7933	36.74/0.9232	36.74/0.9235	36.80/0.9246
2.1.09	31.68/0.7906	33.12/0.8162	35.06/0.8706	35.11/0.8718	35.14/0.8727
2.1.10	26.57/0.6174	27.98/0.6852	30.25/0.8248	30.28/0.8266	30.30/0.8278
平均	26.47/0.6439	27.57/0.6882	29.98/0.8225	30.00/0.8242	30.03/0.8251

表 2 无参考边缘模糊度 NEBM 指标下不同 SR 算法重构图像的比较
Table 2 Comparison of NEBM of reconstructed images with various SR methods

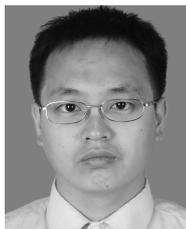
图像序号	双三次插值	CSSS	ASDSAR	sSSIM	mSSIM
2.1.01	20.5486	8.4253	6.5397	6.3086	6.3234
2.1.02	13.9858	6.7281	5.1894	5.0083	4.9873
2.1.03	21.8595	7.4930	5.5971	5.3940	5.3398
2.1.04	16.1872	7.2684	5.3541	5.2511	5.2069
2.1.05	18.2452	8.5107	6.8775	6.6547	6.6201
2.1.06	17.7791	7.0372	5.9246	5.7912	5.7508
2.1.07	20.7583	7.0147	5.7403	5.6281	5.5383
2.1.08	31.8421	9.8763	7.4329	7.2640	7.1046
2.1.09	26.3614	8.0241	6.7781	6.7257	6.3655
2.1.10	21.6806	7.1324	5.7883	5.6247	5.5662
平均	20.9248	7.7510	6.1222	5.9650	5.8803

分利用了图像自身的多尺度结构自相似性, 在重构过程中仅使用待处理低分辨率图像而没有使用任何其他高分辨率图像. 多种图像质量评价指标下多种超分辨率算法的实验结果表明, 本文算法具有更好的重构效果以及更大的分辨率提升能力.

References

- Aly H A, Dubois E. Image up-sampling using total-variation regularization with a new observation model. *IEEE Transactions on Image Processing*, 2005, **14**(10): 1647–1659
- Babacan S D, Molina R, Katsaggelos A K. Total variation super resolution using a variational approach. In: Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, USA: IEEE, 2008. 641–644
- Sen P, Darabi S. Compressive image super-resolution. In: Proceedings of the 43rd Asilomar Conference on Signals, Systems and Computers. Pacific Grove, USA: IEEE, 2009. 1235–1242
- Yang J C, Wright J, Huang T S, Ma Y. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861–2873
- Yang S Y, Sun F H, Wang M, Liu Z Z, Jiao L C. Novel super resolution restoration of remote sensing images based on compressive sensing and example patches-aided dictionary learning. In: Proceedings of the 2011 International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping. Xiamen, China: IEEE, 2011. 1–6
- Protter M, Elad M, Takeda H, Milanfar P. Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*, 2009, **18**(1): 36–51
- Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. Non-local sparse models for image restoration. In: Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 2272–2279
- Suetake N, Sakano M, Uchino E. Image super-resolution based on local self-similarity. *Optical Review*, 2008, **15**(1): 26–30
- Glasner D, Bagon S, Irani M. Super-resolution from a single image. In: Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 349–356
- Dong W S, Zhang L, Shi G M, Wu X L. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Transactions on Image Processing*, 2011, **20**(7): 1838–1857
- Pan Zong-Xu, Huang Hui-Juan, Yu Jing, Hu Shao-Xing, Zhang Ai-Wu, Ma Hong-Bing, Sun Wei-Dong. Super-resolution method based on CS and structural self-similarity for remote sensing images. *Signal Processing*, 2012, **28**(6): 859–872
(潘宗序, 黄慧娟, 禹晶, 胡少兴, 张爱武, 马洪兵, 孙卫东. 基于压缩感知与结构自相似性的遥感图像超分辨率方法. 信号处理, 2012, **28**(6): 859–872)
- Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 2004, **57**(11): 1413–1457
- Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322
- Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 2006, **15**(12): 3736–3745
- Moorthy A K, Bovik A C. A two-step framework for constructing blind image quality indices. *IEEE Signal Processing Letters*, 2010, **17**(5): 513–516
- Marziliano P, Dufaux F, Winkler S, Ebrahimi T. A no-reference perceptual blur metric. In: Proceedings of the 2002

International Conference on Image Processing, Rochester, USA: IEEE, 2002. III-57–III-60



潘宗序 清华大学电子工程系博士研究生. 2010 年获哈尔滨工业大学电子与信息工程学院学士学位. 主要研究方向为图像处理, 模式识别, 压缩感知与稀疏表示, 空间信息处理与应用. 本文通信作者. E-mail: panzxtsinghua@163.com
(**PAN Zong-Xu** Ph. D. candidate in the Department of Electronic Engineering,

Tsinghua University. He received his bachelor degree from Harbin Institute of Technology in 2010. His research interest covers image processing, pattern recognition, compressive sensing and sparse representation, and spatial information processing and application. Corresponding author of this paper.)



禹晶 清华大学电子工程系博士后. 2011 年获清华大学电子工程系博士学位. 主要研究方向为图像处理与模式识别. E-mail: yujing@tsinghua.edu.cn
(**YU Jing** Postdoctoral researcher in the Department of Electronic Engineering, Tsinghua University. She received her Ph. D. degree from Tsinghua Uni-

versity in 2011. Her research interest covers image processing and pattern recognition.)



胡少兴 北京航空航天大学机械工程与自动化学院副教授. 主要研究方向为三维激光扫描技术, 图像处理, 计算机视觉与模式识别. E-mail: husx@buaa.edu.cn
(**HU Shao-Xing** Associate professor at the School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics.

His research interest covers technology of 3-D laser scanning, image processing, computer vision, and pattern recognition.)



孙卫东 清华大学电子工程系教授. 主要研究方向为图像处理, 模式识别, 空间信息处理与应用.

E-mail: wdsun@tsinghua.edu.cn

(**SUN Wei-Dong** Professor in the Department of Electronic Engineering, Tsinghua University. His research interest covers image processing, pattern

recognition, and spatial information processing and application.)