

# 基于自适应多字典学习的单幅图像超分辨率算法

潘宗序<sup>1</sup>, 禹 晶<sup>1</sup>, 肖创柏<sup>2</sup>, 孙卫东<sup>1</sup>

(1. 清华大学电子工程系, 北京 100084; 2. 北京工业大学计算机学院, 北京 100124)

**摘 要:** 自适应字典学习利用图像结构自相似性, 将图像自身作为训练样本, 通过字典学习使图像中的相似块在字典下具有稀疏表示形式. 本文将全局字典学习中利用图像库获取附加信息思想融入到自适应字典学习的过程中, 提出了一种基于自适应多字典学习的单幅图像超分辨率算法, 从低分辨率图像自身与图像库同时获取附加信息. 该算法对低分辨率图像金字塔结构中的图像块进行聚类, 在聚类结果的引导下将图像库中的图像块进行分类, 利用各类中的样本分别构建针对各类的多个字典, 从而确定表达重建图像块的最优字典. 实验表明, 与 ScSR、SISR、NLBP、CSSS 以及 mSSIM 等算法相比, 本文算法具有更好的超分重建效果.

**关键词:** 超分辨率; 稀疏表示; 自适应字典学习; 全局字典学习

**中图分类号:** TN911.73      **文献标识码:** A      **文章编号:** 0372-2112 (2015)02-0209-08

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2015.02.001

## Single Image Super Resolution Based on Adaptive Multi-Dictionary Learning

PAN Zong-xu<sup>1</sup>, YU Jing<sup>1</sup>, XIAO Chuang-bai<sup>2</sup>, SUN Wei-dong<sup>1</sup>

(1. Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;

2. College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China)

**Abstract:** Adaptive dictionary learning uses the low resolution image itself as training samples to make the similar patches have sparse representation over the learned dictionary, so that extra information can be exploited from structural self-similarity by dictionary learning. In this paper, we propose a single image super resolution method based on adaptive multi-dictionary learning. To exploit extra information from both the low resolution image itself, and the image database, the proposed method incorporates the idea of global dictionary learning that the image database can be used to obtain extra information into the process of adaptive dictionary learning. In the proposed method, all patches in the image pyramid of the low resolution image are clustered into several groups, then each patch satisfying a certain condition in the database is classified into one of these groups with the supervision of the clustering results, and multi-dictionary learning is used to learn corresponding dictionaries for different groups. Experimental results demonstrate that our method achieves better result compared with ScSR, SISR, NLBP, CSSS and mSSIM methods.

**Key words:** super resolution; sparse representation; adaptive dictionary learning; global dictionary learning

## 1 引言

在数字图像的采集过程中, 成像设备和拍摄条件会制约所获得图像的空间分辨率, 然而在众多领域中都更需要更高空间分辨率的图像. 在遥感领域, 高分辨率图像的获取有利于地物目标的检测与识别, 但受硬件制造水平、传感器成本、成像平台以及成像环境等多方面因素的影响, 遥感图像的空间分辨率通常难以满足实际需求, 这使得如何提升遥感图像的空间分辨率成为一个亟待解决的问题. 超分辨率 (Super resolution, SR) 方法是利用信号处理手段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 (High resolution, HR) 图像到低分辨率 (Low resolution, LR) 图像的降质过程, 并利用 LR 图像重建 HR 图像. SR 方法的关键是如何有效获取某种类别的额外附加信息来弥补降质过程中损失的细节信息. 根据所采用图像数目的不同, SR 方法可以分为多幅图像 SR 方法和单幅图像 SR 方法. 多幅图像 SR 方法将多幅 LR 图像所提供的互补信息作为附加信息<sup>[1,2]</sup>, 单幅图像 SR 方法通过先验知识或样本学习获取附加信息. 在遥感领域, 由于同一时相、同一区域多幅图像的获取十分困难, 因此单幅图像 SR 方法更具实用性.

基于稀疏表示的 SR 方法利用图像在特定字典下具有稀疏表示这一特性,通过字典学习获取附加信息,是近年来广泛认可的单幅图像 SR 方法.这类方法的关键是确定能够稀疏表示图像块的字典. Yang 等提出了一种基于稀疏表示的超分辨率 (Super-resolution via sparse representation, ScSR) 算法<sup>[3]</sup>,该算法将大量高分辨率图像组成的图像库作为训练样本,这种采用图像库作为训练样本构建字典的过程称为全局字典学习.图像库包含大量具有丰富细节的高分辨率图像,全局字典学习能够获取充足的附加信息,然而由于无法保证所获得附加信息的准确性和可靠性,因此重建图像虽然具有较好的视觉效果,但一般存在较大的均方误差. ScSR 算法的关键是构建能够稀疏表示图像块的全局字典,然而由于图像内容的多样性,构建全局字典需要大量训练样本,这不仅导致字典学习的低效率以及字典构建的不稳定性,而且当图像库不能提供与待处理图像相似的图像块时,将无法保证算法的重建效果.这些问题产生的根源是这种算法试图获取能够稀疏表示所有图像块的全局字典,然而这种全局字典实际上是不存在的.

自适应字典学习利用了图像的结构自相似性,采用 LR 图像自身作为训练样本构建自适应字典.图像结构自相似性是指图像中的大量物体具有相同尺度以及不同尺度相似结构的特性,体现为大量存在的相同尺度以及不同尺度的相似图像块. Protter 等提出的非局部方法利用相同尺度相似图像块构建非局部约束项,通过正则化约束实现超重建<sup>[4]</sup>. Dong 等结合非局部方法与迭代反投影方法,提出了非局部迭代反投影 (Non-local iterative back-projection, NLBP) 算法<sup>[5]</sup>,该算法利用非局部方法构建初始重建图像,并将非局部约束加入到误差反向投影过程中. Glasner 等同时利用图像中相同尺度以及不同尺度的相似图像块,提出了不依赖于其他任何高分辨率图像的单幅图像超分辨率 (Single image super resolution, SISR) 算法<sup>[6]</sup>,该算法将 LR 图像进行金字塔分解,通过在金字塔结构的不同层图像中搜索相似图像块获取多尺度自相似结构中的附加信息.上述算法需要在图像中搜索具有相似结构的图像块,运算复杂度高,缺乏将图像自相似结构中蕴含的附加信息加入到重建图像中的有力工具.自适应字典学习将 LR 图像自身作为训练样本,通过字典学习有效地将蕴含在结构自相似性中的附加信息加入到字典中,并利用自适应字典重建 HR 图像.文献[7]提出了一种基于压缩感知与结构自相似性的单幅图像超分辨率 (Compressive sensing and structural self-similarity based method, CSSS) 算法,该算法结合稀疏表示与图像结构自相似性,将初始重建图像作为字典学习的样本,通过自适应

字典学习将蕴含在相同尺度相似图像块中的附加信息加入到字典中.文献[8]对 CSSS 算法进行了补充和完善,提出了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.该算法对 LR 图像进行金字塔分解并将图像金字塔作为字典学习的样本,通过自适应字典学习将蕴含在相同尺度以及不同尺度相似图像块中的附加信息加入到字典中,同时算法利用相同尺度相似图像块构建非局部约束项,进一步提升了重建图像的质量. CSSS 和 mSSIM 算法利用 LR 图像自身构建自适应字典,解决了全局字典所存在的普遍性问题,然而由于这种算法在重建过程中没有利用其他任何图像,附加信息的获取程度存在一定的局限性.

本文将全局字典学习中利用图像库获取附加信息的思想融入到自适应字典学习的过程中,提出了一种基于自适应多字典学习的单幅图像超分辨率 (Adaptive multi-dictionary learning, AMDL) 算法,从 LR 图像自身以及图像库同时获取附加信息.该算法首先对 LR 图像金字塔结构中的图像块进行聚类,在聚类结果的引导下将图像库中的图像块进行分类,然后再利用各类中的样本分别构建针对各类的多个字典,从而确定表达重建图像块的最优字典.与基于全局字典学习的 SR 方法相比, AMDL 算法将训练样本进行聚类并构建多个字典,这使得字典表达误差降低,并且在每个字典的构建过程中所用的样本数量大幅度减少,字典学习效率得以提高.与基于自适应字典学习的 SR 方法相比, AMDL 算法有效地利用了图像库所提供的附加信息,并通过多字典学习的形式将附加信息加入到多个字典中.

## 2 稀疏表示与字典学习

图像通常在特定字典下具有稀疏表示形式,对于图像块  $x \in R^n$ , 如果存在矩阵  $\Psi$  以及向量  $\alpha$ , 使得:

$$x = \Psi\alpha, \|\alpha\|_0 \ll n \quad (1)$$

则称  $x$  在  $\Psi$  下具有稀疏表示形式,其中  $\Psi = [\psi_1, \dots, \psi_t] \in R^n \times t$  称为字典,字典中的每一列  $\psi_i, i = 1, \dots, t$  称为字典的元素,  $\alpha = [\alpha_1, \dots, \alpha_t]^T \in R^t$  为稀疏表示系数,满足  $\|\alpha\|_0 \ll n$ , 这里  $\|\alpha\|_0$  是指  $\alpha$  中非零元素的个数.

字典学习是指在给定一组样本的条件下,通过机器学习方法构建能够稀疏表示样本的字典.设  $S = [x_1, \dots, x_s]$  为样本组成的矩阵,其中  $s$  为样本个数,  $A = [\alpha_1, \dots, \alpha_s]$  为稀疏表示系数组成的矩阵,  $\alpha_i$  为  $x_i$  对应的稀疏表示系数,字典学习可以表示为:

$$\min_{\Psi, A} \{\|S - \Psi A\|_F^2\}, \text{ s.t. } \forall i \|\alpha_i\|_0 \leq T \quad (2)$$

其中,  $T$  为控制  $\alpha_i$  稀疏程度的参数.常用的字典学习方法

法有 MOD 算法<sup>[9]</sup>和 K-SVD 算法<sup>[10]</sup>等.

根据构建字典时样本选取方式的不同,字典学习方法可以分为全局字典学习和自适应字典学习方法.全局字典学习方法将图像库作为训练样本,通过构建全局字典将蕴含在图像库中的附加信息加入到字典中,尽管这种方法能够获取大量的附加信息,然而对于待处理 LR 图像而言,无法保证从图像库获取的附加信息的针对性和可靠性.自适应字典学习方法将 LR 图像自身作为训练样本,通过构建自适应字典将蕴含在图像多尺度自相似结构中的附加信息加入到字典中,然而由于没有利用其他任何图像,附加信息的获取程度存在一定的局限性.

根据构建字典数目的不同,字典学习方法又可以分为单字典学习和多字典学习方法,单字典学习方法利用样本构建单个字典,多字典学习方法构建多个字典.多字典学习方法将训练样本  $x_i, i = 1, \dots, s$  分成  $K$  类,每一类对应一个字典  $\Psi_k, k = 1, \dots, K$ . 设第  $k$  类样本组成样本矩阵  $S_k = [x_1^k, \dots, x_{s_k}^k]$ ,  $k = 1, \dots, K$ , 其中  $s_k$  为第  $k$  类样本的个数,满足  $\sum_{k=1}^K s_k = s$ ,  $A_k = [\alpha_1^k, \dots, \alpha_{s_k}^k]$ ,  $k = 1, \dots, K$  为稀疏表示系数组成的矩阵,其中  $\alpha_i^k, i = 1, \dots, s_k$  为  $x_i^k$  在字典  $\Psi_k$  下的稀疏表示系数.多字典学习方法通过求解下式所示的最优化问题构建字典  $\Psi_k, k = 1, \dots, K$ :

$$\min_{\Psi_k, A_k} \{ \|S_k - \Psi_k A_k\|_F^2 \}, \text{ s.t. } \|\alpha_i^k\|_0 \leq T_k, i = 1, \dots, s_k \quad (3)$$

其中,  $T_k$  为控制  $\alpha_i^k$  稀疏程度的参数.多字典学习方法将样本进行聚类,并在此基础上利用各类中的样本分别构建针对各类的多个字典,与单字典学习方法相比,多字典学习方法在构建每个字典时所用的训练样本数量更少,并且聚类策略的引入使得训练样本的相似程度更高,各字典表示样本的误差更低.

### 3 基于自适应多字典学习的超分辨率算法

#### 3.1 基于字典学习的超分重建模型

设  $X \in R^N$  表示 HR 图像,  $Y \in R^M$  表示 LR 图像,  $\hat{X} \in R^N$  表示重建图像,可以用如下观测模型来描述 HR 图像到 LR 图像的降质过程,

$$Y = DHX + v \quad (4)$$

其中,  $D$  为降采样矩阵,  $H$  为模糊矩阵,  $v$  为加性噪声.式(4)所示的观测模型说明 LR 图像是由 HR 图像经过模糊、降采样以及加入噪声等过程获取的.为了将图像的稀疏性加入到重建模型中,需要对图像进行分块,设  $x_i \in R^n, i = 1, \dots, p$  表示 HR 图像块,  $\hat{x}_i \in R^n, i = 1, \dots, p$  表示重建图像块,图像块与图像之间的关系为  $x_i = R_i X$ , 其中  $R_i \in R^{n \times N}$  为抽取矩阵.利用多字典学习方法

构建  $K$  个字典  $\Psi_k, k = 1, \dots, K$ , 设  $\Psi_{k_i}, k_i \in \{1, \dots, K\}$  为表达重建图像块  $\hat{x}_i$  的字典,即  $\hat{x}_i = \Psi_{k_i} \alpha_i$ , 其中  $\alpha_i$  为  $\hat{x}_i$  的稀疏表示系数.可以通过对所有重建图像块进行平均处理获得重建图像<sup>[11]</sup>,为了便于书写,引入符号“ $\circ$ ”,则有:

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

将式(5)代入式(4),并加入对  $\alpha_i$  的稀疏性约束项,则有:

$$\hat{\alpha} = \arg \min_{\alpha} \{ \|Y - DH\Psi\alpha\|_2^2 + \lambda \|\alpha\|_{l_m} \} \quad (6)$$

其中,  $\lambda$  为控制图像稀疏性在代价函数中所占比重的参数.式(6)即为基于字典学习的超分重建模型,在式(6)所示最优化问题中,  $m = 0$  或  $m = 1$ , 当  $m = 0$  时式(6)所示的  $l_0$  范数最小化问题是一个非凸的 NP 难问题,当  $m = 1$  时式(6)所示的  $l_1$  范数最小化问题是一个可以求得精确解的凸优化问题.由于当  $\alpha$  足够稀疏时,  $m = 0$  与  $m = 1$  是等价的<sup>[12]</sup>,因此本文采用  $l_1$  范数.

为了进一步保证重建图像的准确性,可以在式(6)所示 SR 模型中加入正则化约束项.本文利用相同尺度相似图像块构建非局部约束项<sup>[4]</sup>,对任意重建图像块  $\hat{x}_i$ , 在初始重建图像  $\hat{X}$  中搜索与  $\hat{x}_i$  相似的图像块  $\hat{x}_i^l, l = 1, \dots, L$ , 计算  $\hat{x}_i$  与  $\hat{x}_i^l$  之间的距离  $e_i^l = \|\hat{x}_i - \hat{x}_i^l\|_2^2$ , 并利用  $\hat{x}_i^l$  的线性组合估计  $\hat{x}_i$ . 设  $\chi_i$  和  $\chi_i^l$  分别为  $\hat{x}_i$  和  $\hat{x}_i^l$  的中心像素灰度值,则  $\hat{\chi}_i = \sum_{l=1}^L \omega_i^l \chi_i^l$ , 其中权值  $\omega_i^l = \exp(-e_i^l/h) / \sum_{l=1}^L \exp(-e_i^l/h)$ ,  $h$  为控制常数,因此可将  $\chi_i$  与  $\hat{\chi}_i$  的差值作为非局部约束项.更进一步的,设  $\omega_i$  表示  $\omega_i^l, l = 1, \dots, L$  组成的向量,  $\chi_i$  表示  $\chi_i^l, l = 1, \dots, L$  组成的向量,则非局部约束项可以写成  $\sum_{\chi_i \in \hat{X}} \|\chi_i - \hat{\chi}_i\|_2^2 = \sum_{\chi_i \in \hat{X}} \|\chi_i - \omega_i^T \chi_i\|_2^2 = \|(I - B)\hat{X}\|_2^2 = \|(I - B)\Psi \circ \alpha\|_2^2$ , 将非局部约束项加入式(6),则有:

$$\hat{\alpha} = \arg \min_{\alpha} \{ \|Y - DH\Psi\alpha\|_2^2 + \lambda \|\alpha\|_1 + \eta \|(I - B)\Psi \circ \alpha\|_2^2 \} \quad (7)$$

其中,  $\eta$  为控制非局部约束项在代价函数中所占比重的参数,  $I$  为单位矩阵,  $B$  为权值矩阵.

#### 3.2 自适应多字典学习

全局字典学习将图像库作为训练样本构建式(7)中的字典  $\Psi$ , 自适应字典学习将 LR 图像自身作为训练样本构建式(7)中的字典  $\Psi$ . 本文结合这两种方法各自的优势,提出一种自适应多字典学习方法,同时利用 LR 图像自身和图像库构建式(7)中的字典  $\Psi$ . 图像金字塔是指将图像进行不同尺度的缩放而形成一系列不同分辨率的图像,在图像金字塔中存在大量相同尺度以及不同尺度的相似图像块,这些多尺度自相似结构为超

分重建提供了必要的附加信息.图1直观地说明了多尺度自相似结构在图像金字塔中的体现,其中  $I_0$  表示 LR 图像,  $I_k$  表示 HR 图像,通过插值可获得  $I_0$  上层的图像,通过模糊以及降采样可获得  $I_0$  下层的图像.

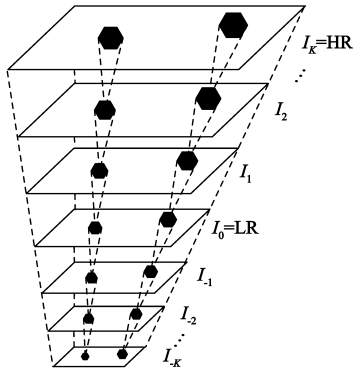


图1 多尺度自相似结构在图像金字塔中的体现

图像金字塔中存在大量多尺度自相似结构,自适应字典学习方法将图像金字塔作为训练样本,通过字典学习的方式使样本中的多尺度相似图像块在字典下具有稀疏表示形式,从而将图像多尺度自相似结构所蕴含的额外附加信息加入到字典中.本文提出的自适应多字典学习算法将全局字典学习中利用图像库获取附加信息的思想融入到自适应字典学习的过程中.自适应多字典学习算法的过程为,首先对 LR 图像进行金字塔分解,将金字塔中的图像块进行聚类并确定类中心,然后通过计算图像块与各类中心的距离将图像库中的图像块加入到相应的类别中,再利用各类中的样本分别构建针对各类的多个字典,最后确定表达重建图像块的字典,进而实现超分重建,以下将对上述步骤分别加以阐述.

### (1) 训练样本聚类

将 LR 图像进行金字塔分解,对金字塔中的图像块进行聚类并确定类中心.在聚类过程中,为了更好地反映图像块的细节信息,本文利用高通滤波提取图像块的高频成分,并仅对高频成分进行聚类<sup>[13]</sup>.设  $x_i, i = 1, \dots, q$  为图像金字塔中的图像块,  $z_i, i = 1, \dots, q$  为  $x_i$  的高频成分,利用 K-means 算法对  $z_i$  进行聚类,产生  $K$  个类别  $C_k = [z_1^k, \dots, z_{q_k}^k], k = 1, \dots, K$ , 其中  $C_k$  为第  $k$  类,  $z_i^k, i = 1, \dots, q_k$  为  $C_k$  中的第  $i$  个样本,  $q_k$  为  $C_k$  中样本的个数,满足  $\sum_{k=1}^K q_k = q$ . 计算  $C_k$  的类中心  $\mu_k = \frac{1}{q_k} \sum_{i=1}^{q_k} z_i^k$  以及半径  $r_k = \max \|z_i^k - \mu_k\|_2, i = 1, \dots, q_k$ .

### (2) 训练样本扩充

通过计算图像块与类中心的距离将图像库中的图像块进行分类.在该过程中,首先计算图像库中图像块的方差并去掉方差较小的图像块,这样做的目的是去

除一些平滑的区域.设  $\tilde{x}_i, i = 1, \dots, s$  为图像库中的图像块,  $\tilde{z}_i, i = 1, \dots, s$  为  $\tilde{x}_i$  的高频成分.计算  $\tilde{z}_i$  与类中心  $\mu_k$  的距离  $d_i^k$ , 设  $d_i^k, k = 1, \dots, K$  的最小值为  $d_i^{k_i}, k_i \in \{1, \dots, K\}$ , 如果  $d_i^{k_i} < \epsilon r_{k_i}$ , 则将  $\tilde{z}_i$  加入到  $C_{k_i}$  中, 否则舍去该图像块, 其中  $\epsilon$  用来控制图像库中图像块与类中心的相似程度.经样本扩充后的类别为  $C_k = [z_1^k, \dots, z_{s_k}^k], k = 1, \dots, K$ , 其中  $s_k$  为扩充后  $C_k$  中的样本个数.

### (3) 多字典学习

分别利用各类中的样本进行多字典学习,从而构建针对各类的多个字典.  $C_k$  是图像块的高频成分所对应的样本矩阵,根据  $x_i$  与  $z_i$  间的对应关系确定图像块所对应的样本矩阵  $S_k = [x_1^k, \dots, x_{s_k}^k], k = 1, \dots, K$ , 其中  $x_i^k$  为  $z_i^k$  对应的图像块,利用式(3)所示的多字典学习方法构建  $S_k$  所对应的字典  $\Psi_k, k = 1, \dots, K$ .

### (4) 多字典稀疏表示

确定表达重建图像块的字典.本文将 LR 图像的插值图像作为初始重建图像  $\hat{X}$ , 对于重建图像块  $\hat{x}_i$ , 设  $\hat{z}_i$  为  $\hat{x}_i$  的高频成分,  $i = 1, \dots, p$ . 根据  $k_i = \arg \min_k \|\hat{z}_i - \mu_k\|_2$  在  $K$  个字典  $\Psi_k, k = 1, \dots, K$  中确定表达  $\hat{x}_i$  的字典  $\Psi_{k_i}, i = 1, \dots, p$ , 该式表明若  $\hat{x}_i$  的高频成分  $\hat{z}_i$  与聚类  $C_{k_i}$  的中心  $\mu_{k_i}$  的距离最小, 则  $\hat{x}_i$  属于类别  $S_{k_i}$ , 因此将  $\Psi_{k_i}$  作为表达  $\hat{x}_i$  的字典.

将通过自适应多字典学习获得的字典  $\Psi = [\Psi_{k_1}, \dots, \Psi_{k_p}]$  应用到式(7)中, 就建立了 AMDL 算法的最终数学模型. 本文通过迭代收缩算法<sup>[14]</sup>求解式(7)所示的最优化问题, 将式(7)的解代入式(5), 可得到重建图像  $\hat{X}$ .

## 4 实验结果与分析

为了验证本文算法的有效性, 本文采用 USC\_SIFI 图像数据库中的 10 幅遥感图像以及其他 2 幅遥感图像进行实验. 在实验中为了便于对重建图像质量进行定量评价, 将原图像作为高分辨率参考图像, 通过高斯模糊以及降采样获取待处理的低分辨率图像, 高斯模糊核的尺寸为  $7 \times 7$ , 行和列方向上的降采样因子为 2. 分别采用 ScSR、SISR、NLBP、CSSS、mSSIM 以及本文的 AMDL 算法对待处理低分辨率图像进行重建, 并对各算法下的重建图像进行对比分析. 实验中相关参数的设置如下: 图像库包含 5 幅与待处理图像内容相近的大幅面高分辨率遥感图像, 从而可获取大量有针对性的额外附加信息, 在样本训练过程中将图像库中的图像进行分块, 图像块之间具有一定的重叠, 从而可获取充足的样本; 图像块的尺寸为  $6 \times 6$ ; 字典  $\Psi_k, k = 1, \dots, K$  中元素的个数为 36; 非局部约束项中, 相似图像块个数  $L$  取值为 12; 迭代收缩算法中的最大迭代次数为 960.

图 2 比较了一幅体育场图像在不同 SR 算法下的重建效果,图 2(a)为高分辨率参考图像,图 2(b)为待处理低分辨率图像.如图 2(c)所示,双三次插值图像边缘模糊,视觉质量很差.图 2(d)为 ScSR 算法的重建图像,ScSR 算法恢复了部分细节,例如体育场的边缘,然而重建图像仍然较为模糊.图 2(e)为 SISR 算法的重建图像,尽管 SISR 算法的锐化效果明显,然而引入了很多虚假细节.图 2(f)为 NLIBP 算法的重建图像,与 ScSR 和 SISR 算法相比,NLIBP 算法重建图像更加清晰,例如右上方的建筑物,且没有引入过多的虚假细节.图 2(g)为

CSSS 算法的重建图像,与 ScSR、SISR 以及 NLIBP 算法相比,CSSS 算法具有更好的重建效果,提高了边缘的锐度,增强了图像中的细节,例如右侧建筑物的边缘.图 2(h)为 mSSIM 算法的重建图像,与 CSSS 算法相比,mSSIM 算法恢复了更多细节,例如体育场的边缘和右侧建筑物的屋顶.图 2(i)为 AMDL 算法的重建图像,与前几种 SR 算法相比,AMDL 算法取得了最好的重建效果,恢复了更多的细节,重建图像更加清晰,例如体育场和右侧建筑物的边缘、右上方的建筑物等.

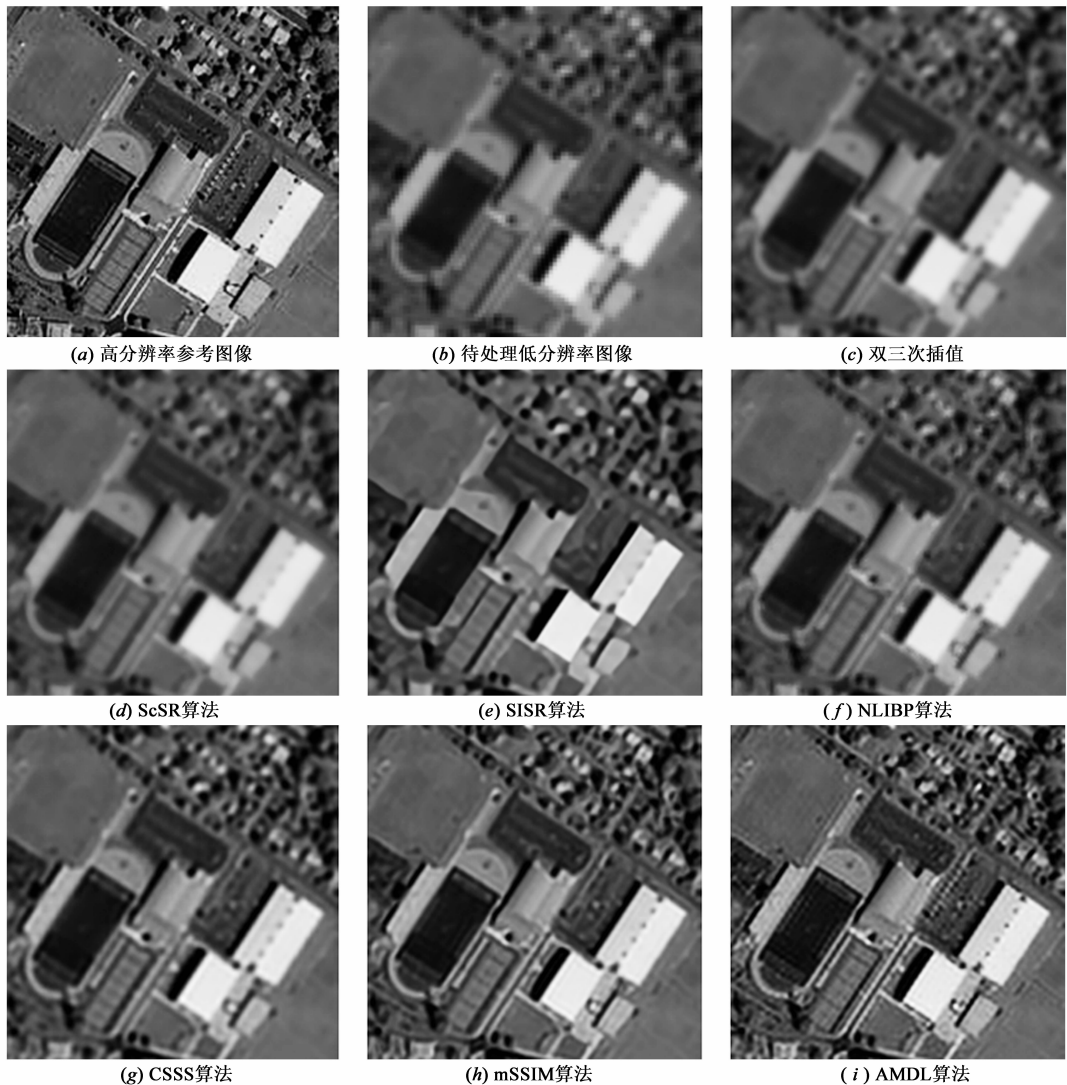


图 2 体育场图像不同SR算法重建效果比较

图 3 比较了一幅市区图像在不同 SR 算法下的重建效果.图 3(a)~图 3(i)分别为高分辨率参考图像、待处理低分辨率图像、双三次插值、ScSR、SISR、NLIBP、CSSS、mSSIM 以及 AMDL 算法的重建图像.如图 3(d)所示,ScSR 算法重建图像较为模糊.如图 3(e)所示,SISR

算法引入了很多虚假细节.与 ScSR 和 SISR 算法相比,NLIBP 算法重建图像更加清晰并且没有引入过多的虚假细节,如图 3(f)所示.从图 3(g)中明显可见,与前几种算法相比 CSSS 算法重建图像更加清晰,如左侧道路和建筑物的边缘,然而由于仅利用了图像中相同尺度

相似图像块所蕴含的附加信息, CSSS 算法重建效果具有一定的局限性. 而 mSSIM 算法同时利用了图像中相同尺度以及不同尺度相似图像块所蕴含的附加信息, 因此恢复了更多细节, 如图 3(h) 所示. 图 3(i) 为 AMDL 算法的重建图像, 与 CSSS 和 mSSIM 算法相比, AMDL 算法取得了更好的超分重建效果并且没有引入明显的虚假细节, 重建图像的边缘更加清晰, 例如左上方的建筑

物以及右侧的车辆. 与仅采用 LR 图像自身作为训练样本的 CSSS 和 mSSIM 算法相比, AMDL 算法同时将 LR 图像自身和图像库作为训练样本, 因此能够获取更多的附加信息. 实验表明, 通过自适应多字典学习, 本文的 AMDL 算法能够更加有效地获取 LR 图像自身以及图像库所提供的额外附加信息.

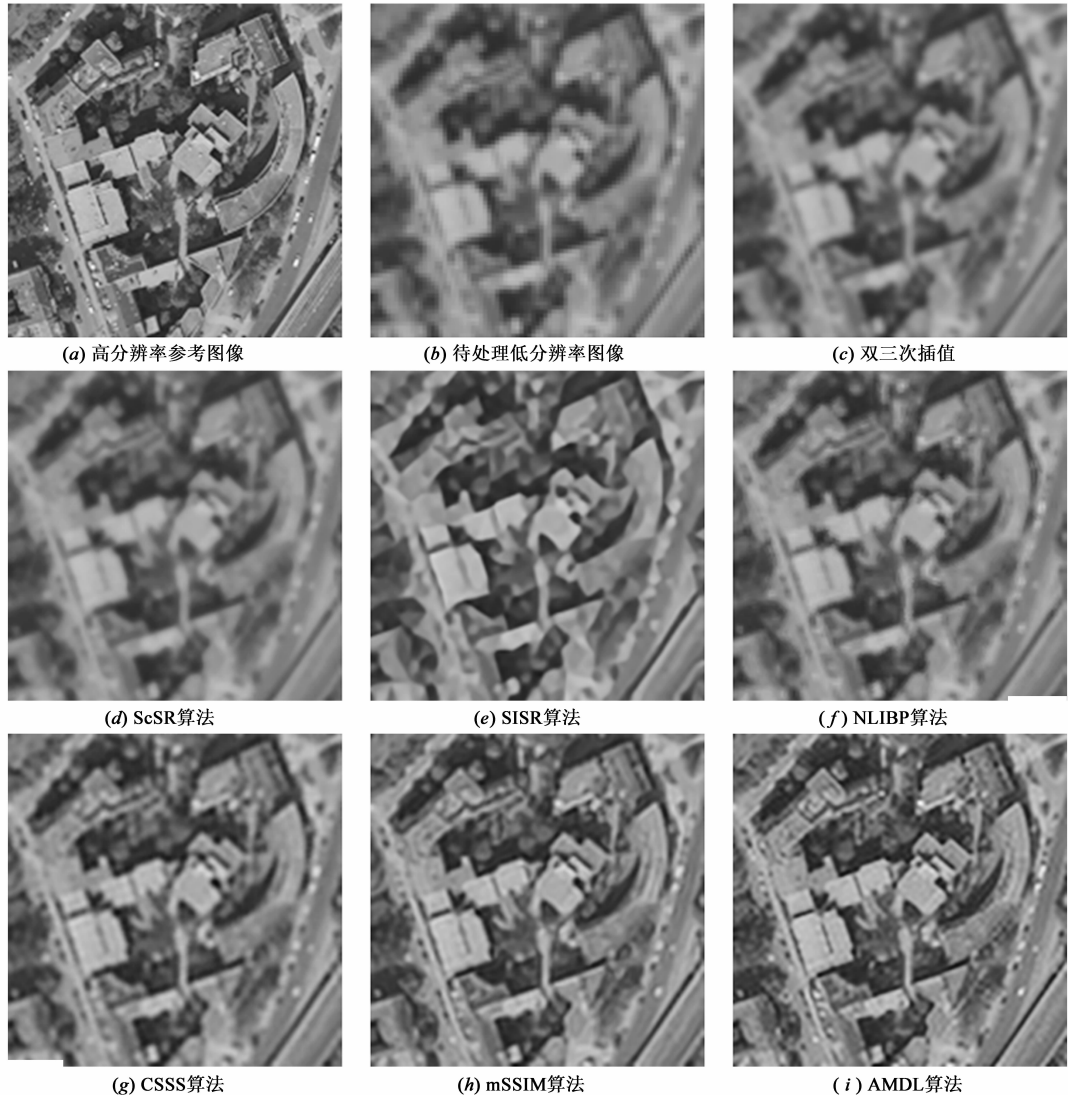


图3 市区图像不同SR算法重建效果比较

本文采用峰值信噪比(Peak signal to noise ratio, PSNR)和结构相似性(Structural similarity, SSIM)对不同 SR 算法的重建图像质量进行定量评价. PSNR 和 SSIM 是两种常用的图像质量评价指标, PSNR 通过计算参考图像与重建图像之间的均方误差来度量重建图像的质量, SSIM 通过计算参考图像与重建图像之间的结构差异来度量重建图像的质量, PSNR 和 SSIM 指标越大表示重建图像的质量越高. 表 1 列出了不同 SR 算法下重建图像的 PSNR 和

SSIM 指标, 其中粗体标出的数值在相应图像质量评价指标下最优. 可以看出, 与 ScSR、SISR、NLBP、CSSS 和 mSSIM 算法相比, 本文的 AMDL 算法在两个评价指标下均具有最好的超分重建效果, 且平均 PSNR 的改善幅度分别为 4.40、4.94、3.27、2.24 和 1.10dB. 对于实验中采用的 12 幅真实遥感图像, ScSR、SISR、NLBP、CSSS、mSSIM 和本文 AMDL 算法的平均运行时间分别为 4293.24s、13086.27s、35.21s、1304.67s、1484.59s 和 10122.90s. 尽管本文算法的

表 1 PSNR 和 SSIM 下不同 SR 算法重建效果比较 (PSNR/SSIM)

序号	ScSR	SISR	NLIBP	CSSS	mSSIM	AMDL
2.1.01	23.96/0.5646	23.87/0.6043	24.91/0.6354	25.74/0.6978	26.65/0.7616	27.53/0.8118
2.1.02	21.14/0.5044	20.97/0.5506	22.04/0.5506	22.83/0.6681	23.86/0.7491	24.82/0.8080
2.1.03	29.16/0.8177	28.74/0.8274	30.19/0.8423	31.16/0.8663	32.23/0.8942	33.14/0.9131
2.1.04	24.47/0.4958	24.14/0.5367	25.33/0.5774	26.09/0.6530	26.94/0.7298	27.83/0.7909
2.1.05	25.25/0.7141	24.84/0.7389	26.96/0.7770	28.50/0.8244	30.10/0.8763	31.71/0.9114
2.1.06	24.70/0.5927	24.49/0.6302	25.93/0.6725	27.03/0.7411	28.23/0.8082	29.36/0.8574
2.1.07	28.16/0.7084	27.84/0.7194	28.82/0.7387	29.27/0.7646	30.03/0.8067	30.78/0.8439
2.1.08	31.68/0.7963	31.79/0.8201	33.34/0.8468	35.28/0.8922	36.80/0.9246	37.97/0.9425
2.1.09	31.94/0.7986	32.05/0.8087	33.00/0.8218	33.97/0.8418	35.14/0.8727	36.14/0.8950
2.1.10	26.85/0.6400	26.81/0.6799	28.00/0.7084	29.04/0.7650	30.30/0.8278	31.66/0.8745
体育场	19.93/0.5350	17.81/0.4466	20.97/0.6206	21.83/0.6957	22.88/0.7668	24.06/0.8283
市区	22.49/0.5915	19.86/0.4801	23.74/0.6777	24.85/0.7483	26.13/0.8158	27.49/0.8683
均值	25.81/0.6466	25.27/0.6536	26.94/0.7058	27.97/0.7632	29.11/0.8195	30.21/0.8621

运行时间仅低于 SISR 算法而高于其他五种算法,然而,本文算法显著提升了重建图像的质量。

## 5 结论

本文提出了一种基于自适应多字典学习的单幅图像超分辨率算法,该算法同时利用了低分辨率图像自身和图像库所提供的额外附加信息.自适应多字典学习对低分辨率图像金字塔中的图像块进行聚类,借助聚类结果将图像库中的图像块进行分类,并利用各类中的样本分别构建针对各类的多个字典.通过自适应多字典学习,该算法在充分利用图像结构自相似性的同时将图像库所提供的附加信息有效加入到多个字典中.与多种超分辨率算法的定性和定量的对比实验表明,本文算法具有更好的超分重建效果和更高的分辨率提升能力。

## 参考文献

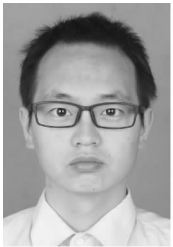
- [1] 邵文泽, 韦志辉. 基于各向异性 MRF 建模的多帧图像变分超分辨率重建[J]. 电子学报, 2009, 37(6): 1256 - 1263.  
Shao Wenzhe, Wei Zhihui. Multi-frame super-resolution reconstruction based on anisotropic Markov random field modeling [J]. Acta Electronica Sinica, 2009, 37(6): 1256 - 1263. (in Chinese)
- [2] 宋锐, 吴成柯, 封颖, 张云锋. 一种新的基于 MAP 的纹理自适应超分辨率图像复原算法[J]. 电子学报, 2009, 37(5): 1124 - 1129.

Song Rui, Wu Chengke, Feng Ying, Zhang Yunfeng. A new MAP based texture adaptive super-resolution image reconstruction algorithm[J]. Acta Electronica Sinica, 2009, 37(5): 1124 - 1129. (in Chinese)

- [3] Yang J, Wright J, Huang T S, Ma Y. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861 - 2873.
- [4] Protter M, Elad M, Takeda H, Milanfar P. Generalizing the nonlocal-means to super-resolution reconstruction [J]. IEEE Transactions on Image Processing, 2009, 18(1): 36 - 51.
- [5] Dong W, Zhang L, Shi G, Wu X. Nonlocal back-projection for adaptive image enlargement[A]. Proceedings of the 2009 IEEE International Conference on Image Processing [C]. Cairo, Egypt: IEEE Computer Society, 2009. 349 - 352.
- [6] Glasner D, Bagon S, Irani M. Super-resolution from a single image[A]. Proceedings of the 12th International Conference on Computer Vision [C]. Kyoto, Japan: IEEE Inc, 2009. 349 - 356.
- [7] Pan Z, Yu J, Huang H, Hu S, Zhang A, Ma H, Sun W. Super-resolution based on compressive sensing and structural self-similarity for remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(9): 4864 - 4876.
- [8] 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[J]. 自动化学报, 2014, 40(4): 594 - 603.  
Pan Zongxu, Yu Jing, Hu Shaoxing, Sun Weidong. Single image

- super resolution based on multi-scale structural self-similarity [J]. Acta Automatica Sinica, 2014, 40(4): 594 – 603. (in Chinese)
- [9] Engan K, Aase S O, Husoy J H. Method of optimal directions for frame design [A]. Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing [C]. Phoenix, AZ, USA: IEEE, 1999. 2443 – 2446.
- [10] Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311 – 4322.
- [11] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries [J]. IEEE Transactions on Image Processing, 2006, 15(12): 3736 – 3745.
- [12] Gribonval R, Nielsen M. Sparse representations in unions of bases [J]. IEEE Transactions on Information Theory, 2003, 49(12): 3320–3325.
- [13] Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution [J]. IEEE Computer Graphics and Applications, 2002, 22(2): 56 – 65.
- [14] 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, 2004, 57(11): 1413 – 1457.

### 作者简介



**潘宗序** 男, 1986 年生于黑龙江省. 2010 年获哈尔滨工业大学电子与信息工程学院学士学位. 现为清华大学电子工程系博士研究生. 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.  
Email: panzxtsinghua@163.com



**肖创柏** 男, 1962 年生于湖南省. 1995 年毕业于清华大学, 获博士学位. 现为北京工业大学计算机学院教授. 主要研究方向为数字信号处理, 音视频信号处理与网络通信, 模式识别等.



**禹晶** 女, 1981 年生于陕西省. 2011 年获清华大学电子工程系博士学位. 现为清华大学电子工程系博士后. 主要研究方向为图像处理和模式识别.



**孙卫东** 男, 1960 年生于山东省. 1988 年获日本东京大学博士学位. 现为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等.