

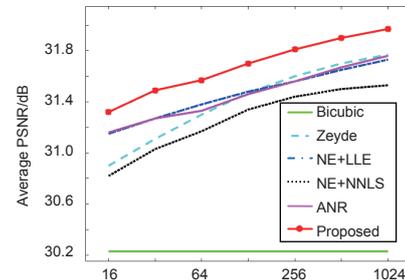


DOI: 10.12086/oe.2018.170537

## 基于聚类和协同表示的超分辨率重建

汪荣贵, 刘雷雷, 杨娟, 薛丽霞\*, 胡敏

合肥工业大学计算机与信息学院, 安徽 合肥 230009



**摘要:** 图像超分辨率重建是利用单幅或多幅降质的低分辨率图像重建得到高分辨率图像, 以提高图像的视觉效果并获得更多可用的信息。本文提出结合图像特征聚类和协同表示的超分辨率重建方法。在训练阶段根据图像的特征信息对图像样本进行聚类并利用图像特征的差异性训练不同的字典, 克服了传统训练单个字典方法对图像特征表示不足的缺点。而且利用协同表示方法求得不同聚类的高、低分辨率图像样本之间的映射矩阵, 提高了图像重建速度。实验表明, 本文方法与其他方法相比, 不仅提高了重建图像的 PSNR 和 SSIM 指标, 而且改善了视觉效果。

**关键词:** 超分辨率重建; 聚类; 协同表示; 映射矩阵

中图分类号: TN919.81

文献标志码: A

引用格式: 汪荣贵, 刘雷雷, 杨娟, 等. 基于聚类和协同表示的超分辨率重建[J]. 光电工程, 2018, 45(4): 170537

## Image super-resolution based on clustering and collaborative representation

Wang Ronggui, Liu Leilei, Yang Juan, Xue Lixia\*, Hu Min

School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China

**Abstract:** Image super-resolution (SR) refers to the reconstruction of a high-resolution (HR) image from single or multiple observed degraded low-resolution (LR) images for the purpose of improving image's visual effects and getting more available information. We propose an image super-resolution algorithm based on collaborative representation and clustering in this paper. In the training stage, the image samples are clustered according to the image features and multiple dictionaries are trained by using the differences of image features, which overcomes the shortcoming of lack of expressiveness of traditional single-dictionary training methods. Moreover, projection matrices between different HR and LR image clustering are computed via collaborative representation, which accelerate the speed of image reconstruction. Experiments demonstrate that compared with other methods, the proposed method not only enhanced PSNR and SSIM metrics for reconstructed images but also improved image's visual effects.

**Keywords:** image super-resolution; clustering; collaborative representation; projection matrices

**Citation:** Wang R G, Liu L L, Yang J, et al. Image super-resolution based on clustering and collaborative representation[J]. *Opto-Electronic Engineering*, 2018, 45(4): 170537

收稿日期: 2017-10-10; 收到修改稿日期: 2018-02-05

基金项目: 国家自然科学基金项目(61672202)

作者简介: 汪荣贵(1966-), 男, 教授, 博士生导师, 主要研究方向为智能视频处理与分析、视频大数据与云计算、智能视频监控与公共安全、嵌入式多媒体技术等领域。E-mail: wangrgui@hfut.edu.cn

通信作者: 薛丽霞(1976-), 女, 副教授, 主要从事数字图像处理、智慧城市、地理信息系统和公共安全等研究。E-mail: xlxxzm@163.com

# 1 引言

超分辨率重建<sup>[1-3]</sup>是图像处理和计算机视觉领域的经典问题之一,它最初是由 Tsai 和 Huang<sup>[4]</sup>提出的,根据重建方法的不同主要可分为基于插值的方法、基于重建的方法和基于学习的方法。

近年来基于学习<sup>[5-6]</sup>的超分辨率重建逐渐成为主流的图像重建方法。Freeman<sup>[7]</sup>等人提出一种基于实例的学习策略,通过利用 Markov 网络对图像块的空间关系进行建模,使得重建得到的图像具有更多的高频细节信息;Chang<sup>[8]</sup>根据流形学习中的局部线性嵌入 (locally linear embedding, LLE)<sup>[9]</sup>技术提出邻域嵌入方法;Yang<sup>[10]</sup>将稀疏编码理论引入超分辨率重建,提出 ScSR 算法;Zeyde<sup>[11]</sup>在 Yang 的基础上加入 K-SVD<sup>[12]</sup>和正交匹配追踪用于字典训练和稀疏系数的求解,加快了高、低分辨率图像字典的训练过程;Yang<sup>[13]</sup>提出 SF(simple function)方法,利用聚类算法将图像特征空间分为多个子空间,然后利用稀疏表示方法对不同聚类图像进行重建;Zhang<sup>[14]</sup>证明了协同表示相对于稀疏表示在人脸识别中的主要作用,协同表示相比较于稀疏表示其求解速度更快,而且不易产生过拟合现象;Timofte<sup>[15]</sup>提出 ANR(anchored neighborhood regression)方法,利用协同表示的方法获得高、低分辨率图像块

之间的映射矩阵,在重建阶段只需通过低分辨率图像块和对应字典原子的映射矩阵相乘即可重建高分辨率图像块,提高了图像重建的速度。

本文提出结合聚类算法和协同表示的图像超分辨率重建方法,主要流程如图 1 所示,该方法首先对训练样本特征进行聚类,将具有相同特征的图像块聚为一类,然后对每个子类学习得到一个完备字典,再对每个子类字典原子依据协同表示的方法得到一个映射矩阵。重建时根据输入的低分辨率图像块,找出其所属类的图像字典,在该类字典中找出与待重建图像块特征向量最近的字典原子,根据该字典原子的映射矩阵和待重建图像块特征向量重建出高分辨率图像块。利用特征聚类方法增强了图像字典的特征表达能力,提高了映射矩阵的准确性,并且利用协同表示离线式计算出图像特征映射矩阵,提高了重建速度。

## 2 相关工作

### 2.1 图像退化模型

图像退化模型表明了高分辨率图像块  $x$  与观测得到的低分辨率图像块  $y$  之间的关系,其数学模型如式 (1)所示:

$$y = DHx + n \quad (1)$$

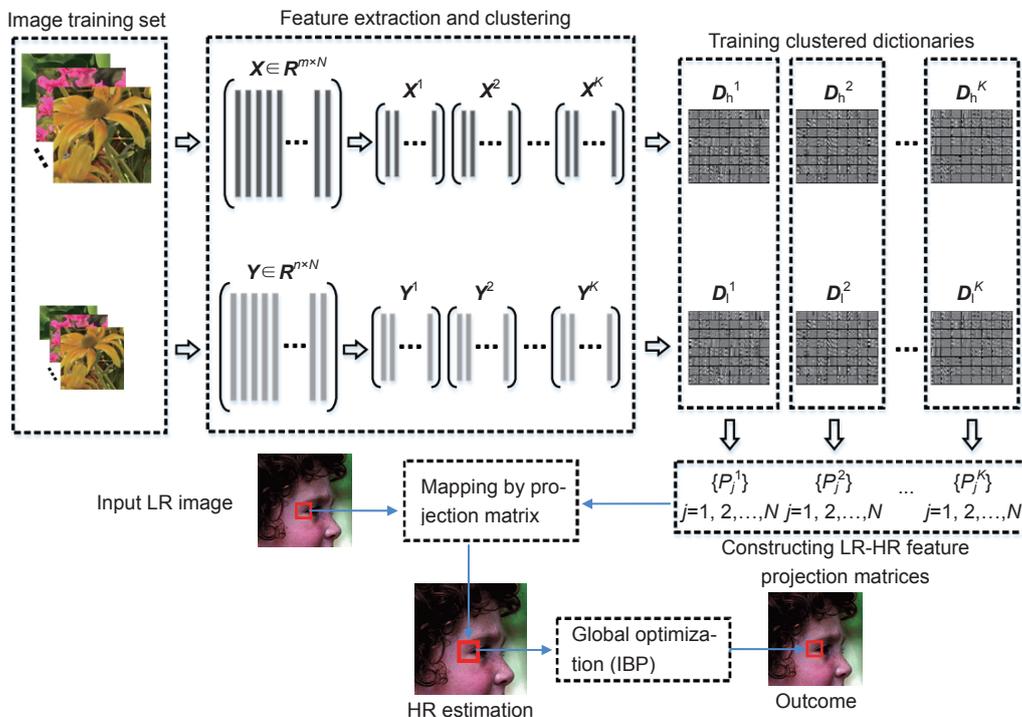


图 1 本文超分辨率重建算法框架图

Fig. 1 Overview of the proposed algorithm

其中： $H$  表示模糊处理， $D$  表示降采样处理， $n$  表示低分辨率图像观测过程中产生的噪声。

### 2.2 协同表示

稀疏表示方法利用  $L_0$  范数或者  $L_1$  范数正则项来约束目标函数，但是  $L_0$  范数是 NP 难问题，而带  $L_1$  范数正则项的最优化问题求解过程十分耗时。协同表示<sup>[14]</sup>则是带  $L_2$  范数正则项的最小二乘回归问题，虽然  $L_2$  范数的稀疏性约束弱于  $L_1$  范数，但是通过岭回归的方法可以很方便地对其进行求解。利用协同表示的方法，低分辨率图像块  $y$  与低分辨率图像完备字典  $D_l$  之间的约束关系如(2)所示：

$$\alpha = \arg \min_{\alpha} \|y - D_l \alpha\|_2^2 + \lambda \|\alpha\|_2^2, \quad (2)$$

其中： $\alpha$  是稀疏系数， $\lambda$  是正则化参数，用于确保求解过程的稳定性。通过岭回归的方法，可以得到式(3)的解为

$$\alpha = (D_l^T D_l + \lambda I)^{-1} D_l^T y, \quad (3)$$

其中  $I$  表示单位矩阵。由于认为高分辨率图像块与其下采样得到的低分辨率图像块之间具有相同的稀疏系数，故与  $y$  相对应的高分辨率图像块  $x$  可以通过式(4)得到：

$$x = D_h \alpha. \quad (4)$$

其中  $D_h$  表示高分辨率图像字典。

结合式(3)和式(4)，可以得到：

$$x = D_h (D_l^T D_l + \lambda I)^{-1} D_l^T y. \quad (5)$$

由式(5)可知，低分辨率图像块  $y$  和与相应的高分辨率图像块  $x$  之间的关系可以通过一个映射矩阵来表示，即  $P = D_h (D_l^T D_l + \lambda I)^{-1} D_l^T$ 。

## 3 基于聚类和协同表示的超分辨率重建

### 3.1 特征提取

记  $X$  表示高分辨率图像，对高分辨率图像  $X$  先下采样  $s$  倍，再采用双三次插值的方法上采样  $s$  倍得到低分辨率图像  $Y_l$ 。对  $Y_l$  进行分块可以得到  $Y_l = \{y_l^1, y_l^2, \dots, y_l^n\}$ ， $n$  表示图像块的总个数。对每一个低分辨率图像块使用梯度算子和拉普拉斯算子提取图像块水平和垂直方向上的边缘特征，特征提取算子为：

$$f_1 = [-1, 0, 1], f_2 = f_1^T, \\ f_3 = [1, 0, -2, 0, 1], f_4 = f_3^T.$$

其中“ $T$ ”表示转置操作。将每一个低分辨率图像块提取得到的四个特征向量组合为一个高维列向量，作

为该图像块特征表示，将所有低分辨率图像块特征表示组合成一个高维特征矩阵。为了提高算法的效率和鲁棒性，使用 PCA 对特征矩阵进行降维，降维后的图像特征用于训练图像字典。

对高分辨率图像首先通过式(6)获得其高频成分  $X_h$ ，然后对其进行分块得到  $X_h = \{x_h^1, x_h^2, \dots, x_h^n\}$ 。利用图像高频成分求得高分辨率图像字典，使得图像的重建过程转化为重建图像高频成分的过程，最后将重建得到的高频成分与原低分辨率图像相加得到最终的图像重建结果。

$$X_h = X - Y_l. \quad (6)$$

### 3.2 聚类字典学习

对低分辨率图像训练集中的所有图像分块后，根据图像块特征使用 k-means 聚类方法将其分为  $K$  个类别，使得具有相似特征的图像块聚为一类，根据每个低分辨率图像块所在的类别将相应的高分辨率图像块也分为  $K$  个类别。之后对每一类低分辨率图像块分别训练图像字典，由于同一类别的图像块均有相似的特征，因此训练得到的图像字典对于该类别的图像特征的表达更为精确，从而提高了具有该类别特征的图像的重建效果，图 2 所示为对图像块进行聚类后的结果，可以看出每一类的图像块均有较为相似的图像特征。

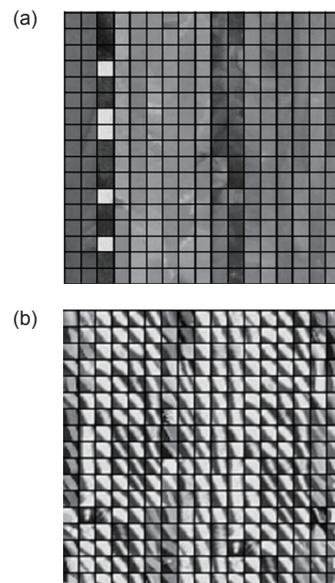


图 2 k-means 聚类后的两类图像块。(a) 类别一；(b) 类别二

Fig. 2 Two categories obtained by k-means clustering. (a) Category one; (b) Category two

低分辨率图像字典的训练使用 K-SVD 算法，每个

字典训练时的字典原子个数和迭代次数相同。记第  $i$  个低分辨率图像类别为  $Y^i = \{y_1^i, y_2^i, \dots, y_p^i\}$ , ( $1 \leq i \leq K$ ), 其中  $p$  表示该类别包含的图像块个数,  $D_l^i$  表示该类别的低分辨率图像字典。聚类字典的训练过程主要分为稀疏表示阶段和字典更新阶段, 在稀疏表示阶段假定稀疏字典  $D_l^i$  是固定的, 通过正交匹配追踪算法(OMP)来估计稀疏表示; 在字典更新阶段通过奇异值分解(singular value decomposition, SVD)的方法对稀疏字典  $D_l^i$  的每个列向量进行逐个修正, 目标公式如(7)所示, 其中  $A^i = \{\alpha_1^i, \alpha_2^i, \dots, \alpha_p^i\}$  为稀疏系数矩阵,  $T_0$  为样本稀疏性最大值。

$$\min_{\{D^i, A^i\}} \|Y^i - D_l^i A^i\|_F^2 \text{ s.t. } \forall q \|\alpha_q^i\| \leq T_0 \quad (7)$$

具体字典学习过程如下:

1) 随机选择  $D_{l(0)}^i$  初始化低分辨率字典。

2) 稀疏表示阶段: 固定字典  $D_{l(0)}^i$ , 采用正交匹配追踪算法(OMP)对训练集  $Y^i = \{y_1^i, y_2^i, \dots, y_p^i\}$  中的每个图像块  $y_p^i$  计算其稀疏系数  $\alpha_q^i$ , 最后将所有的稀疏系数组合得到稀疏系数矩阵  $A^i$ , 求解稀疏系数目标公式如(8)所示:

$$\alpha_q^i = \arg \min_{\alpha_q^i} \|y_p^i - D_l^i \alpha_q^i\|_F^2 \text{ s.t. } \|\alpha_q^i\|_0 \leq T_0 \quad (8)$$

其中  $q=1, 2, \dots, p$ 。

3) 字典更新阶段: 利用上一步得到的稀疏系数矩阵  $A^i$ , 逐列当前字典进行更新优化, 设需要更新当前字典的第  $k$  列  $d_k$ ,  $\alpha_T^k$  为系数矩阵  $A^i$  中第  $k$  行。定义稀疏表示误差矩阵  $E_k$ , 如式(9)所示,  $E_k$  表示去掉原子  $d_k$  的成分对  $p$  个图像样本表示所造成的误差。

$$E_k = Y^i - \sum_{j \neq k} d_j \alpha_j^i, j=1, 2, \dots, p \quad (9)$$

由于误差矩阵  $E_k$  的稀疏性, 在对  $E_k$  进行奇异值分解之前可以去除其中的非零元素以加快字典更新过程。设集合  $\omega_k = \{i | 1 \leq i \leq p, \alpha_T^k(i) \neq 0\}$  由训练集中使用到原子  $d_k$  的所有样本  $y_k^i$  的上标构成, 则可得到  $\alpha_T^k$  中非零点的索引值。根据  $\omega_k$  定义  $\Omega_k$  为  $n \times |\omega_k|$  维的 0-1 矩阵, 求出  $\alpha_R^k = \alpha_T^k \Omega_k$ ,  $E_R^k = E_k \Omega_k$ , 使得  $\alpha_T^k$  和  $E_k$  只保留非零项, 对  $E_R^k$  使用 SVD 得  $E_R^k = U \Delta V^T$ , 为奇异值构成的对角阵, 根据奇异值最大的项来更新字典原子  $d_k$  和  $\alpha_R^k$ 。

对于  $K$  个低分辨率图像类别, 通过不断迭代 2) 和 3) 两步直至达到终止条件即可得到所有的低分辨率图像字典  $\{D_l^i\}_{i=1}^K$ 。为了加快训练过程, 我们不使用 K-SVD 方法训练高分辨率图像字典  $\{D_h^i\}_{i=1}^K$ 。高分辨率图像块  $x_q^i$  的恢复可以表示为  $x_q^i \approx D_h^i \alpha_q^i$ , 因此可以通

过最小化式(10)来计算  $D_h^i$ 。

$$\begin{aligned} D_h^i &= \arg \min_{D_h^i} \sum_q \|x_q^i - D_h^i \alpha_q^i\|_2^2 \\ &= \arg \min_{D_h^i} \|X^i - D_h^i A^i\|_F^2 \end{aligned} \quad (10)$$

其中  $X^i = \{x_1^i, x_2^i, \dots, x_p^i\}$ , 通过式(11)可以得到高分辨率图像字典

$$D_h^i = X^i A^{i+} = X^i A^{i^T} (A^i A^{i^T})^{-1} \quad (11)$$

其中  $A^{i+}$  表示稀疏矩阵  $A^i$  的伪逆矩阵, 对  $K$  个高分辨率图像类别利用上述公式可以得到相应的高分辨率图像字典  $\{D_h^i\}_{i=1}^K$ 。

### 3.3 计算字典原子映射矩阵

由协同表示可知, 低分辨率图像块  $y$  和其对应的高分辨率图像块  $x$  之间的关系可通过一个近似映射矩阵  $P = D_h(D_l^T D_l + \lambda I)^{-1} D_l^T$  来表示, 即  $x = Py$ , 其中  $D_h$  和  $D_l$  分别表示高、低分辨率图像字典。如果对于不同特征的图像块都使用整个字典来求映射矩阵, 则会使得映射矩阵不具有针对性, 而且无法充分利用待重建图像块的特征。本文对低分辨率图像字典的每个原子使用它的  $n$  近邻字典原子来求映射矩阵, 并根据该字典原子的映射矩阵对图像块进行重建, 故对于字典原子  $d$  的映射矩阵表达式改写成:

$$P = N_h(N_l^T N_l + \lambda I)^{-1} N_l^T \quad (12)$$

其中:  $N_l$  表示  $d$  在低分辨率图像字典中的  $n$  个最近邻原子,  $N_h$  表示与  $N_l$  相对应的高分辨率图像字典中的  $n$  个字典原子。

根据以上分析, 在训练学习阶段需要计算出所有聚类字典中每个字典原子的映射矩阵, 对于一个字典原子  $d$ , 使用欧拉距离在该字典中找出它的  $n$  个最近邻字典原子  $N_l$ , 同时可以在高分辨率字典得到相应的  $N_h$ , 根据式(12)即可得到字典原子  $d$  的映射矩阵  $P$ 。

在求  $n$  最近邻字典原子过程中, 由于已经是聚类过后学习得到的字典, 每一类字典都是由特征最为相似的图像块学习得到, 因此通过聚类字典原子计算的映射矩阵更准确, 同时使用映射矩阵也加快了图像的重建过程。

### 3.4 迭代反投影

为进一步减小重建误差, 使用迭代反投影方法(iterative back projection, IBP)<sup>[16]</sup>对高分辨率图像  $X_0$  进行优化, 其目标公式如式(13)所示:

$$X^* = \arg \min X_1 - X_0 \text{ s.t. } Y = DHX_1 \quad (13)$$

其中:  $X_1$  表示对低分辨图像  $Y$  进行插值放大的结果,

$X^*$ 表示迭代终止后得到的高分辨率图像。通过求解式(13)可以得到：

$$X_{t+1} = X_t + ((Y - DHX_t) \uparrow s)^* p \quad (14)$$

其中  $X_t$  表示  $t$  次迭代后得到的高分辨率图像， $\uparrow s$  表示对图像进行  $s$  倍放大， $p$  表示高斯滤波器， $*$ 表示卷积操作。

## 4 实验

本文实验在 Windows 10 操作系统(CPU 四核 3.30 GHz, 内存 8 GB), MATLAB R2013a 上实现。对测试图像进行 2 倍放大重建, 根据重建效果来验证本文方法的有效性, 并与双三次插值方法、Yang<sup>[10]</sup>方法、Zeyde<sup>[11]</sup>方法、NE+LLE<sup>[8]</sup>方法、NE+NNLS<sup>[17]</sup>方法、SF<sup>[13]</sup>方法和 ANR<sup>[15]</sup>方法进行比较。实验训练集中包括 91 幅彩色图像, 每幅图像首先转换为 YCBCR 颜色空间, 以 Y 通道作为训练的图像数据, 低分辨率图像数据则通过对训练集图像进行下采样得到。对高、低分辨率图像分块时提取  $6 \times 6$  的图像块, 相邻图像块之间有 2 个像素的重叠。为了提高重建效果, 对低分辨率图像提取  $3 \times 3$  的图像块, 相邻图像块之间有 2 个像素的重叠。对不同方法的重建效果采用峰值信噪比(peak signal to noise ratio, PSNR)和结构相似性度量

(structural similarity index measurement, SSIM)两个指标来评价。

### 4.1 图像重建质量

图像重建之前, 先对训练集图像进行裁剪, 使图像大小为图像块大小的倍数, 然后提取高、低分辨率图像块, 在实验中将提取的 315210 对图像块分为  $K=9$  类做实验。对每一类低分辨率图像块使用 K-SVD 学习得到低分辨率图像字典, 并求出高分辨率图像字典, 正则化参数  $\lambda=0.01$ , 测试集图片如图 3 所示。

表 1 是图像放大 2 倍时的实验结果, 由表 1 可知, 利用本文方法重建得到的图像的平均 PSNR 和平均 SSIM 指标相比其他方法更高, 其中 Proposed-1 表示本文方法不使用迭代反投影优化的重建效果, Proposed 表示使用迭代反投影优化的重建效果。相比于 ANR 方法本文方法利用了多聚类字典, 重建结果的平均 PSNR 高出 0.22 dB, 平均 SSIM 高出 0.0013。SF 方法对图像特征聚类后直接将每个类别的图像块作为图像字典, 相比于 SF 方法, 本文方法重建结果的平均 PSNR 高出 2.04 dB, 平均 SSIM 高出 0.0217, 表明本文方法聚类后训练得到图像字典优于 SF 方法。

表 2 分别展示了图像放大 2、3 和 4 倍时测试集的平均结果。由此可知, 当增大放大倍数时本文方法相

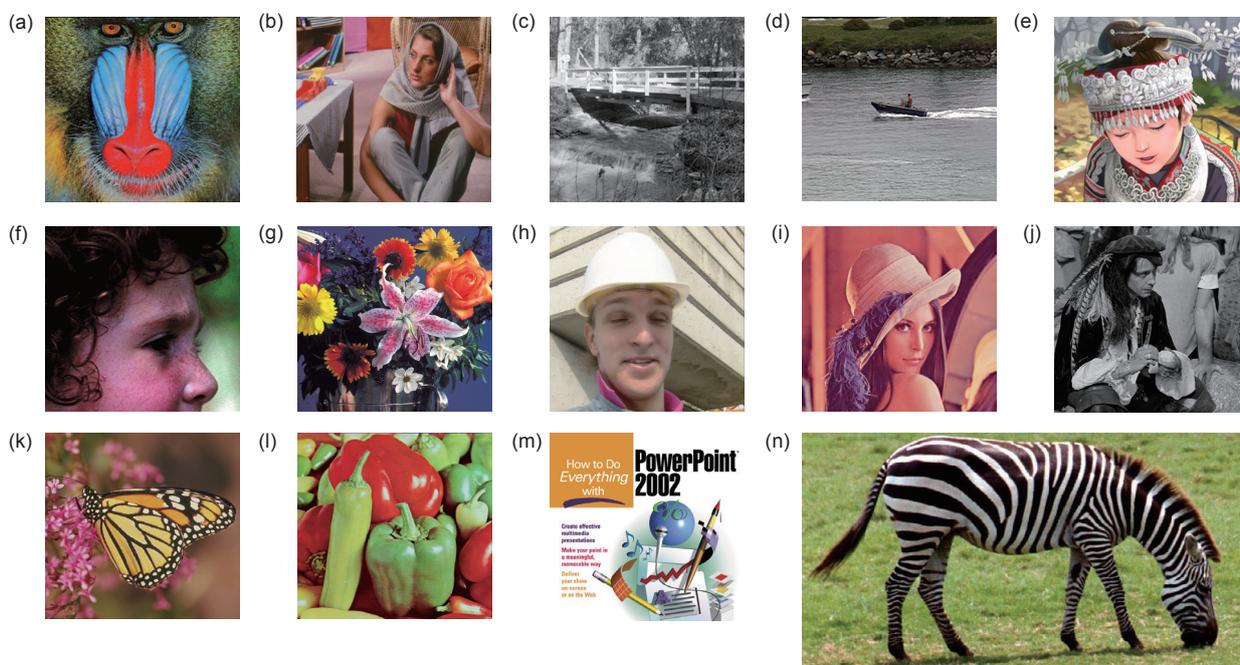


图 3 测试集。(a) 狒狒; (b) 芭芭拉; (c) 桥; (d) 海警; (e) 漫画; (f) 脸; (g) 花; (h) 工长; (i) 莱娜图; (j) 男人; (k) 蝴蝶; (l) 胡椒; (m) 幻灯片; (n) 斑马

Fig. 3 Test image. (a) Baboon; (b) Barbara; (c) Bridge; (d) Coastguard; (e) Comic; (f) Face; (g) Flowers; (h) Foreman; (i) Lenna; (j) Man; (k) Monarch; (l) Pepper; (m) PPT; (n) Zebra

表 1 图片 × 2 倍放大实验结果

Table 1 PSNR(dB) and SSIM values of super resolution images with upscaling factor 2

		Bicubic	Zeyde	NE+LLE	NE+NNLS	SF	ANR	Proposed-1	Proposed
Baboon	PSNR	24.86	25.46	25.52	25.41	25.36	25.54	25.53	25.58
	SSIM	0.6976	0.7513	0.7575	0.7497	0.7465	0.7586	0.7584	0.7639
Barbara	PSNR	28.00	28.67	28.63	28.53	28.25	28.58	28.68	28.69
	SSIM	0.8413	0.8724	0.8733	0.8691	0.8489	0.8729	0.8743	0.8753
Bridge	PSNR	26.58	27.53	27.50	27.33	27.18	27.53	27.60	27.68
	SSIM	0.7929	0.8178	0.8413	0.8343	0.8324	0.8422	0.8426	0.8468
Coastguard	PSNR	29.12	30.42	30.38	30.08	29.59	30.42	30.51	30.59
	SSIM	0.7893	0.8404	0.8435	0.8383	0.8150	0.8446	0.8444	0.8492
Comic	PSNR	26.02	27.61	27.69	27.46	25.67	27.74	27.92	28.02
	SSIM	0.8494	0.8979	0.9006	0.8951	0.8792	0.9011	0.9038	0.9062
Face	PSNR	34.83	35.57	35.60	35.47	34.91	35.65	35.63	35.65
	SSIM	0.8623	0.8818	0.8830	0.8801	0.8591	0.8841	0.8837	0.8851
Flowers	PSNR	30.37	32.22	32.14	31.92	30.11	32.24	32.48	32.57
	SSIM	0.8985	0.9271	0.9274	0.9238	0.9106	0.9287	0.9294	0.9301
Foreman	PSNR	34.14	36.01	36.32	35.92	32.68	36.28	36.63	36.75
	SSIM	0.9518	0.9664	0.9660	0.9656	0.9523	0.9665	0.9681	0.9675
Lenna	PSNR	34.70	36.20	36.28	35.96	35.54	36.29	36.38	36.42
	SSIM	0.9112	0.9262	0.9267	0.9242	0.9108	0.9275	0.9279	0.9280
Man	PSNR	29.25	30.42	30.41	30.23	29.92	30.44	30.59	30.64
	SSIM	0.8458	0.8775	0.8789	0.8740	0.8638	0.8797	0.8807	0.8823
Monarch	PSNR	32.94	35.65	35.45	35.17	33.30	35.61	36.28	36.38
	SSIM	0.9601	0.9726	0.9716	0.9709	0.9583	0.9726	0.9741	0.9730
Pepper	PSNR	34.97	36.55	36.38	36.24	35.76	36.36	36.64	36.72
	SSIM	0.9073	0.9190	0.9184	0.9172	0.9039	0.9190	0.9200	0.9196
PPT	PSNR	26.87	29.29	28.92	29.07	25.27	28.93	29.55	29.66
	SSIM	0.9450	0.9697	0.9648	0.9662	0.9421	0.9651	0.9711	0.9706
Zebra	PSNR	30.63	33.20	33.02	32.60	26.77	33.05	33.33	33.44
	SSIM	0.9086	0.9381	0.9387	0.9342	0.9131	0.9393	0.9400	0.9414
Average	PSNR	30.23	31.77	31.73	31.53	30.02	31.76	31.98	32.06
	SSIM	0.8687	0.8970	0.8994	0.8959	0.8811	0.9001	0.9013	0.9028

表 2 不同放大倍数的实验结果

Table 2 Average PSNR(dB) and SSIM values of different upscaling factor

Scale		Bicubic	Yang	Zeyde	NE+LLE	NE+NNLS	SF	ANR	Proposed
×2	PSNR	30.23	-	31.77	31.73	31.53	30.02	31.76	32.06
	SSIM	0.8687	-	0.8970	0.8994	0.8959	0.8811	0.9001	0.9028
×3	PSNR	27.54	28.31	28.66	28.57	28.47	27.25	28.62	28.79
	SSIM	0.7736	0.7954	0.8078	0.8075	0.8032	0.7856	0.8089	0.8123
×4	PSNR	26.00	25.93	26.81	26.78	26.80	25.22	26.82	26.92
	SSIM	0.7019	0.6958	0.7342	0.7332	0.7323	0.6845	0.7346	0.7359

比其他方法依然能够取得更好的重建结果，验证了本文方法在较大放大倍数情况下的有效性。

为了进一步验证本文算法的有效性，我们对比本文方法和其他方法重建所得图像的视觉效果，图 4 和图 5 表示放大 2 倍时重建图像“Baboon”和“PPT”的局部细节放大图。由图 4 和图 5 的效果图可知，Bicubic 方法会导致图像边缘过于平滑，细节丢失，Zeyde 和 NE+LLE 的方法重建图像边缘部分模糊，SF

方法会产生振铃效应和多余的纹理细节，本文提出的方法相比于 ANR 方法能够重建出更丰富的图像细节，重建图像视觉效果更好。表明本文算法与其他算法相比有更好的超分辨率重建表现，对不同的图像特征能够生成更准确的超分辨率估计。

#### 4.2 聚类个数的影响

本节验证在实验中加入聚类方法对实验效果的影

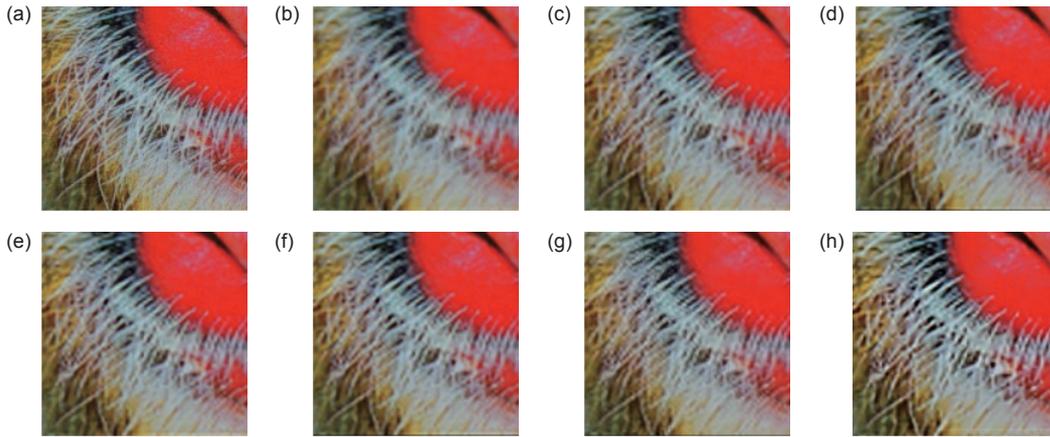


图 4 图片 Baboon 超分辨率重建效果图。(a) 原图; (b) Bicubic 方法; (c) Zeyde 方法; (d) NE+LLE 方法; (e) NE+NNLS 方法; (f) SF 方法; (g) ANR 方法; (h) 本文方法

Fig. 4 SR reconstruction results of image "Baboon". (a) Original; (b) Bicubic; (c) Zeyde; (d) NE+LLE; (e) NE+NNLS; (f) SF; (g) ANR; (h) Proposed



图 5 图片 PPT 超分辨率重建效果图。(a) 原图; (b) Bicubic 方法; (c) Zeyde 方法; (d) NE+LLE 方法; (e) NE+NNLS 方法; (f) SF 方法; (g) ANR 方法; (h) 本文方法

Fig. 5 SR reconstruction results of image "PPT". (a) Original; (b) Bicubic; (c) Zeyde; (d) NE+LLE; (e) NE+NNLS; (f) SF; (g) ANR; (h) Proposed

响,对于 2 倍超分辨率重建实验,从训练集图片中提取的图像块使用 k-means 聚类,聚类类别分别设置为  $K=2, 3, 4, 5, 6, 7, 8, 9, 10$ 。实验结果使用测试集中图片重建结果的平均 PSNR 作为评价指标,不同聚类个数对重建结果的影响如图 6 所示。

由图 6 的实验结果可以看出,随着聚类个数的增加,重建图像的效果也不断增加,表明本文方法加入聚类方法以及训练不同类别字典的有效性,聚类后不仅使得每个子字典相比单个字典的表达能力更强,而且在重建时使用同一类别中的  $n$  近邻字典原子求映射矩阵,也提高了映射矩阵的准确性,故增强了实验效果。由于图像块个数有限,聚类个数的增加会导致有些类别图像块个数较少,训练得到的字典表达性减弱,

因此随着聚类个数的增加 PSNR 指标的增加逐渐趋于平缓。而且聚类个数的增加使得判断待重建图像块所属类别的时间增加,从而增加了图像重建时间。

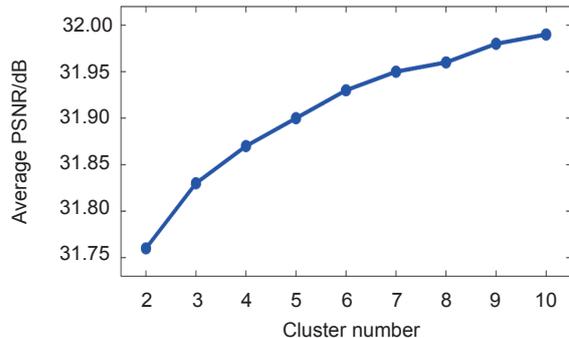


图 6 聚类个数对重建效果的影响

Fig. 6 Influence of clustering number on SR

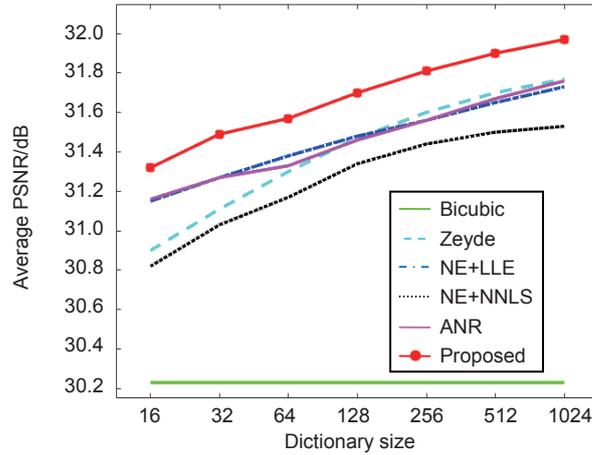


图 7 字典原子大小对图像重建效果的影响  
Fig. 7 Influence of dictionary size on SR

表 3 不同算法的图像重建时间  
Table 3 Time complexity of different methods

Images	Yang	Zeyde	NE+LLE	NE+NNLS	SF	ANR	Proposed
Baboon	69.79	2.36	3.63	16.97	61.06	0.91	6.19
Barbara	73.99	3.95	6.27	29.92	77.94	1.51	11.43
Bridge	82.25	2.45	3.90	18.24	21.86	0.91	6.91
Coastguard	20.54	0.94	1.50	7.14	21.86	0.35	2.41
Comic	29.39	0.80	1.30	6.13	26.11	0.32	2.31
Face	11.77	0.68	1.12	5.14	11.81	0.26	1.98
Flowers	43.25	1.66	2.69	12.45	41.26	0.63	4.88
Foreman	15.50	0.93	1.47	6.93	18.11	0.35	2.58
Lenna	39.31	2.49	3.91	18.98	41.34	0.95	6.59
Man	59.18	2.42	3.89	18.48	57.74	0.91	7.25
Monarch	62.24	3.67	5.87	27.76	53.84	1.37	10.18
Pepper	36.77	2.43	3.87	18.82	38.51	0.92	7.01
Ppt	52.06	2.95	5.01	22.11	52.51	1.19	9.01
Zebra	62.71	2.10	3.36	15.79	57.52	0.78	6.00
Average	47.05	2.13	3.41	16.06	41.53	0.81	6.05

### 4.3 字典大小的影响

从理论上来说，稀疏字典原子越多，图像字典对图像的特征表达能力更好，图像的重建效果越好，但是重建时间也越多，本节探究字典原子个数对图像重建结果的影响。分别选用大小为 16、32、64、128、256、512 和 1024 的字典对图像进行重建，并且与上述提到的方法作对比，图 7 所示为 2 倍放大图像时使用不同字典大小重建图像的效果。从实验结果可以看出在字典原子相同的情况下，本文提出的方法的重建效果更好。

### 4.4 时间复杂度分析

由于本文方法对图像特征聚类后训练得到  $K$  个类别的图像字典，因此在图像重建时需要首先依据  $K$  个聚类中心计算出待重建图像块所属的图像类别，之后再从该类别的图像字典中找出与待重建图像块欧氏距

离最小的字典原子，这是本文算法的主要耗时所在。

表 3 展示了不同算法重建图像耗时的具体数据，本文算法平均耗时为 6.05 s，相比于 SF 方法本文在图像重建效率上有较高的提升，而 ANR 方法只训练单个字典，因此时间效率稍好，但重建效果不如本文算法。

## 5 结 论

本文提出了一种新的基于聚类和协同表示的超分辨率重建方法，本文采用聚类方法对训练集图片提取的图像块根据其特征进行聚类，然后使用聚类后的结果训练得到多个图像字典，从而使得图像字典对图像不同特征的表示更有效，同时根据聚类字典原子计算高、低分辨率图像块映射矩阵，提高了映射矩阵的准确性。而且在重建阶段，对输入的低分辨率图像块根据其特征选择最合适的完备字典和映射矩阵对其进行

重建,使得重建后图像块的特征恢复更准确,从而提高整幅图像的重建效果。实验表明本文利用聚类和协同表示的超分辨率重建算法相比现有算法对图像的重建效果更好,重建出的图像高频信息更多,重建的图像更加接近真实图像。

同时也可以看到本文改进之处,包括利用有监督的聚类算法,提高图像块特征聚类的准确性,也可以在提取图像块特征时引入其他类型的特征,如纹理特征、语义特征等等都是以后有待深入研究的方向。

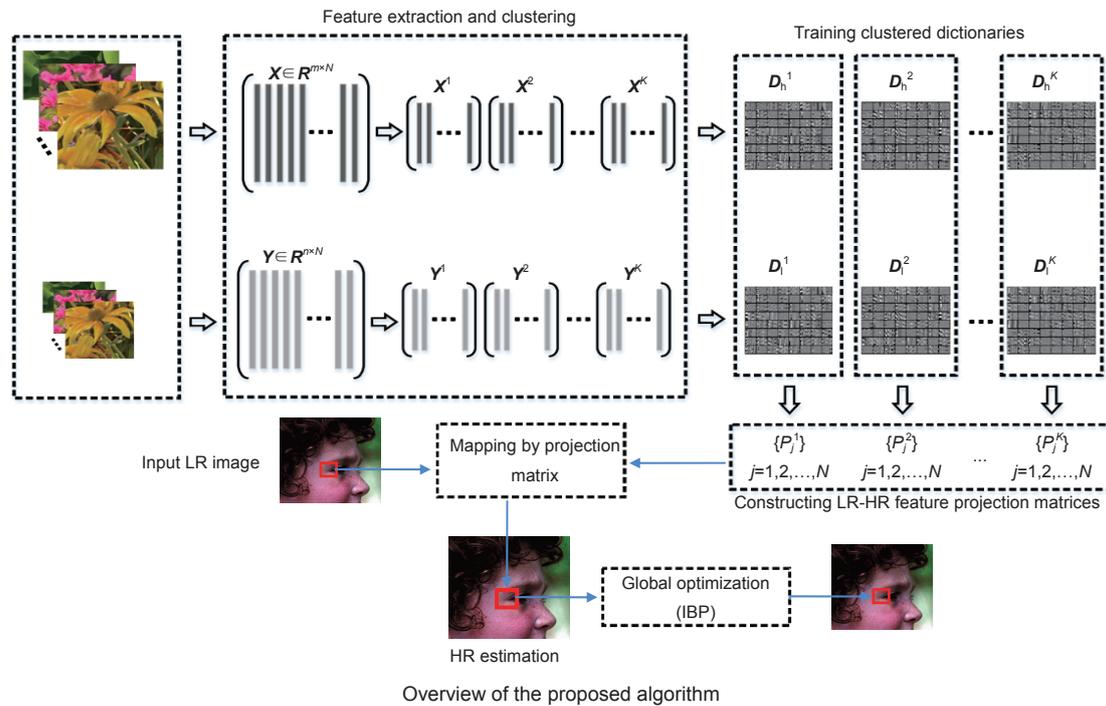
## 参考文献

- [1] Park S C, Park M K, Kang M G. Super-resolution image reconstruction: a technical overview[J]. *IEEE Signal Processing Magazine*, 2003, **20**(3): 21–36.
- [2] Zhan S, Fang Q. Image super-resolution based on edge-enhancement and multi-dictionary learning[J]. *Opto-Electronic Engineering*, 2016, **43**(4): 40–47.  
詹曙, 方琪. 边缘增强的多字典学习图像超分辨率重建算法[J]. *光电工程*, 2016, **43**(4): 40–47.
- [3] Su H, Zhou J, Zhang Z H. Survey of super-resolution image reconstruction methods[J]. *Acta Automatica Sinica*, 2013, **39**(8): 1202–1213.  
苏衡, 周杰, 张志浩. 超分辨率图像重建方法综述[J]. *自动化学报*, 2013, **39**(8): 1202–1213.
- [4] Tsai R Y. Multiframe image restoration and registration[J]. *Advances in Computer Vision and Image Processing*, 1984, **1**(2): 317–339.
- [5] Wu C Z, Hu C S, Zhang M J, et al. Single image super-resolution reconstruction via supervised multi-dictionary learning[J]. *Opto-Electronic Engineering*, 2016, **43**(11): 69–75.  
吴从中, 胡长胜, 张明君, 等. 有监督多字典学习的单幅图像超分辨率重建[J]. *光电工程*, 2016, **43**(11): 69–75.
- [6] Wang R G, Wang Q H, Yang J, et al. Image super-resolution reconstruction by fusing feature classification and independent dictionary training[J]. *Opto-Electronic Engineering*, 2018, **45**(1): 170542.  
汪荣贵, 汪庆辉, 杨娟, 等. 融合特征分类和独立字典训练的超分辨率重建[J]. *光电工程*, 2018, **45**(1): 170542.
- [7] Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution[J]. *IEEE Computer Graphics and Applications*, 2002, **22**(2): 56–65.
- [8] Chang H, Yeung D Y, Xiong Y M. Super-resolution through neighbor embedding[C]//*Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004: 1.
- [9] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, **290**(5500): 2323–2326.
- [10] Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861–2873.
- [11] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]//*International Conference on Curves and Surfaces*, Berlin, Heidelberg, 2010, **6920**: 711–730.
- [12] Aharon M, Elad M, Bruckstein A. rmK-SVD: An algorithm for designing overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322.
- [13] Yang C Y, Yang M H. Fast direct super-resolution by simple functions[C]//*Proceedings of 2013 IEEE International Conference on Computer Vision*, 2013: 561–568.
- [14] Zhang L, Yang M, Feng X C. Sparse representation or collaborative representation: Which helps face recognition[C]//*Proceedings of 2011 IEEE International Conference on Computer Vision*, 2011: 471–478.
- [15] Timofte R, De Smet V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]//*Proceedings of 2013 IEEE International Conference on Computer Vision*, 2013: 1920–1927.
- [16] Irani M, Peleg S. Improving resolution by image registration[J]. *CVGIP: Graphical Models and Image Processing*, 1991, **53**(3): 231–239.
- [17] Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//*Proceedings British Machine Vision Conference*, 2012: 135.

# Image super-resolution based on clustering and collaborative representation

Wang Ronggui, Liu Leilei, Yang Juan, Xue Lixia\*, Hu Min

School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China



**Overview:** Image super-resolution (SR), as an important branch of digital image processing and computer vision, has been widely used in video surveillance, medical imaging and remote sensing image processing in these years. The single-image super-resolution (SISR) is currently a very active area of SR research and we focus on it in this paper. The SISR refers to the reconstruction of a high-resolution (HR) image from an observed degraded low-resolution (LR) image. In this paper, we propose a novel single image super-resolution algorithm by combining clustering and collaborative representation. In the training stage of our method, image patches with similar characteristics are clustered into one sub-class by using k-means clustering algorithm firstly, and then use clustered image patches to learn LR dictionary for each sub-class through K-SVD method. The HR dictionary is computed by using the sparse representation theory which assume that the LR and HR image patches have the same sparse coefficient according to corresponding LR and HR dictionaries. Finally, we compute the projection matrix for each clustered dictionary atom based on the CR with  $l_2$ -norm regularization, which map LR features onto corresponding HR features. In the reconstruction process, for each input LR patch we first find the nearest LR cluster center and then obtain the projection matrix via closet dictionary atom in the corresponding LR dictionary. The HR patch can be estimated by multiplying input LR patch and the projection matrix at last. In conclusion, more expressiveness dictionaries and more effective projection matrices are obtained via our method. Experimental results demonstrate that our proposed algorithm is both quantitatively and qualitatively superior to other classic image SR algorithms. The visual results of HR images show that our proposed algorithm reconstructs more fine details along the dominant edges and suppresses the unexpected artifacts comparing with other methods.

**Citation:** Wang R G, Liu L L, Yang J, *et al.* Image super-resolution based on clustering and collaborative representation [J]. *Opto-Electronic Engineering*, 2018, 45(4): 170537

Supported by National Natural Science Foundation of China (61672202)  
\* E-mail: xlzxm@163.com