

# 基于深度学习特征字典的单帧图像超分辨率重建

赵丽玲<sup>1,2</sup> 孙权森<sup>1</sup> 张泽林<sup>2</sup>

(1. 南京理工大学计算机科学与工程学院, 南京, 210094; 2. 南京信息工程大学信息与控制学院, 南京, 210044)

**摘要:** 在基于字典的单帧图像超分辨率重建算法中, 依赖人工浅层特征设计的字典表达图像特征能力有限。为此, 提出基于深度学习特征字典的超分辨重建方法。该算法首先利用深度网络进行高、低分辨率训练样本图像深层次特征学习; 然后, 在稀疏字典超分辨框架下联合训练特征字典; 最后, 输入单帧低分辨率图像并利用该字典实现超分辨率重建。理论分析表明, 引入深度网络提取图像深层次特征并用于字典训练, 对低分辨率图像的高频信息补充更加有利。实验证明, 与双三次插值以及基于一般人工特征字典的超分辨重建算法相比, 本文算法的主观视觉和客观评价指标均高于对比算法。

**关键词:** 深度学习; 字典学习; 超分辨; 深层次特征提取; 单帧图像

**中图分类号:** TP391      **文献标志码:** A

## Single Image Super Resolution Reconstruction Based on Deep Features Dictionary

Zhao Liling<sup>1,2</sup>, Sun Quansen<sup>1</sup>, Zhang Zelin<sup>2</sup>

(1. School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, 210094, China; 2. School of Information and Control, Nanjing University of Information Science & Technology, Nanjing, 210044, China)

**Abstract:** The ability of image features expression with the dictionary designed by artificial shallow features is limited in dictionary based single image super-resolution reconstruction algorithm. For the reason, an image super resolution reconstruction algorithm based on deep learning and feature dictionary is proposed. Firstly, deep-level feature learning is carried out in high and low resolution training sample images by using deep network. Secondly, the feature dictionary is trained with the combination of sparse coding under the sparse dictionary super resolution frame. Finally, a low resolution image is put in and the super resolution reconstruction is realized by using the dictionary. Theoretical analysis shows that the combination of image deep-level feature extraction and dictionary training by using deep network is more beneficial to high frequency information supplement for low resolution image. Experimental results show that compared with bicubic interpolation and other general artificial feature dictionary based super resolution reconstruction algorithms, the proposed algorithm has better subjective visual and objective evaluation indices.

**Key words:** deep learning; dictionary learning; super-resolution; deep-level feature extraction; single image

## 引言

图像超分辨是指一种由单帧(或多帧)低分辨率退化图像重建出高分辨率图像的第二代图像复原技术<sup>[1]</sup>。20世纪60年代,Harris<sup>[2]</sup>和Goodman<sup>[3]</sup>首次提出了超分辨率算法。超分辨技术发展至今,形成了“单帧图像—多帧图像—新型单帧图像”的3个技术分支。单帧图像超分辨虽然是最早发展起来的,但由于先验信息的引入不足制约了超分辨图像质量的进一步提高<sup>[4-5]</sup>。1984年,Tsai和Huang<sup>[6]</sup>通过比较完美的实验结果,提出了基于多帧图像的超分辨率方法。多帧图像超分辨技术在较长的一段时间内成为主流,取得了相当多的理论和应用成果。然而,多帧图像超分辨由于涉及到图像配准、运动估计等多种图像处理任务,图像重建质量影响因素较多,重建效果不够鲁棒。另外,在许多实际应用领域,如卫星遥感、军事战场和老照片修复等,满足多帧图像超分辨基本数据的获取也比较困难,多帧图像超分辨技术发展遇到瓶颈。随着新的数学理论和方法的不断借鉴引入,以及计算机技术的发展,近年来许多学者展开了“新型单帧图像超分辨”的研究工作<sup>[7-9]</sup>。目前,单帧图像超分辨技术成为了新的研究热点,也是本文的主要研究内容。

目前,根据单帧图像超分辨率是否需要额外数据集,可以分为无外部数据集和有外部数据集两类<sup>[10-11]</sup>。无外部数据集超分辨包括插值方法、迭代反向投影法和凸集投影法等<sup>[12-14]</sup>;有外部数据集超分辨包括基于示例的方法、邻域嵌入法和字典方法<sup>[15-18]</sup>等。在以上算法中,基于字典的超分辨算法可以利用大量隐含在外部数据集中的可用信息,成为新型单帧图像超分辨技术研究的热点。詹曙等<sup>[19]</sup>通过提取图像的一阶、二阶导数等特征,增强了字典表达能力;Dong等<sup>[20]</sup>实现了对图像数据库样本的预先分类,然后针对每一类训练样本图像构建一个特征子字典,提高了字典表达不同结构特征的能力;孙玉宝等<sup>[21]</sup>针对图像的几何、纹理结构等形态,分别提取了几何结构和纹理的分量特征,设计了更加高效并能保持图像几何和纹理结构的多成份子字典,实现了图像超分辨质量的进一步提升;马丽红等<sup>[22]</sup>构建了两种分类字典,分别使用边缘块梯度特征和纹理块局部二值模式特征逼近不同类型结构,提高了超分辨图像的峰值信噪比。另外,多级字典<sup>[23]</sup>、广义主成分分析字典<sup>[24]</sup>、多尺度字典<sup>[25-26]</sup>、压缩感知字典<sup>[27]</sup>和多特征字典<sup>[28]</sup>等多种类型和结构字典的提出及其在超分辨领域的应用,将单帧图像超分辨率算法能力提高到了新的水平。

以上各种字典均是基于图像灰度信息、梯度或纹理结构等浅层人工规则特征构建的,对于蕴含大量丰富特征的复杂图像而言,表达能力仍然有限。因此,如何提高字典的表达能力,使其更准确地描述图像特征是字典超分辨算法中值得继续研究的问题。近年来,深度网络在计算机视觉领域引起了广泛关注,文献<sup>[29]</sup>指出深度卷积神经网络(Convolutional neural networks, CNN)与传统机器学习方法相比具有更强大的特征学习和特征表达能力,在图像分类、物体检测等多个计算机视觉应用领域中取得了令人瞩目的成绩。因此,本文引入深度学习理论,完成高、低分辨率训练样本图像深层次特征提取,并以此构建特征字典,实现图像超分辨重建。实验结果表明,本文提出的基于深度网络构建深层次特征字典的超分辨方法,可以有效提升超分辨图像质量;与双三次插值以及基于一般人工浅层次特征字典的超分辨算法相比,本文算法在主观视觉和客观评价指标中均有一定优势。

## 1 相关理论

### 1.1 ScSR 算法原理及局限性

2010年,Yang<sup>[30]</sup>提出了稀疏表示字典超分辨重建算法(Sparse coding super resolution, ScSR)。目前,该算法在基于字典构建和学习的超分辨算法中最为经典。该算法以自然图像库为训练数据集,训练与高分辨率(High resolution, HR)图像和低分辨率(Low resolution, LR)图像匹配的字典对  $D_h$  (高分辨

率字典)和  $D_l$  (低分辨率字典);然后,对于每个输入的低分辨率块  $y$ ,找到其在  $D_l$  中的稀疏表示系数  $\hat{\alpha}$ ,相应的基于  $D_h$  的高分辨率特征块将通过这些系数联合起来;最后输出高分辨率特征块  $x$ 。该算法的重构原理可以表示为

$$\hat{\alpha} = \min_{\alpha} \|\tilde{D}\alpha - \tilde{y}\|_2^2 + \lambda \|\alpha\|_1 \quad (1)$$

式中:  $\tilde{D} = \begin{bmatrix} FD_l \\ \beta PD_h \end{bmatrix}$ ,  $\tilde{y} = \begin{bmatrix} Fy \\ \beta\omega \end{bmatrix}$ ,  $D$  为过完备字典,  $F$  表示特征提取算子,  $P$  算子用于提取当前待重构 HR 图像块与已重构 HR 图像块重叠的区域以减少块效应,  $\omega$  为重叠区域的像素值,  $\beta$  用于衡量输入 LR 的特征与 HR 重叠区域的匹配情况,  $\lambda$  为拉格朗日算子。利用系数  $\hat{\alpha}$  和字典  $D_h$  得到 HR 图像块,可表示为

$$x = D_h \hat{\alpha} \quad (2)$$

组合各个重构的 HR 图像块就可以获得 HR 图像的估计。然而,对 HR 图像的估计是否准确,与参与超分辨的字典原子是否准确有关,其关键是提取的图像特征对图像结构信息的描述是否足够准确<sup>[31]</sup>。在 ScSR 算法中,对图像特征的提取和描述,选取的是人眼比较敏感的高频特征,提取了与高频特征对应的一阶和二阶梯度,梯度算子为

$$\begin{cases} f_1 = [-1, 0, 1] \\ f_2 = f_1' \\ f_3 = [1, 0, -2, 0, 1] \\ f_4 = f_3' \end{cases} \quad (3)$$

式中  $f_1, f_3$  提取图像水平方向梯度特征,  $f_2, f_4$  提取图像垂直方向的梯度特征,4 个梯度特征按顺序排列成一个向量,作为 LR 图像块的特征。超分辨重构时,ScSR 算法选取梯度特征较大值作为候选字典原子。

然而,对于一些结构信息比较复杂的纹理图像来说,纹理梯度值的变化程度往往小于边缘结构梯度值<sup>[31]</sup>。如果按以上梯度特征选择候选字典原子,所构建的字典会偏重于边缘结构,导致重建图像的纹理结构趋于平滑,超分辨图像质量不高。另外,当待超分辨重建图像与训练集图像类型差异较大时,按照纹理或边缘梯度值选择的候选字典原子,在超分辨重构阶段会产生不适应性,从而影响超分辨重建图像质量。因此,依据人工规则特征构建字典的 ScSR 超分辨算法属于浅层次特征字典,字典在普适性上有一定的局限性。进一步挖掘训练样本图像的深层次特征,从而构建深层次特征字典,提高对图像特征的描述能力,是进一步增强重建图像质量的有效手段。

## 1.2 PCANet 模型

PCANet 是一种具有简单结构的深度学习算法<sup>[32]</sup>,该算法以卷积神经网络(Convolutional neural networks, CNN)和主成分分析(Principal component analysis, PCA)为理论基础,其算法原理如图 1 所示。由文献[31]可知,PCANet 框架由两个 PCA 滤波层、一个哈希层和一个局部直方图计算层 3 部分组成,可以实现对训练样本图像深层次特征的有效提取。但与一般网络不同,PCANet 的滤波器在求解和计算开销中更具优势。PCANet 并不是通过训练过程不断迭代优化得到的,而是通过提取图像局部区域得到映射后直接利用 PCA 提取其主成分得到,每个主成分就是一个滤波器。因此,PCANet 不需要引入规则化参数,也不需要数值优化求解,计算效率比一般深度网络更高。当训练样本图像通过该已知滤波器的深度网络进行特征提取时,也不需要大量的迭代计算来求出最优权值,节省了计算时间。同时,离线训练获得的滤波器,在对测试样本实际操作时也可以减少时间开销。为进一步说明 PCANet 深度网络在特征提取中的优势,本文给出如下举例和理论分析。假设,有  $N$  个输入图像,取大小为  $8 \times$

8, 移动步长为 1 的一个图像块去遍历整个输入图像, 可以得到一个特征映射, 用 PCA 求出这个特征映射的  $L_1$  个主成分, 将  $L_1$  个主成分调整成  $L_1$  个滤波器, 然后像一般 CNN 那样得到  $L_1$  个特征映射; 接下来, 第 2 层和第 1 层的处理一样, 到第 3 层就得到了  $L_2$  个特征映射。这  $L_2$  个特征映射就是 PCANet 提取出来的特征, 可以为后续的图像分析和识别提供数据基础。PCANet 通过学习得到多层网络滤波器的核, 输入图像经过两层滤波器核之后, 可以输出具有较高维度的数据, 作为输入图像的特征参与图像的后续处理和分析。有效的数据越多, 越有利于图像超分辨率重建。

因此, PCANet 进行图像特征提取是对图像各像素进行的直接操作, 且在处理过程中加入了分块处理的步骤, 即在分析图像特征过程中同时结合了像素的邻域信息, 并且克服了 PCA 方法不可避免的降维效果, PCANet 深度网络输出的数据量反而更多。PCANet 通过深度网络学习提取了更加丰富的训练样本图像特征, PCANet 特征可以提供比人工规则特征更加丰富的描述和表达, 为后续的图像处理和识别等准备了更加丰富的先验信息, 对低分辨率图像的高频信息补充更加有利, 更有利于超分辨率重建。

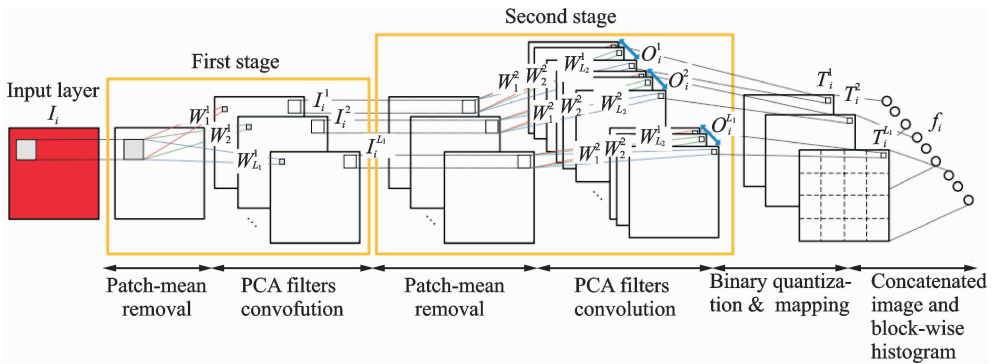


图 1 PCANet 结构框图<sup>[32]</sup>  
 Fig. 1 Framework of PCANet<sup>[32]</sup>

### 1.3 深度学习特征字典

特征是机器学习算法的重要因素, 对最终模型有决定性的影响。如果对训练样本图像有很好的特征表达能力, 通常计算模型就能获得比较满意的结果。近几年, 随着目标图像越来越复杂, 人们开始关注如何构造更加高效的特征字典<sup>[33-36]</sup>, 以实现图像边缘信息、丰富纹理及几何结构等更准确的描述。深度网络则可以挖掘数据的深层次知识, 用深度网络提取训练样本图像更深层次的特征, 能够突破人工规则特征的局限, 从而提高字典的表达能力和适应性。

因此, 本文结合 PCANet 深度网络和 ScSR 稀疏字典两种算法优势, 提出了基于 PCANet 特征字典的单帧图像超分辨率重建算法, 算法框图如图 2 所示。在训练阶段, 首先假定低分辨率与高分辨率图像特征块关于各自字典具有相似的稀疏表达形式, 然后采用 PCANet 深度网络获得训练样本图像深层次特征, 再通过字典联合训练获得一对过完备特征字典  $D_h$  和  $D_l$ 。其中,  $D_h$  表示高分辨率图像特征字典,  $D_l$  表示低分辨率图像特征字典。在超分辨率重建阶段, 对低分辨率图像同样用 PCANet 方法提取深层次特征, 求解待超分辨图像的各低分辨率特征块在字典  $D_l$  上的稀疏表达形式, 并将低分辨率图像特征块关于  $D_l$  的稀疏表示系数直接作用于  $D_h$ , 从而获得对应的高分辨率图像特征块, 最终实现低分辨率图像的超分辨率重建。该算法采用 PCANet 深度网络挖掘训练样本图像特征, 可以获得比非深度网络更加丰富的特征信息。在此基础上构建的深层次特征字典, 同样也提高了特征字典的表达能力, 实现了重建图像质量的有效提升。

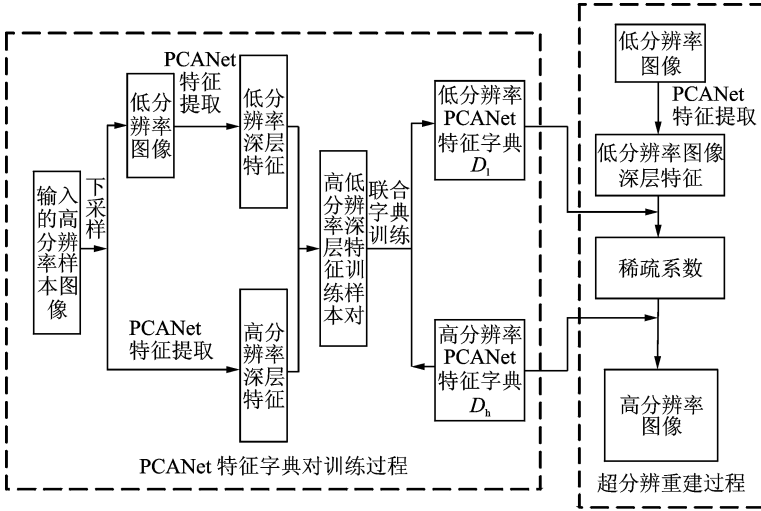


图 2 基于 PCANet 特征字典的单幅图像超分辨率算法概述

Fig. 2 Sketch of single image super resolution algorithm based on PCANet features dictionary

## 2 本文算法

### 2.1 图像预处理

首先,对训练数据中的  $K$  幅高分辨率图像进行模糊下采样,然后将其放大到与高分辨图像尺寸同样大小,获得与高分辨率图像对应的低分辨率图像,构成训练样本对:  $T = \{\mathbf{X}_h, \mathbf{X}_l\}$ , 其中  $\mathbf{X}_h = \{x_h^i\}_{i=1}^K$  为高分辨率特征,  $\mathbf{X}_l = \{x_l^i\}_{i=1}^K$  为低分辨率特征。然后,对于每一个训练样本计算其分块矩阵。以高分辨训练样本为例,选取一个  $k_1 \times k_2$  的滑动窗口(通常选边长为 3, 5 或 7 个像素的正方形窗口),每张  $m \times n$  大小的图片经过滑动窗口提取局部特征之后,就变成  $m \times n$  个  $k_1 \times k_2$  大小的图像块;再对这  $m \times n$  个图像块去平均,完成单张图像的特征提取操作。以上所有  $N$  张高分辨图像均执行以上操作,可以获得一个新的具有  $N \times m \times n$  列的数据矩阵  $\mathbf{X}$ , 该矩阵的每一列表示一个图像块,共有  $k_1 \times k_2$  个元素。

得到的第  $i$  张高分辨训练样本可表示为式(4),总体高分辨样本的分块矩阵表示为式(5),低分辨率样本的分块矩阵计算与高分辨率样本相同,表示为式(6)。

$$\bar{\mathbf{X}}_{hi} = [\bar{x}_{hi,1}, \bar{x}_{hi,2}, \dots, \bar{x}_{hi,mn}] \quad (4)$$

$$\mathbf{X}_h = [\bar{\mathbf{X}}_{h1}, \bar{\mathbf{X}}_{h2}, \dots, \bar{\mathbf{X}}_{hK}] \in \mathbf{R}^{k_1 k_2 \times K_m} \quad (5)$$

$$\mathbf{X}_l = [\bar{\mathbf{X}}_{l1}, \bar{\mathbf{X}}_{l2}, \dots, \bar{\mathbf{X}}_{lK}] \in \mathbf{R}^{k_1 k_2 \times K_m} \quad (6)$$

### 2.2 特征提取

在以上训练样本矩阵的基础上进行 PCANet 特征提取,并以此特征作为 ScSR 模型中的训练样本特征,用于 PCANet 特征字典训练。PCANet 深度学习网络分 3 层完成图像特征提取,下面以高分辨训练图像为例,给出详细的特征提取过程。

#### (1) 第 1 层

第 1 层特征提取过程即为构建 PCA 滤波器并进行卷积的过程,图 1 中第 1 个矩形框(First stage)内所示。假定在第  $i$  层需要的滤波器数量为  $L_1$ ,通过式(7)寻找一系列的标准正交矩阵最小化重构误差,通过提取协方差矩阵  $\mathbf{X}_h$  的前  $L_1$  个最大特征值所对应的特征向量,组成特征映射矩阵,即为 PCA 滤波器,如式(8)所示。

$$\min_{\mathbf{V} \in \mathbf{R}^{k_1 k_2 \times L_1}} \|\mathbf{X}_h - \mathbf{V}\mathbf{V}^T \mathbf{X}_h\|_F^2 \quad \text{s. t. } \mathbf{V}^T \mathbf{V} = \mathbf{I}_{L_1} \quad (7)$$

$$\mathbf{W}_{hl}^1 = \text{mat}_{k_1, k_2} (q_{hl}(\mathbf{X}_h \mathbf{X}_h^T)) \in \mathbf{R}^{k_1 \times k_2} \quad L=1, 2, \dots, L_1 \quad (8)$$

把  $L_1$  个特征向量的每一列重新进行排列,得到 1 个 patch,即获得了  $L_1$  个  $k_1 \times k_2$  的窗口。然后,对每一张图像,都用这  $L_1$  个窗口作一次卷积。通过第 1 层 PCA 滤波器可以保留训练样本图像的主要信息,表示为

$$\mathbf{I}_{hi}^L = \mathbf{I}_{hi} * \mathbf{W}_{hl}^1 \quad i=1, 2, \dots, K \quad (9)$$

(2) 第 2 层

第 2 层的 PCA 映射过程如图 1 第 2 个矩形框(Second stage)内所示。与第 1 层对样本图像分块操作类似,将第 1 层 PCA 映射输出结果作为第 2 层的输入,同样在第 2 层对该矩阵进行块采样、级联和零均值化等操作,结果表示为

$$\mathbf{Y}_h^L = [\bar{\mathbf{Y}}_{h1}^L, \bar{\mathbf{Y}}_{h2}^L, \dots, \bar{\mathbf{Y}}_{hK}^L] \in \mathbf{R}^{k_1 k_2 \times Kmn} \quad (10)$$

$$\mathbf{Y}_h = [\mathbf{Y}_h^1, \mathbf{Y}_h^2, \dots, \mathbf{Y}_h^{L_1}] \in \mathbf{R}^{k_1 k_2 \times L_1 mn} \quad (11)$$

同样,PCA 滤波器由协方差矩阵对应的特征向量组成,滤波器为

$$\mathbf{W}_{hl}^2 = \text{mat}_{k_1, k_2} (q_{hl}(\mathbf{Y}_h \mathbf{Y}_h^T)) \in \mathbf{R}^{k_1 \times k_2} \quad L'=1, 2, \dots, L_2 \quad (12)$$

由于第 1 层有  $L_1$  个滤波器核,第 2 层对上一层输出的每一个特征再进行同样的步骤进行特征提取,可以获得  $L_2$  个特征输出。最终,对于每一张样本图像,PCANet 将输出  $L_1 \times L_2$  个特征矩阵,如式(13)。在结构上,两层 PCA 特征提取过程很相似,PCANet 也可以根据需要被扩展成包含更多层的深度网络结构。

$$\mathbf{O}_{hi}^{L'} = \{\mathbf{I}_{hi}^L * \mathbf{W}_{hl'}^2\}_{L'=1}^{L_2} \quad (13)$$

(3) 输出层

输出层主要针对第 2 层的每个输出矩阵,将这些矩阵都进行二值处理,使矩阵中只包含整数 1 和 0,然后对该矩阵进行二值化哈希编码,如

$$\mathbf{T}_{hi}^L = \sum_{L'=1}^{L_2} 2^{L'-1} H(\mathbf{O}_{hi}^{L'}) = \sum_{L'=1}^{L_2} 2^{L'-1} H(\mathbf{I}_{hi}^L * \mathbf{W}_{hl'}^2) \quad L=1, 2, \dots, L_1 \quad (14)$$

式中:  $\mathbf{T}_{hi}^L$  为高分辨训练图像特征经哈希编码之后的结果,  $2^{L'-1}$  为变换系数使图像中每个像素值都变为 0~255 之间的数值,函数  $H(\cdot)$  类似于一个单位阶跃函数,作用是量化并加大各特征的差异。然后,对上述结果进行直方图编码,完成一幅高分辨样本图像的 PCANet 特征提取,提取结果可表示为

$$\mathbf{F}_{hi} = [\text{Bhist}(\mathbf{T}_{hi}^1), \dots, \text{Bhist}(\mathbf{T}_{hi}^{L_1})]^T \in \mathbf{R}^{(2^L)^{L_1} B} \quad (15)$$

同理可推,经过与高分辨率训练图像相同的步骤,低分辨率训练图像的 PCANet 特征  $\mathbf{F}_{li}$  可表示为

$$\mathbf{F}_{li} = [\text{Bhist}(\mathbf{T}_{li}^1), \dots, \text{Bhist}(\mathbf{T}_{li}^{L_1})]^T \in \mathbf{R}^{(2^L)^{L_1} B} \quad (16)$$

式中:  $\mathbf{F}_{hi}$  和  $\mathbf{F}_{li}$  表示样本图像特征提取结果;Bhist 表示直方图编码过程;B 表示样本图像被分割的图像块数量。

2.3 字典训练

本文采用联合稀疏编码的方式,在 ScSR 框架下训练字典。算法的目标是得到一对可以描述样本复杂特征的字典对  $\mathbf{D}_h$  和  $\mathbf{D}_l$ ,使得  $K$  对图像生成的 PCANet 特征  $\mathbf{F}_{hi}$  和  $\mathbf{F}_{li}$  在  $\mathbf{D}_h$  和  $\mathbf{D}_l$  上具有相同的稀疏表示,而且  $\mathbf{F}_{hi}$  和  $\mathbf{F}_{li}$  具有相同的表示系数,即

$$\{\mathbf{D}_h, \boldsymbol{\alpha}\} = \underset{\mathbf{D}_h, \boldsymbol{\alpha}}{\text{argmin}} \|\mathbf{F}_{hi} - \mathbf{D}_h \cdot \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^K \lambda_i \|\alpha_i\|_1 \quad (17)$$

$$\{\mathbf{D}_1, \boldsymbol{\alpha}\} = \operatorname{argmin}_{\mathbf{D}_1, \boldsymbol{\alpha}} \|\mathbf{F}_{l_1} - \mathbf{D}_1 \cdot \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^K \lambda_i \|\alpha_i\|_1 \quad (18)$$

式中:  $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^K$  为稀疏表示稀疏矩阵,  $\lambda_i$  为平衡系数。为了使高分辨率图像特征和低分辨率图像特征关于各自的字典具有相同的稀疏表达形式, 对式(17, 18)采取联合训练策略, 即

$$\{\mathbf{D}_h, \mathbf{D}_l, \boldsymbol{\alpha}\} = \operatorname{argmin}_{\mathbf{D}_h, \mathbf{D}_l, \boldsymbol{\alpha}} \frac{1}{N} \|\mathbf{F}_{h_1} - \mathbf{D}_h \cdot \boldsymbol{\alpha}\|_2^2 + \frac{1}{M} \|\mathbf{F}_{l_1} - \mathbf{D}_l \cdot \boldsymbol{\alpha}\|_2^2 + \left(\frac{1}{N} + \frac{1}{M}\right) \sum_{i=1}^K \lambda_i \|\alpha_i\|_1 \quad (19)$$

式中:  $N$  和  $M$  分别表示高、低分辨率特征图像块元素值重排成列矢量的维数,  $1/N$  和  $1/M$  用于平衡式(17, 18)中  $\mathbf{D}_h$  和  $\mathbf{D}_l$  两者的代价, 为方便求解, 将式(19)写成

$$\{\mathbf{D}_c, \boldsymbol{\alpha}\} = \operatorname{argmin}_{\mathbf{D}_c, \boldsymbol{\alpha}} \|\mathbf{F}_c - \mathbf{D}_c \cdot \boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^K \lambda'_i \|\alpha_i\|_1$$

$$\mathbf{D}_c = \begin{bmatrix} \frac{1}{\sqrt{N}} \mathbf{D}_h \\ \frac{1}{\sqrt{M}} \mathbf{D}_l \end{bmatrix}, \mathbf{X}_c = \begin{bmatrix} \frac{1}{\sqrt{N}} \mathbf{F}_h \\ \frac{1}{\sqrt{M}} \mathbf{F}_l \end{bmatrix}, \lambda'_i = \left(\frac{1}{N} + \frac{1}{M}\right) \lambda_i \quad (20)$$

式(20)采用迭代方法求解。首先, 通过给定字典  $\mathbf{D}_c$ , 求解每对训练样本数据  $F_{c_i}$  在  $\mathbf{D}_c$  上的稀疏表达系数  $\alpha_i$ , 得到稀疏表示矩阵  $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^K$ , 最后根据  $\boldsymbol{\alpha}$  更新字典  $\mathbf{D}_c$ 。

## 2.4 图像重建

在得到高分辨率和低分辨率图像特征字典对  $\mathbf{D}_h$  和  $\mathbf{D}_l$  后, 对于每一张低分辨率测试图像  $\mathbf{Y}$ , 需要解决经典优化问题

$$\hat{\boldsymbol{\alpha}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \|\tilde{\mathbf{D}}\boldsymbol{\alpha} - \tilde{\mathbf{Y}}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \quad (21)$$

式中:  $\tilde{\mathbf{D}} = \begin{bmatrix} \mathbf{F}\mathbf{D}_l \\ \beta\mathbf{P}\mathbf{D}_h \end{bmatrix}$ ,  $\tilde{\mathbf{Y}} = \begin{bmatrix} \mathbf{F}\mathbf{Y} \\ \beta\boldsymbol{\omega} \end{bmatrix}$ ,  $\mathbf{F}$  为 PCANet 特征,  $\mathbf{P}$  为提取当前高分辨图像块与已重构高分辨图像块的重叠区域, 该策略可以减少块状效应,  $\boldsymbol{\omega}$  为重叠区域像素的个数,  $\beta$  衡量输入低分辨特征块与高分辨重叠区域的匹配情况。求解式(21)即可得到每个低分辨率图像的系数表达系数  $\boldsymbol{\alpha}$ , 将其作用于  $\mathbf{D}_h$ , 便可求得初始高分辨率特征信息图像  $\mathbf{X}_0 = \mathbf{D}_h \boldsymbol{\alpha}$ 。针对  $\mathbf{X}_0$  求解如下优化问题, 使得最终的高分辨率图像满足重构约束条件, 即

$$\mathbf{X}^* = \operatorname{argmin}_{\mathbf{X}} \|\mathbf{H}\mathbf{X} - \mathbf{Y}\|_2^2 + \lambda \|\mathbf{X} - \mathbf{X}_0\|_2^2 \quad (22)$$

式中:  $\mathbf{H}$  为图像的降质算子 ( $\mathbf{H}$  与成像过程有关, 本文实验中该算子仅表示对测试图像的下采样)。对式(22)利用梯度下降法求解, 可得到超分辨率结果图像。

## 2.5 算法步骤

本文算法首先对低分辨率图像进行 PCANet 特征提取, 并基于该深层次特征进行字典训练, 最终实现超分辨重建, 算法如下所示。

**算法 1:** 基于 PCANet 特征字典的单帧图像超分辨重建

- (1) 输入: 低分辨率图像  $\mathbf{Y}$ , 高、低分辨率字典对  $\mathbf{D}_h, \mathbf{D}_l$
- (2) 将低分辨图像进行 2 倍上采样插值  $\mathbf{Y}' = \mathbf{S}(\mathbf{Y})$ ,  $\mathbf{S}$  为上采样算子
- (3) 提取低分辨图像的 PCANet 特征  $\mathbf{F}_l$
- (4) 根据式(21)建立字典  $\mathbf{D}_h, \mathbf{D}_l$  与  $\mathbf{F}_l$  之间的关系
- (5) for  $i=1$  to  $T$  do
- (6) 计算  $\hat{\boldsymbol{\alpha}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \|\tilde{\mathbf{D}}\boldsymbol{\alpha} - \tilde{\mathbf{Y}}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1$ , 得到低分辨率 PCANet 特征关于低分辨率字典  $\mathbf{D}_l$  的稀疏

系数  $\hat{\alpha}$

- (7) 根据稀疏系数  $\hat{\alpha}$  和高分辨率字典  $D_h$ , 估计高分辨率图像特征信息  $X_0 = D_h \alpha$
- (8) 将得到的高分辨率图像特征信息与低分辨率图像合成, 得到初始超分辨结果图像
- (9) end for
- (10) 利用梯度下降法, 计算  $X^* = \operatorname{argmin}_X \|HX - Y\|_2^2 + \lambda \|X - X_0\|_2^2$ , 得到与初始高分辨图像最近的超分辨结果图像  $X^*$
- (11) 输出: 高分辨率图像  $X^*$ , 算法结束。

### 3 实验结果及分析

为了验证图像深层次特征提取对构建特征字典以及实现图像分辨率提高的有效性, 本文选择常用数据集 Set5<sup>[37]</sup>, Set14<sup>[38]</sup> 和 BSD500<sup>[39]</sup> 用于网络训练和测试, 选择 BSD500 中的 Train 文件夹图像训练深度特征字典, 选择 Test 文件夹图像以及 Set5, Set14 作为测试集。实验中, 选择了经典 Bicubic 插值超分辨算法, 和基于人工特征字典超分辨算法, 作为本文基于深度学习特征字典超分辨算法的对比算法。采用主观视觉评价、客观评价指标峰值信噪比 (Peak signal to noise ratio, PSNR)、结构相似度参数 (Structural similarity index measurement, SSIM) 和算法计算时间 (Time) 作为图像质量评价标准。实验所用的图像库包括人物、植物、动物、风景和建筑等, 部分图像如图 3 所示。实验运行环境为: 64 位 Windows 7 SP1 操作系统, Intel(R) Core(TM) i7-3667U 2.0 GHz 处理器, 8 GB 运行内存, Matlab R2012a 版本。



图 3 训练图像

Fig. 3 Training images

#### 3.1 基本参数设置

实验所需的高、低分辨率训练图像集采用如下方法构造, 通过对已知高分辨训练图像 HR 进行两倍的双三次插值下采样得到低分辨率图像, 再对低分辨率图像进行放大获得低分辨率训练图像 LR, 如图 4 所示。对 HR 和 LR 执行预处理操作, 本文实验中训练图像  $K=69$ ;  $k_1=k_2=5$ ;  $m, n$  为每幅训练图像实际尺寸。采用 PCANet 提取训练图像的深层次特征时, 网络参数  $L_1$  取值为 40、 $L_2$  取值为 8; 字典训练过程中, 字典尺寸 512, 稀疏正则项系数 0.2, 块尺寸 5 像素。

#### 3.2 实验结果

图 5 为是各算法在测试图像中的部分实验结果。从图 5 可以看出, 本文算法在超分辨图像细节重建上与 Bicubic (双三次插值) 和 ScSR (稀疏表示超分辨) 两种算法相比的优势。图 5(a) 是降质低分辨率图像, 为了考察超分辨算法能力, 本实验降质过程仅为下采样; 图 5(b) 是双 Bicubic 超分辨算法结果, 可以看出由于先验信息不足, 结果图像中部分纹理平滑效应较大, 图像比较模糊; 图 5(c) 是 ScSR 算法结果, 因算法利用了训练图像集的先验信息, 超分辨结果图像质量比插值算法所提高; 图 5(d) 为本文方法超分辨结果图像, 图像中边界、纹理模糊程度较其他两种方法结果图像有所减弱, 图像细节的视觉效果最为清晰。该实验表明, 本文算法在训练图像中提取的特征最为丰富, 因而训练的到的特征字典可以更好地反映高、低分辨率图像特征, 对构建超分辨的特征字典十分有利。表 1 为各实验超分辨重建结果 PSNR、SSIM 和计算开销 (Time) 的比较。从 PSNR

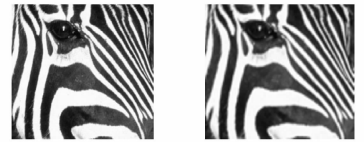


图 4 高分辨率图像及其下采样放大图像

Fig. 4 HR image and its zoomed sub-sampled image

和计算开销 (Time) 的比较。从 PSNR



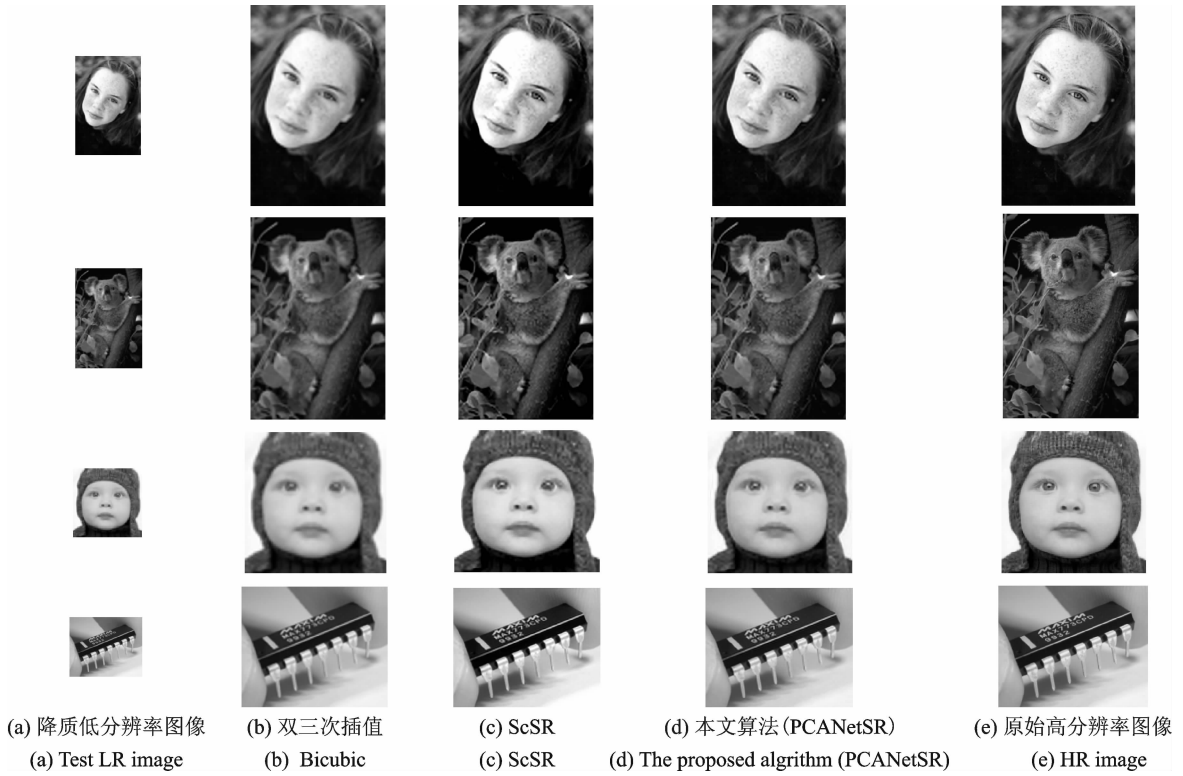


图5 具有两倍放大因子的超分辨率结果图像

Fig. 5 Results of super resolution images with two upscaling factor

表1 不同方法重建图像的客观评价指标

Tab. 1 Objective evaluation of reconstructed images with different methods

图像	PSNR			SSIM			Time/s	
	Bicubic	ScSR	PCANetSR	Bicubic	ScSR	PCANetSR	ScSR	PCANetSR
Girl	34.46	37.43	38.03	0.971 2	0.958 7	0.987 3	35.89	103.67
Koala	33.56	36.89	38.82	0.911 2	0.927 6	0.913 5	28.45	150.20
Baby	33.31	35.36	35.65	0.893 5	0.934 3	0.964 3	33.24	475.48
Microchips	31.56	36.08	38.53	0.901 1	0.901 6	0.975 4	37.25	516.15
Building	33.45	36.90	40.56	0.912 5	0.913 6	0.924 3	36.56	350.38
Butterfly	27.40	30.20	37.79	0.862 3	0.895 6	0.883 2	40.26	204.96
Comic	26.00	28.20	37.36	0.791 6	0.788 5	0.812 5	39.28	107.26
Set5	34.21	37.02	39.68	0.885 3	0.886 1	0.893 5	38.90	155.51
Set14	35.33	36.96	37.57	0.892 1	0.895 7	0.902 3	42.11	450.35
BSD500	32.47	36.23	38.47	0.900 6	0.911 2	0.941 1	45.12	118.14
Average	32.175	35.127	38.246	0.892 14	0.901 29	0.919 74	37.706	263.21

数值结果来看,插值方法 PSNR 均值为 32.17, ScSR 方法 PSNR 均值为 35.12, 而本文算法 PSNR 均值为 38.24, 在实验测试中最高。从 ScSR 和本文算法的 Time 数值结果来看, ScSR 方法 Time 均值为 37.70, 本文算法 Time 均值为 263.21, 相比较而言本文计算开销稍大。通过分析可知, 由于本文算法提取图像深层次特征比一般人工规则特征对图像细节的捕捉能力更强, 因而参与字典学习重建的特征规模更大, 所以算法在运行时间上比 ScSR 运行时间稍长。但从以上客观评价指标 PSNR, SSIM 以及主观

视觉效果来看,获得的超分辨图像质量确有较大提高,因此本文算法在综合效率中仍具有一定优势。

## 4 结束语

本文提出了基于深度学习特征字典的单帧图像超分辨率重建算法,通过挖掘图像深层次特征,增强了高、低分辨率字典表达图像纹理信息和复杂结构的能力,有效保持了超分率重建图像的丰富细节特征。与基于人工规则特征字典的超分辨算法相比,本文算法在图像质量客观评价指标上具有一定优势。下一步的研究方向是,优化算法争取缩短计算时间,并对影响算法效率的网络参数、字典尺寸等诸多因素进行分析和总结,以进一步提升采用该类算法实现图像超分辨的效率。

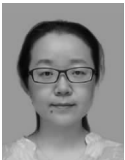
## 参考文献:

- [1] Tian J, Mak K. A survey on super-resolution imaging[J]. *Signal, Image and Video Processing*, 2011, 5(3): 329-342.
- [2] Harris J L. Diffraction and resolving power[J]. *Journal of the Optical Society of America*, 1964, 54(7): 931-936.
- [3] Goodman J W. *Introduction to fourier optics*[M]. New York: McGraw Hill, 1968.
- [4] Zhang Xiangjun, Wu Xiaolin. Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation[J]. *IEEE Transactions on Image Processing*, 2008, 17(6): 887-896.
- [5] Mallat S, Yu G. Super-resolution with sparse mixing estimators[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2010, 19(11): 2889-2900.
- [6] Tsai R Y, Huang T S. Multi frame image restoration and registration[J]. *Advances in Computer Vision and Image Processing*, 1984(1): 317-339.
- [7] Xiao Liang, Huang Lili. Edge and color preserving single image superresolution[J]. *Journal of Electronic Imaging*, 2014, 033002(23): 3.
- [8] Yang J, Wright J, Huang T, et al. Image super-resolution as sparse representation of raw image patches[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, Alaska: IEEE, 2008: 1-8.
- [9] 李晓峰, 曾蕾, 徐进, 等. 基于特征表征的单幅图像超分辨率方法 [J]. *电子科技大学学报*, 2015, 44(1): 22-27.  
Li Xiaofeng, Zeng Lei, Xu Jin, et al. Single image super-resolution based on the feature sign method[J]. *Journal of University of Electronic Science and Technology of China*, 2015, 44(1): 22-27.
- [10] Sun Jian, Zhu Jiejie, Tappen M F. Context-constrained hallucination for image super-resolution[C]// *2010 IEEE Conference on Computer Vision and Pattern Recognition*. Xi'an, China: IEEE, 2010: 231-238.
- [11] Tai Yuwing, Liu Shuaicheng, Brown M S, et al. Super-resolution using edge prior and single image detail synthesis[C]// *2010 IEEE Conference on Computer Vision and Pattern Recognition*. Daejeon, Korea: IEEE, 2010: 2400-2407.
- [12] Tao H J, Tang X J, Tian J W. Super-resolution remote sensing image processing algorithm based on wavelet transform and interpolation[C]// *Proceedings of SPIE*. Hangzhou, China: Society of Photo Optical Instrumentation Engineers, 2003: 259-263.
- [13] Lertrattanapanich S, Bose N K. High resolution image formation from low resolution frames using delaunay triangulation [J]. *IEEE Trans Image Processing*, 2002, 11(12): 1427-1441.
- [14] Dai S, Han M, Wu Y, et al. Bilateral backprojection for single image super resolution[C]// *Proceedings of the 2007 IEEE International Conference on Multimedia and Expo*. Beijing, China: IEEE, 2007: 1039-1042.
- [15] Freeman W T, Jones T R, Pasztor E C. Example-based super resolution[C]// *Proceedings of the 2002 IEEE International Conference on Computer Graphics and Applications*. Los Alamitos, USA: IEEE Computer Society, 2002: 56-65.
- [16] Chang H, Yeung D Y, Xiong Y. Super-resolution through neighbor embedding[C]// *Proceedings of the 2004 IEEE International Conference on Computer Vision and Pattern Recognition*. Washington DC, USA: IEEE Computer Society, 2004, 1: 275-282.
- [17] 练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展[J]. *自动化学报*, 2015, 41(2): 240-260.  
Lian Qiusheng, Shi Baoshun, Chen Shuzhen. Research advances on dictionary learning models, algorithms and applications [J]. *Acta Automatica Sinica*, 2015, 41(2): 240-260.
- [18] 宋慧慧. 基于稀疏表示的图像超分辨率重建算法研究[D]. 合肥: 中国科学技术大学, 2015.  
Song Huihui. Research for image super-resolution reconstruction algorithm based on sparse representation[D]. Hefei: University of Science and Technology of China, 2015.
- [19] 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.  
Zhan Shu, Fang Qi. Image super-resolution based on edge-enhancement and multi-dictionary learning[J]. *Opto-Electronic Engineering*, 2016, 43(4): 40-47.
- [20] Dong W, Zhang L, Shi G, et al. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization[J]. *IEEE Transactions on Image Processing*, 2011, 20(7): 1838.
- [21] 孙玉宝, 韦志辉, 肖亮, 等. 多形态稀疏性正则化的图像超分辨率算法[J]. *电子学报*, 2010, 12: 2898-2903.

Sun Yubao, Wei Zhihui, Xiao Liang, et al. Multimorphology sparsity regularized image super-resolution[J]. *Acta Electronica Sinica*, 2010, 12:2898-2903.

- [22] 马丽红, 黄茵, 黎剑晖. 基于灵活 LBP 纹理字典构造及多特征描述的改进 SCSR 算法[J]. *华南理工大学学报(自然科学版)*, 2015, 43(3):57-65.  
Ma Lihong, Huang Yin, Li Jianhui. Improved SCSR algorithm on the basis of flexible LBP texture dictionary and multi feature description[J]. *Journal of South China University of Technology (Natural Science Edition)*, 2015, 43(3):57-65.
- [23] Thiagarajan J J, Ramamurthy K N, Spanias A. Multilevel dictionary learning for sparse representation of images[C]//*Proceedings of the 2011 IEEE Digital Signal Processing Workshop and Signal Processing Education Workshop*. Sedona, AZ: IEEE, 2011:271-276.
- [24] Vidal R, Ma Y, Sastry S. Generalized principal component analysis (GPCA)[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(12): 1945-1959.
- [25] Mairal J, Sapiro G, Elad M. Learning multiscale sparse representations for image and video restoration[J]. *Multiscale Modeling and Simulation*, 2008, 7(1): 214-241.
- [26] Ophir B, Lustig M, Elad M. Multi-scale dictionary learning using wavelets[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5(5): 1014-1024.
- [27] Rauhut H, Schnass K, Vandergheynst P. Compressed sensing and redundant dictionaries[J]. *IEEE Transactions on Information Theory*, 2008, 54(5):2210-2219.
- [28] 刘燕文. 基于稀疏表示的多成分字典超分辨率图像重建的研究[D]. 厦门: 厦门大学, 2014.  
Liu Yanwen. Research of super-resolution reconstruction based on sparse representation with multi component dictionary [D]. Xiamen: Xiamen University, 2014.
- [29] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. *数据采集与处理*, 2016, 31(1):1-17.  
Lu Hongtao, Zhang Qinchuan. Applications of deep convolutional neural network in computer vision[J]. *Journal of Data Acquisition & Processing*, 2016, 31(1):1-17.
- [30] Yang J, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2010, 19(11):2861-2873.
- [31] 黎剑晖. 基于多特征多尺度字典学习的超分辨率重构算法[D]. 广州: 华南理工大学, 2014.  
Li Jianhui. Super resolution reconstruction algorithm based on dictionary learning using multi-feature and multi-scale description[D]. Guangzhou: South China University of Technology, 2014.
- [32] Chan T H, Jia K, Gao S, et al. PCANet: A simple deep learning baseline for image classification[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2015, 24(12):5017-5032.
- [33] 刘梓, 宋晓宁, 於东军, 等. 基于多成分字典和稀疏表示的超分辨率重建算法[J]. *南京理工大学学报*, 2014, 38(1):1-5.  
Liu Zi, Song Xiaoning, Yu Dongjun, et al. Super-resolution reconstruction algorithm based on multi-component dictionary and sparse representation[J]. *Journal of Nanjing University of Science and Technology*, 2014, 38(1):1-5.
- [34] 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.
- [35] Lian Qiusheng, Zhang Junqin, Chen Shuzhen. Single image super-resolution algorithm based on two-stage and multi-frequency -band dictionaries[J]. *Acta Automatica Sinica*, 2013, 39(8): 1310-1320.
- [36] Zelnik-Manor L, Rosenblum K, Eldar Y C. Dictionary optimization for block-sparse representations[J]. *IEEE Transactions on Signal Processing*, 2012, 60(5): 2386-2395.
- [37] Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//*Proceedings of the 2012 British Machine Vision Conference*. Guildford: British Machine Vision Association, 2012:1-10.
- [38] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[J]. In *Curves and Surfaces*, 2012:711-730.
- [39] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//*IEEE International Conference on Computer Vision (ICCV)*. [S. l.]: IEEE, 2001:416-423.

#### 作者简介:



赵丽玲(1978-),女,博士研究生,研究方向:图像处理与超分辨技术, E-mail: zhaoliling@nuist.edu.cn.



孙权森(1963-),男,教授,研究方向:模式识别、图像处理、遥感信息系统。



张泽林(1993-),男,硕士研究生,研究方向:深度学习理论及应用。