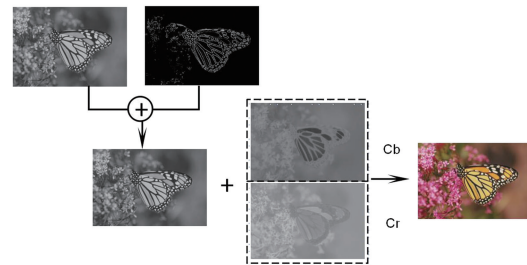


DOI: 10.12086/oe.2018.170542

融合特征分类和独立字典训练的超分辨率重建

汪荣贵, 汪庆辉*, 杨娟, 胡敏

合肥工业大学计算机与信息学院, 安徽 合肥 230009



摘要: 超分辨率重建技术在重构图像细节、改善图像视觉效果方面具有重要作用。为进一步提高图像的重建质量, 提出了一种有效的超分辨率重建方法。首先提取图像块的几何特征来构造决策树, 以期通过监督的方式进行图像块分类。然后针对不同类型的图像块训练集, 分别基于 K-SVD 独立训练相应的高分辨率字典和低分辨率字典。最后为了保证图像块的准确和快速重建, 对高分辨率训练集和低分辨率训练集的系数求解映射矩阵, 其用于在重建阶段将低分辨率稀疏系数映射为高分辨率稀疏系数以达到重建目的。实验结果表明, 本文的方法与其他经典的超分辨率重建方法相比, 在重建效果方面具有明显提高。

关键词: 超分辨率; 字典训练; 稀疏表示; 决策树; 特征分类

中图分类号: TP391

文献标志码: A

引用格式: 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

Image super-resolution reconstruction by fusing feature classification and independent dictionary training

Wang Ronggui, Wang Qinghui*, Yang Juan, Hu Min

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

Abstract: Super-resolution reconstruction plays an important role in reconstructing image detail and improving image visual effects. A new effective super-resolution method is proposed. Firstly, we extract the geometric features of the image patch to construct the decision tree, which will be used in patch classification in a supervised way. Then, we train the high-resolution and low-resolution dictionaries based on K-SVD independently for different types of training sets. Finally, we solve the mapping matrix for the coefficients between the high-resolution and low-resolution training set, which are used to map the low-resolution coefficients to high-resolution coefficients during the reconstruction phase to ensure accurate and fast reconstruction of the image patches. The experimental results show that the proposed method has a significant improvement in the reconstruction effect compared with other classic methods.

Keywords: super-resolution; dictionary training; sparse coding; decision tree; feature classification

Citation: 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.

收稿日期: 2017-10-14; 收到修改稿日期: 2017-12-20

基金项目: 国家自然科学基金项目 (61672202)

作者简介: 汪荣贵 (1966-), 博士, 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: 汪庆辉 (1992-), 硕士研究生, 主要从事数字图像处理、数据挖掘、机器学习等研究。E-mail: wqh1649759610@163.com

1 引言

图像超分辨率重建技术是指从输入的一幅或多幅低分辨率图像重建出高分率图像的技术。由于高分率图像具有更加丰富而细腻图像细节信息,故超分辨率重建技术被广泛用于卫星遥感^[1], 医疗诊断^[2]等领域。

超分辨率重建算法主要可以分为三类: 基于插值、重建和学习的算法。基于插值^[3-4]的算法主要利用待重建像素和邻域像素之间的关系来推测待重建像素的值。插值算法可以实现图像的实时重建,但是重建出的图片往往过于平滑,很难得到较好的效果。基于重建^[5-6]的算法一般基于重建模型加入先验知识,使得基于重建的算法在边缘保持和人工效应的抑制方面有着较好的效果,但是高频细节信息往往还是不能有效地被重建出来,低分辨率图像和高分率图像缩放尺寸越大,重建效果越差。与前两种方法相比,由于基于学习的方法能够产生低分辨率图像中所没有的高频细节信息,因此,近年来此类方法吸引了众多研究者的关注。在此方法中,通过学习低分辨率图像和对应的高分辨率图像之间的关系,来预测目标高分率图像。Yang^[7]等首先提出了利用字典进行超分辨率重建的方法,即分别利用高低训练集训练相应的高低分辨率字典,然后用不同的字典原子组合来表示任何一种图像块的几何结构。在这类超分辨率重建的方法中,非常关键的阶段就是字典训练过程,

许多研究者对此展开研究并作出相应改进。Yang^[8]等提出了联合字典训练,其利用单一字典训练的模式将低分辨率和高分率字典放在一起进行训练得到相应的字典对。Yang^[9]等利用低分辨率字典对低分辨率训练集进行稀疏表示并优化高分率字典,然后利用低分辨率训练集自身的几何信息优化低分辨率字典,不断迭代直到获得较好的高低分辨率字典对。Xu^[10]等基于 K-SVD 方法在稀疏编码阶段利用高低分辨率字典对相应训练集轮流稀疏编码的方法,迭代式地训练出高低分辨率字典对。Zeyde^[11]和 Timofte^[12-13]等利用 K-SVD^[14]方法训练出低分辨率字典,然后基于系数相等的原则,利用最小二乘的方式求解得到高分率字典。

以上字典训练方式均是基于高低分辨率系数不变性原则,这样训练的字典往往对于字典原子的特征表达能力具有一定的限制性。事实上,这两种稀疏系数可能并不是严格相等的。为了能够以一种更加准确并且灵活的方式训练字典,让字典原子能够充分发挥相应的表达能力,本文基于 0 范数原则,利用 K-SVD 方法进行独立字典训练以期能够重建出质量更好的图像。

另一方面,经典的字典训练是利用整个训练集直接训练高低分辨率字典对,这两类字典中的原子无差别地排列在字典之中,这样会造成各种类型的特征原子在重建图像块的过程中相互影响。所以本文依据图像块自身固有的几何特征进行分类,采用分类

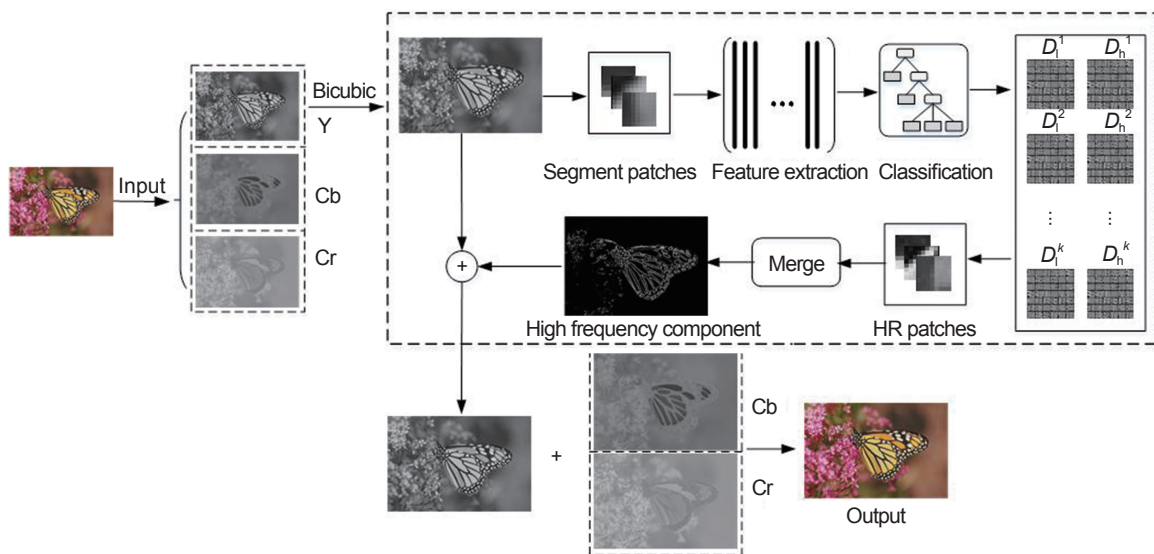


图 1 超分辨率重建过程图

Fig. 1 The flowchart of super-resolution reconstruction

重建的思想进一步优化。Yang^[15], Yang^[16] 和 Dong^[17] 利用 K-means 算法将训练集进行分类获得 K 对字典, 从实验结果上看, 利用 K 对字典进行重建取得了更加理想的重建结果。然而由于 K-means 的非监督性, 难以取得更加理想的分类效果。本文将图像的几何特征加入到分类算法中, 设计了一种基于决策树的监督型分类算法。实验表明, 与其他经典的重建算法相比, 本文提出的方法在视觉感官和评价指标方面都具有明显的提高。

图 1 展示了本文方法的全局重建流程, 首先提取图像 Y 通道的分量, 并在插值重建后, 分割图像块, 提取相应特征。然后将特征输入到决策树中进行分类, 针对每一类别训练字典并重建图像获得高频分量, 最后将重建出的高频分量与 Cb 和 Cr 通道的分量进行融合, 获得最终的高分辨率重建图像。

2 数据准备与训练集分类

2.1 数据预处理

首先, 与色度和饱和度相比, 人眼对亮度具有更强的敏感性, 所以只对亮度通道进行超分辨率重建, 即对原始图像进行提取 Y 通道得到图像 X 。

其次, 需要得到两类图像用于字典训练——高分辨率和低分辨率图像, 所以对图像 X 以采样因子 $s(s \geq 1)$ 进行下采样得到图像 X_s , 然后对 X_s 以采样因子 s 进行上采样得到低分辨率图像 X_l , 将图像 X_l 按照从上到下, 从左到右的次序进行分块得到 $X_l = \{x_l^1, x_l^2, \dots, x_l^N\}$ 。另一方面, 提取图像 X 的高频作为高分辨率图像 X_h , 提取方法为: $X_h = X - X_l$, 然后按照同样的方式对 X_h 进行分块得到 $X_h = \{x_h^1, x_h^2, \dots, x_h^N\}$ 。

为了后期能够将训练集用于分类和字典训练中, 即能够更好地计算块间距离和增加稀疏编码的准确度, 根据下面的滤波器, 对每个低分辨率图像块提取四组特征数据, 并将这 4 个特征数据首尾依次相接形成该图像块的一个特征向量: $f_1 = [1, 0, 0, -1]$, $f_2 = [1, 0, 0, -2, 0, 0, 1]$, $f_3 = f_1^T$, $f_4 = f_2^T$, 式中 f_1 和 f_2 是分别用于求取图像水平方向的一阶和二阶导数, f_3 和 f_4 是其对应的转置, 分别用于求取图像垂直方向的一阶和二阶导数。

最后, 特征向量的首尾相接造成了低分辨率特征向量的维度过大, 为了加快超分辨率重建的效率, 在保持原图像特征块 99.99% 的能量下, 利用主成分分析 (principal component analysis, PCA) 方法进行降维,

并将结果作为最终的低分辨率训练集。

2.2 利用决策树分类训练集

在对训练集中的图像进行分块之后, 提取图像块自身的几何特征并依此构造分类决策树。本文将图像块大致归为 3 类: 平滑块、方向主导块和随机方向块, 其中平滑块是指训练集中几何结构平滑、块中像素值相差不大的块; 方向主导块是指训练集中具有某个明确方向的块, 这种图像块多出现于图像的纹理过渡部分; 随机方向块介于上述两种块中间, 它既没有前者的平滑的纹理, 也没有后者明确的方向, 往往具有多个方向或者异于块中整体纹理的随机点, 具体分析如下:

1) 平滑块最大的特点在于块内像素值差别不大, 这导致平滑块的方差 τ 与另外两者相比较小, 因此将方差作为一个划分训练集类别的图像特征。

2) 随机方向块和方向主导块最大的不同在于块内是否具有明显的边缘方向, 很难通过简单的统计信息对二者进行区分, 文中借鉴了 Feng^[18] 中分解梯度场的思想来提取块内纹理的方向。若用 $f(x, y)$ 表示一个图像, 则此图像在点 (x_i, y_j) 处的梯度为:

$$g_k = \nabla f(x_i, y_j) = [f'_x(x_i, y_j), f'_y(x_i, y_j)], \quad (1)$$

式中: $f'_x(x_i, y_j)$ 和 $f'_y(x_i, y_j)$ 分别代表图像 $f(x, y)$ 在 x 和 y 方向的偏导, $1 \leq k \leq N$, N 为图像块中像素个数, 则对于每个图像块的局部梯度场 $G \in R^{n \times 2}$ 为

$$G = [g_1, g_2, \dots, g_N]^T. \quad (2)$$

对此梯度场 G 进行奇异值分解 (SVD) 可以得到: $G = USV^T$, 矩阵 S 中的奇异值 σ_1 和 σ_2 代表了梯度场主次方向的能量值, V 中第 1 列 v_1 和第 2 列 v_1 分别为梯度场的主次方向。根据能量值 σ_1 和 σ_2 定义用于图像块的梯度特征 γ 为

$$\gamma = \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}, \quad (3)$$

由式 (3) 可得 $\gamma \in [0, 1]$, 若两种能量值相差越大, 图像特征 γ 越大, 则图像块为方向主导块的概率越大; 反之, 图像块为随机方块的概率越大。

3) 对于方向主导块来讲, 将梯度场的主方向旋转 $\pi/2$ 即为块中纹理的主方向, 本文进一步利用图像块主方向与水平方向的角度 θ 对方向主导块进行分类, 同时将角度 θ 作为建立决策树的特征, 角度 θ 计算公式为:

$$\theta = \arctan \frac{v_1(y)}{v_2(x)} + \frac{\pi}{2}, \quad (4)$$

式中: $v_1(x)$ 和 $v_1(y)$ 分别为梯度场主方向的相应坐标值。

图 2 为建立的经验决策树, 其中在利用角度对方向主导块更细粒度的分类时, 依据步长 $\pi/6$ 将区间 $[0, \pi]$ 进行划分, 使得相应方向主导块能够落入相应的类别下, 具体分类方式为:

假设训练集为 $X = \{x_1, x_2, \dots, x_N\}$, 则对于每个图像块 x_i 都对应着一条特征记录 $x_i = [\tau_i, \gamma_i, \theta_i]$ 。从树中可以看出, 首先利用方差 τ_i 对训练集进行分类, 若 $\tau_i \leq \tau^*$, 则图像块 x_i 为平滑块, 其中 τ^* 为某个常数临界方差值; 否则利用分类信息 γ_i 进一步分类随机方向块和方向主导块, 若 $\gamma_i \leq \gamma^*$, 则方向块 x_i 为随机方向块, γ^* 为特征 γ 的分类临界值; 否则利用角度 θ_i 将方向主导块依据角度划分到不同的类别。

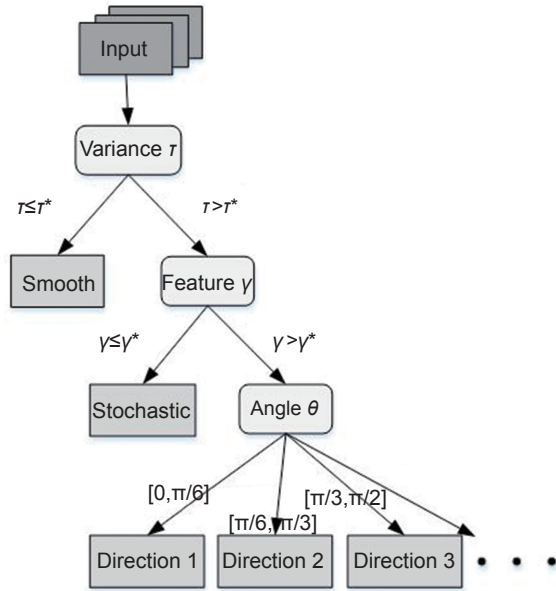


图 2 分类决策树
Fig. 2 The decision tree of classification

3 独立字典训练与图像重建

3.1 字典独立训练

许多经典的字典训练方式都是基于系数不变性的假设, 然而这往往对字典训练具有一定限制作用, 从而影响图像的重建。所以本文提出了字典独立训练的方式。

确切地说, 利用 K-SVD 方法分别训练每类训练集对—— $X_l^k = \{x_1^{l,k}, x_2^{l,k}, \dots, x_N^{l,k}\}$ 和 $X_h^k = \{x_1^{h,k}, x_2^{h,k}, \dots, x_N^{h,k}\}$,

以期获得低分辨率子字典和高分辨率子字典, 相应公式为:

$$D_l^k, A_l^k = \arg \min_{D_l^k, A_l^k} \|x_i^{l,k} - D_l^k a_i^{l,k}\|_2^2 \text{ s.t. } \|a_i^{l,k}\|_0 \leq T, \quad (5)$$

$$i = 1, 2, \dots, N$$

$$D_h^k, A_h^k = \arg \min_{D_h^k, A_h^k} \|x_i^{h,k} - D_h^k a_i^{h,k}\|_2^2 \text{ s.t. } \|a_i^{h,k}\|_0 \leq T, \quad (6)$$

$$i = 1, 2, \dots, N$$

式中: $D_l^k \in R^{d_l \times n}$ 表示第 k 类训练集的低分辨率字典, $A_l^k \in R^{n \times N}$ 表示利用当前字典获得的本类低分辨率训练集的系数矩阵, $D_h^k \in R^{d_h \times n}$ 和 $A_h^k \in R^{n \times N}$ 分别表示相应的高分辨率字典及其系数矩阵, $a_i^{l,k}$ 表示低分辨率块 $x_i^{l,k}$ 的稀疏系数, $a_i^{h,k}$ 表示高分辨率块 $x_i^{h,k}$ 的稀疏系数, T 是稀疏度约束, 是一个整数常量。 d_l 表示低分辨率字典原子维度, d_h 表示高分辨率字典原子维度, n 表示字典原子个数, N 表示训练集中图像块的个数。

式 (5) 和式 (6) 表明在满足稀疏度 T 的条件下, 训练出的字典需要能够最优近似训练集中的所有块, 使得均方误差项达到最小。为了达到这个目的并训练出特征表达能力更准确、更完善的字典, 低分辨率字典 D_l^k 和高分辨率字典 D_h^k 分别依据式 (5) 和式 (6) 通过 K-SVD 进行独立训练得到。下面以低分辨率字典 D_l^k 为例描述一下训练细节。

1) 初始化: 随机选择训练集中的样例来初始化低分辨率字典 D_l^k 。

2) 稀疏编码阶段

对于给定的训练集 $X_l^k = \{x_1^{l,k}, x_2^{l,k}, \dots, x_N^{l,k}\}$, 利用当前低分辨率字典 D_l^k 对每个训练项基于式 (7) 进行稀疏编码, 最后将稀疏系数按列依次组合得到系数矩阵 A_l^k 。本文利用 OMP 追踪算法来解决式 (7) 中 0 范数的稀疏表示问题, 具体的稀疏编码模型为:

$$a_i^{l,k} = \arg \min_{a_i^{l,k}} \|x_i^{l,k} - D_l^k a_i^{l,k}\|_2^2 \text{ s.t. } \|a_i^{l,k}\|_0 \leq T, \quad (7)$$

$$i = 1, 2, \dots, N$$

$$a_i^{h,k} = \arg \min_{a_i^{h,k}} \|x_i^{h,k} - D_h^k a_i^{h,k}\|_2^2 \text{ s.t. } \|a_i^{h,k}\|_0 \leq T, \quad (8)$$

$$i = 1, 2, \dots, N$$

3) 字典更新阶段

利用稀疏编码阶段获得的系数矩阵 A_l^k , 按列更新当前字典中的每一项原子 d_i ($d_i \in D_l^k, 1 \leq i \leq n$), 以及更新与该原子相关的系数 A_l^i ($A_l^i \in A_l^k, 1 \leq i \leq n$), A_l^i 表示系数矩阵 A_l^k 中的第 i 行。字典的训练目标是 minimized 训练集和字典表示结果的均方误差:

$$\begin{aligned} D_1^k &= \arg \min \|X_1^k - D_1^k A_1^k\|_F^2 \\ &= \arg \min \left\| X_1^k - \sum_{j=1}^n d_j A_1^j \right\|_F^2. \end{aligned} \quad (9)$$

为了更新原子 d_i 及其相关系数 A_1^i , 定义误差矩阵 E_i 来指示原子 d_i 应该更新的内容, 即 E_i 则代表着子项 $d_i A_1^i$ 需要表示的成分:

$$E_i = X_1^k - \sum_{j=1, j \neq i}^n d_j A_1^j. \quad (10)$$

然后对 E_i 进行奇异值分解 (SVD), 选择具有最大能量的奇异值来最优逼近 E_i 。但是存在的问题是对 E_i 直接进行奇异值分解, 在更新与 d_i 相关的系数 A_1^i 时往往会造成系数的非稀疏性。所以建立索引 ω_i 用来指示 A_1^i 的非零项:

$$\omega_i = \{j \mid 1 \leq j \leq N, A_1^i(j) \neq 0\}, \quad (11)$$

根据式 (11) 建立一个 0~1 映射矩阵 $\Phi_i \in R^{N \times |\omega_i|}$, 使得系数 A_1^i 被映射之后只保留非零项, 即 $A_{R_i}^i = A_1^i \Phi_i$, 其中 $A_{R_i}^i$ 是 A_1^i 中的非零元素依次排列得到。利用相同的思想进行处理 E_i 可以得到: $E_{R_i}^i = E_i \Phi_i$, 然后对 $E_{R_i}^i$ 进行 SVD 分解后得到:

$$E_{R_i}^i = U \Lambda V^T = \sum_{j=1}^k u_j \sigma_j v_j^T, \quad (12)$$

其中: U 为左奇异向量 u 组成的矩阵, V 为右奇异向量 v 组成的矩阵, Λ 为 k 个奇异值 σ 组成对角阵。奇异值 σ_j 越大, 代表子项 $u_j \sigma_j v_j^T$ 与矩阵 $E_{R_i}^i$ 差别越小, 表示所占能量越大, 因此选择奇异值最大的那一项 $u_m \sigma_m v_m^T$ 来更新字典原子及其系数, 即用 u_m 更新 d_i , 用 $\sigma_m v_m^T$ 依据相应索引 ω_i 更新 A_1^i 。

不断重复步骤 2) 和步骤 3), 直到满足迭代停止条件为止。此时获得了低分辨率字典 D_1^k 和相应的稀疏系数矩阵 A_1^k , 以同样的方式依据式 (6) 训练高分辨训练集得到高分辨率字典 D_h^k 和系数矩阵 A_h^k , 其中在进行稀疏编码时利用式 (8)。

3.2 多类映射函数学习

经过决策树算法分类训练集之后, 获得了 K 个训练集对 $\{X_1^k, X_h^k\}_{k=1}^K$, 针对每类训练集训练字典, 可以得到一个用来表达该类的字典对 $\{D_1^k, D_h^k\}$ 和相应的系数矩阵 $\{A_1^k, A_h^k\}$ 。但是由于独立训练的关系, 导致低分辨率字典 D_1^k 表达一个低分辨率图像 Y 的稀疏系数与高分辨率字典 D_h^k 表达 Y 对应的高分辨率图像 X 的系数不一致。虽然字典对各自的训练集能够进行有效表示, 但如果继续利用低分辨率系数矩阵 A_1^k 来恢复高分辨率图像, 往往会造成重建结果不理想, 一

面是因为欠拟合和过拟合因素, 另一方面因为原子选择的因素。因此有必要基于系数矩阵 A_1^k 和 A_h^k 去求解它们的关系映射以达到更好的重建效果。

同一类训练集中的图像块具有类似的几何结构, 所以它们的稀疏系数也具有一定的相似性。在基于质量和效率的考虑下, 线性映射足以用来求解这个关系映射函数。本文利用最小二乘方法 (least square, LS) 来求解, 则第 k 类训练集的映射求解公式可写为

$$M_k = \arg \min_{M_k} \|A_h^k - M_k A_1^k\|_F^2. \quad (13)$$

式 (13) 的解析解为

$$M_k = A_h^k A_1^{kT} (A_1^k A_1^{kT})^{-1}. \quad (14)$$

对每类训练集中的系数矩阵应用式 (14) 即可获得相应的映射矩阵 $\{M_k\}_{k=1}^K$ 。

3.3 图像重建

假设给定待重建图像 $Y = \{y_1, y_2, \dots, y_N\}$, 利用决策树对每个待重建的块 y_i 进行分类, 若图像块 y_i 根据相应的分类特征划分到了第 k 类训练集, 则利用相应的低分辨率字典 D_1^k 对图像块 y_i 进行稀疏编码得到稀疏系数 a_i^1 , 此问题可以表示成:

$$a_i^1 = \arg \min \|y_i - D_1^k a_i^1\|_2^2 + \lambda \|a_i^1\|_0. \quad (15)$$

式 (15) 可以通过 OMP 算法进行逼近求解得到块 y_i 的稀疏系数 a_i^1 。根据预先计算的本类训练集映射矩阵 M_k , 利用 a_i^1 来计算高分辨率字典 D_h^k 表示 y_i 的高分辨率版本 x_i 的系数 a_i^h :

$$a_i^h = M_k a_i^1. \quad (16)$$

重建模型为

$$x_i = D_h^k a_i^h, \quad (17)$$

则结合式 (14)、式 (16) 和式 (17), 可得到:

$$x_i = D_h^k M_k a_i^1 = D_h^k A_h^k A_1^{kT} (A_1^k A_1^{kT})^{-1} a_i^1. \quad (18)$$

由式 (18) 可知, 只需通过一次稀疏编码求出低分辨率稀疏系数 a_i^1 即可获得块 y_i 的重建块 x_i 。最后将所有重建的块 $x_1, x_2, \dots, x_N$ 进行拼接组合得到重建的高分辨率图像 X , 其中块重叠区域取像素均值。

4 实验结果与分析

为验证本文方法的有效性, 本文方法与 Yang^[8] 和 Zeyde^[11] 的基于稀疏编码的方法、Chang^[19] (NE+LLE) 和 Bevilacqua^[20] (NE+NNLS) 的基于最小二乘回归的方

法以及 Timofte^[12] 的锚定邻域回归 (ANR) 方法进行实验对比。其中, 在训练集分类阶段, 本文选用的方差阈值 τ 为 0.001, 当图像块放大 2 倍时, 对应高分辨率图像块尺寸为 6×6 , 分类特征 γ 阈值设置为 0.3; 当图像块放大 3 倍时, 对应高分辨率图像块尺寸为 9×9 , 分类特征 γ 阈值设置为 0.2; 当图像块放大 4 倍时, 对应高分辨率图像块尺寸为 12×12 , 分类特征 γ 阈值设置为 0.17。另外, 为保证比较好的重建结果, 设置稀疏度 T 为 20, 字典原子个数为 1024。用于测试的图像集如图 3 所示。

表 1 展示了图像在放大 3 倍时的重建结果, 本文选用了 PSNR、SSIM^[21]、VIF^[22] 和 FSIM^[23] 作为评价指标, 其数值均是越大, 表示图像重建效果越好。在 PSNR 方面, 本文的重建结果比 Yang^[8]、Zeyde^[11]、NE+LLE、NE+NNLS 以及 ANR 方法平均依次要高出 0.76、0.36、0.48、0.61、0.45。相应地, 在 SSIM 方面, 本文的重建结果平均依次要高出 0.0164、0.0060、0.0071、0.0108、0.0062; 在 VIF 和 FSIM 指标方面, 本文的重建结果依然比其他方法具有明显优势。综上所述, 表 1 的实验数据直接表明了本文方法的有效性。

为进一步验证本文算法的有效性, 表 2 展示了在图像放大倍数分别为 2、3、4 的情况下的重建结果, 其中每一项数据均为图 3 中所有图片重建结果的平均值。表中数据表明, 本文的重建结果在不同放大倍数情况下均比其他方法的重建结果质量高。具体来讲, 与 ANR 方法相比, 当放大倍数为 2 的时候, 本文方法的重建效果在 PSNR 方面比 ANR 要高出 0.21, 在 SSIM 方面要高出 0.0009; 在 VIF 方面要高出 0.0017; 在 FSIM 方面要高出 0.0019。同样地, 当

放大倍数为 3 的时候, 在 PSNR 方面要高出 0.45, 在 SSIM 方面要高出 0.0062, 在 VIF 方面要高出 0.0138, 在 FSIM 方面要高出 0.0067; 当放大倍数为 4 的时候, 在 PSNR 方面要高出 0.37, 在 SSIM 方面要高出 0.0077, 在 VIF 方面要高出 0.0109, 在 FSIM 方面要高出 0.0082。这基本可以说明在放大倍数较大时, 本文算法的可用性和稳定性相比其他方法更好一些。

另一方面, 为展示本文设计的决策树分类算法的优势, 分别利用决策树以及 K-means 算法分类训练集, 然后训练字典并重建图像。同时为保证对比的公平性, 同样利用 K-means 算法将训练集分为 8 类。表 3 展示了本次实验的比较结果, 从表中数据可以看出, 利用 K-means 分类重建的平均 PSNR 为 32.88, 平均 SSIM 为 0.9285, 本文方法的重建结果在 PSNR 方面比 K-means 重建结果高 0.34, 在 SSIM 方面比 K-means 重建结果高 0.0025, 在 VIF 方面高出 0.0087, 在 FSIM 方面高出 0.0030。实验数据对比结果表明本文的决策树分类效果比 K-means 更准确。

最后, 表 4 展示了关于时间方面和其他算法的对比试验, 表中数据为当图像放大倍数为 3 时, 重建图 3 中所有图片所需要的平均时间, 单位为 s。从表中数据可以看出, 本文方法的平均时间为 7.62 s, 这个时间快于 Yang^[8] 和 NE+NNLS, 但慢于其他方法, 这是因为本文算法需要对每个图像块进行稀疏编码而非像 ANR 那样利用离线式映射矩阵重建图像。

5 结 论

本文结合特征分类和独立字典训练提出了一种有效的超分辨率重建方法。利用了分类重建的思想,

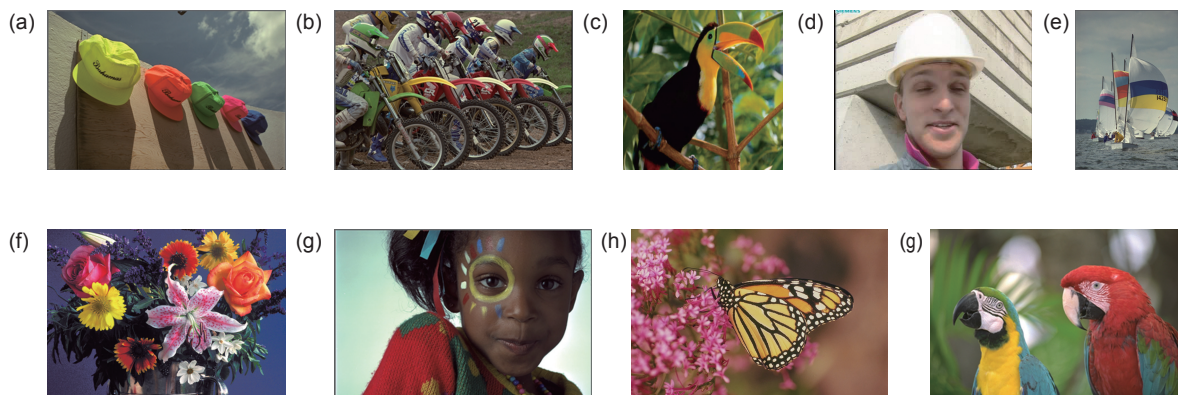


图 3 图像集。(a) 帽子; (b) 比赛; (c) 鸟; (d) 工长; (e) 帆船; (f) 花; (g) 女孩; (h) 蝴蝶; (i) 鹦鹉
Fig. 3 Image set. (a) Hats; (b) Match; (c) Bird; (d) Foreman; (e) Sailboats; (f) Flowers; (g) Girl; (h) Butterfly; (i) Parrots

表 1 图像放大 3 倍的重建结果数据表
Table 1 The experimental results with magnification 3

Image	Evaluation	Bicubic	Yang	Zeyde	NE+LLE	NE+NNLS	ANR	Proposed
hats	PSNR	32.07	33.11	33.50	33.29	33.24	33.24	33.88
	SSIM	0.8799	0.8900	0.9031	0.9004	0.8991	0.9002	0.9083
	VIF	0.4633	0.4688	0.5723	0.5671	0.5574	0.5664	0.5956
	FISM	0.8886	0.9055	0.9196	0.9161	0.9142	0.9153	0.9257
match	PSNR	23.39	24.17	24.19	24.22	24.14	24.26	24.51
	SSIM	0.6479	0.7081	0.7052	0.7091	0.7006	0.7117	0.7229
	VIF	0.2965	0.3287	0.3565	0.3596	0.3501	0.3634	0.3755
	FISM	0.7823	0.8246	0.8215	0.8247	0.8183	0.8265	0.8347
sailboats	PSNR	28.16	28.94	28.95	28.83	28.74	28.84	29.17
	SSIM	0.8390	0.8559	0.8635	0.8597	0.8582	0.8598	0.8686
	VIF	0.3759	0.3656	0.4465	0.4437	0.4354	0.4458	0.4599
	FISM	0.8513	0.8717	0.8758	0.8705	0.8684	0.8691	0.8800
girl	PSNR	30.37	31.33	32.21	31.82	31.71	31.76	32.31
	SSIM	0.8734	0.8694	0.8906	0.8868	0.8821	0.8867	0.8927
	VIF	0.4424	0.4358	0.5315	0.5292	0.5203	0.5333	0.5483
	FISM	0.8950	0.8944	0.9215	0.9148	0.9114	0.9126	0.9206
parrots	PSNR	30.66	31.23	31.40	31.40	31.31	31.43	31.77
	SSIM	0.9042	0.9056	0.9181	0.9182	0.9154	0.9190	0.9216
	VIF	0.5170	0.4733	0.5914	0.5952	0.5820	0.5981	0.6037
	FISM	0.9323	0.9350	0.9452	0.9453	0.9421	0.9457	0.9482
bird	PSNR	32.58	34.11	34.51	34.50	34.23	34.53	35.08
	SSIM	0.9256	0.9391	0.9476	0.9474	0.9429	0.9485	0.9522
	VIF	0.5504	0.5842	0.6555	0.6576	0.6399	0.6637	0.6725
	FISM	0.9283	0.9449	0.9498	0.9493	0.9450	0.9502	0.9548
flowers	PSNR	27.23	28.25	28.42	28.35	28.21	28.45	28.68
	SSIM	0.8013	0.8297	0.8376	0.8377	0.8327	0.8396	0.8436
	VIF	0.3683	0.3988	0.4467	0.4463	0.4365	0.4516	0.4613
	FISM	0.8445	0.8713	0.8760	0.8745	0.8699	0.8764	0.8818
foreman	PSNR	31.18	32.04	33.16	33.10	32.94	33.15	33.72
	SSIM	0.9058	0.9132	0.9294	0.9284	0.9251	0.9292	0.9355
	VIF	0.5309	0.5326	0.6348	0.6392	0.6222	0.6449	0.6605
	FISM	0.9123	0.9287	0.9385	0.9383	0.9349	0.9388	0.9455
butterfly	PSNR	29.43	30.71	31.09	30.89	30.74	31.02	31.63
	SSIM	0.9198	0.9290	0.9380	0.9357	0.9339	0.9371	0.9423
	VIF	0.4600	0.4576	0.5461	0.5457	0.5318	0.5513	0.5658
	FISM	0.9658	0.9750	0.9820	0.9808	0.9799	0.9816	0.9856
average	PSNR	29.45	30.43	30.83	30.71	30.58	30.74	31.19
	SSIM	0.8552	0.8711	0.8815	0.8804	0.8767	0.8813	0.8875
	VIF	0.4450	0.4495	0.5312	0.5315	0.5195	0.5354	0.5492
	FISM	0.8889	0.9057	0.9144	0.9127	0.9094	0.9129	0.9196

表 2 不同放大倍数实验结果数据表
Table 2 The experimental results with different magnifications

Scale	Evaluation	Bicubic	Yang	Zeyde	NE+LLE	NE+NNLS	ANR	Proposed
× 2	PSNR	32.33	26.08	34.35	34.31	34.07	34.36	34.57
	SSIM	0.9205	0.9128	0.9421	0.9421	0.9402	0.9432	0.9441
	VIF	0.6484	0.5985	0.7574	0.7605	0.7496	0.7639	0.7656
	FISM	0.9406	0.9184	0.9613	0.9612	0.9596	0.9614	0.9633
× 3	PSNR	29.45	30.43	30.83	30.71	30.58	30.74	31.19
	SSIM	0.8552	0.8711	0.8815	0.8804	0.8767	0.8813	0.8875
	VIF	0.4450	0.4495	0.5312	0.5315	0.5195	0.5354	0.5492
	FISM	0.8889	0.9057	0.9144	0.9127	0.9094	0.9129	0.9196
× 4	PSNR	27.86	27.84	28.98	28.86	28.77	28.92	29.29
	SSIM	0.8028	0.7986	0.8297	0.8278	0.8248	0.8295	0.8372
	VIF	0.3189	0.2881	0.3888	0.3893	0.3800	0.3935	0.4044
	FISM	0.8490	0.8527	0.8755	0.8732	0.8706	0.8744	0.8826

表 3 与 K-means 分类对比的实验数据
Table 3 The experimental results compared with K-means

Type	Evaluation	Hats	Pparrots	Bird	Foreman	Butterfly	Average
Kmeans	PSNR	33.52	31.47	34.73	33.34	31.36	32.88
	SSIM	0.9035	0.9191	0.9492	0.9311	0.9396	0.9285
	VIF	0.5784	0.5999	0.6675	0.6460	0.5626	0.6109
	FISM	0.9210	0.9465	0.9520	0.9417	0.9838	0.9490
Proposed	PSNR	33.88	31.77	35.08	33.72	31.63	33.22
	SSIM	0.9083	0.9216	0.9522	0.9355	0.9371	0.9310
	VIF	0.5956	0.6037	0.6725	0.6605	0.5658	0.6196
	FISM	0.9257	0.9482	0.9548	0.9455	0.9856	0.9520

表 4 本文算法关于时间方面实验
Table 4 The experiment about time of the proposed algorithm

	Bicubic	Yang	Zeyde	NE+LLE	NE+NNLS	ANR	Proposed
Time	0.00	40.53	1.56	3.03	14.23	0.64	7.62

并针对 K-means 分类的非监督性，提取分类特征并设计了基于决策树的分类方式，达到了比较好的分类效果。另一方面，本文提出的这种字典训练方式是按照相应训练集类别进行独立字典训练，并根据二者的映射关系实现了具有质量保证的图像重建，同时在重建期间只使用了一次稀疏编码，从而保证了效率。实验结果直接表明本的方法比其他经典方法具有更好的重建质量。

参考文献

[1] Wang L Z, Lu K, Liu P. Compressed sensing of a remote sensing

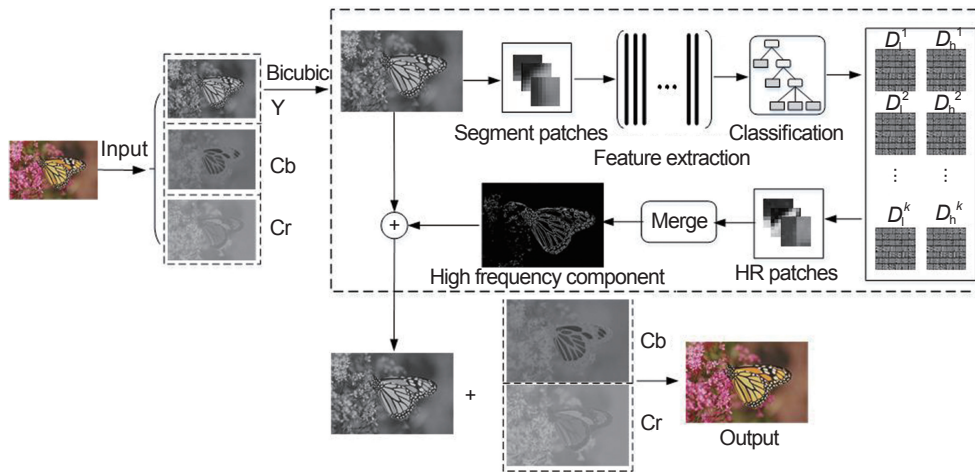
image based on the priors of the reference image[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, **12**(4): 736–740.
 [2] Wang Y H, Li J B, Fu P. Medical image super-resolution analysis with sparse representation[C]//*Proceedings of the 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2012: 106–109.
 [3] Li X, Orchard M T. New edge-directed interpolation[J]. *IEEE Transactions on Image Processing*, 2001, **10**(10): 1521–1527.
 [4] Zhang L, Wu X L. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. *IEEE Transactions on Image Processing*, 2006, **15**(8): 2226–2238.
 [5] Dai S Y, Han M, Xu W, et al. SoftCuts: a soft edge smoothness prior for color image super-resolution[J]. *IEEE Transactions on Image Processing*, 2009, **18**(5): 969–981.
 [6] Sun J, Shum H Y. Image super-resolution using gradient profile prior: US 2010/0086227 A1[P]. 2010-04-08.

- [7] Yang J C, Wright J, Huang T, *et al.* Image super-resolution as sparse representation of raw image patches[C]//*Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2008: 1–8.
- [8] 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.
- [9] Yang J C, Wang Z W, Lin Z, *et al.* Coupled dictionary training for image super-resolution[J]. *IEEE Transactions on Image Processing*, 2012, **21**(8): 3467–3478.
- [10] Xu J, Qi C, Chang Z G. Coupled K-SVD dictionary training for super-resolution[C]//*Proceedings of the 2014 IEEE International Conference on Image Processing*, 2014: 3910–3914.
- [11] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]//*Proceedings of the 7th International Conference on Curves and Surfaces*, 2010: 711–730.
- [12] Timofte R, De V, Van G L. Anchored neighborhood regression for fast example-based super-resolution[C]//*Proceedings of the 2013 IEEE International Conference on Computer Vision*, 2013: 1920–1927.
- [13] Timofte R, De S V, Van G L. A+: adjusted anchored neighborhood regression for fast super-resolution[C]//*Proceedings of the 12th Asian Conference on Computer Vision*, 2014: 111–126.
- [14] 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.
- [15] Yang W M, Tian Y P, Zhou F, *et al.* Consistent coding scheme for single-image super-resolution via independent dictionaries[J]. *IEEE Transactions on Multimedia*, 2016, **18**(3): 313–325.
- [16] Yang S Y, Liu Z Z, Wang M, *et al.* Multitask dictionary learning and sparse representation based single-image super-resolution reconstruction[J]. *Neurocomputing*, 2011, **74**(17): 3193–3203.
- [17] Dong W S, Li X, Zhang L, *et al.* Sparsity-based image denoising via dictionary learning and structural clustering[C]//*Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, 2011: 457–464.
- [18] Feng X G, Milanfae P. Multiscale principal components analysis for image local orientation estimation[C]//*Conference Record of the 36th Asilomar Conference on Signals, Systems and Computers*, 2002, 1: 478–482.
- [19] 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.
- [20] Bevilacqua M, Roumy A, Guillemot C, *et al.* Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//*Proceedings of British Machine Vision Conference (BMVC)*, 2012.
- [21] Wang Z, Bovik A C, Sheikh H R, *et al.* Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612.
- [22] Sheikh H R, Bovik A C. Image information and visual quality[J]. *IEEE Transactions on Image Processing*, 2006, **15**(2): 430–444.
- [23] Zhang L, Zhang L, Mou X Q, *et al.* FSIM: a feature similarity index for image quality assessment[J]. *IEEE Transactions on Image Processing*, 2011, **20**(8): 2378–2386.

Image super-resolution reconstruction by fusing feature classification and independent dictionary training

Wang Ronggui, Wang Qinghui*, Yang Juan, Hu Min

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



The flowchart of super-resolution reconstruction

Overview: Super-resolution reconstruction plays an important role in reconstructing image detail and improving image visual effects. It aims at reconstructing a high-resolution image from one or multiple low-resolution images. Because the high-resolution image contains more details than the low-resolution image, the high-resolution image possesses more values in remote sensing, medical diagnostic and so on. The super-resolution reconstruction can be divided into three categories: interpolation-based methods, reconstruction-based methods and learning-based methods. Because the learning-based methods can generate high-frequency details which are not available in the low-resolution, it attracts the attention of many researchers, and gradually becomes the main method of super-resolution reconstruction. We proposed a new effective super-resolution method based on the classification reconstruction and independent dictionary training. Most of the current classification of reconstruction methods are based on kmeans, but it tend to be difficult to further improve quality of reconstruction images due to the way of unsupervised clustering. So we analyze different image patches and three features including variance, and gradient feature, which are used to generate the decision tree for classification. Variance is used to distinguish smooth patches and other patches, gradient feature is used to distinguish stochastic patches and direction patches, and the last one is used to further distinguish directions. The experimental results show that our classification method has better classification results than kmeans. On the other hand, we train the high and low-resolution dictionaries based on K-SVD independently for different types of training sets. However, there is an obvious difference between the low-resolution coefficient and the high-resolution coefficient due to the independent dictionary training. Although the dictionary can represent respective training set well, the reconstructed images tend to be inferior if we insist on replacing the high-resolution coefficient using the low-resolution one. So it is necessary to solve their mapping function in order to achieve the better reconstruction effect. The mapping function is used to map the low-resolution coefficients to high-resolution coefficients during the reconstruction phase to ensure accurate and fast reconstruction of the image patches. We show the experimental results in detail when the magnification is 3, the experiment with different magnification and the comparative experiment with K-means. In general, the experimental results show that the proposed method has a significant improvement in the reconstruction effect compared with other classic methods.

Citation: 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.

Supported by National Natural Science Foundation of China (61672202)

*E-mail: wqh1649759610@163.com