

稀疏字典编码的超分辨率重建*

李 民^{1,3+}, 程 建^{1,2}, 乐 翔², 罗环敏¹

¹(电子科技大学 地表空间信息技术研究所, 四川 成都 611731)

²(电子科技大学 电子工程学院, 四川 成都 611731)

³(桂林空军学院 科研部, 广西 桂林 541003)

Super-Resolution Based on Sparse Dictionary Coding

LI Min^{1,3+}, CHENG Jian^{1,2}, LE Xiang², LUO Huan-Min¹

¹(Institute of Geo-Spatial Information Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, China)

²(School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

³(Department of Scientific Research, Guilin Airforce Academy, Guilin 541003, China)

+ Corresponding author: E-mail: gllm126@163.com

Li M, Cheng J, Le X, Luo HM. Super-Resolution based on sparse dictionary coding. *Journal of Software*, 2012, 23(5): 1315-1324. <http://www.jos.org.cn/1000-9825/3989.htm>

Abstract: Learning-Based super-resolution methods usually select several objects with similar features from some examples according to the low-resolution image, then estimate super-resolution result using optimization algorithm. But the result is usually limited by the quality of matching objects and only geometric construction of the images is selected as matching feature, so matching accuracy is relatively low. This paper presents a sparse dictionary model for image super-resolution, which unifies the feature patches of high-resolution (HR) and low-resolution (LR) images for sparse coding. To break through the aforementioned limitations, this method builds a sparse association between HR and LR images, and realized simultaneous matching and optimization methods. The study uses a MCA method to improve the accuracy for feature extraction and carry out super-resolution reconstruction and denoise simultaneously. Sparse K-SVD algorithm is adopted as optimization method to reduce the computation time of sparse coding. Some experiments with real images show that this method outperforms other learning-based super-resolution algorithms.

Key words: super resolution; sparse dictionary; learning-based; morphological component analysis (MCA); sparse K-SVD

摘 要: 基于学习的超分辨率方法通常根据低分辨率图像从样本库中选取若干特征相似的匹配对象,再使用优化算法进行超分辨率估计,但其结果受匹配对象的质量限制,并且匹配特征一般只选择图像的几何结构信息,匹配准确性较低.提出了稀疏字典编码的超分辨率模型,将高、低分辨率图像特征块统一进行稀疏编码,建立高、低分辨率图

* 基金项目: 国家重点基础研究发展计划(973)(2007CB714406); 中国博士后基金(200902609); 电子科技大学青年科技基金(JX0804)

收稿时间: 2010-06-06; 定稿时间: 2011-01-31

CNKI 网络优先出版: 2011-05-26 13:38, <http://www.cnki.net/kcms/detail/11.2560.tp.20110526.1338.004.html>

像的稀疏关联,同步实现匹配搜索和优化估计,突破了上述方法的限制.应用形态分量分析法提取图像的特征数据,提高了特征匹配的准确性,并同步实现超分辨率重建和降噪功能.优化方法采用稀疏K-SVD算法以提高稀疏字典编码的计算速度.采用自然图像进行实验,与其他基于学习的超分辨率算法相比,重建所得到的图像质量更优.

关键词: 超分辨率;稀疏字典;基于学习;形态分量分析;稀疏K-SVD

中图法分类号: TP391 **文献标识码:** A

高分辨率图像在现实生活中具有很大的应用价值,如医学图像诊断、遥感影像、高分辨率视频等领域.但由于成像器件物理条件的限制,有时获取高分辨率图像难度很大、成本很高.目前,获得高分辨率图像的主要方法之一是超分辨率图像重建(简称超分).近年来,基于重建的多帧超分辨率方法和基于学习的单帧超分辨率方法^[1-3]得到更多的关注.多帧超分重建的基本思想是,应用信号处理技术,从多帧低分辨率图像中提取高频信息来重建高分辨率图像^[4,5].重建方法建立图像降质的物理模型,假设原始高分辨率图像经几何变换、模糊和降采样操作后得到低分辨率图像,最终利用多帧低分辨率图像来反演高分辨率图像.由于降质模型中的参数很难确定,特别是图像间亚像素级的配准参数,因此这类方法提高图像分辨率的能力有限^[4].基于学习的方法是近几年的研究热点,它是通过对高、低分辨率图像特征块的学习,建立它们之间的关系,将这种关系作为先验知识来指导超分辨率重建.

然而,当前基于学习的超分算法缺乏有效的学习方法来表示这种先验知识,从先验知识库中找出相匹配的信息非常困难.为突破这个问题,Yang等人提出一种基于稀疏编码(sparse coding,简称SC)的学习算法,建立高、低分辨率图像块之间的稀疏关联,取得了非常好的效果^[6].这种自适应稀疏表示的字典采用显式字典表示,非常灵活,但缺乏正则性和有效性.Rubinstein等人提出一种参数字典模型来平衡效率和自适应性,并定义为稀疏字典(sparse dictionary,简称SD)^[7],它是将字典的每个原子通过固定的基字典再次稀疏分解.这种模型结构简单,并保持了足够灵活性.本文在这种稀疏字典的基础上,结合稀疏表示的超分模型,提出一种基于稀疏字典编码的超分重建算法,在重建效果和算法效率方面都有较大的提升.本文主要贡献如下:

- 采用稀疏字典编码方法,建立高低分辨率图像特征块间的稀疏关联关系.在保持稀疏字典自适应性和灵活性的同时,加强了其正则性和有效性.与其他基于学习的超分辨率算法相比,提高了图像的超分辨率质量.采用稀疏K-SVD算法提高稀疏字典编码的计算速度,加快了稀疏字典库的建立.
- 训练样本的构成中:(1)采用高分辨率图像块的高频分量.与Yang等人的稀疏编码超分算法中直接使用原始高分辨率图像块相比,所需样本数更少,超分重建效果更好;(2)应用形态分量分析法(morphological component analysis,简称MCA)获取低分辨率图像特征,更有效地利用了图像的几何结构和纹理特征信息,提高了匹配的准确性,同时实现了超分辨率重建和降噪功能.

本文第1节介绍基于学习的超分方法的背景知识及相关研究.第2节详细说明基于稀疏字典的超分模型及算法.相关实验结果在第3节列出.第4节对本文进行总结,并说明下一步的研究内容.

1 基于学习的超分辨率方法

基于学习的超分方法最早是由Freeman等人^[1,2]提出的,假设图像由3种形态分量组成:高频分量 h 、中频分量 m 和低频分量 l ,并假定低分辨率图像是其对应的高分辨率图像损失了高频分量的结果.基于学习的超分方法的目的就是从低频和中频分量信息中恢复高频信息,即最大化概率 $p(h|m,l)$, h 与 l 相对独立,所以简化为 $p(h|m)$.

在人脸图像超分辨率方面,Baker和Kanade^[4]第一次提出“虚幻脸(face hallucination)”的思想,引入图像的梯度先验信息,即图像的拉普拉斯金字塔、高斯金字塔的一阶和二阶梯度作为特征空间进行训练.这种方法可以说是第2类基于学习的超分辨率算法,主要针对人脸图像.随后,Liu等人也提出相似的思路,他们的算法结合了全局的人脸结构图像和局部的特征图像^[8].

Hertzmann等人^[9]和Efros等人^[10]应用局部特征变换方法,根据训练图像间的局部相似性来指导输入图像

推理出类似图像.这种图像类推(image analogies,简称 IA)^[9]的思想也可应用于超分辨率重建,可以说是第 3 类基于学习的超分辨率算法.

继文献[2,4,9]之后,出现了基于学习的超分辨率算法的热潮.Rajaram 等人^[11]提出一种无参数信任追踪算法,将图像块定义为无方向图模型,其中的关联函数采用无参数核密度来表示.Chang 等人^[12]将多个优选结果进行加权合成,采用局部线性嵌入(local linear embedding,简称 LLE)的流形学习算法对加权系数进行估计,取得了较好的超分辨率效果.Sun 等人^[13]从样本图像提取轮,并将其作为先验信息来增强输入图像的高频细节.Ni 等人^[14]提出支持向量机(support vector machine,简称 SVM)的超分辨率方法,在 DCT 域内将超分辨率问题转换图像正确性准则的估计问题.Kim 等人^[15]将稀疏回归模型应用于单帧图像的超分辨率重建,结合稀疏编码和核岭回归(kernel ridge regression,简称 KRR)的方法,更好地对高、低分辨率样本图像对进行组织,取得了较好的效果.国内学者在此方面也做了大量研究.中国科学院软件研究所的吴恩华、古元亭^[16]借鉴图像类推的思想,利用图像局部信息在不同尺度方面的传递特性,为低分辨率图像补充结构信息.中山大学的潘琪等人^[17]提出的学习算法针对医学图像,应用了图像的局部几何相似性,并结合了非局部平均的思想.

然而,当前大多数算法的思路是:首先从样本库中匹配搜索与输入图像块特征相似的 1 个或 K 个相似块;再从 K 个相似块优选 1 个或对这 K 个相似块进行加权合成.但这类算法性能受 K 个优选样本的质量影响,并且没有充分利用其他样本中所包含的先验信息,限制了对输入图像和样本进行建模的灵活性.为突破这种限制, Yang 等人^[6]提出基于稀疏编码的超分辨率重建模型,将匹配搜索和加权合成两个步骤统一到一个框架中,最大限度地利用了样本库中所有样本的先验信息.这种方法中字典超完备,结构灵活,自适应能力强;超分重建的最终效果好,计算效率高.Wang 等人^[18]提出一种类似算法,整合了 Yang^[6]和 Freeman^[1]的思路,将中频和低频图像块进行稀疏编码,对图像损失的高频信息进行估计.

2 稀疏字典编码的超分辨率算法

稀疏编码的基本思想是假设自然信号能够被压缩表示,或者可由预定义的原子线性组合表示.设信号 $x \in \mathbb{R}^N$,其稀疏表示问题可以表述为 $\alpha = \arg \min_{\alpha} \|\alpha\|_0, \text{ s.t. } \|x - D\alpha\|_2 \leq \varepsilon$.其中, α 为 x 的稀疏表示, $D = \{d_1, d_2, \dots, d_L\} \in \mathbb{R}^{N \times L}$ ($L \geq N$) 称为字典, d_i 称为原子.这个模型中最根本的问题就是字典 D 的表示,一般有两种方法:基于分析的方法与基于学习的方法(或称基于合成).在基于分析的方法中,字典是隐式的,它通过一定的数据模型来表示,结构性较好,并有快速数值计算方法^[7],如小波(wavelet)、曲线波(curvelet)、轮廓波(contourlet)、条带波(bandelet)等.但这种字典结构一般是固定的,对不同类型的数据自适应能力较差;在基于学习的方法中,字典是从一些样本中通过机器学习推理而来,通常由一显式矩阵表示.这种方法自适应能力更强,能够得到更精细的字典表示,有更好的性能应用.但得到的字典缺乏正则性和有效性,存在冗余,计算效率较低,且复杂的约束条件限制了字典的大小和处理信号的维数.Rubinstein 等人^[7]提出一种稀疏字典的方法,有效地整合了上述两种方法,并继承了各自的优点.它是将字典 D 中每个原子通过一基字典 Φ 来表示的一种方法: $D = \Phi A$.这种字典模型通过 A 的调整,自适应性更强;通过 Φ 的选择,稀疏结构更有效、紧凑,减小了字典训练学习过程中的过拟合和不稳定性.

本文在稀疏字典表示模型的基础上,结合基于稀疏编码的超分辨率思想^[6],提出稀疏字典表示的超分辨率模型和算法.同时,应用 MCA 方法提取图像的特征数据作为字典训练对象,提高了特征匹配的准确性,并且同时实现了超分辨率重建和降噪功能.

2.1 稀疏字典学习

根据训练样本 $X = \{x_1, x_2, \dots, x_r\}$,通过学习稀疏字典 A 来稀疏表示 X ^[7]

$$A = \arg \min_{A, Z} \|X - \Phi AZ\|_2^2, \text{ s.t. } \|z_i\|_0 \leq t, \|a_j\|_0 \leq p, \|\Phi a_j\|_2 = 1 \quad (1)$$

其中, Z 为 X 的稀疏表示, D 中每个原子(列向量) Φa_j 的 l_2 范数约束避免原子的尺度歧义性, Φ 为字典 D 稀疏表示的隐式字典(本文采用超完备 DCT).

为对高、低分辨率图像特征块统一进行稀疏关联学习,定义训练样本对 $P = \{X^h, Y^l\}$, $X^h = \{x_1, x_2, \dots, x_n\}$, $Y^l = \{y_1,$

y_2, \dots, y_n }, 每个样本对 (x_i, y_i) 都由图像特征块的列向量表示, x_i 表示高分辨率图像块的高频分量, y_i 表示低分辨率图像特征块, 特征的选择计算见第 2.3 节. 稀疏编码的目的就是通过对 P 的学习来估计稀疏字典, 并且将高分辨率图像块的高频分量和低分辨率图像特征块统一到一个稀疏字典编码框架中, 使它们具有相同的稀疏表示. 目标函数表示为

$$\min_{A^h, A^l, Z} \frac{1}{N} \|X^h - \Phi^h A^h Z\|_2^2 + \frac{1}{M} \|Y^l - \Phi^l A^l Z\|_2^2, \text{ s.t. } \|z_i\|_0 \leq t, \|a_j^h\|_0 \leq p, \|a_j^l\|_0 \leq p, \|\Phi^h a_j^h\|_2 = 1, \|\Phi^l a_j^l\|_2 = 1 \quad (2)$$

A^h 表示高分辨率图像块的高频分量的稀疏字典, A^l 对应低分辨率图像特征块的稀疏字典. N 和 M 分别是高、低分辨率图像块向量表示方式的维数, 它们作为公式(2)中的系数是为了平衡高、低分辨率图像块尺度问题对整个编码框架的影响. 公式(2)可简化表示为

$$\min_{A^h, A^l, Z} \|X - \Phi A Z\|_2^2, \text{ s.t. } \|z_i\|_0 \leq t, \|a_j\|_0 \leq p, \|\Phi a_j\|_2 = 1 \quad (3)$$

其中, $X = \begin{bmatrix} \frac{1}{\sqrt{N}} X^h \\ \frac{1}{\sqrt{M}} Y^l \end{bmatrix}$, $\Phi = \begin{bmatrix} \Phi^h & 0 \\ 0 & \Phi^l \end{bmatrix}$, $A = \begin{bmatrix} \frac{1}{\sqrt{N}} A^h \\ \frac{1}{\sqrt{M}} A^l \end{bmatrix}$. 本文采用稀疏 K-SVD 算法^[7]解公式(3), 迭代过程主要分两步:

稀疏编码阶段, 假定当前字典估计值固定, 估计稀疏表示 z_i , 算法与基本的 K-SVD 一致, 可采用任意的追踪算法实现; 字典修正阶段, 对稀疏字典的每个列向量 a_j 逐个修正. 具体算法流程见算法 1.

算法 1. 稀疏字典学习算法(稀疏 K-SVD^[7]).

1. 输入: 训练样本库 P , 稀疏字典的隐式基字典 Φ , 原子稀疏性最大值 p , 训练样本稀疏性最大值 t , 迭代次数 k .
2. 输出: 估计稀疏字典表示 A^h, A^l .
3. 初始化: 字典表示初始化值 $A^{h(0)}, A^{l(0)}$, $A = A^{(0)} = \begin{bmatrix} \frac{1}{\sqrt{N}} A^{h(0)} \\ \frac{1}{\sqrt{M}} A^{l(0)} \end{bmatrix}^T$.
4. 逐个对高、低分辨率图像特征样本对 (x_i, y_i) 计算:
 - 稀疏编码阶段: 固定 A , 可采用任何追踪算法估计 $\forall i: z_i = \arg \min_{z_i} \|p_i - \Phi A z_i\|_2^2, \text{ s.t. } \|z_i\|_0 \leq t$.
 - 字典修正阶段: 逐个对 A 中每列 a_j 进行修正:
 - ◇ $I = \{ \text{只应用到 } a_j \text{ 原子表示的的样例索引} \}$;
 - ◇ 标准化 $g^T = Z_{j,I}^T, g^T = g^T / \|g^T\|_2$;
 - ◇ 根据文献[7]中的引理 1 计算 a_j ;
 - ◇ 标准化 $\|\Phi a_j\|_2 = 1 \Rightarrow a_j = a_j / \|\Phi a_j\|_2$;
 - ◇ 修正 $g = (P_I^T - \Phi A z_i) \Phi a_j$.

2.2 稀疏字典编码的超分辨率模型

基本思路与文献[6]类似, 对输入低分辨率图像特征块 y (特征信息详见第 2.3 节), 根据字典 A^l 估计其稀疏表示 α , 再根据高分辨率图像块稀疏字典 A^h 和 α 估计高频分量 x, y 的稀疏表示问题可以表示为

$$\min \|\alpha\|_0, \text{ s.t. } \|y - \Phi^l A^l \alpha\|_2^2 \leq \varepsilon \quad (4)$$

虽然公式(4)是 NP-Hard 问题^[19], 著名学者 Donoho 已证明, 只要 α 足够稀疏, 可以转换为 l^1 范数的最小化问题^[20]. 公式(4)对每个图像块单独进行处理, 并没有考虑相邻图像块间的匹配问题. 本文使用与文献[2]类似的 one-pass 算法, 对公式(4)进行改进:

$$\min \|\alpha\|_1, \text{ s.t. } \|y - \Phi^l A^l \alpha\|_2^2 \leq \varepsilon_1, \|R \Phi^h A^h \alpha - w\|_2^2 \leq \varepsilon_2 \quad (5)$$

R 用于提取当前估计的高频分量块与其相邻块的重叠区域, w 表示已估计高频分量在重叠区域的估计值. 最终的优化问题表示为

$$\alpha^* = \arg \min_{\alpha} \lambda \|\alpha\|_1 + \|\tilde{y} - \tilde{D}\alpha\|_2^2, \tilde{y} = \begin{bmatrix} \beta w \\ y \end{bmatrix}, \tilde{D} = \begin{bmatrix} \beta R \Phi^h A^h \\ \Phi^l A^l \end{bmatrix} \quad (6)$$

β 用于权衡低分辨率图像特征与估计的高频分量间的匹配权重.公式(6)可以通过任意追踪算法估计出 y 的最优稀疏表示 α^* ,最后通过 $x^* = \Phi^h A^h \alpha^*$ 得到高频分量的最终估计值,再与相应的中、低频分量线性组合得到超分辨率图像的初始估计 X_0 .

上述模型将 y 特征块的稀疏字典表示作为先验信息,在局部区域内指导高频分量的重建.为使最终估计 X_0 满足全局重建约束 $Y=DHX$ (其中, D 表示降采样操作, H 表示模糊滤波器^[21]),定义目标函数:

$$X^* = \arg \min_X \|X - X_0\|, \text{ s.t. } Y = DHX \quad (7)$$

上述优化问题可以采用反向投影算法(back-projection,简称 BP)^[6,21],将 BP 算法的结果 X^* 作为最终的超分辨率图像估计值.具体流程见算法 2.

算法 2. 稀疏字典超分辨率算法.

1. 输入:训练字典 A^l, A^h ,稀疏字典表示采用的隐式字典 Φ ,低分辨率观测图像 Y_0 .
2. 输出:高分辨率图像 X^* .
3. 初始化:计算 Y_0 的特征图像 Y .
4. 从 Y 的左上角逐个取 3×3 图像块 y ,并保留 1 个像素宽的重叠区域,根据公式(6),计算稀疏表示的最优估计值 α^* ,计算高频分量的最优估计值 $x^* = \Phi^h A^h \alpha^*$.
5. y 的尺度放大图像与估计出的高频分量 x^* 线性叠加,得到超分辨率重建的初始估计 X_0 .
6. 根据公式(7),计算最终高分辨率图像的最优估计值 X^* ,使其满足全局重建约束.

2.3 基于MCA的图像特征

从视觉角度分析,人类对图像中的高频信息非常敏感;并且,低分辨率图像中的高频信息对预测超分辨率图像的高频信息非常重要.在超分辨率重建过程中,通常选择高频信息作为图像特征.通用的方法有拉普拉斯变换^[2]、高斯微分滤波器^[13]、提取一阶和二阶梯度特征^[6,12]等.但这些图像特征主要是利用了图像的几何结构信息,并没有考虑到图像的纹理特征,因此,表示图像的视觉特征信息并不全面,也没有考虑噪声对特征的影响.本文使用 MCA 方法改进上述问题.

MCA 模型是采用两种弱相关性的字典来表示图像中两种重要的视觉信息:几何结构信息和纹理信息.在 MCA 框架中,一个重要的假设是图像的几何结构分量和纹理分量在某个特定的基库或过完备子字典下是类内稀疏的,而各形态分量稀疏表示的基库或过完备子字典之间具有弱相干性.正是这种结构分量和纹理分量的分类稀疏表示的强稀疏性,使图像形态分量达到有效的分离.

设有图像 I ,选择两个字典,一个表示几何结构分量 I_A ,一个表示纹理分量 $I_B, I = I_A + I_B$,其中, $I_A = T_A \alpha_A (T_A \in \mathbb{R}^{N \times L}, L \gg N), I_B = T_B \alpha_B (T_B \in \mathbb{R}^{M \times Q}, Q \gg M)$.目标函数表示为

$$\langle \alpha_A^*, \alpha_B^* \rangle = \arg \min_{(\alpha_A, \alpha_B)} \|\alpha_A\|_1 + \|\alpha_B\|_1 + \mu \|I - T_A \alpha_A - T_B \alpha_B\|_2^2 + \gamma TV(T_A \alpha_A) \quad (8)$$

其中,整体变分 $TV(T_A \alpha_A)$ 表示的几何结构分量 I_A 梯度的 l^1 范数,保证图像几何结构分量 $T_A \alpha_A$ 的梯度是稀疏的,即 $T_A \alpha_A$ 更接近分段平滑图像分量.

对 T_A 和 T_B 的选择,采用已有的变换方法.为计算方便,一般要求 T_A 和 T_B 字典具有快速正向变换、逆变换和伴随变换.对几何结构分量,可以选用能够稀疏表示平滑部分的字典,如小波(wavelet)、曲线波(curvelet)、轮廓波(contourlet)等变换.对纹理分量部分选用离散余弦变换(DCT)、Gabor 变换等^[22].本文分别选用 Curvelet 和 DCT.有关 MCA 的详细内容请见文献[23].

类似于 Chang 等人^[11]及 Yang 等人^[6]的方法,选择低分辨率图像的几何结构分量 I_A 的一阶和二阶梯度信息作为特征,4 个用于提取图像梯度信号的滤波器定义为^[6]

$$f_{11} = [-1, 0, 1], f_{12} = f_{11}^T, f_{21} = [1, 0, -2, 0, 1], f_{22} = f_{21}^T \quad (9)$$

将 4 个滤波器应用于 I_A 的图像块,将产生 4 个特征向量,并与 I_B 的图像块合成 1 个向量表示低分辨率图像块的特征向量。

参考 Freeman 等人^[2]的方法,本文选择高分辨率图像块的高频分量作为高分辨率图像块的特征向量,并与低分辨率图像块的特征向量合成为 1 个向量作为稀疏字典编码的 1 个训练样本,如图 1 所示(每个样本由 1 个向量表示,由 3 部分组成:高频分量 H^H 图像块、 L_A 的一阶和二阶梯度图像特征块、 L_B 图像块)。图中的 H^H 是 H 的高频分量; L_A, L_B 分别为低分辨率图像 L 经 MCA 分解后的几何结构分量和纹理分量。

这种方法充分利用了图像几何结构分量的高频信息,并整合了图像的纹理信息,提高了各图像块之间的匹配精度,同时实现了超分辨率重建和降噪功能。

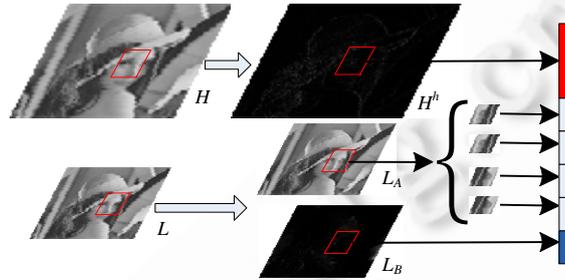


Fig.1 Training examples

图 1 训练样本

3 实验

选择一些自然图像作为训练集^[6],进行尺度放大 3 倍的超分辨率实验.并与 K-NN 方法、LLE 方法、SC 方法进行比较.对彩色图像,人类视觉对图像亮度的变化更敏感,所以只取相应的亮度通道进行处理.低分辨率图像取 3×3 图像块,并保留与周围图像块 1 个像素的重叠.对应的高分辨率图像取 9×9 图像块,保留 3 个像素的重叠.在具体实施过程中,我们采用稀疏编码的超分辨率算法中的方法,应用 Bicubic 插值算法,从低分辨率图像的 2 倍放大图像中提取特征数据,即 6×6 图像块。

3.1 质量对比

从训练图像集中随机选取 50 000 对高、低分辨率图像特征块进行训练.为保证超分图像与原始图像进行对比,我们对实验用原始图像进行裁剪,以使图像大小为 3 的倍数.为平衡计算效率和结果质量,稀疏字典大小固定为 1 024.选择公式(6)中的参数 $\lambda=0.01, \beta=1$.图 2 比较了本文方法(SD)与其他几种基于学习的超分算法的重建效果(采用算法从左至右下依次为低分辨率图像(face),Bicubic,K-NN,LLE,SC,SD).视觉效果方面,K-NN 方法有振铃现象,LLE 产生平滑效果,SD 和 SC 相对较好.通过多幅图像实验,本文方法的 PSNR 值相对于其他方法有不同程度的提高,见表 1.其中,face 和 lena 是彩色图像,其余为灰度图像。



Fig.2 Super-Resolution results by a factor of 3

图 2 超分辨率 3 倍尺度放大效果图

Table 1 PSNR of images

表 1 各图像 PSNR 对比

图像	Bicubic	K-NN	LLE	SC	SD
Face	28.4562	28.7122	29.1903	30.0732	30.8350
Barbara	22.3522	24.4539	25.8044	27.0761	27.3849
Cameraman	30.6361	30.6297	30.5916	30.8729	30.8737
Castle	29.3790	29.4851	29.4884	29.2843	29.4862
Lena	26.5976	26.9802	27.0076	27.5716	27.9941
Peppers	28.1233	28.4308	28.8751	29.1698	29.2158
House	28.8657	29.2370	29.5562	30.2369	30.2682

3.2 样本数量的影响

以 face 图像进行实验.随机选择 50 000,30 000,10 000,5 000 和 2 000 个样本进行训练学习,用本文方法和 Yang 等人的方法进行超分放大 3 倍实验,对比结果如图 3 所示.

可见,在同样样本个数条件下,本文方法的 PSNR 指标更好.当训练样本个数小于 5 000 时,两种方法的 PSNR 均下降很快.为保证结果质量,一般选择样本个数不小于 10 000 个.

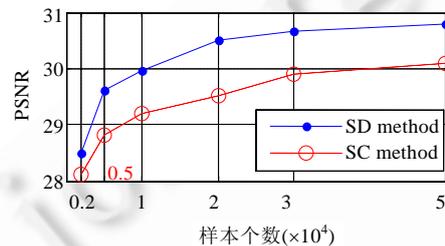


Fig.3 Influence on the super-resolution quality of the number of examples

图 3 样本个数对超分辨率质量的影响

3.3 字典大小的影响

根据上一节的样本,随机选择 30 000 个图像块来训练.稀疏字典大小分别选择 256,512,1 024 和 2 048.对 face 图像分别采用本文方法和 Yang 等人的方法进行超分放大 3 倍实验,对比结果如图 4 所示.可见,在稀疏字典大小相同条件下,本文方法 PSNR 更好.理论上讲,稀疏字典越大,结果质量越好,但计算量也越大.由图中可见,当稀疏字典逐步增大时,PSNR 曲线的上升趋于平缓.为兼顾超分重建质量和计算效率,稀疏字典的大小一般选为 1 024.

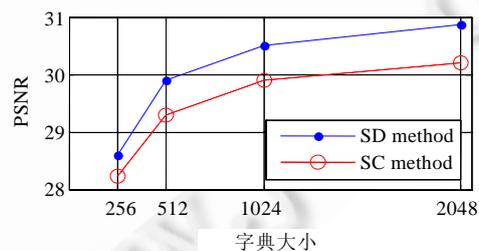


Fig.4 Influence on the super-resolution quality of size of sparse dictionary

图 4 稀疏字典大小对超分辨率质量的影响

3.4 噪声影响

大多数超分辨率算法的处理对象是无噪图像.处理有噪图像一般分两步:先进行降噪,再进行超分辨率重建.稀疏编码的超分方法通过调整 λ (文献[6]中的公式(8))来实现降噪,但会产生明显的平滑效果,降噪功能很有

限.本文方法同步进行降噪和超分重建,在提取图像特征时采用了 MCA 分解方法,在此过程中,通过公式(8)中参数 μ 的选择对图像特征块进行降噪预处理;同时,通过调解公式(6)中的参数 λ 进一步去噪.选择 $\mu=4\sigma$ (σ 是噪声方差)^[22], $\lambda=0.1\sigma$ ^[6].选择 30 000 个训练样本,稀疏字典大小为 1 024,对不同噪声强度的 lena 彩色图像进行实验,并与 Bicubic 和 SC 方法对比,结果如图 5 和图 6 所示(从左到右,噪声强度 σ 分别为 1,2,4,6).当噪声强度较弱时,本文方法降噪能力较强,结果质量较好;当噪声渐强时,本文方法与 SC 方法 PSNR 指标下降较快,说明超分辨率反演模型严重病态.

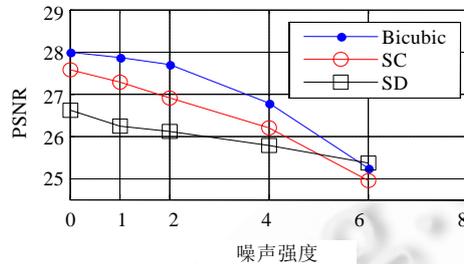


Fig.5 Influence on the super-resolution quality of the number of examples

图 5 样本个数对超分辨率质量的影响



Fig.6 Influence on the super-resolution quality of Gaussian noise

图 6 高斯噪声对超分辨率质量的影响

4 总 结

本文提出一种基于稀疏字典的超分辨率模型,建立高、低分辨率图像间的稀疏关联,同步实现匹配搜索和优化估计.应用形态分量分析方法提取图像的几何结构分量和纹理分量,将图像块的几何结构分量的一阶和二阶梯度特征信息及纹理分量组成样本进行训练,提高了特征匹配的准确性,同步实现了超分辨率重建和降噪功能.对比 Yang 等人的稀疏编码方法,稀疏字典更紧凑、有效;在重建质量相同情况下,所需样本更少.与其他基于学习的超分辨率方法相比,本文所提出的方法的结果质量和计算效率均有较大提高.尽管如此,这种方法仍存在一些改进之处,比如,进一步提高算法对不同类型图像的自适应性;减小字典大小、稀疏性、最大编码错误等经验参数对重建结果的影响;增强样本间的不相关性以训练更有效的稀疏字典,加快稀疏字典编码的速度.另外,将本方法推广应用到多帧影像的超分重建也是一个很好的研究方向.

致谢 在此,我们向对本文工作给予支持和建议的同行,尤其是电子科技大学自动化学院的王富同学表示感谢.

References:

- [1] Freeman WT, Pasztor EC, Carmichael OT. Learning low-level vision. *Int'l Journal of Computer Vision*, 2000,40(1):25-47. [doi: 10.1023/A:1026501619075]
- [2] Freeman WT, Jones TR, Pasztor EC. Example-Based super-resolution. *IEEE Computer Graphics and Applications*, 2002,22(2): 56-65. [doi: 10.1109/38.988747]

- [3] Capel D, Zisserman A. Super-Resolution from multiple views using learnt image models. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2001). Kauai: IEEE Computer Society, 2001. 627–634. [doi: 10.1109/CVPR.2001.991022]
- [4] Baker S, Kanade T. Limits on super-resolution and how to break them. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,24(9):1167–1183. [doi: 10.1109/TPAMI.2002.1033210]
- [5] Park SC, Park MK, Kang MG. Super-Resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 2003,20(3):21–36. [doi: 10.1109/MSP.2003.1203207]
- [6] Yang JC, Wright J, Huang T, Ma Y. Image super-resolution as sparse representation of raw image patches. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). Anchorage: IEEE Computer Society, 2008. 1–8. [doi: 10.1109/CVPR.2008.4587647]
- [7] Rubinstein R, Zibulevsky M, Elad M. Double sparsity: Learning sparse dictionaries for sparse signal approximation. *IEEE Trans. on Signal Processing*, 2010,58(3):1553–1564. [doi: 10.1109/TSP.2009.2036477]
- [8] Liu C, Shum HY, Zhang CS. A two-step approach to hallucinating faces: Global parametric model and local nonparametric model. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2001). Kauai: IEEE Computer Society, 2001. 192–198. [doi: 10.1109/CVPR.2001.990475]
- [9] Hertzmann A, Jacobs CE, Oliver N, Curless B, Salesin DH. Image analogies. In: Proc. of the Computer Graphics Proceedings, Annual Conf. Series. Los Angeles: ACM Press, 2001. 327–339. [doi: 10.1145/383259.383295]
- [10] Efros AA, Freeman WT. Image quilting for texture synthesis and transfer. In: Proc. of the 28th Annual Conf. on Computer Graphics and Interactive Techniques. Los Angeles: ACM Press, 2001. 341–346. [doi: 10.1145/383259.383296]
- [11] Rajaram S, Gupta MS, Petrovic N, Huang TS. Learning-Based nonparametric image super-resolution. *EURASIP Journal on Applied Signal Processing*, 2006,2006(1):1–11. [doi: 10.1155/ASP/2006/51306]
- [12] Chang H, Yeung DY, Xiong YM. Super-Resolution through neighbor embedding. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), Vol.2. Washington: IEEE Computer Society, 2004. 275–282. [doi: 10.1109/CVPR.2004.1315043]
- [13] Sun J, Zheng NN, Tao H, Shum HY. Image hallucination with primal sketch priors. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR). Madison: IEEE Computer Society, 2003. 729–736. [doi: 10.1109/CVPR.2003.1211539]
- [14] Ni KS, Nguyen TQ. Image superresolution using support vector regression. *IEEE Trans. on Image Processing*, 2007,16(6):1596–1610. [doi: 10.1109/TIP.2007.896644]
- [15] Kim KI, Kwon Y. Single-Image super-resolution using sparse regression and natural image prior. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010,32(6):1127–1233. [doi: 10.1109/TPAMI.2010.25]
- [16] Gu YT, Wu EH. Image-Analogies based super-resolution. *Journal of Software*, 2008,19(4):851–860 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/851.htm> [doi: 10.3724/SP.J.1001.2008.00851]
- [17] Pan Q, Luo XN, Zhu JW. Learning-Based medical image magnification algorithm by local geometric similarity. *Journal of Software*, 2009,20(5):1146–1155 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3546.htm> [doi: 10.3724/SP.J.1001.2009.03546]
- [18] Wang JJ, Zhu SH, Gong YH. Resolution enhancement based on learning the sparse association of image patches. *Pattern Recognition Letters*, 2010,31(1):1–10. [doi: 10.1016/j.patrec.2009.09.004]
- [19] Amaldi E, Kann V. On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. *Theoretical Computer Science*, 1998,209(1-2):237–260. [doi: 10.1016/S0304-3975(97)00115-1]
- [20] Donoho DL. For most large underdetermined systems of equations, the minimal l^1 -norm near-solution approximates the sparsest near-solution. *Communications on Pure and Applied Mathematics*, 2006,59(7):907–934. [doi: 10.1002/cpa.20131]
- [21] Irani M, Peleg S. Motion analysis for image enhancement: Resolution, occlusion, and transparency. *Journal of Visual Communications and Image Representation*, 1993,4(4):324–335. [doi: 10.1006/jvci.1993.1030]
- [22] Starck JL, Elad M, Donoho DL. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. on Image Processing*, 2005,14(10):1570–1582. [doi: 10.1109/TIP.2005.852206]

- [23] Starck JL, Elad M, Donoho DL. Redundant multiscale transforms and their application for morphological component separation. *Advances in Imaging and Electron Physics*, 2004,132:287-348. [doi: 10.1016/S1076-5670(04)32006-9]

附中文参考文献:

- [16] 古元亭,吴恩华.基于图像类推的超分辨率技术.软件学报,2008,19(4):851-860. <http://www.jos.org.cn/1000-9825/19/851.htm> [doi: 10.3724/SP.J.1001.2008.00851]
- [17] 潘琪,罗笑南,朱继武.基于学习的局部几何相似性的医学图像放大.软件学报,2009,20(5):1146-1155. <http://www.jos.org.cn/1000-9825/3546.htm> [doi: 10.3724/SP.J.1001.2009.03546]



李民(1974-),男,山西闻喜人,博士,副教授,主要研究领域为计算机视觉,图像处理.



程建(1978-),男,博士,副教授,主要研究领域为图像处理,模式识别,计算机视觉,多源遥感信息智能挖掘.



乐翔(1986-),男,硕士,主要研究领域为图像处理,计算机视觉.



罗环敏(1972-),男,博士,主要研究领域为雷达信号处理,图像处理.