

统一最小二乘规则的单幅图像超分辨率算法

赵小乐¹, 吴亚东^{1*}, 田金沙¹, 张红英²

(1. 西南科技大学 计算机科学与技术学院, 四川 绵阳 621010; 2. 西南科技大学 信息工程学院, 四川 绵阳 621010)

(* 通信作者电子邮箱 wyd028@163.com)

摘要:基于机器学习的超分辨率方法是一个很有发展前景的单幅图像超分辨率方法,稀疏表达和字典学习是其中的研究热点。针对比较耗时的字典训练与恢复精度不高图像重建,从减小低分辨率(LR)和高分辨率(HR)特征空间之间差异性的角度提出了一种使用迭代最小二乘字典学习算法(IIS-DLA),并使用锚定邻域回归(ANR)进行图像重建的单幅图像超分辨率算法。迭代最小二乘法整体优化过程极大地缩短了低分辨率字典/高分辨率字典的训练时间,它采用了与锚定邻域回归相同的优化规则,有效地保证了字典学习和图像重建在理论上的一致性。实验结果表明,所提算法的字典学习效果比K-均值奇异值分解(K-SVD)和Beta过程联合字典学习(BPJDL)等算法更高效,图像重建的效果也优于许多优秀的超分辨率算法。

关键词:迭代最小二乘法;锚定邻域回归;稀疏表达;字典学习;图像超分辨

中图分类号: TP391.1 **文献标志码:** A

Single image super-resolution based on unified iterative least squares regulation

ZHAO Xiaole¹, WU Yadong^{1*}, TIAN Jinsha¹, ZHANG Hongying²

(1. School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang Sichuan 621010, China;

2. School of Information engineering, Southwest University of Science and Technology, Mianyang Sichuan 621010, China)

Abstract: Machine learning based image super resolution (SR) has been proved to be a promising single image SR technology, sparseness representation and dictionary learning has become the hotspot of it. Aiming at the time-consuming dictionary training and low-accuracy SR recovery, a SR algorithm has been proposed from the perspective of reducing the inconsistency between Low Resolution (LR) feature and High Resolution (HR) feature spaces as far as possible. The proposed method adopted Iterative Least Squares Dictionary Learning Algorithm (IIS-DLA) to train LR/HR dictionaries and Anchored Neighborhood Regression (ANR) to recover HR images. IIS-DLA is able to train LR/HR dictionaries in relatively short time because of its integral optimization procedure, and it adopts the same optimization strategy with ANR which reduces the diversity between LR/HR dictionaries effectively. A large number of experiments show that the proposed method achieves superior dictionary learning to K-means Singular Value Decomposition (K-SVD) and Beta Process Joint Dictionary Learning (BPJDL) algorithms etc., and provides better SR results than other state-of-the-art SR algorithms.

Key words: Iterative Least Squares (IIS); Anchored Neighborhood Regression (ANR); Sparseness Representation (SR); dictionary learning; super resolution

0 引言

单幅图像的超分辨率处理技术是数字图像处理中的研究热点之一,大量国内外研究人员对其进行了广泛而深入的研究。数字图像的超分辨率处理技术大致经历了插值方法、重建方法和机器学习方法三个阶段。基于插值的传统超分辨率处理技术一般将离散的图像信号看成连续信号,利用连续性特征估计未知采样点的像素值。然而,自然图像中包含了大量的奇异性特征,比如边缘、角点、高频纹理区域等^[1],因此,基于连续性假设的插值算法处理效果较差,往往会形成明显的振铃和混叠效应。基于重建的方法一般模拟数字图像的形成过程,通过对数字成像设备采集图像信号的过程建模来求解最终结

果。这种方法往往需要为特定模型假设某些参数,比如摄像机的点扩散函数(Point Spread Function, PSF),环境噪音等。这些假设的参数通常情况下与实际参数不符,甚至差别很大,这会造成超分辨率结果出现严重偏差^[2-3]。另外,对放大因子十分敏感也是重建方法的主要缺陷之一。

相对而言,机器学习方法比前两者更具吸引力,无论是从理论研究的角度还是从实践运用的角度都具有更大的提升空间。基于机器学习的方法最早体现在Freeman等^[4]的样本学习方法中。该方法利用置信传播算法求解马尔可夫域模型(Markov Random Field, MRF),然后根据低分辨率(Low Resolution, LR)图像块来预测高分辨率(High Resolution, HR)图像块。这种方法需要一个由大量LR/HR图像块对组

收稿日期:2015-08-12;修回日期:2015-10-04。

基金项目:国家自然科学基金资助项目(61303127);四川省科技支撑计划项目(2014SZ0223);四川省教育厅重点项目(13ZA0169);中国科学院“西部之光”人才培养计划项目(13ZS0106);西南科技大学创新基金资助项目(15yex053)。

作者简介:赵小乐(1987-),男,四川南部人,硕士研究生,CCF会员,主要研究方向:数字图像处理; 吴亚东(1979-),男,河南周口人,教授,博士,CCF会员,主要研究方向:数字图像处理、信息可视化、人机交互; 田金沙(1988-),女,河北衡水人,硕士研究生,主要研究方向:数字图像处理; 张红英(1976-),女,四川德阳人,教授,博士,主要研究方向:数字图像处理。

成的外部数据库,这造成了严重的时耗问题。为了提高算法执行效率,Chang 等^[5]基于局部线性嵌入理论^[6]提出了邻域嵌入法(Neighbor Embedding, NE),通过假设低分辨率特征空间与高分辨率特征空间具有相似结构来减少重建样本数量,一定程度上提高了算法执行效率。许多研究人员对这种方法进行了深入研究并提出了相应的改进方案^[7-9],但是固定数量的邻域始终会引起过拟合或欠拟合问题。无论如何,不恰当的拟合都会丢失近似精度。针对这一问题,Yang 等^[10]首先提出用稀疏表达来处理超分辨率问题。由于稀疏近似过程中并不固定邻域数量,而是使近似误差尽可能小,从而有效地避免了不恰当的拟合问题。

与稀疏表达密切相关的另一个问题是字典学习,字典学习的任务是从训练集中提取一个原子项集合形成过完备字典,使得基于该字典的稀疏表达能够提供尽可能小的近似误差。信号处理领域的研究表明,过完备字典能够灵活而有效地表达数字信号,通过特定字典学习算法所训练的字典可以产生比传统表达方式更精确的近似结果^[11]。目前已经出现了许多字典训练算法,比如迭代最小二乘字典学习算法(Iterative Least Squares Dictionary Learning Algorithm, ILS-DLA)^[12],K-均值奇异值分解(K-means Singular Value Decomposition, K-SVD)算法^[13],递归最小二乘字典学习算法(Recursive Least Squares Dictionary Learning Algorithm, RLS-DLA)^[14]以及在线字典学习算法(Online Dictionary Learning, ODL)^[15-16]等。这些算法训练字典的原理各不相同,有着不同的应用场合,但它们的最终目标都是最小化对特定训练集的近似误差。然而,所有的这些算法都着眼于精确而有效地学习单一特征空间上的字典。在许多信号处理问题中通常存在两个对偶的特征空间,而且这两个对偶特征空间之间往往存在未知而复杂的非线性关系^[17]。虽然文献[11,17]都尝试着寻找对偶字典之间的映射关系,如图1虚线部分所示,但复杂的计算过程严重降低了算法效率。为了进一步提升算法执行效率,Timofte 等^[18-19]提出的锚定邻域回归算法(Anchored Neighborhood Regression, ANR)利用最小二乘规则将对偶字典合并成一个映射矩阵,然后将重建阶段的 L_1 范式约束的最优化问题修改为 L_2 范式约束的最优化问题。这样便将高分辨率图像块的重建操作就转化为矩阵乘法操作,极大地降低了时间消耗。

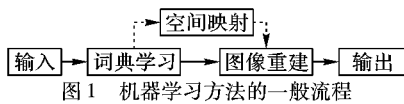


图1 机器学习方法的一般流程

利用稀疏表达与字典学习进行图像超分辨率重建一般有三个步骤,即字典学习、空间映射和图像重建。由于传统机器学习方法并没有对空间映射进行特殊处理,可以认为这类方法经过字典学习后直接进行图像重建,如图1中实线部分所示。我们将原始 ANR 算法称为 SVD-ANR 算法,文中提出一种基于 ILS-DLA 字典训练和 ANR 重建框架的超分辨率算法,我们称为 ILS-ANR 算法。与文献[11]或文献[17]不同,所提算法并不试图直接寻找 LR 特征空间和 HR 特征空间之间的映射关系,而是尽可能地让字典学习、空间映射以及图像重建基于相同的规则,从而减少不同特征空间之间的差异性以提升超分辨率效果。由于 ILS-DLA 算法的整体优化策略,字典学习阶段的时间消耗也明显小于 K-SVD 算法,在训练集较大的情况下

能有效提高字典训练速度。

1 统一最小二乘规则的超分辨率重建

1.1 特征提取与处理框架

所提算法与典型的基于机器学习的超分辨率算法一样,也包括字典学习、空间映射和图像重建三个部分,这分别对应了图2中的 I、II、III 三个过程。值得说明的是,这里的字典学习单指学习 LR 特征空间内的字典,HR 特征空间内的字典由空间映射得到。

要采用机器学习进行超分辨率重建,首先要从训练图像集提取训练数据。这就涉及到训练数据收集或特征提取策略的问题。首先,将训练图像转换为 YCrCb 颜色空间并提取其 Y 通道内的数据作为训练的原始数据,Cr 和 Cb 通道的数据直接采用插值算法进行处理即可。这主要是因为人眼对亮度的敏感程度比对色度和饱和度的敏感程度更强,所以大多数算法都只对亮度通道进行处理。将得到的 Y 通道图像 HRI 进行下采样操作 $\downarrow \alpha$ (α 为下采样因子) 便可得低分辨率图像 LRI,再用双立方插值算法将 LRI 图像上采样至原始大小得到中等级别分辨率图像 MRI。低分辨率特征数据就是从 MRI 直接提取特征块得到,高分辨率特征数据就是从 HRI-MRI 中直接提取图像块得到。对于低分辨率特征数据,我们提取 MRI 图像在 x 方向和 y 方向上的一阶、二阶导数一共四组特征数据,将这四组特征数据串接在一起形成对应图像块的一个特征向量,所有特征向量按列组成低分辨率特征数据集 X ; 对于高分辨率特征数据,我们提取残差图像 HRI-MRI 的原始数据块,将每个图像块转换成列向量并将其组织到高分辨率特征数据集 Y 中,如图2所示。

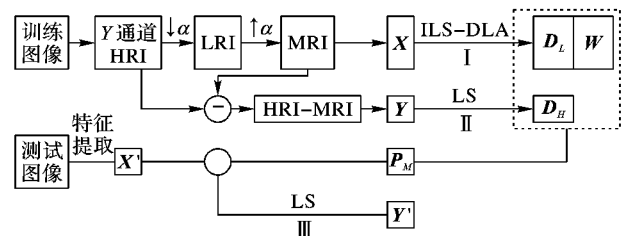


图2 所提算法整体处理框架

当所有训练数据都准备好后,首先利用 ILS-DLA 字典训练算法学习低分辨率特征空间内的过完备字典 D_L 和对应的系数矩阵 W ,然后根据直接空间映射和最小二乘规则直接计算高分辨率特征空间的过完备字典 $D_H = Y W^T (W W^T)^{-1}$,此时便可根据 D_L 、 D_H 和 W 计算锚定邻域映射矩阵 P_M ,这样就完成了整个字典训练阶段。实际上,映射矩阵 P_M 也是由其他不同的目标函数根据最小二乘规则计算得到的。

根据锚定邻域回归的处理过程可知,图像重建阶段不需要再求解 L_0/L_1 范式的最优化问题,我们采用与 LR 数据集相同的策略从输入图像中提取 LR 特征数据集 X' ,其中的每一个特征向量与对应的映射矩阵相乘便可得相应的高分辨率特征向量,再将这些高分辨率特征向量组织在一起就形成高分辨率特征向量集 Y' 。

1.2 ILS-DLS 字典学习

首先考虑低分辨率字典 D_L ,假设字典原子项的维度为 N ,总共有 K 个原子项,那么 D_L 可以表示为一个 $N \times K$ 的矩阵。在稀疏表达中,可以用字典原子项的线性组合来表达或近似

任意给定的一个 $N \times 1$ 向量 \mathbf{x} 。假设 \mathbf{x}_a 表示 \mathbf{x} 的近似值,则:

$$\mathbf{x}_a = \mathbf{D}_L \mathbf{w} \approx \mathbf{x} \quad (1)$$

其中: \mathbf{w} 是一个 $K \times 1$ 的系数向量,它的每一个元素是对应字典原子项的系数(或权重)。由于 $\mathbf{x}_a = \mathbf{D}_L \mathbf{w}$,所以近似误差为 $\mathbf{r} = \mathbf{x} - \mathbf{x}_a = \mathbf{x} - \mathbf{D}_L \mathbf{w}$ 。若系数向量 \mathbf{w} 中的非零元素个数为 s ,那么上述稀疏表达的稀疏因子就为 s/N 。在测试向量 \mathbf{x} 和 LR 字典 \mathbf{D}_L 都给定的情况下寻找系数向量 \mathbf{w} 就是一个稀疏近似问题,通常将这一问题描述为:

$$\arg \min_{\mathbf{w}} \|\mathbf{w}\|_p + \lambda \|\mathbf{x} - \mathbf{D}_L \mathbf{w}\|_2^2 \quad (2)$$

其中: $p \in \{0, 1\}$ 是施加在系数向量 \mathbf{w} 上的约束条件。参数 λ 用于平衡近似误差,以使目标解更加稳定。当 $p = 0$ 时上述问题是一个 NP 难题,可以使用正交匹配追踪算法(Orthogonal Matching Pursuit, OMP) 对其进行求解。这样得到的效果是最好的,但在许多实际应用中并没有必要求解最优解;当 $p = 1$ 时上述问题更容易计算,一般使用最小角回归算法(Least Angle Regression, LARS) 求解,但最终解的精度有所下降。

字典学习是与稀疏近似密切相关的一个概念。假设 LR 训练集是由 L 个 N 维向量组成的 $N \times L$ 矩阵 \mathbf{X} ,字典学习的目标就是从该训练集中学习一个 $N \times K$ 的字典 \mathbf{D}_L 和对应的系数矩阵 \mathbf{W} ,使得近似误差 $\mathbf{R} = \mathbf{X} - \mathbf{D}_L \mathbf{W}$ 尽可能小,这一问题一般被描述为:

$$\{\hat{\mathbf{D}}_L, \hat{\mathbf{W}}\} = \arg \min_{\mathbf{D}_L, \mathbf{W}} \sum_{i=1}^L \|\mathbf{w}_i\|_p + \lambda \|\mathbf{X} - \mathbf{D}_L \mathbf{W}\|_2^2 \quad (3)$$

其中: $\|\mathbf{d}_i\|_2 = 1$ 为字典 \mathbf{D}_L 中的原子项。式(3)对应的最优化问题不是一个凸优化问题,但是当将 \mathbf{D}_L 或 \mathbf{W} 固定时,对另一个变量的求解过程就是一个凸优化过程,大多数字典训练算法都采用迭代训练模式,即先固定字典利用合适的匹配追踪算法求解权重矩阵,然后固定权重矩阵根据特定的规则更新字典,见如下 ILS-DLA 字典训练算法的详细描述。

ILS-DLA 字典训练算法。

1) 初始化字典 \mathbf{D}_L ,一般从训练集中随机选取 K 个向量作为初始字典;

2) 固定 \mathbf{D}_L ,使用 OMP 算法寻找权重矩阵 \mathbf{W} ,这实际上是由 L 个独立的式(2)组成的最优化问题;

3) 保持 \mathbf{W} 固定,利用最小二乘规则求解 LR 特征空间内的字典 $\mathbf{D}_L = (\mathbf{X} \mathbf{W}^T) (\mathbf{W} \mathbf{W}^T)^{-1}$;

4) 归一化原子项,即将每个原子项投影为单位向量;

5) 返回第2)步,直到满足算法终止条件。

在训练字典之前,要先对样本数据进行降维,这里使用了主成分分析算法(Principal Component Analysis, PCA) 来执行降维操作,这样可以极大地减少计算量。ILS-DLA 是一个简单而有效的字典学习算法,上述算法的终止条件一般是迭代次数或近似精度。第三步是在权重矩阵固定的情况下求解字典,实际上是求解下列最优化问题:

$$\hat{\mathbf{D}}_L = \arg \min_{\mathbf{D}_L} \|\mathbf{X} - \mathbf{D}_L \mathbf{W}\|_2^2 \quad (4)$$

根据最小二乘规则,对式(4)的目标函数求导并将其设为0便可得 $\mathbf{D}_L = \mathbf{X} \mathbf{W}^T (\mathbf{W} \mathbf{W}^T)^{-1}$ 。注意上述算法每次迭代都对字典 \mathbf{D}_L 进行整体更新,这与 K-SVD 算法不同。K-SVD 算法每次迭代都会遍历整个字典逐个更新字典原子项(实际上也同时更新了部分权重系数),所以 ILS-DLA 算法比 K-SVD 算法更高效。

对于高分辨率特征空间内的过完备字典 \mathbf{D}_H ,可以直接使用 ILS-DLA 算法训练所得的系数矩阵,并基于最小二乘规则进行计算,即:

$$\mathbf{D}_H = \mathbf{Y} \mathbf{W}^T (\mathbf{W} \mathbf{W}^T)^{-1} \quad (5)$$

1.3 锚定邻域回归

大多数稀疏表达方法都用 L_1 范式来约束目标函数,但是 L_1 范式约束的最优化问题的求解过程十分耗时。锚定邻域回归算法用 L_2 范式来约束目标函数,这样便可以利用岭回归方法对其进行求解。在稀疏表达阶段,需要根据低分辨率过完备字典 \mathbf{D}_L 求解下列最优化问题来确定稀疏系数:

$$\hat{\mathbf{w}}_i = \arg \min_{\mathbf{w}_i} \|\mathbf{x}_i - \mathbf{D}_L \mathbf{w}_i\|_2^2 + \lambda \|\mathbf{w}_i\|_2^2; i = 1, 2, \dots, L \quad (6)$$

其中: \mathbf{x}_i 表示 \mathbf{X} 的第 i 个列向量。根据最小二乘规则求解式(6),需要将其中目标函数的导数设为0,即:

$$\mathbf{D}_L^T (\mathbf{D}_L \mathbf{w}_i - \mathbf{x}_i) + \lambda \mathbf{w}_i = 0$$

由此可得系数向量的近似解:

$$\mathbf{w}_i = (\mathbf{D}_L^T \mathbf{D}_L + \lambda \mathbf{I})^{-1} \mathbf{D}_L^T \mathbf{x}_i \quad (7)$$

其中: \mathbf{I} 表示单位矩阵。我们采用直接空间映射将这里得到的稀疏系数运用到高分辨率特征空间,从而可得与 \mathbf{x}_i 对应的高分辨率特征向量 \mathbf{y}_i :

$$\mathbf{y}_i = \mathbf{D}_H \mathbf{w}_i = \mathbf{D}_H (\mathbf{D}_L^T \mathbf{D}_L + \lambda \mathbf{I})^{-1} \mathbf{D}_L^T \mathbf{x}_i \quad (8)$$

假设:

$$\mathbf{P}_M = \mathbf{D}_H (\mathbf{D}_L^T \mathbf{D}_L + \lambda \mathbf{I})^{-1} \mathbf{D}_L^T \quad (9)$$

则根据低分辨率特征向量计算高分辨率特征向量的过程就可以表示为下列直接映射形式:

$$\mathbf{y}_i = \mathbf{P}_M \mathbf{x}_i \quad (10)$$

如果上述映射过程中始终整体使用 \mathbf{D}_L 和 \mathbf{D}_H ,那么根据低分辨率特征计算高分辨率特征的过程就是一个全局回归过程。如果对两个对偶字典中的原子项进行分组,对每一组原子项都按照上述方式计算各自的映射矩阵,就是所谓的锚定邻域回归过程。具体来说,就是对每一个原子项,在字典中寻找它的 k 个最近邻域,这就会在 \mathbf{D}_L 和 \mathbf{D}_H 上分别产生一个大小为 $N \times (k+1)$ 的子字典 $\bar{\mathbf{D}}_L$ 和 $\bar{\mathbf{D}}_H$,则根据式(9)可以计算该分组的子映射矩阵:

$$\bar{\mathbf{P}}_M = \bar{\mathbf{D}}_H (\bar{\mathbf{D}}_L^T \bar{\mathbf{D}}_L + \lambda \mathbf{I})^{-1} \bar{\mathbf{D}}_L^T$$

字典中的每一个原子项都可以计算其对应的子映射矩阵,在进行超分辨重建时便可以直接用该子映射矩阵计算高分辨率特征向量。这样,映射矩阵就可以根据字典训练结果预先计算出来,而根据字典进行超分辨重建的整个过程就演化成矩阵乘法问题,从而极大地减少了图像重建过程中的时间消耗。

由上述字典训练的整个过程可以看出, \mathbf{D}_L 和 \mathbf{D}_H 的学习过程在理论上都是基于最小二乘规则进行的(参见图2)。从锚定邻域回归算法的推导过程也可以看出, ANR 实际上也是基于最小二乘规则来求解 L_2 范式最优化问题。由此可见,所提 ILS-ANR 算法的训练和重建阶段都以最小二乘规则为理论基础,这也是将其称为统一最小二乘规则的原因。

2 实验结果与分析

这里从字典学习和图像重建两个角度分别进行了对比试验。我们对 ILS-DLA 和 K-SVD 算法在相同条件下字典

学习的精度和时间消耗,以及根据不同字典进行图像重建的超分辨率效果。所有算法都在 ACER 4738ZG 电脑上运行,CPU 为 Intel Pentium P6100,主频 2 GHz,内存 2.0 GB,64 位 Win 7 操作系统。在其他运行平台上所得到的数据会略有差异,但数据之间的相对关系不会改变。

2.1 ILS-DLA 与 K-SVD 字典学习性能比较

ILS-DLA 和 K-SVD 字典学习算法与 RLS-DLA,ODL 等算法都属于单个特征空间上的字典学习算法,文献[12-16]已经对它们之间的性能和效果做了更为广泛的对比。而文献[11]和[17]属于对偶字典学习算法,它们更关注于如何精确表达对偶特征空间上的映射关系,本质上与单个特征空间上的字典学习算法不同。由于 K-SVD 和 ILS-DLA 都属于单个特征空间上的字典学习算法,而且都被应用在 ANR 的学习框架下,所以这里只对它们进行相应对比。

训练图像来自文献[10]的训练图像集,其中包含了 91 幅自然图像。特征提取策略如图 2 所示,提取数据块时采用

了与 ANR 算法^[18-19]不同的方式,我们对训练图像进行全覆盖提取,而没有进行修边,这样既可以充分利用每一幅图像又可以简化图像重建的操作。图像块大小为 9×9 ,步长为 6×6 (即两个方向上的重叠像素个数都为 6)。按照图 2 中的特征提取策略从上述训练图像集中提取到 142 418 个训练向量。图 3 标注中的 $\times 2, \times 3, \times 4$ 分别表示放大 2 倍,3 倍和 4 倍的情况。由于在进行字典学习之前使用了 PCA 算法对原始数据进行降维,而不同的缩放因子会得到不同的降维效果,所以训练向量的维度各不相同。从图 3(a)~(c) 可以看出,ILS-DLA 算法的训练精度随着字典大小增加而增加,K-SVD 算法虽然总体上也有这样的趋势,但显然存在一些异常情况。ILS-DLA 的训练精度变化相对平稳,K-SVD 算法的波动幅度却比较明显。除了 $\alpha = 3$ 和字典大小为 2 048 情况,ILS-DLA 算法的训练精度都要高于 K-SVD 算法。

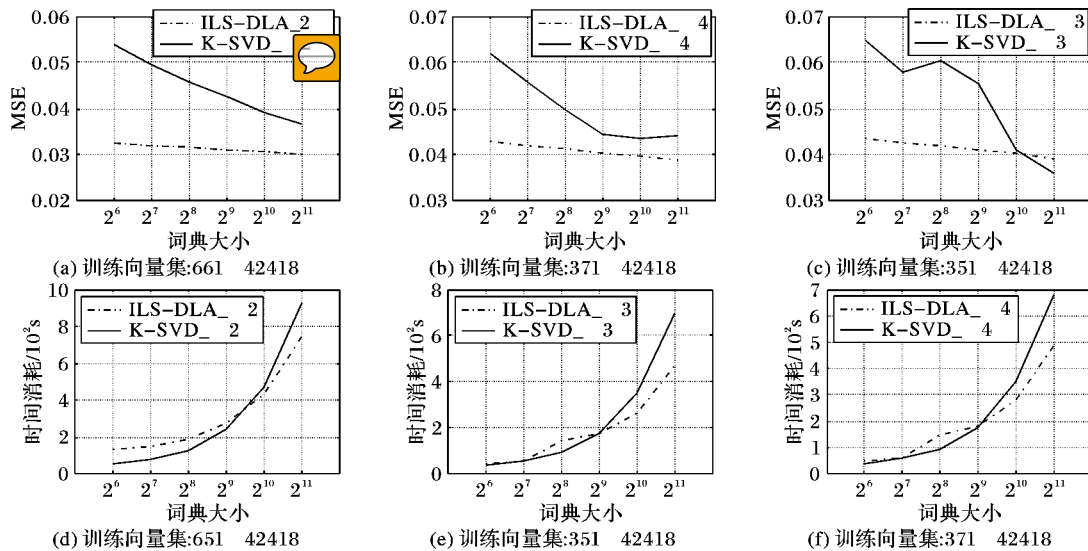


图3 ILS-DLA 与 K-SVD 字典学习效果对比

2.2 字典大小与放缩因子对重建效果的影响

图 4 给出了字典大小对图像重建质量的影响。测试图像为文献[18]中的 Baboon, Flower, Lena, Monarch 和 Tower 五幅图像,图 4 中的 PSNR 和 SSIM 值是测试图像集上的平均值。由图 4 可以看出,在相同缩放比例的情况下,字典越大对应的重建效果整体上越好,但这种效果上的增加并不明显,而且随着缩放比例增大,字典大小对重建效果的影响越小。从图 4 还可以明显看到,相同条件下,重建效果随着缩放比例增大而变差。

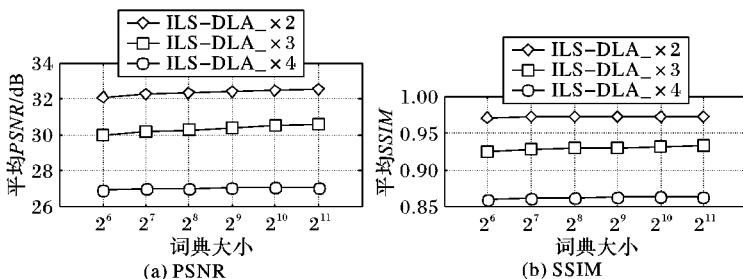


图4 不同字典大小对图像重建效果的影响

事实上,图 4 给出的结果很容易分析。一方面,字典越大,即字典包含的原子项越多,它的过完备性也越大,表达能

力也就越强,所以相同条件下重建效果会相对较好。极端的情况下,最大的字典就是训练集本身,训练过程中的表达误差为 0。另一方面,随着缩放比例的增大,图像自身丢失的信息越多,重建所需的必要信息也就越少,所以重建质量也就越来越差。

2.3 不同超分辨率算法之间的重建效果比较

这里选用文献[18-19]中表现相对较好的几种算法和文献[11]所提方法进行比较,其中 SVD-GR 表示使用 K-SVD 字典学习的全局回归算法,NE-LLE 表示基于局部线性嵌入的邻域嵌入算法。BPJDL (Beta Process Joint Dictionary Learning)^[11]考虑了对偶字典之间的映射关系,其余算法则根据 LR 字典采用直接映射方式求得 HR 字典。表 1、表 2 和表 3 分出了字典大小为 512、1 024 以及 2 048 时各种算法的 PSNR 和 SSIM 值,第一列的 $\times 2, \times 3$ 和 $\times 4$ 分别表示放大 2 倍、放大 3 倍和放大 4 倍的情况,每一行中加粗的值表示这一行 PSNR 或 SSIM 的最大值。从这几个表中可以看出,采用 ILS-DLA 字典训练算法学习字典后,图像重建的效果提升非常明显。无论是放大 2 倍、3 倍还是 4 倍的情况下,

ILS-DLA 的平均 PSNR 和 SSIM 对其他几种优秀超分辨率算法都有较大幅度的提升。另外,可以看到每个表的测试中会出现个别表现优于所提算法的值,这主要是由于图像的边缘和纹理等结构有所差异。由于 ILS-ANR 算法统一了处理过程中的优化规则,使 LR 特征空间内的字典和 HR 特征空间内的字典具有更相似的稀疏表达特性,所以所提算法整体上优于其他几种方法。

图5~图7分别展示了表1~表3中几种算法的视觉效果对比。这里不再给出双立方插值的效果图,因为基于连续性假

设的双立方插值算法所得到的结果模糊比较严重,与参考图像差别很大。图5对应的字典大小为512,缩放因子为2,测试图像为 Tower,大小为 477×321 ;图6对应的字典大小为1024,缩放因子为3,测试图像为 Monarch,大小为 762×504 ;图7对应的字典大小为2048,缩放因子为4,测试图像为 Flower,大小为 477×321 。为了更好地排版,所有图像都进行了适当的缩放。这几种算法都更接近参考图像,但可以明显地观察到所提算法更加精确地恢复了图像的边缘和纹理特征,比如 Tower 图像中的拱形边缘和 Monarch 图像中蝴蝶翅膀上的纹理结构。

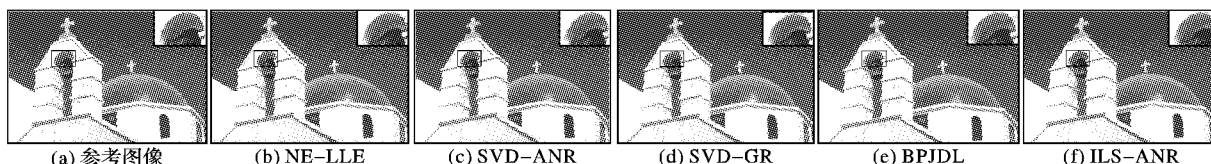


图5 视觉效果对比(字典大小512,SR×2)

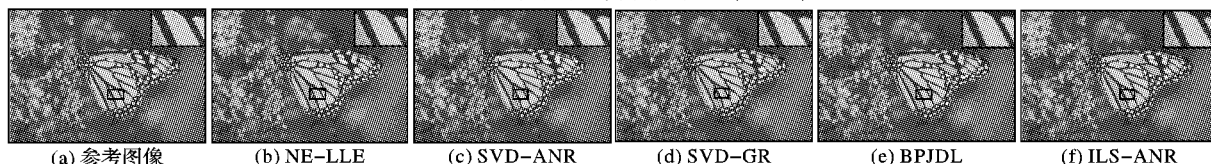


图6 视觉效果对比(字典大小1024,SR×3)

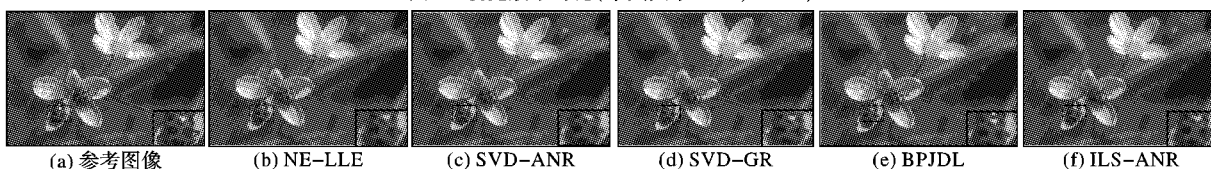


图7 视觉效果对比(字典大小2048,SR×4)

表1 几种典型超分辨率算法重建结果的 PSNR 与 SSIM 对比(字典大小:512)

图像×2	Bicubic		NE-LLE		SVD-ANR		SVD-GR		BPJDL		ILS-ANR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Baboon	23.9357	0.8933	24.7675	0.9551	24.8727	0.9619	<u>24.9693</u>	0.9629	24.8993	0.9621	24.9009	<u>0.9721</u>
Flower	33.8755	0.9506	36.4065	0.9621	36.6744	0.9690	36.5459	0.9689	36.6459	0.9690	<u>36.7041</u>	<u>0.9692</u>
Lena	31.2259	0.9645	34.2724	0.9865	34.5272	0.9892	34.5752	0.9896	34.5832	0.9898	<u>34.6031</u>	<u>0.9914</u>
Monarch	29.9467	0.9801	33.1756	0.9913	<u>33.4602</u>	0.9935	33.3379	0.9937	33.3762	0.9939	33.4424	<u>0.9942</u>
Tower	29.1069	0.9228	32.0744	0.9394	32.1801	0.9465	31.9898	0.9443	32.1874	0.9453	<u>32.3226</u>	<u>0.9470</u>
AVERAGE	29.6182	0.9423	32.1393	0.9669	32.3429	0.9720	32.2836	0.9724	32.3384	0.9720	<u>32.3946</u>	<u>0.9742</u>

表2 几种典型超分辨率算法重建结果的 PSNR 与 SSIM 对比(字典大小:1024)

图像×3	Bicubic		NE-LLE		SVD-ANR		SVD-GR		BPJDL		ILS-ANR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Baboon	23.1430	0.8026	23.4116	0.8431	23.4575	0.8471	23.4566	<u>0.8489</u>	23.4383	0.8452	<u>23.4597</u>	0.8474
Flower	33.5219	0.9360	34.4625	0.9411	34.6815	0.9478	34.2810	0.9430	34.5074	0.9438	<u>34.7033</u>	<u>0.9480</u>
Lena	31.7198	0.9537	32.7378	0.9641	<u>32.9309</u>	<u>0.9672</u>	32.6124	0.9651	32.9019	0.9649	32.9217	0.9670
Monarch	29.4632	0.9708	30.6670	0.9777	30.8513	0.9806	30.4472	0.9788	30.8375	0.9796	<u>30.9184</u>	<u>0.9808</u>
Tower	29.0496	0.9034	30.5105	0.9113	30.5329	0.9175	29.8427	0.9010	30.5629	0.9183	<u>30.6308</u>	<u>0.9191</u>
AVERAGE	29.3795	0.9133	30.3579	0.9275	30.4908	0.9320	30.1280	0.9274	30.4496	0.9304	<u>30.5268</u>	<u>0.9325</u>

表3 几种典型超分辨率算法重建结果的 PSNR 与 SSIM 对比(字典大小:2048)

图像×4	Bicubic		NE-LLE		SVD-ANR		SVD-GR		BPJDL		ILS-ANR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Baboon	21.8813	0.6360	22.1005	0.6902	22.1409	0.6950	22.1521	0.6963	22.1427	0.6952	<u>22.1771</u>	0.7010
Flower	28.6687	0.8714	30.6310	0.8917	30.7740	0.9012	30.8051	0.9029	30.7651	0.9004	<u>30.9706</u>	0.9043
Lena	29.0665	0.9075	29.4649	0.9122	29.6478	0.9187	29.2267	0.9127	29.3945	0.9176	29.6345	0.9189
Monarch	25.2417	0.9153	26.6090	0.9341	26.6140	0.9382	26.6295	0.9380	26.6395	0.9379	<u>26.6564</u>	0.9383
Tower	23.9539	0.8248	26.0690	0.8491	26.1258	0.8593	26.0973	0.8531	26.1073	0.8575	<u>26.2033</u>	<u>0.8605</u>
AVERAGE	25.7624	0.8310	26.9749	0.8555	27.0605	0.8625	26.9821	0.8606	27.0098	0.8617	<u>27.1284</u>	<u>0.8646</u>

3 结语

基于机器学习的单幅图像超分辨率算法通常将字典学习(包括空间映射)和图像重建看成两个独立的过程。字典可以利用字典学习算法对外部数据库进行学习预先得到,然后在所得字典的基础上进行图像重建。为了提高重建效果,先前的工作大多致力于提高字典学习的精度或精确表达低分辨率字典和高分辨率字典之间的映射关系。事实上,字典学习和图像重建都是不定解问题,单方面提升字典学习精度并不能最大限度上提升重建效果。与几种对偶字典学习算法试图显式表达对偶特征空间之间映射关系的做法不同,本文从字典学习和图像重建的角度统一整个处理过程中的优化规则,使低分辨率特征空间内的字典和高分辨率特征空间内的字典具有更想相似的稀疏表达特性,从而提高图像超分辨率处理的效果。

参考文献:

- [1] FREEDMAN G, FATTAL R. Image and video upscaling from local self-examples [J]. ACM transactions on graphics, 2011, 30(2): 1 – 10.
- [2] MICHAELI T, IRANI M. Nonparametric blind super-resolution [C]// ICCV 2013: Proceedings of the 2013 IEEE International Conference on Computer Vision. Washington, D. C.: IEEE Computer Society, 2013: 945 – 952.
- [3] MICHAELI T, IRANI M. Blind deblurring using internal patch recurrence [M]// FLEET D, PAJDLA T, SCHIELE B, et al. Computer Vision—ECCV 2014, LNCS 8691. Berlin: Springer, 2014: 783 – 798.
- [4] FREEMAN W T, JONES T R, PASZTOR E C. Example-based super-resolution [J]. IEEE Computer graphics and applications, 2002, 22(2): 56 – 65.
- [5] CHANG H, YEUNG D-Y, XIONG Y. Super-resolution through neighbor embedding [C]// CVPR 2004: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, D. C.: IEEE Computer Society, 2004, 1: 275 – 282.
- [6] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(5500): 2323 – 2326.
- [7] CHAN T M, ZHANG J, PU J, et al. Neighbor embedding based super-resolution algorithm through edge detection and feature selection [J]. Pattern recognition letters, 2009, 30(5): 494 – 502.
- [8] BEVILACQUA M, ROURNY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding [EB/OL]. [2015-04-09]. https://www.researchgate.net/publication/260351242_Low-Complexity_Single-Image_Super-Resolution_based_on_Nonnegative_Neighbor_Embedding.
- [9] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Super-resolution using neighbor embedding of back-projection residuals [C]// DSP 2013: Proceedings of the 2013 18th International Conference on Digital Signal Processing. Piscataway, NJ: IEEE, 2013: 1 – 8.
- [10] YANG J C, WRIGHT J, HUANG T, et al. Image super-resolution as sparse representation of raw image patches [C]// CVPR 2008: Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, D. C.: IEEE Computer Society, 2008: 1 – 8.
- [11] HE L, QI H, ZARETZKI R. Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution [C]// CVPR 2013: Proceedings of the 2013 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, D. C.: IEEE Computer Society, 2013: 345 – 352.
- [12] ENGAN K, SKRETTING K, HUSØY J H. Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation [J]. Digital signal processing, 2007, 17(1): 32 – 49.
- [13] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE transactions on signal processing, 2006, 54(11): 4311 – 4322.
- [14] SKRETTING K, ENGAN K. Recursive least squares dictionary learning algorithm [J]. IEEE transactions on signal processing, 2010, 58(4): 2121 – 2130.
- [15] MAIRAL J, BACH F, PONCE J, et al. Online dictionary learning for sparse coding [C]// ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning. New York: ACM, 2009: 689 – 696.
- [16] YEGANLI F, OZKARAMANLI H. Improved online dictionary learning for sparse signal representation [C]// Proceedings of the 2014 22nd Signal Processing and Communications Applications Conference. Piscataway, NJ: IEEE, 2014: 1702 – 1705.
- [17] YANG J, WANG Z, LIN Z, et al. Coupled dictionary training for image super-resolution [J]. IEEE Transaction on image processing, 2012, 21(8): 3467 – 3478.
- [18] TIMOFTE R, DE V, van GOOL L. Anchored neighborhood regression for fast example-based super-resolution [C]// ICCV 2013: Proceedings of the 2013 IEEE International Conference on Computer Vision. Washington, D. C.: IEEE Computer Society, 2013: 1920 – 1927.
- [19] TIMOFTE R, de SMET V, van GOOL L. A +: adjusted anchored neighborhood regression for fast super-resolution [M]// CREMERS D, REID I, SAITO H, et al. Computer Vision—ACCV 2014, LNCS 9006. Berlin: Springer, 2014: 111 – 126.

Background

This work was supported by the project of Natural Science Foundation of China (61303127), the project of science and technology support program of Sichuan province (2014SZ0223), the major project of Education Office of Sichuan province (13ZA0169), and the program of the “the power of science” talent cultivation plan of Sichuan province (13ZS0106), the innovation fund project of Southwest University of Science and Technology (15ycx053).

ZHAO Xiaole, born in 1987, M. S., candidate. His main research interests include digital image processing, machine learning, and data mining.

WU Yadong, born in 1979, Ph. D., professor. His research interests include digital image processing, information visualization, man-machine interaction.

Tian Jinsha, born in 1988, M. S., candidate. Her main research interests include digital image processing.

ZHANG Hongying, born in 1976, Ph. D., professor. Her main research interests include digital image processing and pattern recognition.