

基于预测稀疏编码的快速单幅图像超分辨率重建

沈 辉*, 袁晓彤, 刘青山

(江苏省大数据分析技术重点实验室(南京信息工程大学), 南京 210044)

(*通信作者电子邮箱 shenhui2013@qq.com)

摘 要:针对经典的基于稀疏编码的图像超分辨率算法在重建过程中运算量大、计算效率低的缺点,提出一种基于预测稀疏编码的单幅图像超分辨率重建算法。训练阶段,该算法在传统的稀疏编码误差函数基础上叠加编码预测误差项构造目标函数,并采用交替优化过程最小化该目标函数;测试阶段,仅需将输入的低分辨率图像块和预先训练得到的低分辨率字典相乘就能预测出重建系数,从而避免了求解稀疏回归问题。实验结果表明,与经典的基于稀疏编码的单幅图像超分辨率算法相比,该算法能够在显著减少重建阶段运算时间的同时几乎完全保留超分辨率视觉效果。

关键词:图像超分辨率;预测稀疏编码;字典学习;交替优化

中图分类号: TP391.41 **文献标志码:** A

Fast super-resolution reconstruction for single image based on predictive sparse coding

SHEN Hui*, YUAN Xiaotong, LIU Qingshan

(Jiangsu Key Laboratory of Big Data Analysis Technology (Nanjing University of Information Science and Technology), Nanjing Jiangsu 210044, China)

Abstract: The classic super-resolution algorithm via sparse coding has high computational cost during the reconstruction phase. In view of the disadvantages, a predictive sparse coding-based single image super-resolution method was proposed. In the training phase, the proposed method imposed a code prediction error term to the traditional sparse coding error function, and used an alternating minimization procedure to minimize the resultant objective function. In the testing phase, the reconstruction coefficient could be estimated by simply multiplying the low-dimensional image patch with the low-dimensional dictionary, without any need to solve sparse regression problems. The experimental results demonstrate that, compared with the classic single image super-resolution algorithm via sparse coding, the proposed method is able to significantly reduce the reconstruction time while maintaining super-resolution visual effect.

Key words: image super-resolution; predictive sparse coding; dictionary learning; alternative optimization

0 引言

图像超分辨率是指将低分辨率图像转换为高分辨率图像,是图像处理领域的热门研究方向之一。考虑到硬件成本以及硬件技术的局限性,现有的图像超分辨率工作研究主要集中在软件算法领域,主要解决的问题是将输入的单幅或者多幅低分辨率图像通过算法处理重建出一幅高分辨率图像^[1]。目前,图像超分辨率技术在卫星遥感探测、医学图像、军事监控、公共安全等领域发挥着重要的应用。

单幅图像的超分辨率主要分为基于插值的方法、基于重建的方法、基于学习的方法三大类。基于插值的算法包括经典的二次插值、双三次插值^[2]等方法,具有实现简单的优点。然而,采用插值法重建出来的图像往往存在锯齿效应,尤其是当图像放大显示时,锯齿效应更为明显。基于重建^[3-4]的方法主要利用图像的先验知识建立超分辨率模型。由于先验模型选取受主观因素影响较大,这类方法缺乏自适应性,重建出的结果也往往不令人如意。

近年来,随着机器学习理论和算法的突飞猛进,基于学习

的图像超分辨率算法成为主流方法,此类方法依托相似的训练集训练字典,能够预测出低分辨率图像中丢失的细节信息。早期,Freeman等^[5]提出的基于实例的方法,利用马尔可夫模型建立高、低分辨率图像之间的联系,但是这种算法需要数以百万计的高、低分辨率图像块进行训练,计算量非常巨大。之后,Chang等^[6]提出的局部线性嵌入的流形学习方法,学习出低分辨率图像块到高分辨率图像块的局部特征映射模型,使得高分辨率图像块能够用近邻的线性组合表示。然而,在重建的过程中,使用固定的 K 个近邻会产生过拟合或者欠拟合效应从而导致成像模糊。稀疏编码模型是近年来稀疏学习的研究热点之一。给定一组输入数据,稀疏编码的主要目标就是学习一个字典(码本)以捕捉原始数据的高维特征,使得输入信号能够在该字典上构造一组稀疏表示系数(编码)来线性重建该信号。与主成分分析方法不同的是,稀疏表示能够学习一个过完备的基并且增加了一个惩罚项来确保稀疏程度。目前,稀疏编码在特征表示上面已经有了很多成功的应用,例如图像建模方面^[7]、音频分类方面^[8]、自然语言处理方

收稿日期:2014-12-29;修回日期:2015-03-22。

基金项目:国家自然科学基金资助项目(61402232);江苏省自然科学基金资助项目(BK20141003)。

作者简介:沈辉(1990-),男,江苏如东人,硕士研究生,主要研究方向:稀疏表示、并行计算;袁晓彤(1980-),男,江苏南通人,教授,博士,CCF会员,主要研究方向:机器学习、计算机视觉;刘青山(1975-),男,安徽庐江人,教授,博士,主要研究方向:图像分析、视频分析、机器学习。

面^[9]等。文献[10]中,Yang等利用稀疏编码的方法进行单幅图像超分辨率重建,取得了不错的结果。该方法的基本思想是在训练阶段对图像分块处理,假设高、低分辨率图像块共享一个稀疏编码,通过一组高、低分辨率图像数据集训练出对应的高、低分辨率字典对。在重建过程中,对于输入的低分辨率图像块,首先利用稀疏表示模型求解出最佳的稀疏编码;然后利用高分辨率字典和低分辨率稀疏编码的乘积得到高分辨率输出图像块;最后将高分辨率图像块组合成高分辨率图像。然而,该算法在重建阶段需要对每一个图像块求解一个稀疏表示问题(等价于稀疏约束二次规划),可想而知,当图像块较多时,算法的运算量较大,重建效率低。Dong等^[11]提出了自适应选择字典和正则项的重建方法,虽然该方法提高了稀疏表示的有效性,但是重建后的图像多块状效应,并且整个重建过程也比较费时。

针对稀疏编码图像超分辨率算法重建效率低的缺点,本文提出一种新颖的基于预测稀疏编码的图像超分辨率算法。相比经典的稀疏编码算法,本文算法在原有的稀疏表示误差项的基础上增加了编码线性预测误差项,使得学习得到的低分辨率字典不仅具有稀疏编码功能,同时具有线性预测解码功能。本文采用交替优化的方法优化所提出的编解码误差函数,得到最优的高分辨率和低分辨率字典对。重建阶段,给定一个低分辨率图像块,只需要将低分辨率字典和图像块相乘就可以直接得到稀疏编码系数,然后直接利用高分辨率字典和该系数相乘重建出高分辨率块。

1 基于稀疏表示的图像超分辨率

稀疏表示理论认为一个信号可以通过一组基的线性组合表示,而且其中大多数基的系数为零或者接近于零,只有少数非零系数起到了关键作用。基于稀疏表示的图像超分辨率模型^[10]将高分辨图像和低分辨率图像分别按空间划分成一系列子块,采用联合训练的策略^[12]为高分辨率图像和低分辨率图像块分别训练两个字典 D_h 和 D_l ,并且假设了它们重建时的稀疏系数相同:

$$X_h \approx D_h A, X_l \approx D_l A \quad (1)$$

式中: X_h 是高分辨率图像块, X_l 是对应的低分辨率图像块(比如可以通过对 X_h 作下采样加模糊处理得到)。训练阶段,给定一组高低分辨率图像块样本对 X_h 和 X_l ,字典可以通过求解如下优化问题得到:

$$\min_{\{D_h, D_l, A\}} [\|X_h - D_h A\|_2^2 / N + \|X_l - D_l A\|_2^2 / M + \lambda(1/N + 1/M) \|A\|_1] \quad (2)$$

$$\text{s. t.} \quad \|D_{h,i}\|_2 \leq 1, \|D_{l,i}\|_2 \leq 1$$

式中 N 和 M 分别是高、低分辨率图像块以向量形式表达的维数。测试阶段,给定输入低分辨率图像块 y_l 和优化得到的低分辨率字典 D_l ,其稀疏重建系数可以通过求解下面的稀疏表示问题得到:

$$\min_A \|y_l - D_l A\|_2^2 + \lambda \|A\|_1 \quad (3)$$

当求得最佳的稀疏系数 A 后,重建的高分辨率图像块就可以表示为:

$$y_h = D_h A \quad (4)$$

可以看出,上述稀疏编码图像超分辨率方法在测试阶段

需要对每一个图像块求解一个稀疏编码问题以得到重建系数,该过程的计算开销是很大的。比如一幅 512×512 的原始图像,采用文献[10]提供的代码进行高分辨率重建,运行时间需要几十分钟,其中很大一部分计算时间都花在上千个图像块的稀疏表示系数求解上。

2 基于预测稀疏编码的图像超分辨率

2.1 线性预测稀疏编码

稀疏编码过程往往需要迭代求解,因此在大数据环境下耗时较多。作为稀疏编码的一个变种,预测稀疏编码在采用一个非迭代逼近步骤来取代迭代编码过程,可以显著减少测试阶段的计算量。预测稀疏编码已经在图像和视频物体识别中取得了很好的应用^[13-14]。本文中,给定采样的训练块对 $\{X_h, X_l\}$,考虑如下训练准则实现高、低分辨率字典训练和编码线性预测:

$$\begin{aligned} \min_{\{D_h, D_l, A\}} f(D_h, D_l, A) \quad (5) \\ \text{s. t.} \quad \|D_{h,i}\|_2 \leq 1, \|D_{l,i}\|_2 \leq 1 \\ f(D_h, D_l, A) = \|X_h - D_h A\|_2^2 / N + \\ \|D_h^T X_h - A\|_2^2 / N + \|X_l - D_l A\|_2^2 / M + \\ \|D_l^T X_l - A\|_2^2 / M + \lambda(1/N + 1/M) \|A\|_1 \end{aligned} \quad (6)$$

可以看出该准则函数就是在经典的稀疏编码图像超分辨率训练准则基础上分别增加了高、低分辨率线性解码项,从而可以使得训练出来的字典具有编码线性预测的功能。

2.2 优化过程

稀疏编码的优化算法已经得到广泛研究^[15-16]。与式(2)类似,本文的准则函数式(5)是一个凸问题。但是如果固定 D_h 、 D_l 和 A 其中一个,优化另一个,就成为凸优化问题。当 D_h 和 D_l 固定,关于 A 的优化过程变为一个稀疏约束二次规划问题,本文采用线性规划的方法解决。在固定 A ,更新 D_h 和 D_l 的过程中,本文采用梯度投影下降法求解如下二次优化问题:

$$\begin{aligned} \min_{\{D_h, D_l, A\}} [\|X_h - D_h A\|_2^2 / N + \|D_h^T X_h - A\|_2^2 / N + \\ \|X_l - D_l A\|_2^2 / M + \|D_l^T X_l - A\|_2^2 / M] \\ \text{s. t.} \quad \|D_{h,i}\|_2 \leq 1, \|D_{l,i}\|_2 \leq 1 \end{aligned} \quad (7)$$

整个优化过程如算法1所示。

算法1 字典优化。

- 1) 令 $t = 0$ 。采用高斯随机矩阵对字典 D_l 和 D_h 初始化处理,并将字典的每一个列都作单位归一化。
- 2) 固定 $D_h^{(t)}$ 和 $D_l^{(t)}$,采用线性规划的方法,更新 $A^{(t)}$:
 $A^{(t)} = \arg \min_A f(D_h^{(t)}, D_l^{(t)}, A)$
- 3) 固定 $A^{(t)}$,采用梯度投影下降法,更新 $D_h^{(t+1)}$ 和 $D_l^{(t+1)}$:
 $D_h^{(t+1)}, D_l^{(t+1)} = \arg \min_{D_h, D_l} f(D_h^{(t)}, D_l^{(t)}, A)$
s. t. $\|D_{h,i}\|_2 \leq 1, \|D_{l,i}\|_2 \leq 1$
- 4) 令 $t = t + 1$,迭代2)和3),直到收敛。

2.3 重建

对于给定的低分辨率输入块 y_l ,可以简单地采用如下线性方式预测稀疏系数:

$$A \approx D_l^T y_l \quad (8)$$

该线性预测函数只涉及一次矩阵和向量相乘,大大简化了编码的求解过程,显著减少了重建所需要的时间。同时,不用考虑传统算法在求解过程中稀疏正则化参数对重建结果带

来的影响。高分辨率块 y_h 的重建过程可以表示如下:

$$y_h \approx D_h A \approx D_h (D_l^T y_l) \quad (9)$$

算法2给出了整个超分辨率过程。

算法2 基于预测稀疏编码的图像超分辨率算法。

输入 预训练得到的字典 D_h 和 D_l , 一幅低分辨率图像 Y 。

输出 超分辨率图像 X^* 。

1) For 对于 Y 中的每个 5×5 小块 y_l , 从左到右, 从上到下, 在每个方向上保证4个像素重叠:

① 计算每个块中的平均像素值, 每个块的像素值减去均值;

② 计算高分辨率块 $y_h = D_h (D_l^T y_l)$;

③ 计算 $y_h + m$, 填入高分辨率图像 X_0 。

End

2) 在解空间中, 找到与 X_0 最匹配的图像:

$$X^* = \arg \min_X \|SHX - Y\|_2^2 + c \|X - X_0\|_2^2$$

H 是一个模糊算子, S 是一个下采样算子, 参见文献[10]。

3 实验结果与分析

本文采用了 Yang 等^[10] 提供的 69 张自然图像作为训练集, 进行图像尺度放大实验, 放大倍数为 2。在训练时, 选取了 100 000 个块, 每个块的大小是 5×5 , 相邻块之间重叠 4 个像素。对于彩色图像, 实验只在亮度通道内进行处理。图 1 是训练时的部分高、低分辨率图像块。整个实验的训练平台是 Intel Core i7-3770 CPU @ 3.40 GHz, 采用的 Matlab 版本是 R2011b。

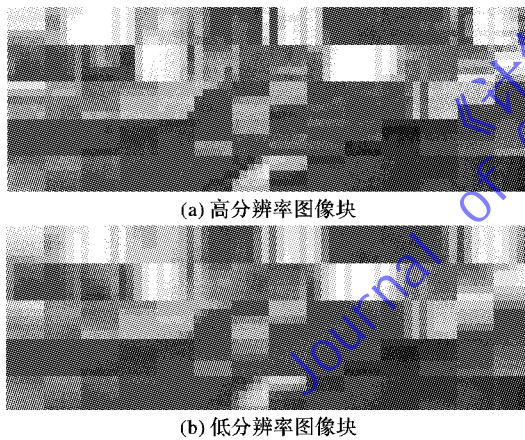


图1 高、低分辨率图像块

3.1 重建质量对比

首先运用本文提出的算法对常见的图像作了实验。所有实验采用的字典均是由文献[10]提供的训练集训练得到。图2、图3对比了本文算法和插值法、Yang 算法^[10]、Dong 算法^[11]的实验结果。

从图像的视觉效果来看, 插值法得到的结果边缘模糊, 多锯齿状, 一些细节信息丢失, 例如图2(a)中睫毛部分、眼角的纹理部分、眼睛黑白区分的地方, 图3(a)中车轴的圈纹也呈现模糊状。而 Yang 算法明显增加了很多细节信息, 图2(b)中睫毛还有眼角的纹理都能很好地重现出来, 在图3(b)中车轴的纹理能够得到清晰的体现。Dong 算法实验结果整体偏模糊, 并且有明显的块状效应, 图像的局部细节完全没有反映出来。本文算法得到的实验结果与 Yang 算法相比, 在视觉效果上相当。仔细比较图2中(b)和(d), 本文算法不但很好地

体现出了睫毛、眼角部位的细节信息, 而且与 Yang 算法的结果相比, 整体上更显干净、清晰。对比图3中(b)子图和(d), 本文算法也能得到和 Yang 算法相当的实验结果。

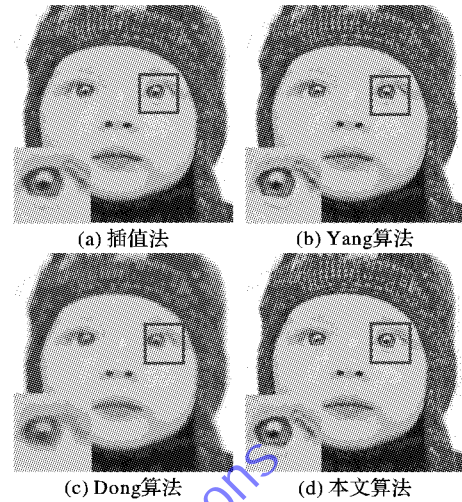


图2 Baby 图片重建结果对比

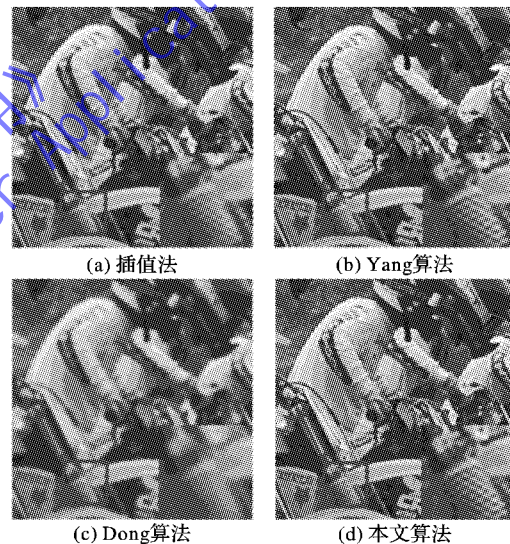


图3 Bike 图片重建结果对比

除了视觉效果这种主观评价标准^[18-19], 超分辨率图像重建结果也可以利用峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 和结构相似度^[17] (Structural Similarity, SSIM) 等客观标准进行质量评价。表1列出了本文所考虑的不同超分辨率方法重建结果的 PSNR 和 SSIM 数值。从表1中可知, 本文算法和插值法、Dong 算法相比, 在 PSNR 和 SSIM 上均略有提高。与 Yang 算法相比, 本文算法的 PSNR 和 SSIM 相当或者略有下降。

3.2 实验时间对比

本文着重强调重建时间的对比。在实验时, 本文选取了多幅大小不同的图像作为输入, 分别用本文算法、Yang 算法、Dong 算法进行实验, 每幅图像实验三次, 记录下三次的平均时间, 实验时间对比如表2所示。从表2可以看出, 本文算法在重建过程中所需要的时间明显少于 Yang 的方法和 Dong 的算法。由于本文方法在重建阶段无需再对每一个块去求解一个稀疏系数问题, 因而省去了大量的计算开销, 特别是当图像维数大的时候, 本文算法优势更加明显。可以看到, 当输入图

像大于 $128 * 128$ 时,本文算法所需时间比 Yang 的算法快了一个数量级,重建所需时间大概仅为 Yang 的十分之一。

表1 PSNR 和 SSIM 值

图像	插值法		Yang 算法		Dong 算法		本文算法	
	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM
Baby	32.5147	0.9017	32.2385	0.9280	33.2874	0.9147	33.8208	0.9247
Bike	25.6556	0.8432	28.0642	0.9058	27.0043	0.8736	27.2201	0.8867
Linda	32.7946	0.9012	34.8122	0.9277	33.4841	0.9076	34.0388	0.9194
Parrots	31.3751	0.9246	33.8079	0.9469	32.2306	0.9323	32.9729	0.9404

表2 不同输入图像大小的实验时间对比

方法	150 * 174	256 * 256	288 * 288	512 * 512	1024 * 1024	2048 * 2048
Yang 算法	26.8332	100.5228	125.8511	375.8540	1524.1078	6197.8422
Dong 算法	106.8517	247.4872	319.3603	1142.2380	2603.2970	12622.5849
本文算法	4.9830	10.4747	12.4691	36.2313	143.0850	629.9974

4 结语

本文提出了一种基于预测稀疏编码的快速单幅图像超分辨率重建算法。本文算法在经典的稀疏编码算法基础上增加了编码线性预测误差项,并且根据新的误差函数设计了字典训练方法。在重建过程中,本文直接采用低分辨率字典和输入图像块的乘积得到近似稀疏编码系数,省去了经典算法在重建过程中需要计算编码系数的运算开销。从实验效果来看,本文算法在几乎完全保留了经典稀疏编码图像超分辨率算法重建效果的同时,显著减少了重建时间。在未来的工作中,将寻求更好的优化策略,进一步减少重建时间;同时,采用分布式并行的策略快速训练出字典也将是下一步的研究方向。

参考文献:

- [1] SHAH A J, GUPTA S B. Image super resolution - a survey [C]// Proceedings of the 2012 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking. Piscataway: IEEE, 2012: 1-6.
- [2] HOU H, ANDREWS H. Cubic splines for image interpolation and digital filtering [J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1978, 26(6): 508-517.
- [3] TSAI R Y, HUANG T S. Multiple frame image restoration and registration [J]. Advances in Computer Vision and Image Processing, 1984, 1(2): 317-339.
- [4] IRANI M, PELEG S. Super resolution from image sequences [C]// Proceedings of the 10th International Conference on Pattern Recognition. Piscataway: IEEE, 1990: 115-120.
- [5] FREEMAN W T, JONES T R, PASZTOR E C. Example-based super-resolution [J]. Computer Graphics and Applications, 2002, 22(2): 56-65.
- [6] 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. Piscataway: IEEE, 2004, 1: 275-282.
- [7] COATES A, NG A Y. The importance of encoding versus training with sparse coding and vector quantization [C]// ICML 2011: Proceedings of the 28th International Conference on Machine Learning. New York: ACM, 2011: 921-928.
- [8] GROSSE R, RAINA R, KWONG H, *et al.* Shift-invariant sparse coding for audio classification [EB/OL]. [2014-12-01]. <http://axon.cs.byu.edu/Dan/778/papers/Sparse%20Coding/ng3.pdf>.
- [9] BRADLEY D M, BAGNELL J A. Differentiable sparse coding [C]// Proceedings of the Twenty-second Neural Information Processing Systems Conference. Cambridge: MIT Press, 2008: 113-120.
- [10] YANG J, WRIGHT J, HUANG T S, *et al.* Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [11] 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): 1838-1857.
- [12] YANG J, WANG Z, LIN Z, *et al.* Coupled dictionary training for image super-resolution [J]. IEEE Transactions on Image Processing, 2012, 21(8): 3467-3478.
- [13] FARABET C, LECUN Y, KAVUKCUOGLU K, *et al.* Large scale fpga-based convolutional networks [EB/OL]. [2014-12-03]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.226.3808&rep=rep1&type=pdf>.
- [14] KAVUKCUOGLU K, SERMANET P, BOUREAU Y-L, *et al.* Learning convolutional feature hierarchies for visual recognition [C]// Proceedings of the 2010 Neural Information Processing Systems Conference. Cambridge: MIT Press, 2010: 1090-1098.
- [15] MURRAY J F, KRNNETH-DELGADO K. Learning sparse over-complete codes for images [J]. The Journal of VLSI Signal Processing, 2006, 45(1/2): 97-110.
- [16] LEE H, BATTLE A, RAINA R, *et al.* Efficient sparse coding algorithms [EB/OL]. [2014-12-04]. <http://robotics.stanford.edu/~hlee/nips06-sparsecoding.pdf>.
- [17] 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.
- [18] YANG J, LIN Z, COHEN S. Fast image super-resolution based on in-place example regression [C]// CVPR2013: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2013: 1059-1066.
- [19] YANG C-Y, MA C, YANG M-H. Single-image super-resolution: a benchmark [C]// ECCV2014: Proceedings of the 13th European Conference on Computer Vision. Berlin: Springer, 2014: 372-386.