

利用位置字典对的人脸图像超分辨率方法

李玉花, 齐春

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对基于稀疏表示的人脸超分辨率算法存在的字典尺寸大、训练时间长等问题, 提出一种基于位置字典对的超分辨率重建方法. 由于同一位置的人脸图像块具有相似的结构和内容, 更有可能用相同的字典原子进行线性组合表示, 因此把训练人脸图像按位置分块, 首先为每个位置训练一个位置字典对, 利用获得的多个位置字典对, 对低分辨率测试人脸图像进行基本重建, 然后应用残差补偿方法对位置块进行补偿. 实验结果表明, 由所提方法重建的人脸图像具有更好的视觉效果, 与应用原始图像块进行稀疏表示的图像超分辨率算法相比, 平均图像结构相似度指标值提高了0.082, 同时字典训练时间缩短了约5倍.

关键词: 人脸图像; 超分辨率; 稀疏表示; 位置字典

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 0253-987X(2012)06-0007-05

A Super-Resolution Method for Face Images Using Position-Dictionary Pairs

LI Yuhua, QI Chun

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A novel method using position-dictionary pairs is proposed to reconstruct a high resolution frontal face image from a single low resolution version. Since the contents and structures in the same position of face image patches are similar and are more likely to be better represented by the combination of the same dictionary atoms, face images used for training are blocked into overlapped position-patches. Then a small position-dictionary pair is trained for each position. A basic high resolution image can be recovered for each low resolution input by using these position-dictionary pairs. Finally, the visual quality of the reconstructed image is improved by using a residual compensation method. Experimental results show that face images reconstructed by the proposed method have better visual effect. A comparison with the algorithm of image super-resolution via raw image patches based sparse representation show that the SSIM of the proposed method is 0.082 higher; while the training time is shortened about 5 times.

Keywords: face image; super-resolution; sparse representation; position-dictionary

图像超分辨率重建是指利用退化降质的低分辨率图像, 运用相应算法来获得一幅或多幅清晰的高分辨率图像的技术. 人脸图像超分辨率重建是图像超分辨率重建技术在人脸这一特定对象上的应用, 在公安刑侦、视频监控系统等领域有着广泛的应用

需求.

目前人脸图像超分辨率重建算法主要有两类: 基于重建的方法^[1-2]和基于机器学习的方法^[3-6]. 基于重建的方法是利用多幅低分辨率图像融合得到一幅高分辨率图像, 但该类方法存在固有的局限性, 即

当抽取率较大时,获取准确的亚像素运动信息是不现实的^[7].基于机器学习的方法学习高、低分辨率人脸图像间的统计先验关系,并将其用作超分辨率重建过程中的约束条件,该方法可以获得更多的图像先验信息,较基于重建的方法有很大优势.

稀疏表示作为一种新兴的信号处理和数据分析手段,已在基于机器学习的图像超分辨率重建算法^[8-9]中得到应用.其中,文献[8]提出的利用图像稀疏表示模型的图像超分辨率算法比较具有代表性.该算法通过迫使对应的高、低分辨率训练图像块在各自字典中具有相同的稀疏表示系数来离线训练一个高、低分辨率字典对,对于任一低分辨率测试图像块,求解其在低分辨率字典中的稀疏表示系数,由高分辨率字典与该系数相乘得到重建的高分辨率图像块.文献[8]算法虽然可以较好地重建高分辨率图像,但存在一些缺点,譬如由该方法重建得到的图像视觉效果不够理想,以及字典尺寸较大导致训练时间较长等.

近年来,文献[10-11]的研究结果表明:在字典训练过程中,并非所有训练图像块都具有同等的重要性,往往相似的图像块在同一字典中的稀疏表示系数更有可能具有相同的支撑集,即可以由相同的字典原子进行线性组合表示.对于人脸图像这一特定对象,由于人脸具有特殊的结构性,每幅人脸都由眼、眉、鼻子、嘴、脸颊等几部分组成,而且这些部分具有特定的相对位置,同一位置的人脸图像块的结构和内容相似,更有可能用相同的字典原子进行线性组合表示.因此,本文把人脸图像按位置分块,并为每个位置学习一个位置字典对,利用获得的多个位置字典对,对低分辨率测试人脸图像进行基本重建,然后对残差图像运用残差补偿,以进一步改善图像的视觉效果.

1 基于位置字典对的图像超分辨率原理

1.1 人脸图像的基本重建

将对齐的原始高分辨率训练人脸图像下采样得到低分辨率训练图像.把各个训练图像按位置分为若干重叠的位置块,同一位置的高分辨率块构成高分辨率训练小组,与之对应的低分辨率块构成低分辨率训练小组,对应高、低分辨率训练小组作为一个训练小组对.针对每个训练小组对,运用文献[8]中的字典对训练方式,通过迫使对应的高、低分辨率训练图像位置块在各自的位置字典中具有相同的稀疏表示系数,来获取一个位置字典对,可表示为

$$\min_{\mathbf{A}_k, \mathbf{D}_k^h, \mathbf{D}_k^l} \|\mathbf{A}_k\|_1 \quad \text{s. t.} \quad \begin{cases} \|\mathbf{F}\mathbf{X}_k^h - \mathbf{F}\mathbf{D}_k^l \mathbf{A}_k\|_F^2 \leq \epsilon_1 \\ \|\mathbf{X}_k^h - \mathbf{D}_k^l \mathbf{A}_k\|_F^2 \leq \epsilon_2 \end{cases} \quad (1)$$

式中: $\mathbf{D}_k^h$ 和 $\mathbf{D}_k^l$ 分别是第 k 个高、低分辨率位置字典; $\mathbf{A}_k$ 是稀疏系数矩阵; $\mathbf{X}_k^h$ 和 $\mathbf{X}_k^l$ 分别是第 k 个高、低分辨率训练小组; $\mathbf{F}$ 是高频特征提取算子(本文使用一阶梯度算子 $\mathbf{h}_1=[1,0,-1]$ 和 $\mathbf{h}_2=\mathbf{h}_1^T$,二阶梯度算子 $\mathbf{h}_3=[1,0,-2,0,1]$ 和 $\mathbf{h}_4=\mathbf{h}_3^T$); ϵ_1 和 ϵ_2 是较小的正实数.利用拉格朗日乘数法,式(1)可等价

$$\min_{\mathbf{D}_k, \mathbf{A}_k} \lambda_1 \|\mathbf{A}_k\|_1 + \frac{1}{2} \|\mathbf{D}_k \mathbf{A}_k - \mathbf{X}_k\|_F^2 \quad (2)$$

式中: $\mathbf{D}_k = \begin{bmatrix} \mathbf{F}\mathbf{D}_k^l \\ \beta \mathbf{D}_k^h \end{bmatrix}$; $\mathbf{X}_k = \begin{bmatrix} \mathbf{F}\mathbf{X}_k^l \\ \beta \mathbf{X}_k^h \end{bmatrix}$; β 和 λ_1 是正实数.

针对式(2)这一特定形式,目前有多种求解方式,本文运用文献[12]提出的求解方法对式(2)进行求解,得到一个位置字典对(包含一个高分辨率位置字典和一个低分辨率位置字典).离线训练位置字典对的示意图见图1.

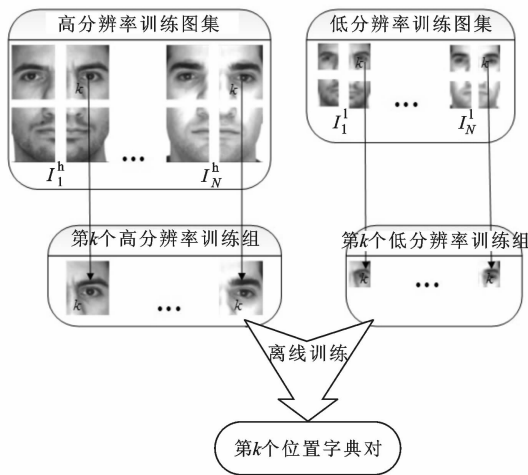


图1 离线训练位置字典对示意图

对于任一原始低分辨率测试人脸图像,将其按位置分为若干重叠的位置块,对于每个位置块,利用式(3)求解其在对应的低分辨率位置字典中的稀疏表示系数

$$\min \|\alpha_k\|_1 \quad \text{s. t.} \quad \|\mathbf{F}\mathbf{D}_k^l \alpha_k - \mathbf{y}_k\|_2^2 \leq \epsilon_1 \quad (3)$$

式中: ϵ_1 是较小的正实数; $\mathbf{y}_k$ 是第 k 个低分辨率测试图像位置块; α_k 是 $\mathbf{y}_k$ 在 $\mathbf{D}_k^l$ 中的表示系数.

由于对应的高、低分辨率人脸图像位置块在各自的位置字典中具有相同的稀疏表示系数,所以结合该系数及对应的高分辨率位置字典 $\mathbf{D}_k^h$ 可以得到基本重建的人脸图像位置块

$$\tilde{\mathbf{x}}_k^h = \mathbf{D}_k^h \mathbf{a}_k \quad (4)$$

式中: $\tilde{\mathbf{x}}_k^h$ 是原始高分辨率人脸图像位置块 $\mathbf{x}_k^h$ 的估计。

最后,把所有基本重建的人脸图像位置块按原位置放置,块与块之间的重叠部分取均值,得到基本重建的人脸图像 $\tilde{\mathbf{X}}^h$ 。

1.2 位置块残差补偿

一幅特定人脸图像一般包括两类信息:普通的人脸信息(共有)和特定的人脸信息(特有)。为丰富人脸图像的特有信息,利用上节所述方法进一步对位置块进行残差补偿。首先,利用图像集 II (与训练位置字典对所使用的图像集 I 不同)按训练位置字典对的方式训练多个残差位置字典对。其中,高、低分辨率残差训练图像由以下方式获得:对于每幅原始低分辨率训练图像,运用人脸图像的基本重建方法进行重建得到基本重建的图像;将原始高分辨率训练图像与相应的基本重建图像之差作为高分辨率残差训练图像^[13];将基本重建图像下采样得到近似低分辨率图像,由原始低分辨率训练图像减去相应的近似低分辨率图像得到低分辨率残差训练图像^[13]。然后,把低分辨率残差测试图像分为重叠的位置块,利用上节所述基本重建方法对低分辨率残差位置块进行重建。将重建得到的所有残差位置块按原位置放置,重叠区域取均值,得到一幅残差图像。最后,由基本重建的人脸图像和残差图像相加得到高分辨率人脸图像。

2 实验结果与分析

本文使用 AR 图像库进行人脸图像超分辨率重建。选取 2 056 幅图像,按两眼中心及嘴巴中心进行对齐,将人脸部分切割出来且统一到 112×88 像素,保存为灰度图像。其中 1 430 幅作为高分辨率训练图像集 I, 556 幅作为高分辨率训练图像集 II, 剩余的 70 幅作为高分辨率测试图像集。本文参照与文献[8]对所选图像做相同的处理:首先把原始低分辨率训练、测试图像进行 2 倍双 3 次插值,将得到的图像分别作为低分辨率训练图像和低分辨率测试图像;然后按以下分块原则分块,把每幅低分辨率训练图像按位置分为 8×8 像素的位置块,相邻块之间有 2 像素重叠;对应的高分辨率训练图像分为 16×16 像素的位置块,相邻块之间有 4 像素的重叠。为说明本文方法的有效性,分别采用本文方法、文献[8]方法、文献[9]方法及双 3 次插值方法,对低分辨率测试人脸图像进行 2×2 倍超分辨率重建。本文应用图像结

构相似度指标(SSIM)^[14]来衡量重建图像的质量

$$Q_{\text{SSIM}(x,y)} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (5)$$

式中: x 和 y 是非负图像信号; μ_x 和 μ_y 分别是 x 和 y 的均值; σ_x^2 和 σ_y^2 分别是 x 和 y 的方差; σ_{xy} 是 x 与 y 的协方差; $c_1 = (k_1 L)^2$ 、 $c_2 = (k_2 L)^2$ 是变量, L 是像素值的动态范围, $k_1 = 0.01$, $k_2 = 0.03$ 。

2.1 位置字典对的贡献

图 2 给出了一组不同方法对低分辨率图像重建的实验结果,可以看出,由本文方法重建的图像具有更丰富的细节信息和比较清晰的边缘,可以提供更好的视觉效果,而且图像的 SSIM 值更高;文献[8]方法的结果图比较模糊,细节信息欠缺。图 3 给出了实验结果图放大 3 倍后的局部对比,可以看出,由于本文方法利用了位置字典对进行稀疏表示和基于位



(a)实际高分辨率图像



Q_{SSIM} 0.903 4 0.906 2 0.900 4

(b)本文方法重建结果



Q_{SSIM} 0.838 8 0.819 1 0.806 1

(c)文献[8]方法重建结果



Q_{SSIM} 0.819 1 0.849 7 0.809 3

(d)双 3 次插值方法重建结果

图 2 采用 3 种方法进行重建的实验结果

置块的残差补偿,所以本文方法重建的人脸图细节信息更丰富,视觉效果更好.表 1 对比了本文方法和文献[8]方法中字典的训练时间.表 1 结果表明,本文方法在迭代 10 次、每个位置字典原子个数为 40 的情况下,63 个位置字典对的训练时间仅为 47.725 min,相对于文献[8]方法,字典的训练时间大大缩短了.在不使用残差补偿的情况下,对低分辨率人脸图像进行超分辨率重建.图 4 给出了 70 幅测试图像实验结果数据的盒图,可见本文方法提供的平均 SSIM 值最高.就 70 幅测试图像而言,本文方法的峰值信噪比(PSNR)平均值为 28.653 3,而文献[8]方法的 PSNR 平均值为 26.925 0,可见本文方法有效地提高了 PSNR 值.

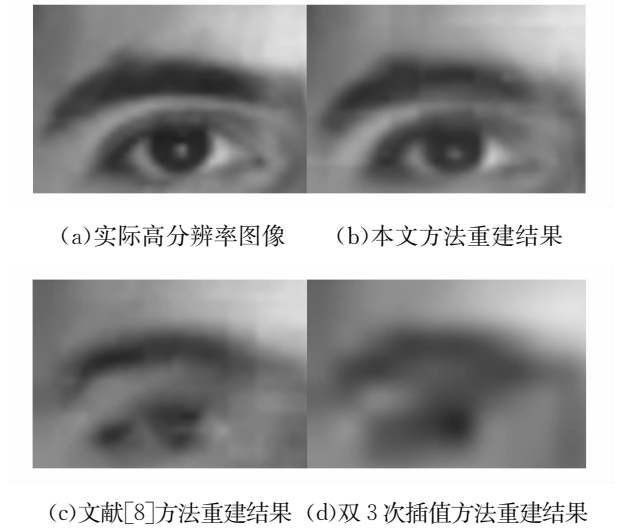


图 3 采用 3 种方法对图像细节重建结果的比较(×3)

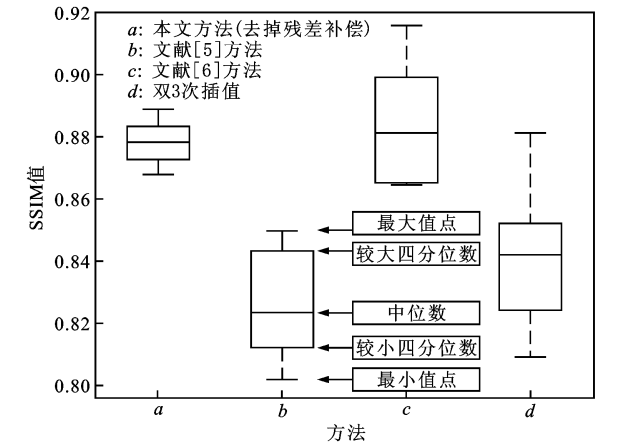


图 4 实验结果盒图

2.2 残差补偿的贡献

图 5 和图 6 对仅使用位置字典对进行人脸图像超分辨率和使用位置字典对及残差补偿进行人脸图像超分辨率这两种方式得到的图像质量进行了比

方法	参数设置		时间/min
	迭代次数	字典原子总数	
本文	10	2 520	47.725
文献[8]	10	500	317.168
文献[8]	10	800	439.139

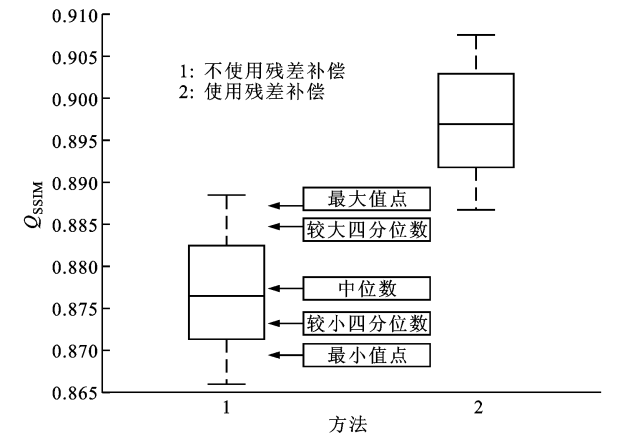
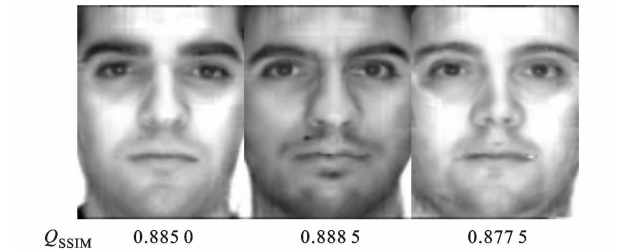
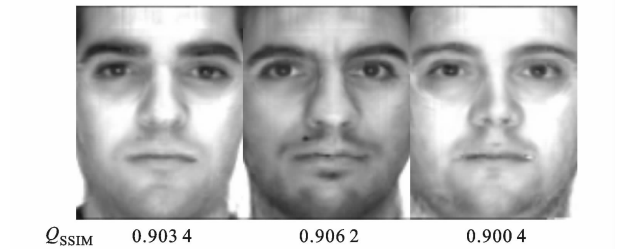


图 5 位置字典对和残差补偿的贡献比较



(a) 仅利用位置字典对时的部分结果



(b) 利用位置字典对及残差补偿时的部分结果

图 6 使用位置字典与不使用位置字典情况下的结果对比

较,可见由后一种方式得到的图像视觉效果更好且 SSIM 值更高,但差距并不十分显著.这说明了 SSIM 值的提高主要来源于位置字典的贡献,而残差补偿的作用相对要小一些.实验结果证明了本文方法的有效性.

3 结 论

本文针对文献[8]方法直接用于人脸图像时存在重建图像视觉效果差、字典尺寸较大导致字典训练时间较长等缺点,提出了位置字典对训练方法.该方法提高了重建图像的质量,且在很大程度上缩减了字典训练时间.此外,基于位置块的稀疏残差补偿方法在一定程度上提高了重建图像的视觉效果.实验结果验证了本文算法的有效性.

参考文献:

- [1] FARSIU S, ROBINSON M D, ELAD M, et al. Fast and robust multiframe super resolution [J]. IEEE Transaction on Image Processing, 2004, 13(10):1327-1344.
- [2] IRANI M, PELEG S. Super resolution from image sequences [C]//10th International Conference on Pattern Recognition. Piscataway, NJ, USA: IEEE,1990:115-120.
- [3] GLASNER D, BAGON S, RANI M. Super-resolution from a single image [C] // IEEE 12th International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2009: 349-356.
- [4] PICKUP L C. Machine learning in multi-frame image super resolution [EB/OL]. (2009-03-12)[2011-12-10]. <http://www.robots.ox.ac.uk/~vgg/publications/2008/Pickup08/>.
- [5] MAIRAL J, ELAD M, SAPIRO G. Sparse representation for color image restoration [J]. IEEE Transaction on Image Processing, 2007, 17(1):53-69.
- [6] HUANG Hua, WU Ning. Fast facial image super-resolution via local linear transformations for resources limited applications [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2011, 21(10): 1363-1377.
- [7] ROBINSON D, MILANFAR P. Fundamental performance limits in image registration [J]. IEEE Transaction on Image Processing, 2004, 13(9):1185-1199.
- [8] YANG Jianchao, WRIGHT J, MA Yi. Image super-resolution as sparse representation of raw image patches [C]//2008 Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2008: 1-8.
- [9] DONG Weisheng, ZHANG Lei, SHI Guangming. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization [J]. IEEE Transaction on Image Processing, 2011, 20(7): 1838-

1857.

- [10] HUANG Junzhou, ZHANG Tang, METACAS D. Learning with structured sparsity [C]//ICML'09 Proceedings of the 26th Annual International Conference on Machine Learning. New York, USA: ACM, 2009: 417-424.
- [11] HUANG Junzhou, HUANG Xiaolei, METAXAS D. Learning with dynamic group sparsity [C] // 2009 IEEE 12th International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2009:64-71.
- [12] LEE H, BATTLE A, RAINA R, et al. Efficient sparse coding algorithms [J]. Advances in Neural Information Processing Systems,2007,19(2):801-808.
- [13] 马祥,齐春.全局重建和位置块残差补偿的人脸图像超分辨率算法[J].西安交通大学学报,2010,44(4):9-12.
MA Xiang, QI Chun. A face image super resolution algorithm based on global reconstruction and position patch residue compensation [J]. Journal of Xi'an Jiao-tong University, 2010, 44(4):9-12.
- [14] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transaction on Image Processing, 2004,13(4):600-612.

[本刊相关文献链接]

- 蒋舒卉,钱学明.一种重要区域重采样的人脸检索方法. 2012,46(2):119-123.
- 穆为磊,高建民,陈富民,等.符合人眼视觉特性的焊缝射线数字图像增强方法. 2012,46(3):90-93.
- 赵一博,韩妙飞,闫相国.几种漫射光学成像图像重建算法的比较研究. 2012,46(4):106-111.
- 史金钢,齐春.一种双约束稀疏模型图像修复算法. 2012,46(2):6-10.
- 吴攀超,王宗义,林欣堂.采用局部差分模型描述的彩色图像配准技术. 2011,45(10):30-37.
- 徐胜军,韩九强,赵亮,等.用于图像分割的局部区域能量最小化算法. 2011,45(8):7-12.
- 郝元宏,韩静,齐春.一种新的浮选泡沫图像识别方法. 2011,45(4):104-108.
- 陈蓉,孙剑,徐宗本.彩色图像分割中基于图上半监督学习算法研究. 2011,45(2):11-14.
- 杨艺,韩崇昭,韩德强.利用特征子空间评价与多分类器融合的高光谱图像分类. 2010,44(8):20-24.
- 马祥,齐春.全局重建和位置块残差补偿的人脸图像超分辨率算法. 2010,44(4):9-12.
- 温浩,郭崇慧.利用粒子群优化的人脸特征提取识别算法. 2010,44(4):48-51.

(编辑 刘杨)