

Image Super Resolution Based on Patch Rotation and Sparse Representation

Jingman Xia¹, Wei Li¹, Jie Tang^{1,2}, Rong Liu¹, Xingcan Li²

¹Chongqing Huafu Industrial Co., Ltd., Chongqing

²Chongqing Changpeng Industrial Group Co., Ltd., Chongqing

Email: 515742230@qq.com

Received: May 3rd, 2015; accepted: May 15th, 2015; published: May 22nd, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the intelligent vehicle applications, high resolution image has become an integral part of automobile function module. However, the feature of the training image patch of the traditional sparse representation based reconstruction method is unitary, which leads to a large number of sample patches to train a dictionary. In order to reduce the number of training samples, this paper proposed a method based on patch rotation and sparse representation. By rotating the patch for different angle, the number of the patches is reduced, the feature of the patch is increased and the dictionary type becomes rich. During the reconstruction process, the adaptive weighted method is used to obtain the high resolution image. Experiments show that, the proposed method compared with the traditional method, not only has significant improvement in subjective quality, but also greatly improves the objective quality.

Keywords

Rotation, Adaptive Weighted, High Resolution, Sparse Representation

基于块旋转和稀疏表示的图像超分辨率重建

夏静满¹, 厉伟¹, 汤捷^{1,2}, 刘荣¹, 李星灿²

¹重庆华福车船电子设备制造有限公司, 重庆

²重庆长鹏实业(集团)有限公司, 重庆

Email: 515742230@qq.com

收稿日期：2015年5月3日；录用日期：2015年5月15日；发布日期：2015年5月22日

摘要

在智能车应用领域，高分辨率的图像已经成为汽车功能模块中不可或缺的一部分。然而传统的基于稀疏表示的高分辨率图像重建方法中所用的训练样本块特征单一，这就导致需要大量的样本块来训练字典。为了减少训练样本块，本文提出一种基于块旋转策略和稀疏表示的超分辨率重建算法。通过将图像块旋转不同的角度，从而减少样本块，增加特征数量，丰富训练字典的类型。在重建过程中，采用自适应加权求和的方式求得高分辨率图像。实验证明，所提出的方法较传统的方法，不仅在主观质量上有明显的提升，在客观质量上也有较大幅度的提高。

关键词

旋转，自适应加权，高分辨率，稀疏表示

1. 引言

在现代智能车应用领域，越来越需要高分辨率的图像。然而，在一些条件下，由车载成像设别所拍摄的图像分辨率不高，这就明显的给后续处理和智能控制带来不便。为了在不改变硬件水平的条件下，增加图像的分辨率，超分辨率重建技术越来越受到关注。图像的超分辨率重建方法主要可以分为三个主要的方面：基于插值的方法、基于重建的方法和基于学习的方法。

基于插值的假设低分辨率图像是由高分辨率图像直接下采样得到的，因此对低分辨率图像进行上采样可获得高分辨率图像。在上采样的过程中，去混叠能力是非常重要的。插值方法又可以分为最近邻插值，双线性插值和双三次插值。最近邻插值相对简单，但是容易出现块效应。双线性插值虽然比最近邻插值复杂，但是有振铃效应。双三次插值最复杂，但是容易出现过拟合现象。

基于重建的方法是建立在对图像退化模型的合理建模上。这种方法还要结合相关的一些先验信息来重建高分辨率图像。对多张低分辨率图像融合时，最重要的是图像之间的运动估计与配准，精确的运动估计和配准有利于重建出高质量的高分辨率图像[1]。目前广泛研究的基于重建的算法有迭代反向投影算法、凸集投影(POCS)算法、最大后验概率(MAP)算法，混合 MAP/POCS 法等[2]-[6]。基于重建的方法需要大量的先验知识来保证重建图像的质量，但是这就极大的增加了重建所需要的时间，而且对重建过程中初始值的选取有很大的依赖，在重建倍数较大时，重建结果较差。

最近，基于学习的方法越来越受到关注。Freeman 等[7]首次提出了基于学习的超分辨率重建概念，其基本思想是利用机器学习的方法训练得到高、低分辨率图像块之间的对应关系，然后利用这种相互之间的关系来预测高分辨率图像块。在文献[7]中 Freeman 等先分别从高、低分辨率图像库中采集对应的高、低分辨率图像块训练集，采用 Markov 网络模型估计低分辨率和高分辨率图像块间的关系，对于要重建的低分辨率图像将其分块后，用高、低分辨率图像块之间的关系去预测每个待重建图像块在 Markov 网络的位置，这样就能寻找到和每个待重建图像块对应的高分辨率块，将它们进行组合之后就可以得到一幅高分辨率图像。受流形学习方法的启发，Chang 等[8]提出了基于局部邻域嵌入(Locally Linear Embedding, LLE)的单幅图像超分辨率重建方法。为了克服大量的样本块对重建时间的影响，Yang 等[9]提出的基于稀疏表示的图像超分辨率重建算法。这种方法通过大量对应的高低分辨率图像块的联合学习，强迫它们具有同

样的稀疏表示，得到一对高、低分辨率冗余字典。然后，利用稀疏分解可获得待重建低分辨率图像块在低分辨率冗余字典下的稀疏表示，这种稀疏表示可以对高分辨率冗余字典中的原子块进行线性组合，从而得到重建后的高分辨率图像块，然后高分辨率图像就可以通过将把这些图像块在考虑块之间的重叠的条件下进行组合便可得到高分辨率图像。

但是这种传统的基于稀疏表示的方法在训练字典的时候，需要大量的样本，而且不能保证这些样本代表图像的大部分特征。为了克服这个问题，本文提出一种基于块旋转策略和稀疏表示的超分辨率重建算法。通过将图像块旋转不同的角度，从而减少样本块，增加特征数量，丰富训练字典的类型。在重建过程中，采用不同类别自适应加权求和的方式求得高分辨率图像。实验表明，所提出的方法与传统的方法相比，不仅在主观质量上有明显的提升，在客观质量上也有较大幅度的提高。

2. 传统的基于稀疏表示的图像超分辨率重建

2.1. 字典训练

在传统的基于稀疏表示超分辨率重建方法中，首先要训练对应的字典，即高分辨率字典和低分辨率字典。在训练字典之前，需要从训练图像中选取对应的高分辨率图像块和低分辨率图像块。Yang 等提出联合字典训练模型[9]：

$$\{D_h, D_l\} = \arg \min_{D_h, D_l} \frac{1}{N} \|X_h - D_h \alpha\|_2^2 + \frac{1}{M} \|Y_l - D_l \alpha\|_2^2 + \lambda \left(\frac{1}{N} + \frac{1}{M} \right) \|\alpha\|_1 \quad (1)$$

式中 N 和 M 分别代表高分辨率图像块和低分辨率图像块变换为列向量形式时的维度， α 是稀疏表示系数， $X_h = \{X_1, \dots, X_n\}$ 表示高分辨率图像块构成的集合， $Y_l = \{Y_1, \dots, Y_n\}$ 表示低分辨率图像块构成的集合。

更进一步的，可以将式(1)简化为如下的形式：

$$D = \arg \min_D \|X - D\alpha\|_2^2 + \hat{\lambda} \|\alpha\|_1 \quad (2)$$

$$\text{式中, } X = \begin{bmatrix} \frac{1}{\sqrt{N}} X_h \\ \frac{1}{\sqrt{M}} Y_l \end{bmatrix}, \quad D = \begin{bmatrix} \frac{1}{\sqrt{N}} D_h \\ \frac{1}{\sqrt{M}} D_l \end{bmatrix}, \quad \hat{\lambda} = \lambda \left(\frac{1}{N} + \frac{1}{M} \right).$$

2.2. 约束重建

对低分辨率图像的重建是分别对每一个低分辨率图像块进行重建的。对于每一个待重建的低分辨率图像块 Y_l^i ，首先利用公式(3)求得在低分辨率字典 D_l 下的稀疏表示系数：

$$\alpha_i = \min_{\alpha_i} \left\{ \|Y_l^i - D_l \alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \right\} \quad (3)$$

强迫高分辨率图像块与低分辨率图像块共享同一个稀疏表示系数，那么高分辨率图像块就可以根据公式(4)来求得：

$$\hat{X}_h^i = D_h \alpha_i \quad (4)$$

融合所有重建的高分辨率图像块，就得到最终的高分辨率图像输出：

$$X = \left(\sum_i R_i^T R_i \right)^{-1} \left(\sum_i R_i^T \hat{X}_h^i \right) \quad (5)$$

3. 所提出的方法

3.1. 训练块旋转策略

传统的基于稀疏表示的方法，虽然选取了大量的训练样本块来保证训练字典的质量，但是，当选取的样本块很多时，计算的复杂度明显上升，而且，这些图像块还不一定能够包含多种图像块的特征。所以为了克服这个缺点，本文提出基于块旋转策略的训练样本构成方法。

假设在某个图像中，选取的训练用高分辨率图像为 X_i ，则将这个图像块按照旋转算子作旋转处理。本文中，图像块分别做 45 度，135 度，225 度和 315 度的旋转操作，旋转之后，将这些图像块划归到同一个尺度，以保证训练过程的正确进行，即对这些块做列向量的处理。那么这些虽然在内容上是相同的，但是，进行列向量处理之后，相同的内容就有了不同的表现形式，那么训练得到的字典也就不相同。对于图像块 X_i ，经过上述的旋转过程之后，构成了旋转集合 $\hat{X}_i = \{X_i, X_i^{45}, X_i^{135}, X_i^{225}, X_i^{315}\}$ 。同样的，在低分辨率图像的对应位置，选取了低分辨率图像块集合 $Y_i = \{Y_i, \dots, Y_n\}$ 。

3.2. 字典训练模型

在 2.1 节中，将图像块做旋转之后，高分辨率图像块和低分辨率图像块的集合中所包含的样本数量是明显不同的。而且，高分辨率图像块的样本数是低分辨率图像块样本数的 5 倍。对于同一个低分辨率图像块 Y_i 来说，就会有 5 个高分辨率的图像块与之相对应。那么，在训练字典的时候，就有必要分别对每一种对应情况做字典训练。也就是说，共形成 5 种训练样本集： $\{Y_i, X_i\}_{i=1}^n$ ， $\{Y_i, X_i^{45}\}_{i=1}^n$ ， $\{Y_i, X_i^{135}\}_{i=1}^n$ ， $\{Y_i, X_i^{225}\}_{i=1}^n$ ， $\{Y_i, X_i^{315}\}_{i=1}^n$ 。这样，当选取的高分辨率图像块一定时，通过旋转处理，就得到不同的特征表现形式。对每一种情况做字典训练，就得到如下的联合字典训练模型：

$$\begin{aligned} & \{D_h, D_h^{45}, D_h^{135}, D_h^{225}, D_h^{315}, D_l\} \\ & = \arg \min_{D_h, D_l} \frac{1}{N} \|X_h - D_h \alpha\|_2^2 + \frac{1}{N} \|X_h^{45} - D_h^{45} \alpha\|_2^2 + \frac{1}{N} \|X_h^{135} - D_h^{135} \alpha\|_2^2 \\ & \quad + \frac{1}{N} \|X_h^{225} - D_h^{225} \alpha\|_2^2 + \frac{1}{N} \|X_h^{315} - D_h^{315} \alpha\|_2^2 + \frac{1}{M} \|Y_i - D_l \alpha\|_2^2 + \lambda \left(\frac{5}{N} + \frac{1}{M} \right) \|\alpha\| \end{aligned} \quad (6)$$

式中， $D_h, D_h^{45}, D_h^{135}, D_h^{225}, D_h^{315}$ 分别是未经旋转，旋转 45 度，旋转 135 度，旋转 225 度，以及旋转 315 度时所获得的字典。在训练字典的过程中，主要采用文献[9]中的联合字典训练的方法得到对应的字典。

通过旋转处理来训练字典一个明显的好处是，减少了从原始图像中提取的图像块的数量，以及保证了特征的多种表现形式。只需要选取少量的图像块，就可以实现字典表现的完整性。

3.3. 加权重建模型

对于每一个待重建的低分辨率图像块 Y_i ，首先用训练得到的低分辨率字典 D_l 求得其稀疏表示系数 α_i 。然后，分别用训练得到的五个高分辨率字典得到对应的图像块：

$$\hat{X}_i = D_h \alpha_i \quad (7)$$

$$\hat{X}_i^{45} = D_h^{45} \alpha_i \quad (8)$$

$$\hat{X}_i^{135} = D_h^{135} \alpha_i \quad (9)$$

$$\hat{X}_i^{225} = D_h^{225} \alpha_i \quad (10)$$

$$\hat{X}_i^{315} = D_h^{315} \alpha_i \quad (11)$$

这五类高分辨率的图像块得到之后，一个可行的办法就是将这五类图像块进行融合，得到最终的高

分辨率图像块。本文提出一种自适应加权的块融合方法。公式(7)就是传统的基于稀疏表示的图像重建公式。虽然由公式(7)重建的图像块可以作为高分辨率图像块的近似，但是，这种方法重建的结果中容易出现振铃和瑕疵。将此结果作为初始结果，本文提出的自适应加权的重建模型如下：

$$X_i^* = \hat{X}_i \exp\left(-\frac{\|\hat{X}_i - \hat{X}_i\|_2^2 + \varepsilon}{h}\right) + \hat{X}_i^{45} \exp\left(-\frac{\|\hat{X}_i^{45} - \hat{X}_i\|_2^2}{h}\right) + \hat{X}_i^{135} \exp\left(-\frac{\|\hat{X}_i^{135} - \hat{X}_i\|_2^2}{h}\right) + \hat{X}_i^{225} \exp\left(-\frac{\|\hat{X}_i^{225} - \hat{X}_i\|_2^2}{h}\right) + \hat{X}_i^{315} \exp\left(-\frac{\|\hat{X}_i^{315} - \hat{X}_i\|_2^2}{h}\right) \quad (12)$$

式中， ε 是非常小的数， h 是指数参数。通过融合所有的图像块，就得到初始的高分辨率图像 X_0 。

3.4. 后处理模型

低分辨率图像的获取模型一般表示为：

$$Y = SHX \quad (13)$$

为了使得重建的图像满足低分辨率图像的获取模型，以及提升重建图像的质量，采用全局约束后处理模型来优化最终的结果。这种全局优化模型表示为：

$$X = \arg \min_x \|SHX - Y\|_2^2 + \beta \|X - X_0\|_2^2 \quad (14)$$

此公式可以用简单的梯度下降算法来求解。

4. 实验验证分析

4.1. 实验设置

所有的实验都是在 CPU 为 2.80 GHz 和 2.90 GB 内存的台式电脑上实现。我们选择 30 幅图像作为训练图像，测试图像如图 1 所示。选取 30,000 个图像块来训练字典。字典大小设置为 512。块大小设置为 5×5 ，



Figure 1. Test images

图 1. 测试图像

Table 1. PSNR and SSIM by different methods
表 1. 不同方法的 PSNR 和 SSIM 值

Image	Measures	Methods			
		Bicubic	LLE	SCSR	所提方法
Butterfly	PSNR	25.71	25.26	27.11	27.75
	SSIM	0.889	0.870	0.910	0.921
Peppers	PSNR	29.28	28.87	30.28	30.79
	SSIM	0.898	0.882	0.910	0.919
Lena	PSNR	29.79	29.95	30.76	31.01
	SSIM	0.824	0.816	0.847	0.853
Parrot	PSNR	30.70	29.99	31.86	32.01
	SSIM	0.880	0.864	0.897	0.900
Koala	PSNR	29.43	29.01	30.32	30.73
	SSIM	0.806	0.787	0.841	0.845
Moth	PSNR	25.99	25.95	27.16	27.71
	SSIM	0.825	0.806	0.852	0.868
Tiger	PSNR	24.53	23.85	25.49	25.74
	SSIM	0.805	0.763	0.848	0.856

重叠 4 个像素。用峰值信噪比(PSNR)和结构相似度(SSIM)来评价重建图像的客观质量。实验中，将双三次插值(Bicubic)、局部线性嵌入(LLE)和传统的基于稀疏表示的重建方法(SCSR)与本文方法对比。

4.2. 实验结果

表 1 中展示了不同算法重建的图像其 PSNR 和 SSIM 值。从表 1 中可以看出 SCSR 方法相比 Bicubic、LLE 方法有明显提高，而且与学习方法中的 LLE 方法相比也有较大幅度的提高。本文所提出的方法与 SCSR 方法相比，在 PSNR 和 SSIM 上均有提高。这是由于将训练图像块做旋转处理之后，增加了特征数量，而且字典所表征的特征也更多。

为了展示不同算法的视觉效果，我们选取了 Parrot 的重建结果以及局部放大图像，如图 2 所示。从图中我们可以观察到双三次插值不能很好地重建高频细节，而且图像出现模糊。虽然 LLE 方法能够恢复一些高频细节，但是在重建结果中出现明显的块效应，重建结果不真实。SCSR 方法能够复原大量的高频细节和尖锐的边缘，但是在重建结果中出现模糊的边缘和不准确的几何结构。本文所提的方法不仅能够重建较好的边缘，纹理和细节，而且视觉效果更好。

5. 结论

为了克服智能车应用领域，获取图像的分辨率较低的问题，本文提出一种基于块旋转策略和稀疏表示的图像超分辨率重建方法。在稀疏表示模型下，块旋转策略主要是为了减少样本量，增加样本特征数据。为了重建高分辨率图像，提出一种自适应加权融合的图像重建方法。实验结果证明了所提方法在汽

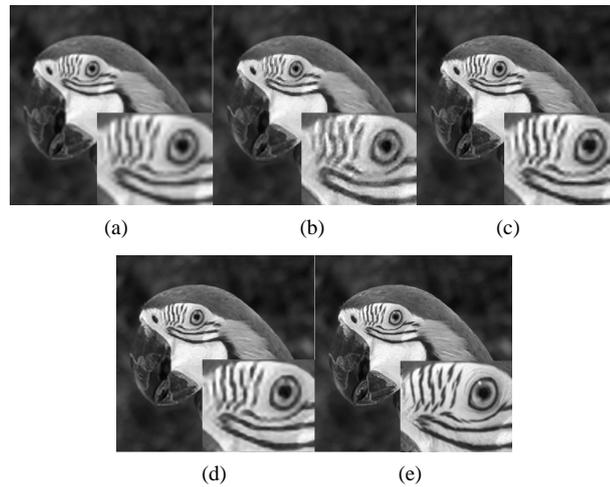


Figure 2. Parrot images recovered by different methods with a factor of 3; (a) Bicubic; (b) LLE; (c) SCSR; (d) Proposed method; (e) Original HR image

图 2. 不同方法 3 倍重建的 Parrot 图像; (a) Bicubic; (b) LLE; (c) SCSR; (d) 所提方法; (e) 原始高分辨率图像

车功能模块的研发设计中具有一定的应用潜力。

基金项目

重庆市应用开发计划项目(cstc2013yykfC60006), 重庆市应用开发计划项目(cstc2013yykfB60006)。

参考文献 (References)

- [1] 刘良辰 (2012) 基于整体到局部的分布式人脸超分辨率重建策略. 重庆大学硕士学位论文, 重庆.
- [2] Irani, M. and Peleg, S. (1991) Improving resolution by image registration. *CVGIP: Graphical Models and Image Proceedings*, **53**, 231-239.
- [3] Takeda, H. and Kskoui, P. (1989) High resolution image recovery from image-plane array using convex projections. *Journal of the Optical Society of America A*, **6**, 715-726.
- [4] 孟庆武 (2004) 预估计混叠度的 MAP 超分辨率处理算法. *软件学报*, **2**, 207-214.
- [5] 杨妮 (2011) 基于混合 MAP/POCS 的序列图像超分辨率重建算法研究. 昆明理工大学, 昆明.
- [6] Su, B.H. and Jin, W.Q. (2003) POCS-MAP based super-resolution image restoration. *Acta Photonica Sinica*, **32**, 502-504.
- [7] Freeman, W.T., Jones, T.R. and Pasztor, E.C. (2002) Example-based super-resolution. *IEEE Computer Graphics and Applications*, **22**, 56-65.
- [8] Chang, H., Yenng, D.Y. and Xiong, Y. (2004) Super-resolution through neighbor embedding. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, 275-282.
- [9] Yang, J.C., Wright, J. and Huang, T. (2010) Image super-Resolution via sparse representation. *IEEE Transaction on Image Processing*, **19**, 2861-2873.