

基于压缩感知和 SVR 的自学习单幅图像超分辨率重建

秦绪佳^{1,2} 单扬洋¹ 肖佳吉¹ 郑红波¹ 张美玉¹

(浙江工业大学计算机科学与技术学院 杭州 310032)¹

(浙江省可视媒体智能处理技术研究重点实验室 杭州 310023)²

摘要 针对依靠外部图像库的超分辨率(SR)重建算法训练耗时长、容易出现错误高频细节的问题,提出了一种基于压缩感知(CS)理论和支持向量回归(SVR)的单幅图像超分辨率重建方法。对降质图像本身训练 SVR 模型,充分挖掘图像自身的自相似特点。训练过程中先对输入图像边缘进行检测并对图像块进行分类,然后稀疏编码图像块,再根据图像的标签向量和稀疏表示矩阵训练得到 SVR 模型,并在测试过程中利用该模型预测高分辨率(HR)图像。实验结果表明,与基于外部库方法重建图像的方法相比,该算法所得结果的细节更加真实;与双三次插值方法相比该算法所得结果的边缘更加清晰。

关键词 超分辨率重建,压缩感知,支持向量回归,双三次插值

中图法分类号 TP391 文献标识码 A

Self-learning Single Image Super-resolution Reconstruction Based on Compressive Sensing and SVR

QIN Xu-jia^{1,2} SHAN Yang-yang¹ XIAO Jia-ji¹ ZHENG Hong-bo¹ ZHANG Mei-yu¹

(School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310032, China)¹

(Key Laboratory of Visual Media Intelligent Processing Technology of Zhejiang Province, Hangzhou 310023, China)²

Abstract For the long learning time and the easiness to occur wrong and high frequency details of super-resolution(SR) reconstruction algorithm which traditionally depends on external image database, this paper presented a single image SR reconstruction method based on compressive sensing(CS) and support vector regression(SVR). SVR model is training for degraded image itself to make full use of the self similarity of the image. In training stage, we firstly detected image edge and classified image patch into low and high frequency blocks. Then we did image block sparse coding, and trained a SVR model using image's label vector and sparse representation matrix. Finally, we predicted the high resolution (HR) image with SVR using this model in the testing stage. Experiments show that the proposed method is more realistic than the method based on external library, and the edge is more clear than bicubic interpolation method.

Keywords Super-resolution reconstruction, Compressive sensing, Support vector regression(SVR), Bicubic interpolation

1 引言

图像超分辨率(Super-Resolution, SR)重建技术在医学诊断、生物信息识别等方面具有重大实际意义,其主要内容是对一幅或多幅低分辨率(Low Resolution, LR)图像进行重建得到一幅高分辨率(High Resolution, HR)图像,从而提高图像的视觉质量。图像之间是否包含互补信息是实现 SR 重建的前提条件。根据互补信息的来源方式可大致将 SR 重建技术分为三大类型:基于插值的方法、基于重建的方法和基于学习的方法。基于插值的方法的原理通俗易懂,且运算量较小,常应用在实时 SR 重建中,常用的插值方法有最近邻插值(NN 插值)、双线性插值(Bilinear 插值)、双三次插值(Bicubic 插值)等。基于重建的方法直接从多幅空间上的 LR 图像中来获取它们的互补信息,可将其细分为基于运动的、基于尺度的和基于模糊的等。基于重建的方法也存在一些不足,该类方法一般利用正则化模,HR 图像的先验知识就是这些正则化

约束项(通常定义为平滑约束项),然而这些正则化约束项会随着放大倍数的增加使重建图像过于平滑。LR 图像所能提供的互补信息会随着分辨率倍数的提高而减少,单依靠增加 LR 图像数量来产生高频细节难以取得满意的重建效果。研究表明基于重建的超分辨率重建效果在大于 2 倍后就无法满足实际需求^[1]。

Donoho 等人于 2006 年提出新的信号采样理论,即压缩感知理论(CS 理论),其主要内容是若信号在变换域是稀疏的,则可以利用与变换矩阵非相干的观测矩阵来对稀疏表示系数进行测量,从而得到低维的线性投影,最后通过解决一个优化问题来得到近似值的原始信。Yang 等人^[2]于 2008 年第一次在 SR 领域应用了压缩感知的理论,将信号的稀疏表示与图像超分辨率相结合。他们通过学习外部图像库,训练 LR 稀疏表示字典 D_l 和 HR 稀疏表示字典 D_h 并建立它们之间的对应关系,对 LR 图像 Y 从左上角提取 3×3 大小的图像块 y (水平、垂直方向各重叠一个像素),把每一个 LR 图像块 y

本文受国家自然科学基金(61672462, 61672463),浙江省科技计划(2016C33165)资助。

秦绪佳(1968—),男,博士,教授,博士生导师,CCF 会员,主要研究方向为计算机图形学, E-mail: qxj@zjut.edu.cn;单扬洋(1993—),男,硕士生,主要研究方向为图像处理;肖佳吉(1990—),男,硕士生,主要研究方向为图像处理;郑红波(1977—),女,博士,讲师,主要研究方向为图像处理, E-mail: zhb@zjut.edu.cn(通信作者);张美玉(1965—),女,硕士,教授,主要研究方向为图像处理。

表示成 D_i 和稀疏向量 α 的乘积 ($y=D_i\alpha$), 然后求解以下最优优化问题:

$$\begin{aligned} \min \quad & \|\alpha\|_1 \\ \text{s. t.} \quad & \|FD_i\alpha - Fy\|_2 \leq \epsilon_1 \\ & \|PD_h\alpha - w\|_2 \leq \epsilon_2 \end{aligned} \quad (1)$$

将求得的最优解 α^* 与 D_h 相乘即可得到高分辨率图像块 x , 并将其填充到 HR 网格, 遍历完成后便得到最终的 HR 图像。将其与传统的反投影方法、Freeman 等人基于学习的方法^[3]和 Dai S 等人的软边缘方法^[4]的重建效果做了充分比较, 实验结果表明其重建后的图像边缘更加清晰, 主观视觉效果优于 Freeman 等人基于学习的方法。该文献引起了许多学者的关注, 在此之后各国大量学者投入到基于 CS 理论的图像 SR 重建的研究中。美国 Rice 大学于 2009 年成功研制出的单像素相机^[5]也是基于 CS 理论。Yang 等人^[6]于 2013 年提出了一种通过单幅图像自学习来进行 SR 重建的框架, 该框架不但不需要训练数据的集合, 而且也不需要事先保证图像块存在自相似性。他们根据贝叶斯理论证明该框架不仅可以学习还可选择最优的 SVR 模型, 从而达到 SR 重建误差最小的目的。Zhang 等人^[7]利用局部和非局部的先验知识并引入多尺度字典的 SR 重建方法, 其中局部先验使用 steering 核回归从一个小的局部区域中预测目标像素值来达到抑制伪影的目的, 非局部先验把大邻域的加权平均作为目标像素值的预测来丰富视觉细节。Yu 等人^[8]基于高斯混合模型和 MAP-EM 算法提出一个分段线性估计 (PLE) 的通用框架, 该框架主要用来解决图像反问题。Michael 等人^[9]于 2014 年提出了一种基于 LR 和 HR 图像块稀疏表示的统计预测模型, 该模型通过误差平方和均值最小估计 (MMSE) 来得到 HR 图像块的预测, 并可以使用前馈神经网络理论来进行合理解释。他们在计算复杂度、数值标准和主观视觉 3 个方面与已有的基于高低分辨率字典对的方法做了充分的比较。实验结果表明他们的方法在计算复杂性和重建质量上取得了令人满意的权衡结果。

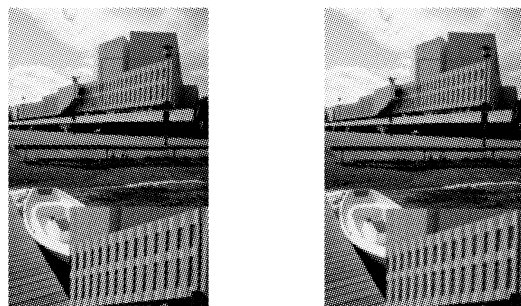
支持向量回归 (Support Vector Regression, SVR) 是支持向量机的推广, 它能够在高维特征空间拟合数据而无需事先知道数据的分布情况, 其泛化能力使得它能较好地预测未知输出。Li 等人^[10]早在 2005 年就把 SVR 运用于盲图像进行卷积运算, 并取得了比最大似然估计更好的效果。随后, 基于 SVR 的超分辨率重建方法也不断被提出^[11-12]。

本文基于文献^[12]的思想, 提出了一种结合图像稀疏表示和 SVR 的自学习 SR 重建方法。本文第 2 节介绍图像内部的自相似性及自学习的优势; 第 3 节详细介绍所提算法的改进及算法的具体步骤; 第 4 节为实验结果及讨论, 并将已有方法和本文方法进行比较分析; 最后总结全文。

2 图像内部的自相似性

Daniel Glasner 等人于 2009 年的研究^[13]表明一幅自然图像内部有多处地方重复出现图像块, 并且在不同尺度放大后的图像之间也会重复出现。Glasner 等人提出了一种基于重建和示例学习的 SR 重建结合的统一框架。应用该方法进行图像重建的图像质量比 2002 年 Freeman 等人的实验结果更高。Maria Zontak 和 Michal Irani^[14]于 2011 年的研究表明, 与普通的外部图像统计信息相比, 能够提供更有力的图像先验的却是待处理图像本身的内部统计信息。文献^[2]对采用待处理图像本身进行 SR 重建与采用 200 幅来源于外部图

像库的图像进行 SR 重建进行对比, 在直观上说明根据图像本身的自相关信息得到的图像在视觉效果上更好。图 1 给出了采用图像自身信息的 SR 重建结果 (见图 1(a)) 和采用外部图像库的 SR 重建结果 (见图 1(b))。



(a) 利用图像本身重建

(b) 利用外部库图像重建

图 1 利用不同统计信息的 SR 重建效果

对比两幅图中建筑物的窗户部分, 可以清楚地看到图 1(a) 采用待处理图像自身统计信息重建得到的图像窗户边缘更加锋利, 但在图 1(b) 中的对应部分却出现了比较严重的振铃效应。

基于外部库学习的方法的缺点如下: 1) 因为 HR 图像块来源于外部, 图像细节并不真实, 所以它在很大程度上属于一种幻想出的图像; 2) 如果外部图像库与当前待处理图像属于不同类型图像, 则容易产生错误高频细节问题; 3) 一般需要很大的外部训练库且训练时间较长, 而基于待处理图像自身的方法则不需要采用这种庞大的外部训练库, 其时间复杂度较低, 图像所学习到的信息与当前图像高度相关。

3 基于 CS 和 SVR 的自学习图像 SR 重建

3.1 算法的整体处理流程

图 2 给出了改进的 SR 重建算法的流程图。SR 重建算法的处理过程主要由 3 个部分构成: 图像块分类、图像块稀疏表示、SVR 模型学习和预测。

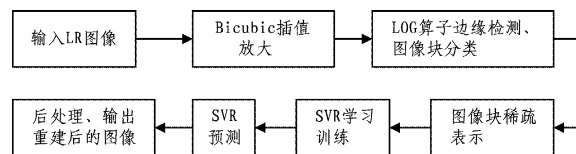


图 2 改进方法的流程图

3.2 图像块分类

待处理图像的高频图像块需要学习一个 SVR 模型, 低频图像块也需如此, 即需对不同的图像特征进行 SVR 学习和预测, 因此根据图像当前像素点的位置信息 (即属于边缘或背景), 再对其周围一定尺寸的图像块进行分类保存。具体方法如下:

首先, 将输入的低分辨率图像记为 $inputImg$, 采用 bicubic 插值方法将其放大 2 倍, 获得高分辨率插值图像 $interpImg$ 。为实现对插值图像 $interpImg$ 进行高、低频分块, 我们需要对插值图像 $interpImg$ 进行边缘检测, 将边缘检测后的二值图像记为 $HRBndyImg$ 。

文献^[12]采用 Mean Shift 算法来过分割图像。相关文献表明, Mean Shift 算法在确定初始密度中心时, 若是随机地选取特征空间的初始位置, 则可能无法得到较好的聚类结果, 最终使得分割后的图像不一致。另外, 过分割图像是指对图像的过度分割, 即把原本属于一个整体的目标分成了多个。经

过分割处理后,可能将本是背景的像素当成了某个物体的边缘,从而导致在后面的图像块分类处理中会把原本应该归类为背景集合的图像块归类到边缘集合中,可能使之后训练得到的 SVR 模型不准确,因此我们考虑使用其他的图像边缘检测算法。基于一阶导数的 Rober 和 Sobel 算子、基于二阶导数的 Laplacian 算子等是常用的图像锐化算子。一阶导数容易出现宽边缘,二阶导数对细节响应更为敏感,因此采用基于二阶导数的 Laplacian 算子对插值图像 interpImg 进行边缘检测。鉴于输入的 inputImg 图像可能含有噪声,而 Laplacian 算子对噪声点的响应更加明显,因此需要在锐化之前对其进行平滑处理。采用高斯平滑算子来完成这一过程,把高斯平滑算子和 Laplacian 算子相结合即为高斯拉普拉斯算子(Laplacian of a Gaussian, LoG)。我们使用 Matlab 自带的 edge 函数,以 LoG 算子作为参数,得到二值图像 HRBndyImg 。将 person512 缩小 1 倍得到的 256×256 像素的图像作为测试输入图像(见图 3(a)),其灰度值范围为 $0 \sim 255$,图 3(b)为输入图像插值放大 2 倍的图像,图 3(c)为插值放大图像的边缘检测结果。

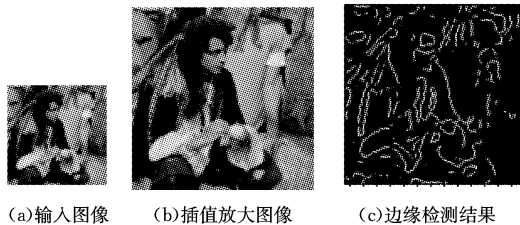


图 3 图像插值放大及边缘检测

然后,对插值放大图像块进行分类。将图像中的第 j 行、第 i 列的像素位置记为 $p(j, i)$,其灰度值为 $f(j, i)$ 。对于插值图像的高频部分,选取 HRBndyImg 的左上角像素点进行遍历,如果 HRBndyImg 的 $f(j, i)$ 为 1,那么插值图像 interpImg 上对应的像素点就属于边缘部分。用 255 除该像素的灰度值,就可以将插值图像 interpImg 归一化到 $0 \sim 1$ 范围,然后用一个标签向量 training_label 保存该灰度值。插值图像 interpImg 中所有归类为图像边缘的像素的灰度值都存储在该标签向量中。同时,在 interpImg 中取出以 $p(j, i)$ 为中心的大小为 5×5 的图像块,再对图像块内的每个像素的灰度都除以 255 以归一化到 $0 \sim 1$,将其转换为行向量并保存到矩阵 training_instance 中。矩阵 training_instance 的每一行都保存的是 interpImg 中以归类为边缘点的像素为中心的 5×5 大小的图像块。

类似地,对于插值图像的低频部分,即图像 HRBndyImg 第 j 行、第 i 列的像素灰度为 0,插值图像 interpImg 上对应的当前像素点是低频背景的一部分,采用上述方法对其进行归一化处理并将该灰度记录到一个标签向量 $\text{training_label_low}$ 中。在 interpImg 中取出以 $p(j, i)$ 为中心的大小为 5×5 的图像块,归一化处理并转换为行向量后保存到矩阵 $\text{training_instance_low}$ 中。

需要注意的是,我们只处理灰度图像,如果输入的 LRImg 是 RGB 彩色图像,则先将该图像的颜色分量从 RGB 空间切换到 YIQ 空间,然后只提取 YIQ 空间图像的亮度值 Y 分量并对其进行处理,I 和 Q 分量不变,最后再将处理后的图像由 YIQ 空间切换到 RGB 空间。

3.3 图像块稀疏表示

K-SVD 等传统的字典构造算法都是迭代批处理的,每次

迭代时它们都会存取整个训练样本集,因此无法高效地应对大型的训练样本集或视频序列数据。针对该问题,我们应用 Julien Mairal 等人^[16]提出的在线字典学习算法,该算法每次只处理一个信号,这在图像和视频处理的应用中非常关键,其运行效率更高、更节省内存。对于图像块向量矩阵(即图像信号集合) training_instance 和 $\text{training_instance_low}$,通过这种在线字典学习算法即可得到高频字典 D_h 和低频字典 α_l 。然后通过高效的 LARS-Lasso 算法稀疏编码高频和低频的图像信号集合,得到的稀疏系数向量分别为 D_h 和 α_l 。最小角回归(Least Angle Regression, LARS)^[17]是一种解决 Lasso 或者 Elastic-net 问题的算法。以高频图像块集合 training_instance 为例,所提稀疏编码问题可以表示为:

$$\min \|\alpha_h\|_1 \quad \text{s. t.} \quad \|y_h - D_h \alpha_h\|_2 \leq \epsilon \quad (2)$$

其中, y_h 是高频图像块集合 training_instance 中的一个图像块(转化为列向量的形式), D_h 是由前文的在线字典学习算法以 y_h 为输入信号通过学习得到的字典, ϵ 是衡量重建图像块和输入图像块之间误差的阈值。将式(2)变换为无条件极值的形式:

$$\min \frac{1}{2} \|y_h - D_h \alpha_h\|_2^2 + \lambda \|\alpha_h\|_1 \quad (3)$$

其中, λ 用于平衡第一项重建误差和第二项 α_h 的稀疏度。我们使用上述的 LARS 算法来求解式(3)。

对低频图像块集合 $\text{training_instance_low}$ 也做类似的处理,得到的稀疏系数向量 α_h 和 α_l 将用于学习 SVR 模型。

3.4 SVR 学习和预测

我们运用支持向量回归(Support Vector Regression, SVR)理论并通过 V-SVR 模型找到 HRImg 图像块和 LRImg 图像块中心像素标签之间的映射函数。因为支持向量的样本数比输入的训练样本数少,所以该模型的复杂度较低。

在训练阶段,学习 SVR 模型的训练集(高、低频图像块)的属性以得到的稀疏系数 α_h 和 α_l 为准,而学习 SVR 模型的训练集的标签则是以将块分类后得到的低频标签向量 $\text{training_label_low}$ 和高频标签向量 training_label 为准。这些输入的特征和 HRImg 图像中对应的像素标签之间的映射函数就是被 SVR 模型所学习的,而且还会得到一个模型文件。具体步骤是:1)对每个块分类后得到的高、低频标签向量建立回归模型;2)用高频标签向量和稀疏矩阵 α_h 训练学习得到 SVR 高频模型文件,用低频标签向量和稀疏矩阵 α_l 训练学习得到 SVR 低频模型文件。

在 SR 重建测试阶段也会对测试图像稀疏编码得到稀疏系数 α_h' 和 α_l' ,然后根据上述训练阶段所得到的模型文件来预测测试图像的像素标签。最后进行稀疏编码,得到稀疏系数向量 α_h' 和 α_l' 。具体步骤是:1)将 HRImg 图像高、低频分块带入相应位置的回归模型,用高、低频模型文件以及稀疏矩阵 α_h' 和 α_l' 预测得到相应的新标签向量;2)将预测后的标签向量变换到 $0 \sim 255$ 范围,将其作为重建 SR 图像的像素值。

3.5 保持边缘的后处理

为了更好地拟合数据,SUR 模型更倾向于应用光滑的函数,但会造成函数在快速变化的地方变得较为光滑。但是对于一幅图像,同样的方法会在一定程度上模糊图像的边缘,与此同时还可能会产生振铃效应。为了解决该问题,根据图像边缘不连续性的图像先验知识^[16],对重建后的图像进行后处理,去除振铃效应,使图像边缘更加锋利。自然图像先验和重

建约束两部分组成先验知识,其表达式如下:

$$\Pr(\{x\}|\{y\}) = \frac{1}{C} \prod_{(i,j)} \underbrace{\exp\left(-\frac{1}{2} \left(\frac{|D_i * \tilde{x}_j|}{\sigma_N}\right)^2\right)}_{\text{Natural image prior}} \cdot \prod_{i=1}^{MN} \underbrace{\exp\left(-\frac{1}{2} \left(\frac{|W * \tilde{x}_i| - y_i}{\sigma_R}\right)^2\right)}_{\text{Reconstruction constraint}} \quad (4)$$

其中, $\{y\}$ 表示 LR 图像, $\{x\}$ 表示 HR 图像, 自然图像先验约束是等号右边的第一部分, 重建约束是等号右边的第二部分。条件概率如何达到最大值是该算法的目的。将式(4)转化为一个因子图的表示形式, 采用 max-product 置信度传播算法计算随机变量的最大后验概率 MAP 的估计, 具体的算法原理和流程请参见文献[16]。实验表明该图像先验知识的后处理过程能减弱振铃效应, 改善被模糊的图像边缘。

3.6 算法的整体描述

本文 SR 算法的主要流程分为 3 个部分: 图像块分类、图像块稀疏表示、SVR 模型学习和预测。算法的具体步骤如下:

Step 1 输入训练的低分辨率图像 LRImg, 采用 bicubic 插值对其放大 2 倍后记为 interpImg。

Step 2 用 LoG 算子对 interpImg 进行边缘检测, 将得到的边缘二值图像记为 HRBndyImg。

Step 3 对 interpImg 图像块进行分类并保存标签向量 training_label 和 training_label_low 以及图像块矩阵 training_instance 和 training_instance_low。

Step 4 采用在线字典学习算法对 training_instance 进行学习得到高频字典 D_h , 采用 training_instance 和 D_h 进行稀疏编码得到 α_h 。同理对 training_instance_low 进行学习得到低频字典 D_l , 采用 training_instance_low 和 D_l 进行稀疏编码得到 α_l 。

Step 5 采用标签向量 training_label 和稀疏矩阵 α_h 进行训练得到 SVR 模型文件 model。同理对标签向量 training_label_low 和稀疏矩阵 α_l 进行训练得到模型文件 modelLow。

Step 6 采用模型文件 model 和稀疏矩阵 α_h 进行预测得到标签向量 temp_predicted_label。同理, 用模型文件 modelLow 和稀疏矩阵 α_l 进行预测得到标签向量 temp_predicted_label_low。

Step 7 将预测后的标签向量变换到 0~255 范围, 生成 HR 图像。

Step 8 对 HR 图像进行后处理并输出。

4 实验结果及讨论

实验选择的图像数据库是 USC-SIPI, 从中选取 person512, susana256, tree256 和 House256 等若干图像作为实验对象。采用峰值信噪比 (PSNR) 来衡量重建的图像与原 HR 图像的误差。重建实验的开发环境为 MATLABR2010b, 使用的处理器为 Intel Core i3, 2.53GHz, 内存为 2GB, 操作系统是 32 位 Win7 系统。学习和预测 SVR 模型采用台湾大学林智仁开发的 LIBSVM 软件包¹⁾, 学习过完备字典和稀疏编码则采用 Julien Mairal 等人^[15] 开发的稀疏建模工具箱 SPAMS, 在 Bicubic 插值过程中用到的函数是 Matlab 的函数 imresize。比较本文方法和 Bicubic 插值、文献[2]方法的 PSNR 值, 如表 1 所列(实验中对 LR 图像放大 2 倍)。

表 1 不同 SR 重建方法 PSNR 值的对比

重建方法	Person512	House256	susana256	tree256
Bicubic 插值	27.4538	32.1757	33.7426	26.8716
文献[2]方法	25.4414	26.0577	27.0632	23.0975
本文方法	27.8113	32.2243	34.7123	27.0591

由表 1 可知, 本文方法的 PSNR 值与 Yang 等人^[2] 的方法相比提高很多, PSNR 值平均提高 5.04dB。Yang 等人在公开源代码中默认放大 3 倍, 我们则考察放大 2 倍的效果, 因此根据其说明文档改变放大倍数后需要重新训练字典。对此也做了实验, 但是实验训练耗时长, 其主要原因在于由花草树木组成的文献[2]的外部库与本文实验图像在特征上有明显不同, 因而得到了存在误匹配的字典和存在一些明显的虚假细节的图像, 而不是真实的丢失的细节, 这就是文献[2]的 PSNR 值反而比 Bicubic 插值低的原因。本文方法比 Bicubic 插值方法的 PSNR 值更高, 平均提高了 0.39dB。图 4 给出了 Person512 图像在几种方法下的实验结果对比图, 其中图 4(a) 为 HR 原图像, 图 4(b) 为 bicubic 插值得到的图像, 图 4(c) 为文献[2]得到图像, 图 4(d) 为本文方法得到图像。

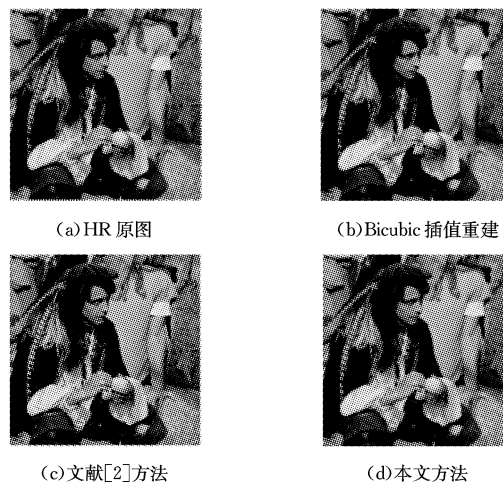


图 4 Person512 图像实验结果的对比图

图 5 给出了图 4 中各幅图像的局部放大图。

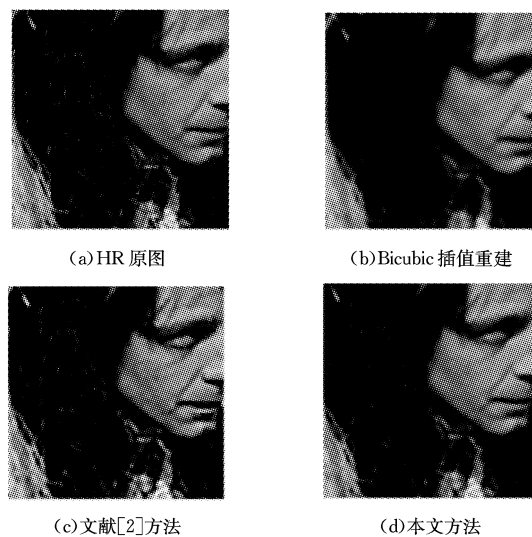


图 5 Person512 图像实验结果的放大图

图 6 给出了 House256 图像在几种方法下的实验结果对比图。

¹⁾ <http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html>

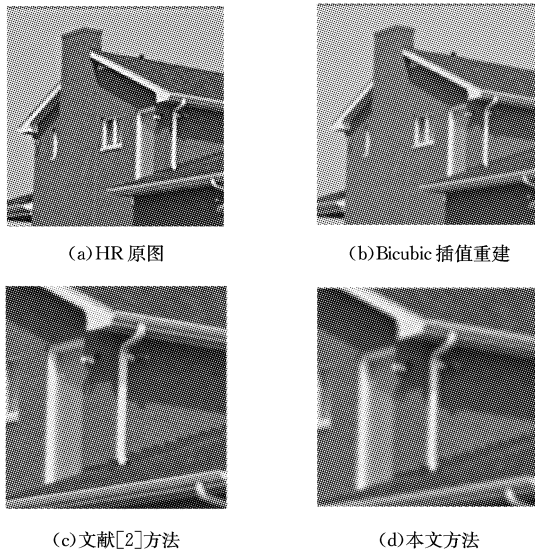


图 6 House256 图像实验结果的对比图

图 7 给出了图 6 中各幅图像局部放大图。

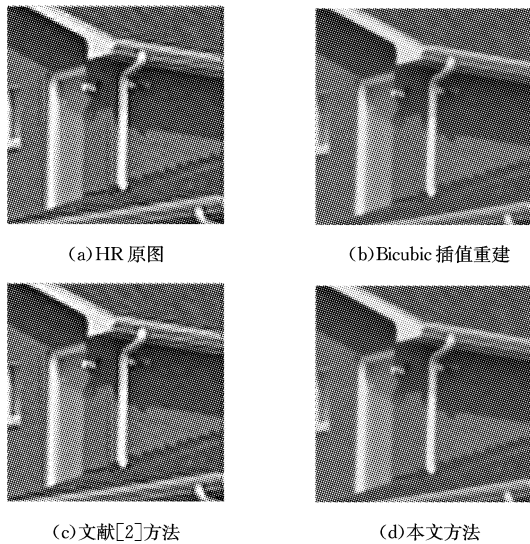


图 7 House256 图像实验结果放大图

图 8 给出了 susana256 图像在几种方法下的实验结果对比图。

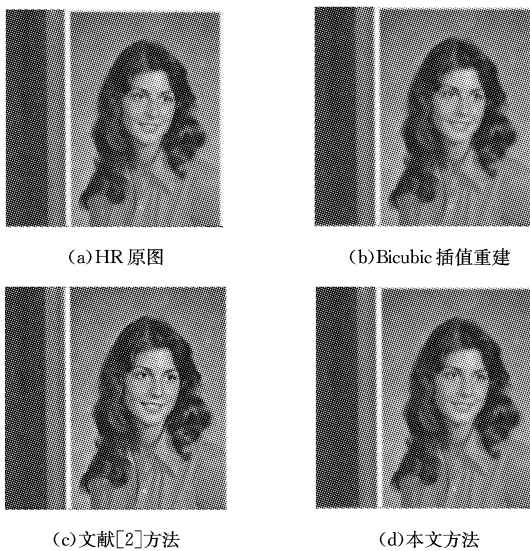


图 8 Susana256 图像实验结果的对比图

图 9 给出了图 8 中各幅图像的局部放大图。

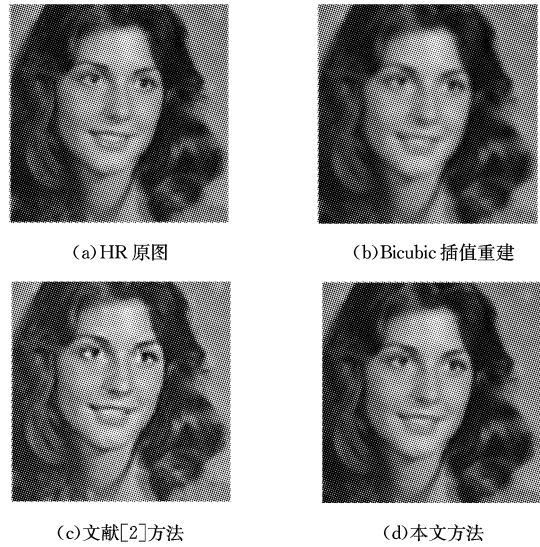


图 9 Susana256 图像实验结果的局部放大图

图 10 给出了 tree256 图像在几种方法下的实验结果对比图。

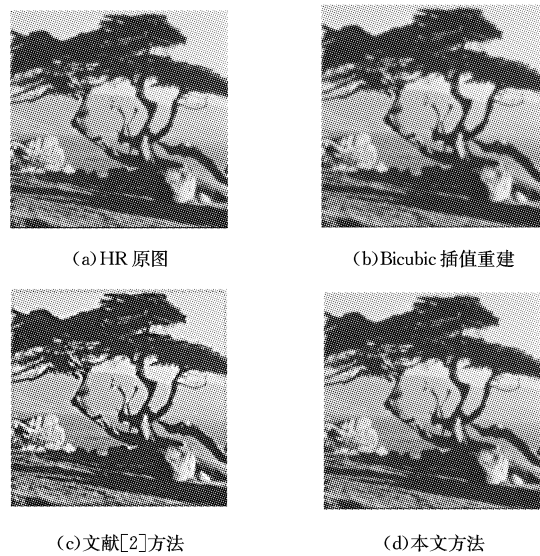


图 10 Tree256 图像实验结果的对比图

图 11 给出了图 10 中各幅图像的局部放大图。

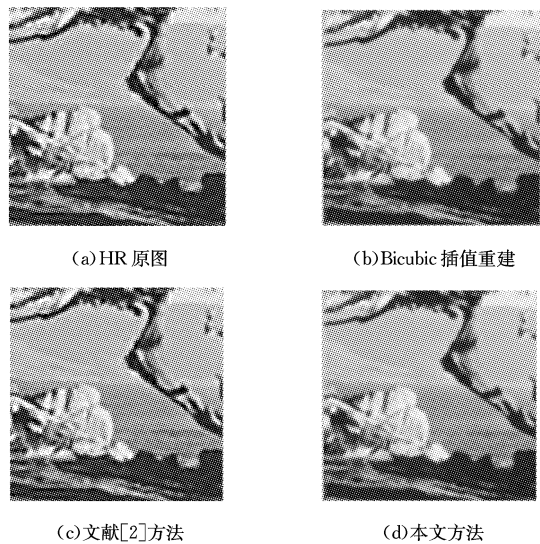


图 11 Tree256 图像实验结果的局部的放大图

从上述4组实验图像中可以明显看到,在视觉效果上,图11(b)比图11(d)的图像细节保持得更清晰,边缘更锋利。Bicubic会产生与Bilinear相似的平均化的滤波作用,会固有地模糊图像。利用图像边缘不连续性的图像先验知识,对重建后的图像进行后处理,去除振铃效应,边缘保持得更加清晰。图4(c)的脸颊处,图6(c)的白色斜屋檐、水管、窗框处,图8(c)的衣领处和图11(c)的树干周围相比于相应图中的图(d)都出现了比较明显的模糊。上述图在视觉效果上验证了本文的后处理方法产生了效果。

各组图中的图(d)相比于图(c)重建的高频信息更加真实,因为文献[2]的训练字典采用的是外部图像库,所以重建的图像会产生错误的高频细节问题,而不是真实的图像细节。图4(c)和图5(c)的鼻子、嘴巴和眉毛处、图6(c)和图7(c)的白色斜屋檐、水管、窗框处,图8(c)的眉毛、眼睛处和图10(c)的树干周围都存在明显的虚假细节。对图像进行降质,分类图像块,区别属于图像不同区域的像素进行训练,以学习更准确的SVR模型。此外,若预测出的细节和HR原图像越接近,则最后得到的重建图像的PSNR值比文献[2]方法的就更好。

图12给出了本文算法的SR重建结果,图12(a)为3幅 128×128 的低分辨率图像,图12(b)为对应的 256×256 的高分辨率SR重建图像。从视觉效果看,本文算法可达到较好的重建效果。



(a)低分辨率原图 (b)SR重建结果

图12 本文算法的SR重建

结束语 采用外部图像库训练字典重建后的图像很容易产生错误的高频细节问题,而字典的训练耗时较长,为了弥补

这两个缺陷,本文提出一种通过待处理图像自身来训练字典并进一步训练SVR模型的自学习方法。针对Mean Shift算法在分割图像时的分割结果具有不稳定性,可能会使之后通过训练得到的SVR模型产生不准确的问题,采用LoG算子来提取图像的边缘。学习字典采用效率较高的在线字典学习算法,稀疏编码则采用LARS稀疏分解算法,并采用SVR模型学习预测HR图像,根据图像边缘的先验知识对重建后的图像进行后处理,去除振铃效应,使得图像边缘更加清晰,视觉效果更好。

参考文献

- [1] LIN Z C, SHUM H Y. Fundamental limits of reconstruction-based super-resolution algorithms under local translation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 83-97.
- [2] YANG J C, WRIGHT J, HUANG T. Image super-resolution as sparse representation of raw image patches[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition, Alaska; IEEE Computer Press, 2008: 1-8.
- [3] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[J]. International Journal of Computer Vision, 2000, 40(1): 25-47.
- [4] DAI S, HAN M, XU W, et al. Soft edge smoothness prior for alpha channel super resolution[C]//Proceedings of IEEE International Conference on Computer Vision, Washington, DC; IEEE Computer Press, 2007: 5-8.
- [5] DURARLE M F, DAVENPORT M A. Single-Pixel Imaging via Compressive Sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 83-91.
- [6] YANG M C, WANG Y C F. A Self-Learning Approach to Single Image Super-Resolution[J]. IEEE Transactions on Multimedia, 2013, 15(3): 498-508.
- [7] ZHANG K B, GAO X B, TAO D C. Multi-scale dictionary for single image super-resolution[C]//IEEE Conference on Computer Vision and Pattern Recognition, Alaska, IEEE Computer Press, 2012: 1114-1121.
- [8] YU G S, GUILLERMO S, STÉPHANE M. Solving Inverse Problems With Piecewise Linear Estimators: From Gaussian Mixture Models to Structured Sparsity[J]. IEEE Transactions on Image Processing, 2012, 21(5): 2418-2499.
- [9] PELEG T, ELAD M. A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution[J]. IEEE Transactions on Image Processing, 2014, 23(6): 2569-2582.
- [10] LI D L, MERSEREAU R M, SIMSKE S J. Blind Image Deconvolution Using Support Vector Regression[C]//Proceedings IEEE International Conference on Acoustics, Speech and Signal Process, Washington, DC; IEEE Computer Press, 2005: 113-116.
- [11] LI D L, RUSSELL S S. Single Image super-resolution based on support vector regression[C]//Proceedings of International Joint Conference on Neural Networks, Washington, DC; IEEE Computer Press, 2006: 2898-2901.

画,但是对于重叠程度较小的多笔画,本文算法能够更好地识别笔画的连接性,完成笔画的聚类、拟合;文献[14]则将其识别为多条笔画,如图9中的长直线以及两段圆弧。本文算法更注重多笔画之间的联系,判断其从属于单一几何图元的可能性,然后完成其聚类、拟合。

结束语 针对手绘草图中的间断绘制多笔画问题,提出了一种基于笔画容差带的手绘草图间断多笔画判定与拟合方法,在解决问题的同时提供了一种解决问题的思路,即将多笔画间断绘制问题转化成了重复绘制问题,扩展了已有的技术手段的应用范围。同时,利用已有笔画的走势推测线元的预期走势,并构建假想笔画,用于间断绘制多笔画的判定;将间断绘制笔画的位置关系作为判定的标准,并对判定的结果进行拟合,得到标准的几何线元,为后续的草图识别工作奠定了基础。该方法具有以下特点:

1)利用笔画容差带将间断绘制多笔画问题转化为重复绘制问题,扩大了笔画容差带的应用范围,简化了间断绘制多笔画的判定过程。

2)由于间断绘制多笔画间的重叠程度较低,现有的空间关系不足以判定多笔画之间的联系,提出了利用假想笔画构建间断绘制多笔画之间联系的方法,增强了多笔画间原有的联系,并且符合设计人员观察手绘草图时对间断绘制多笔画的判断过程。

3)针对二次曲线的假想笔画构建问题,综合利用多笔画的信息,避免了因单一笔画信息不足造成的多解以及求解过程复杂的问题;同时假想笔画能够较好地反映预期笔画的走势,满足多笔画判定的要求。

该方法将间断绘制多笔画问题转化为重复绘制问题,取得了一定的效果。但是该方法应用于重复绘制多笔画问题时,实际上并不需要绘制假想笔画。如何根据多笔画的重叠程度区分间断绘制多笔画和重复绘制多笔画,并根据判断的结果选择相应的笔画判定拟合方法,是今后研究的方向。

参 考 文 献

- [1] 孙正兴,冯桂焕,周若鸿. 基于草图的人机交互技术研究进展[J]. 计算机辅助设计与图形学学报, 2005, 17(9): 1889-1899.
- [2] 宋保华,叶军,于明玖,等. 笔输入草图的分层识别[J]. 计算机辅助设计与图形学学报, 2004, 16(6): 753-758.
- [3] 谢强,冯桂焕,孙正兴. 基于上下文的在线草图识别[J]. 计算机科学, 2007, 34(3): 216-219.
- [4] 李路,周良,丁秋林. 基于贝叶斯网络的草图符号识别研究[J]. 计算机科学, 2011, 38(6): 262-265.
- [5] 马嵩华,田凌. 捕捉设计意图的二维草图识别技术[J]. 计算机辅助设计与图形学学报, 2012, 24(10): 1337-1345.
- [6] PETERSON E J, STAHOVICH T F, DOI E, et al. Grouping Strokes into Shapes in Hand-Drawn Diagrams[C]//Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10). Association for the Advancement of Artificial Intelligence, 2010: 974-979.
- [7] KU D C, QIN S F, WRIGHT D K. Interpretation of Overtracing Freehand Sketching for Geometric Shapes[C]//Proceedings of the 14th International Conference in Central Europe on Computer Graphics. Aire-la-Ville: Eurographics Association Press, 2006: 263-270.
- [8] 王淑侠,高满屯,齐乐华. 基于二次曲线的在线手绘图识别[J]. 西北工业大学学报, 2007, 25(1): 37-41.
- [9] WANG S X, GAO M T, QI L H. Online Freehand Sketching Recognition Using Fuzzy Theory[J]. Pattern Recognition and Artificial Intelligence, 2008, 21(3): 317-325.
- [10] 王淑侠,王关峰,高满屯,等. 基于时空关系的在线多笔画手绘折线段曲线识别[J]. 计算机科学, 2012, 39(9): 269-274.
- [11] WANG S X, WANG G F, GAO M T, et al. Recognition of online multi-stroke freehand conic sections based on time-space relationship[J]. Pattern Recognition and Artificial Intelligence, 2011, 24(1): 82-89.
- [12] WANG S X, QIN S F, GAO M T. New grouping and fitting methods for interactive overtraced sketches[J]. The Visual Computer, 2014, 30(3): 285-297.
- [13] 王淑侠,王守霞,王关峰,等. 基于几何特征的在线手绘草图分割[J]. 计算机辅助设计与图形学学报, 2015, 27(9): 1686-1693.
- [14] 王淑侠,王守霞,王关峰,等. 在线多笔画重复绘制草图的判定与聚类[J]. 计算机辅助设计与图形学学报, 2015, 27(10): 1980-1987.
- [12] YANG M C, CHU C T, WANG Y C F. Learning Sparse Image Representation With Support Vector Regression For Single Image Super Resolution[C]//IEEE International Conference on Image Processing. Washington, DC: IEEE Computer Press, 2010: 1973-1976.
- [13] GLASNER D, BAGON S, IRANI M. Super-Resolution From a Single Image[C]//Proceedings of 2009 IEEE 12th International Conference on Computer Vision. Kyoto: IEEE Computer Press, 2009: 349-356.
- [14] ZONTAK M, IRANI M. Internal Statistics of a Single Natural Image[C]//Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition. America: Colorado Springs, 2011: 977-984.
- [15] MAIRAL J, BACH F, PONCE J. Online Learning for Matrix Factorization and Sparse Coding[J]. Journal of Machine Learning Theory, 2010, 11(1): 19-60.
- [16] TAPPEN M F, RUSSELL B C, FREEMAN W T. Exploiting the Sparse Derivative Prior for Super-Resolution and Image Demosaicing[C]//proceedings of IEEE workshop on Statistical and Computational Theories of Vision. Fort Collins: IEEE Computer Press, 2003: 1-24.
- [17] EFRON B, HASTIE T, JOHNSTONE I. Least Angle Regression[J]. The Annals of Statistics, 2004, 32(2): 407-499.

(上接第 174 页)