

基于分级子空间回归的压缩人脸图像复原

刘心宇, 干宗良, 刘 峰

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘要:人脸图像具有自然图像不具备的对称特征和几何结构相似性。由于人脸图像往往具备非常复杂的角点和纹理特征,因此很难找到一种全局模型将压缩图像映射到原始未压缩图像。针对此问题,提出一种新颖的基于分级子空间回归的压缩人脸复原算法,该算法包括训练和复原两个部分。在训练部分,利用压缩人脸图像的边缘方向分布规律,将压缩-未压缩图像块对划分到多个浅层子空间中。然后对每个基于边缘方向分类的浅层子空间,利用 K-means 聚类算法得到它的深层子空间,并在每个深层子空间中训练得到相应的线性映射。在复原阶段,对每个输入的压缩图像块分析得到它的边缘方向,从而选择合适的线性映射,得到复原后的输出图像块。实验结果表明,该算法在 PSNR 和 SSIM 上均优于现有的常用复原算法,并且能够有效地去除压缩失真和锯齿效应,提高视觉效果。

关键词:压缩人脸图像;子空间回归;边缘方向;浅层子空间;深层子空间

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2019)06-0159-05

doi:10.3969/j.issn.1673-629X.2019.06.033

Compressed Face Image Restoration Based on Hierarchical Subspace Regression

LIU Xin-yu, GAN Zong-liang, LIU Feng

(School of Telecommunications & Information Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract: Face images have the character of symmetry and similar geometry structures that are not available in natural images. Since face images often have very complex corner points and texture feature, it is difficult to find a global model to map compressed images to raw uncompressed images. Aiming at this problem, we propose a novel compressed face restoration algorithm based on hierarchical subspace regression, which includes two parts: training and restoration. In the training, the rule of the face edge-orientation distribution is used to classify the compressed-uncompressed patch pairs into shallow subspaces. Then, the K-means clustering is used to cluster the deep subspaces of each shallow subspace, and corresponding linear mapping training is performed for each deep subspace. In the restoration, an appropriate linear mapping selected based on the edge orientation of compressed input image patch is applied to generate the restored output image patch. The experiment shows that the PSNR and SSIM are better than the existing popular algorithm, and it can effectively remove the blocking artifact and zigzag effect, so as to improve the visual effect.

Key words: compressed face images; subspace regression; edge-orientation; shallow subspace; deep subspace

0 引言

为了进一步节约带宽和设备的存储空间,在视觉通信和计算机系统中通常采用有损的图像压缩方式,例如 JPEG^[1] 和 HEVC-MSP^[2]。不可避免地,有损压缩将引起块效应、振铃效应和模糊等压缩失真。它们通常是由处理图像块时的不连续性,高频分量的粗量化等原因导致的。这些失真不仅会降低视觉的感知质

量,而且对很多将压缩图像作为输入的图像处理算法造成不利影响^[3]。

早期的算法^[4-5]使用滤波的方法来去除简单的失真。随后,基于数据的方法被提出来以避免不准确的压缩降质模型。基于稀疏的图像复原方法^[6-11]可以产生锐化的图像,但复原后的图像往往存在模糊的边缘和不自然的平滑区域。文献[12]中,杨航等在傅里叶

收稿日期:2018-07-31

修回日期:2018-11-14

网络出版时间:2019-03-06

基金项目:国家自然科学基金(61471201)

作者简介:刘心宇(1994-),男,硕士研究生,研究方向为图像处理与图像通信;干宗良,博士,副教授,研究方向为分布式视频编码和图像视频信号处理等;刘 峰,博士,教授,研究方向为图像处理与网络视频通信等。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190306.1017.086.html>

域得到图像的初步估计,再利用结构字典学习的方法得到复原图像。最近,Dong 等^[3]通过对 SRCNN 模型的特殊应用^[13],将深度学习方法^[14]引入到这个问题。虽然这些算法在 PSNR 上取得了很好的效果,但并没有包含太多特定问题的先验知识^[15]。较深的网络以及较小的步长都会导致计算量大大增加^[16],这导致深度学习算法在低配置的硬件上往往难以实现。

A+^[17]通过改进 ANR^[18]的线性映射学习来进一步提高性能。当完成字典训练后,A+以相关性为准则,在 500 万个候选图像特征向量里挑选 2 048 个相关性最高的特征向量来训练驻点回归模型,极大地改善了图像质量。

文中提出一种基于分级子空间回归的压缩图像复原方法,应用于 Webp 压缩人脸图像。在训练部分,首先根据压缩人脸图像的边缘方向,将压缩-未压缩图像块对划分到不同的浅层子空间中,其次对每个浅层子空间使用 K 均值聚类进一步划分其深层子空间。最后,通过将 A+回归模型应用到每个深层子空间中,可以获得一组彼此独立的子空间回归映射矩阵。在复原阶段,分析每个压缩图像块的边缘方向,判断它所属的浅层子空间,从而选择合适的深层子空间线性映射来复原压缩图像块。

通过提取压缩人脸图像最明显的特征分布约束信息,该算法克服了特征选择,多尺度特征域以及光照强度的影响,具有良好的鲁棒性。

1 HOG 特征提取

1.1 计算梯度大小和方向

分别计算压缩人脸图像像素点 (x, y) 的梯度幅值和梯度方向:

$$G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \quad (1)$$

$$\alpha(x, y) = \arctan \frac{G_x(x, y)}{G_y(x, y)} \quad (2)$$

其中, $G_x(x, y)$ 和 $G_y(x, y)$ 分别是像素点 (x, y) 水平和垂直方向上的梯度; $G(x, y)$ 表示该像素点的梯度幅值; $\alpha(x, y)$ 表示梯度方向。

1.2 生成 HOG 特征描述

将梯度方向划分为 M 个方向,即将角度划分为 M 个 bin。通过计算每个图像块内像素点分布在不同 bin 区间的梯度幅值之和,生成 M 维特征向量。对 M 维的特征向量标准化,就得到该图像块的 HOG 特征描述。

2 基于边缘方向的分级子空间回归

2.1 边缘方向分析

众所周知,同一个类别内信号的差异性比一般的差异性要小。Yang 等^[19]将输入的降质空间划分成足够精确的子空间,并在不同的子空间中学习回归。正如文献[20-21]论述的那样,如果将图像恰当地划分成更小图像块,并根据特征将每个图像块适当地划分成多个子空间,则每个图像子空间都可以学习到压缩图像块到未压缩图像块的有效线性映射。

通过调研压缩人脸图像时发现,位于边缘的像素点梯度分布遵循一定的规律,而且和在压缩自然图像中明显不同。测试从 VGGface2 数据集^[22]提取的人脸图像和通用的自然图像数据集 set115 来验证是否存在这一规律,结果如图 1 所示。在图 1 中发现,位于边缘的像素点在压缩人脸图像中的分布要比自然图像更加对称,而且在 45 度到 135 度的范围内,分布的数量明显更多。

这一特征表明,可以通过边缘方向来区分压缩的人脸图像和自然图像,并进一步利用这一特征来复原压缩人脸图像。

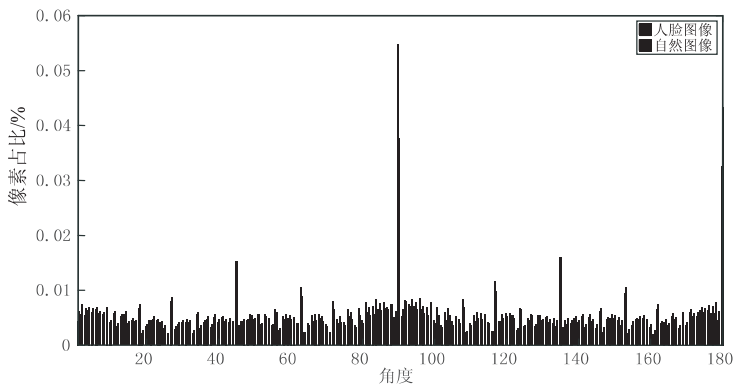


图 1 人脸图像和自然图像边缘点梯度方向对比

因为矢量方向和它的相反方向可以被认为是一个方向,所以在 1.2 节,采用 180 度的角度区间。同时为了避免式 2 在计算角度时分母为零的情况,将水平方向设置成一个非常小的值,这样导致 90 度附近的像

素点数量会非常多,直接影响到复原结果。为了更准确有效的计算,把 90 度作为角度计算的起始点,即把第一象限映射到第三象限来获得闭区间的角度分布,同时基于像素点数目的分布将 180 度量化成 M 个区

间。这样,对每个位于边缘的图像块,就以它 HOG 特征描述中的最大值为这个图像块的边缘方向,并赋予相应的索引。

2.2 基于边缘方向的浅层子空间聚类

当压缩图像具有较大的噪声时,将相似的压缩图像块聚类到不同的浅层子空间中,这并不能保证它们对应的原始图像块也是相似的。为了进一步提高子空间聚类的准确性和增强算法的抗噪能力,获取一个与目标块 y_m 相同位置中心但具有更大尺寸的图像块 y_n ,以 y_n 的边缘方向作目标块 y_m 浅层子空间划分的标准。

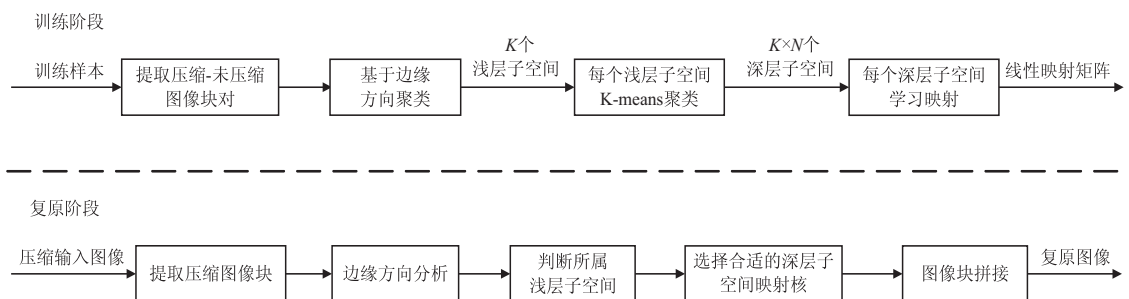


图2 算法总框架

2.3 分级子空间回归

文中压缩图像块由它的特征表示,这些特征是对压缩图像在水平方向和垂直方向的一阶梯度和二阶梯度特征的级联表示。此外,在训练阶段,其目标在于找到一个广义的线性映射 P_c ,它能将所属于浅层子空间 C 的压缩图像块映射到原始的未压缩版本,如

$$X_c = P_c Y_c \quad (3)$$

其中, X_c 表示处于同一浅层子空间 C 的原始图像块 $x_{i,c}$ 的矢量化集合; Y_c 表示对应的压缩块 $y_{i,c}$ 的矢量化集合。

在每个浅层子空间中使用 K 均值聚类算法获得每个浅层子空间下的深层子空间,同时得到的 K 均值聚类中心就作为深层子空间中心的代表。在这个过程中,应用主成分分析 (PCA) 来减少计算量。虽然同一浅层子空间中的图像块都具备相近的边缘方向,但仍旧存在偏差。这种偏差也反映在每个深层子空间的中心和其他深层子空间中的图像块差异很大,所以当完成深层子空间的划分时,每个深层子空间的中心就作为学习回归时的线性映射核。

对于浅层子空间 C 中的每个深子空间,都可以看做是 l_2 范数正则化的最小二乘回归^[18]。

$$P_c = \operatorname{argmin}_p \|X_c - PY_c\|_2^2 + \lambda \|P\|_2 \quad (4)$$

可以得到它的最优估计为:

$$P_c = X_c Y_c^T (Y_c Y_c^T + \lambda I)^{-1} \quad (5)$$

其中, I 是单位矩阵; λ 是正则化参数。

根据 A^+ ^[17] 可知,驻点周围的流形和驻点的相关

为了进一步加快算法的运行速度并判断目标块 y_m 是否位于图像边缘上,采用 Canny 算子来提取压缩图像的边缘点,如果 y_n 包含任何边缘点,则计算 y_n 的边缘方向并分配以 y_m 相应的子空间索引。如果未包含任何边缘点,对该子空间分配索引 0。

通过这种方式,利用边缘方向信息将图像块聚类到适当的浅层子空间中,每个子空间都由一组具备相似边缘方向的压缩图像块和未压缩图像块对组成。对每个基于边缘方向聚类的浅层子空间,都可以学习到一个压缩图像块到未压缩图像块的有效线性映射矩阵,方便在复原阶段使用。图2为文中算法总框架。

性很大,准确度要优于整个样本集。在这里用深层子空间中心周围的流形来代表整个深层子空间。

$$P_{d_c} = S_{x,d_c} S_{y,d_c}^T (S_{y,d_c} S_{y,d_c}^T + \lambda I)^{-1} \quad (6)$$

其中, P_{d_c} 是对应每个深层子空间中心 d_c 的映射矩阵; S_{y,d_c} 是同一浅层子空间中与 d_c 相关性最高的 N 个压缩图像块; S_{x,d_c} 为 S_{y,d_c} 对应的原始图像块。

使用同一浅层子空间下的压缩-未压缩图像块对,可以通过式3~6学习到相应浅层子空间的广义线性映射模型 P_c 。

在复原阶段,为了复原压缩人脸图像,只需要预先训练完所有深层子空间下的线性映射。

一旦输入压缩人脸图像,首先将其分成一组相同位置中心不同尺度大小的图像块,以分析确定目标块的边缘方向,从而得到对应的浅层子空间索引。其次,通过计算相关性,对每个目标块 $y_{j,c}$ 搜索同一浅层子空间中距离它最近的线性映射核 d_c 。然后,通过将锚定到 d_c 的线性映射 P_{d_c} 和 $y_{j,c}$ 相乘得到复原后的图像块 $x_{j,c}$,最后拼接得到完整的复原后的图像。

$$x_{j,c} = P_{d_c} y_{j,c} \quad (7)$$

3 实验结果与分析

为了公平地比较,使用 VGGFace2 数据集中不重复的4000幅图像作为训练集,使用从网络获得的50幅中国明星人脸图像和从 VGGface2 另外获得的50幅外国人脸图像作为测试的100幅图像。由于人类视觉系统对亮度分量的敏感度高于色度分量,因此将

RGB 图像转换为 YCbCr 图像,并且仅在 Y 通道上进行复原。使用图像质量因子 $Q = 1$ 对训练图像执行 WebP 编解码,以生成相应的压缩图像。然后将每个压缩图像划分为 7×7 重叠的图像块 y_m ,并由具有相同中心位置,大小为 13×13 的图像块 y_n 的边缘方向来进行子空间划分,根据经验将边缘方向量化为 16 个方向。

将文中算法与去噪算法 BM3D^[23] 和稀疏算法 A+ 进行比较。为了测试算法性能,测试集的平均 PSNR

和 SSIM 如表 1 所示。可以看到,与对比算法相比,文中算法具有更高的平均 PSNR 和 SSIM。由于空间限制,只列出了其中两幅具有代表性的图像,如图 3 和图 4 所示。

表 1 三种复原算法平均 PSNR,SSIM 比较

指标	压缩图像	BM3D	A+	文中算法
PSNR	32.25	32.17	32.97	33.14
SSIM	0.882 9	0.880 6	0.893 5	0.895 9



图 3 三种复原算法结果对比(VGG14)



图 4 三种复原算法结果对比(CN10)

可以看出,BM3D 可以降低图 3(c) 中的模糊程度,但它也忽略了图 4(c) 中的部分高频信息,这是因为 BM3D 使用一个统一的 sigma 参数来处理图像。从图 3(d) 可以看出,A+ 在边缘存在锯齿效应,并且在图 4(d) 中存在一些模糊,主要原因是 A+ 学习的字典过于紧凑,无法从一些不常见的训练图像块中获取细节。而文中算法产生了更清晰的边缘,并且产生了更清晰的纹理,如图 3(e) 和图 4(e)。

4 结束语

提出了一种基于分级子空间回归的压缩图像复原算法。利用人脸图像的先验知识和结构特征,将人脸图像块分类到浅层子空间中,再通过 K 均值聚类算法分类到合适的深层子空间中,确保邻域选择的快速准确,从而获得更有效的线性映射矩阵。实验结果表明,相对于传统复原方法,该算法提高了图像的峰值信噪比,锐化了图像的纹理和边缘,减少了模糊和锯齿效应,改善了图像视觉效果。

参考文献:

[1] 刘方敏,吴永辉,俞建新. JPEG2000 图像压缩过程及原理

概述[J]. 计算机辅助设计与图形学学报,2002,14(10): 905-911.

[2] SULLIVAN G J,OHM J,HAN W J,et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2012,22(12):1649-1668.

[3] DONG Chao,DENG Yubin,CHEN C L,et al. Compression artifacts reduction by a deep convolutional network[C]// IEEE international conference on computer vision. Santiago, Chile:IEEE,2016:576-584.

[4] 吴雪垠,吴 谨,张 鹤. 逆滤波法在图像复原中的应用[J]. 信息技术,2011(10):183-185.

[5] 张德丰,张葡青. 维纳滤波图像恢复的理论分析与实现[J]. 中山大学学报:自然科学版,2006,45(6):44-47.

[6] CHANG Huibin,NG M K,ZENG Tiejong. Reducing artifacts in JPEG decompression via a learned dictionary[J]. IEEE Transactions on Signal Processing,2014,62(3):718-728.

[7] JUNG C,JIAO L,QI H,et al. Image deblocking via sparse representation[J]. Signal Processing:Image Communication, 2012,27(6):663-677.

[8] ROTHE R,TIMOFTE R,VAN L. Efficient regression priors for reducing image compression artifacts[C]//IEEE interna-

tional conference on image processing. Quebec City, Canada; IEEE,2015:1543–1547.

[9] 陈凤娥. 基于稀疏表示的图像复原算法研究[D]. 武汉:武汉大学,2014.

[10] 杨亚威,胡双演,张士杰,等. 基于字典对联合学习的退化图像复原方法[J]. 计算机辅助设计与图形学学报,2015,27(3):406–413.

[11] 徐焕宇,孙权森,李大禹,等. 基于投影的稀疏表示与非局部正则化图像复原方法[J]. 电子学报,2014,42(7):1299–1304.

[12] 杨航,吴笑天,王宇庆. 基于结构字典学习的图像复原方法[J]. 中国光学,2017,10(2):207–218.

[13] DONG C, CHEN C L, HE K, et al. Learning a deep convolutional network for image super-resolution[M]. [s. l.]: Springer International Publishing, 2014:184–199.

[14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//International conference on neural information processing systems. Lake Tahoe, Nevada:IEEE,2012:1097–1105.

[15] ZHANG Kai, ZUO Wangmeng, GU Shuhang, et al. Learning deep CNN denoiser prior for image restoration[C]//IEEE conference on computer vision and pattern recognition. Honolulu, USA:IEEE,2017:2808–2817.

[16] HE Kaiming, SUN Jian. Convolutional neural networks at constrained time cost[C]//IEEE conference on computer vision and pattern recognition. Boston, MA, USA:IEEE,2015:5353–5360.

[17] TIMOFTE R, SMET V D, GOOL L V. A+: adjusted anchored neighborhood regression for fast super-resolution[M]. [s. l.]: Springer International Publishing, 2014:111–126.

[18] TIMOFTE R, DE V, GOOL L V. Anchored neighborhood regression for fast example-based super-resolution[C]//IEEE international conference on computer vision. Sydney, NSW, Australia:IEEE,2013:1920–1927.

[19] YANG C Y, YANG M H. Fast direct super-resolution by simple functions[C]//IEEE international conference on computer vision. Sydney, NSW, Australia:IEEE,2014:561–568.

[20] 薛纪令,陈华华. 基于自适应协稀疏正则化的图像复原[J]. 杭州电子科技大学学报,2015,35(2):25–28.

[21] 赵志辉,赵瑞珍,岑翼刚,等. 基于稀疏表示与线性回归的图像快速超分辨率重建[J]. 智能系统学报,2017,12(1):8–14.

[22] CAO Qiong, SHEN Li, XIE Weidi, et al. VGGFace2: a dataset for recognising faces across pose and age[C]//IEEE international conference on automatic face & gesture recognition. Xi'an, China:IEEE,2018:67–74.

[23] DABOV K, FOI A, KATKOVNIK V, et al. BM3D image denoising with shape-adaptive principal component analysis[C]//Proceedings of workshop on signal processing with adaptive sparse structured representation. Saint-malo, France:CCSD,2009.

《计算机技术与发展》投稿要求

1. 文稿具有创新性,内容科学可信,结构严谨,图表清晰,文稿必须使用 word 排版,应在 4 页以上(A4 纸,Word03 默认页面,5 号宋体,包括图表在内),文稿包括:题名(20 字以内),作者姓名,作者单位(包含地区名和邮编),摘要(须达到 300 汉字以上,包括:目的、方法、结果、结论,英文译文同步、准确),关键词(4~8 个),中图分类号,正文,参考文献(总数应在 15 篇以上,外文文献不少于 5 篇),与中文对照的英文(包括:题名、作者姓名、作者单位、摘要、关键词)。提供所有作者的简介:姓名、性别、出生年、学位(职称)、研究方向。
2. 文中的图表要符合国家规范,一般不应超过 6 个,图形尽量用 visio 处理,版面要清晰、紧凑、美观,图中文字采用六号宋体或 Times New Roman,打印清晰,不用彩色,坐标图中要注计量单位、符号,表格一律采用三线表(不画竖线)。
3. 文中所有物理量和单位应符合国家标准或国际标准,外文字母、单位、符号的大小写、正斜体、上下角标及易混字母应标注清楚,标量用白斜体字母(可以加下标)表示,矢量、矩阵和张量用黑斜体字母(可以加下标)表示。数学公式必须用数学编辑器处理。
4. 文稿章节编号采用三级标题:一级如 1,2,3,⋯;二级如 1.1,1.2,1.3,⋯;三级如 1.1.1,1.1.2,1.1.3,⋯。
5. 参考文献尽量选用近几年计算机专业期刊论文,未公开发表的资料请勿引用。参考文献序号按文中引用顺序排列。