

图神经网络回归的人脸超分辨率重建*

呼延康, 樊鑫, 余乐天, 罗钟铤

(大连理工大学 软件学院, 辽宁 大连 116621)

通讯作者: 樊鑫, E-mail: xin.fan@dlut.edu.cn



摘要: 人脸超分辨率(super-resolution, 简称 SR)即将输入模糊的低分辨率(low-resolution, 简称 LR)人脸图像通过一系列算法处理得到较为清晰的高分辨率(high-resolution, 简称 HR)人脸图像的过程. 相比自然图像, 不同人脸图像的不同位置通常具有相似的结构. 针对人脸图像的局部结构一致性特点, 提出一种新的基于图结构的人脸超分辨率神经网络回归方法. 将输入低分辨率图像表示为图结构, 进而为图结构中每一个节点的局部表示训练一个浅层神经网络进行超分辨率回归. 与基于规则矩形网格的方法相比, 图结构在描述一个像素的局部信息时, 不仅考虑到图像坐标的相关性, 同时也关注了纹理的相似性, 能够更好地表达图像局部特征. 训练过程中, 利用已收敛的相邻节点的神经网络参数初始化当前节点的神经网络参数, 不仅加快了神经网络的收敛速度, 而且提高了预测精度. 与包括深度卷积神经网络在内的基于学习的超分辨率最新算法比较, 实验结果表明, 所提算法取得了更高的准确率. 提出的图神经网络(graph neural networks, 简称 GNN)并不局限于解决人脸超分辨率问题, 还可以用于处理其他具有不规则拓扑结构的数据, 解决不同的问题.

关键词: 人脸图像; 超分辨率; 图; 神经网络; 回归

中图法分类号: TP391

中文引用格式: 呼延康, 樊鑫, 余乐天, 罗钟铤. 图神经网络回归的人脸超分辨率重建. 软件学报, 2018, 29(4): 914-925. <http://www.jos.org.cn/1000-9825/5405.htm>

英文引用格式: HUYAN K, FAN X, YU LT, LUO ZX. Graph based neural network regression strategy for facial image super-resolution. Ruan Jian Xue Bao/Journal of Software, 2018, 29(4): 914-925 (in Chinese). <http://www.jos.org.cn/1000-9825/5405.htm>

Graph Based Neural Network Regression Strategy for Facial Image Super-Resolution

HUYAN Kang, FAN Xin, YU Le-Tian, LUO Zhong-Xuan

(School of Software, Dalian University of Technology, Dalian 116621, China)

Abstract: Facial image super-resolution (SR) generates a high-resolution (HR) facial image from a low-resolution (LR) facial image. Compared with natural images, facial images are so highly structured that local patches at similar locations across different faces share similar textures. In this paper, a novel graph-based neural network (GNN) regression is proposed to leverage this local structural information for facial image SR. Firstly, the grid representation of an input face image is converted into its corresponding graph representation, and then a shallow neural network is trained for each vertex in the graph in order to regress the SR image. Compared with its grid-based counterpart, the graph representation combines both coordinate affinity and textural similarity. Additionally, the NN weights of a vertex are initialized with those converged ones from its neighbors, resulting fast convergence for training and accurate regression. Experimental comparison with the state-of-the-art SR algorithms including those based on deep convolutional neural networks (DCNN) on two benchmark face sets demonstrate the effectiveness of the proposed method in terms of both qualitative inspections and quantitative metrics. The proposed GNN is not only able to deal with facial SR, but also has the potential to apply to data examples with any irregular

* 基金项目: 国家自然科学基金(61572096, 61432003)

Foundation item: National Natural Science Foundation of China (61572096, 61432003)

本文由“多媒体大数据处理与分析”专题特约编辑赵耀教授、李波教授、华先胜研究员、文继荣教授、蒋刚毅教授、常冬霞副教授推荐.

收稿时间: 2017-04-28; 修改时间: 2017-06-26; 采用时间: 2017-10-13; jos 在线出版时间: 2017-12-01

CNKI 网络优先出版: 2017-12-04 06:46:50, <http://kns.cnki.net/kcms/detail/11.2560.TP.20171204.0646.006.html>

topology structure.

Key words: facial image; super-resolution; graph; neural network; regression

近年来,随着计算机性能,特别是 GPU 性能的提升,以及视频图像等海量数据的支持,卷积神经网络取得了前所未有的成就.它在目标检测^[1]、图像分类识别^[2]、艺术风格转换^[3]、视频图像分析^[4],特别是图像超分辨率等研究领域都得到了广泛应用.在图像超分辨率,特别是在针对人脸图像这种特殊场景下的超分辨率问题中,当描述人脸上一个像素的局部信息时,往往希望它的局部描述信息中应尽量包含人脸信息,而排除背景等与人脸不相关的信息.然而,基于经典信号处理、机器学习,乃至最新卷积神经网络(CNN)的超分辨率方法,通常使用具有规则矩形结构的操作提取图像中每个像素的局部模式^[5],忽略了相邻像素间在内容上的相关性,最终导致在一个像素局部描述信息中既包含与该像素相关的有效信息,同时也包含与该像素无关的无效信息.因此,针对类似人脸的结构化物体图像,在对下采样倍数较大的图像进行超分辨率提取时,这类方法一般会生成严重影响视觉效果效果的伪影.

若将 2D 图像考虑成图结构,图像中每一个像素为图结构中的一个节点,从而每个节点既包含该像素的位置信息,也包含纹理信息.在描述一个节点的局部信息时不再简单地采用规则的矩形网格,而是以节点之间的相关性为度量,计算与该节点最相关的 k 个近邻来描述该节点,即能很好地克服上述问题.在去噪、去块效应等应用场景中,采用图结构刻画图像的方法能够获得更为理想的结果.最近,针对图这类具有任意拓扑结构的数据^[6,7],研究者们提出了多种图卷积神经网络(graph convolutional network,简称 GCN)^[8-10],具有类似标准 CNN 的深层结构,在图上定义卷积、池化(pooling)操作以及参数共享机制等.GCN 在图像颜色转化、风格转化和去噪等方面进行了初步的尝试,取得了优于传统深度 CNN 的结果^[9].

本文受此类方法启发,试图用图结构描述 2D 人脸图像,探讨图表示与传统的基于规则网格表示的差异,并针对 2D 人脸图像具有局部一致性的特点,提出了一种新的基于图结构的人脸超分辨率神经网络回归算法.与 GCN 不同,提出的回归算法无需在图结构上定义池化以及参数共享等操作,而是为图结构中每个节点的局部信息训练一个相对简单的浅层子神经网络,从而将对整幅人脸图像的超分辨率回归问题划分为多个针对局部信息可解子问题.同时,考虑到图像像素各个图结构回归之间的相关性,进一步提出一种网络间参数传递共享机制,不但加快了每个子网络的收敛速度,同时还提高了回归准确率.在最新的图像分类文献^[11]中,He 等人利用集成学习的思想提出了 ResNeXt,通过为同一目标训练多个具有相同结构的简单卷积网络达到提升分类准确率的目的,实现比增加网络宽度和深度更有效的提升网路性能的方法.ResNet^[12]与 ResNeXt^[11]更像是决策树和随机森林的关系,而本文方法从问题本身出发,将一个复杂问题划分为多个相对简单的子问题,整合多个具有简单结构的神经网络共同解决整张人脸图像的超分辨率问题,每个子网络反映不同局部结构的属性,建立不同的回归映射.

本文的主要贡献为:将低分辨率(LR)人脸图像转化为图结构,以更好地描述每个像素的局部信息,并将整张人脸图像的超分辨率回归问题分解成多个针对局部信息进行回归的子问题,为每个像素训练一个简单的神经网络.并在神经网络训练过程中,建立参数传递机制,从而加快每个网络的收敛速度且提高了预测精度.本文提出的图神经网络(GNN)的应用范围不局限于二维人脸图像超分辨率问题,还可以应用于具有不同拓扑结构的数据,解决不同的问题.

本文第 1 节介绍图像超分辨率研究现状.第 2 节详细叙述基于图结构的人脸超分辨率回归方法.第 3 节为验证实验和对比实验.最后为总结.

1 超分辨率相关工作

图像超分辨率方法可分为基于函数插值、基于重建和基于学习 3 类.

基于函数插值方法最早用于图像缩放,利用固定的核函数通过 LR 图像中的几个相邻像素来估计高分辨率(HR)图像中缺失的像素值,常见的插值方法包括最近邻插值(nearest neighbor interpolation)、双线性插值(bilinear interpolation)和双三次插值(bicubic interpolation).这类方法生成 HR 图像时只需要一张 LR 图像,不需要训练,且速度

较快,但是,对整幅图像的所有位置采用相同结构的核函数进行操作,往往导致 HR 图像的过度平滑,并且伴有锯齿和振铃效应.文献[13]以方向梯度直方图(histograms of oriented gradients,简称 HoG)为特征,提出了一种稀疏的近邻选择方法,在近邻选择时不仅仅依靠物理坐标上的相似,还考虑到纹理特征,有效提升了插值过程中的细节表现.文献[14,15]在插值过程中分别引入软边平滑和梯度曲线等先验知识,从而达到更好的效果.

在基于重建的方法中,往往通过在相同场景中的一组 LR 图像来重构一张 HR 图像^[16-19].这类方法通常用于解决视频序列的超分辨率问题^[20].在重建过程中,LR 图像序列的配准与降噪是其中最需要解决的问题.

近年来,基于学习的超分辨率方法成为研究的热点.在该类算法中,它们假设 HR 图像和 LR 图像之间隐含的映射关系可以通过利用大量的训练样本学习得到.根据面向对象的不同,这类方法又可以分为两类,一类为自然图像的超分辨率算法,它用于处理各种场景下的超分辨率问题^[21,22].随着深度学习算法的广泛应用,基于卷积神经网络的自然图像超分辨率算法在预测精度和细节表现上都有了极大的飞跃^[23,24];另一类方法用于专门解决某一类问题,例如人脸超分辨率和艺术作品的超分辨率.本文主要就人脸超分辨率问题展开研究.

人脸图像具有结构化的特点,基于全局脸的人脸超分辨率方法,通过主成分分析^[25,26]、典型相关分析^[27]或者最小二乘分解^[28]等方法对训练样本进行子空间分解,得到一组能够表征 LR 和 HR 人脸图像特征的基向量.将输入 LR 人脸图像投影到该 LR 人脸图像子空间,得到输入图像在此子空间上的投影系数,并利用该投影系数在 HR 子空间中进行超分辨率重建^[29,30].这类方法可以很好地保持人脸的全局信息,并且拥有较强的去噪能力,但在人脸局部细节信息的恢复上,表现欠佳.为了获得更多的高频细节信息,基于局部像素块人脸超分辨率方法最近获得了更多的关注,研究者们将图像划分为多个小块,学习每个低分辨率和高分辨率图像块之间的映射关系.其中,针对高低分辨率图像块流形空间具有局部一致性的假设,提出了很多基于局部线性约束或者稀疏表示的方法^[31-35].这类方法的基本假设是高低分辨率图像流形空间具有局部几何一致性,通过将低分辨率图像块之间的近邻关系保持到高分辨率空间,来实现超分辨率重建.但在图像降采样过程中,存在图像信息丢失的现象.因此,对于一张 LR 图像,可能存在多张 HR 图像与之对应,所以高低分辨率图像在流形空间具有局部几何一致性的假设并不一定成立^[36].为此,很多基于回归的方法被提出.文献[37]使用高斯过程回归,利用 LR 图像中每个像素的近邻逐像素地回归 HR 图像,通过大量训练数据直接建立 LR 图像到 HR 图像的映射,取得了很好的结果.文献[38]结合插值、回归以及重建 3 类方法,首先利用 LR 图像块与 HR 图像块之间的映射关系回归 LR 图像块的局部结构先验(local structure prior,简称 LSP),然后利用 LSP 来预测 HR 图像.

2 基于图结构的人脸超分辨率回归方法

低分辨率图像往往被建模成由高分辨率图像通过模糊、噪声和降采样等一系列操作作用的结果,表示为

$$I_{LR} = DBI_{SR} + N \quad (1)$$

其中, I_{LR} 表示低分辨率图像, I_{SR} 表示高分辨率图像, D 表示降采样, B 表示模糊, N 表示添加噪声.本文针对人脸图像讨论其中的一种特殊情况,即忽略模糊和噪声的影响.

$$I_{LR} = DI_{SR} \quad (2)$$

本节主要介绍我们提出的基于图结构的人脸超分辨率神经网络回归方法.图 1 给出基于图结构的人脸超分辨率神经网络回归方法的整个流程.

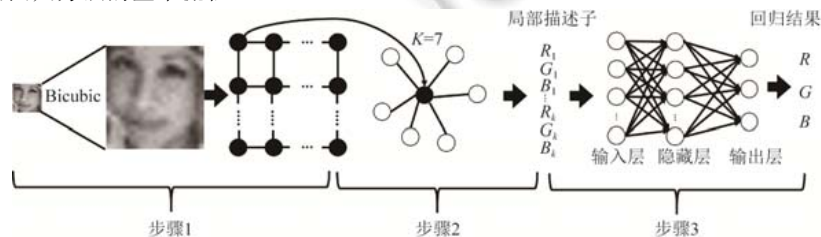


Fig.1 Overview of our graph based face image super-resolution regression method

图 1 基于图结构的人脸超分辨率神经网络回归方法流程图

图 1 主要分为 3 个部分:(1) 将二维图像一般化为传统图结构.(2) 计算图结构中每个节点的 k 近邻,从而获得每个节点的局部描述子.(3) 为图结构中的每个节点训练一个神经网络,利用节点的局部描述子回归人脸超分辨率图像.

2.1 人脸图像到图结构转换

本文将二维人脸图像考虑成一个无向图 $G=(v,e,w)$,其中, v 表示无向图中的节点,2D 人脸图像中的每一个像素点被认为是一个节点; e 表示节点之间的边,假设每一个节点到其他所有节点(包括此节点本身)都存在边; w 为邻接矩阵,保存节点到节点间每条边的权重.对于 G 中的每一个节点 v 都有 $v=[index,x,f]^T$,其中, $f=[R,G,B]$ 分别表示该节点所对应的像素的 R,G,B 值, $x=[u,v]$ 表示像素的归一化图像坐标,而 $index$ 表示节点的序号.

在将输入的低分辨率人脸图像转化为图结构时,首先,与其他超分辨率方法一样,采用 Bicubic 插值算法将低分辨率人脸图像插值成与目标分辨率大小相同的图像,如图 1 中步骤 1 所示.而后,按照固定顺序遍历图像中的每个像素,获得图结构中每个节点的信息.

在遍历图像中每个像素构造节点信息时,每个节点的 $index$ 值需要满足: $index$ 值相邻的两个节点,所对应的两个像素在图像坐标系中也是相邻的.如图 2 所示,为了达到这个目的,采用顺时针遍历图像中每个像素.图中方框表示像素点,方框中的数字表示遍历每个像素的先后顺序,也表示生成的节点中的 $index$ 值.由于图像中局部的两个相邻像素在很大程度上较为相似,可以认为相邻像素的网络权重也较为接近,因此,在逐个训练每个节点的神经网络时,只随机初始化 $index=1$ 这个节点的网络权重,之后的每个节点的训练都由它的前驱节点的网络权重进行初始化.随后的实验对比了随机初始化和依靠相邻节点初始化两种网络权重初始化方法,提出的方法不仅加快了网络的收敛速度,同时可使网络更好地收敛,在训练时间和准确率上都比随机初始化有更好的表现.

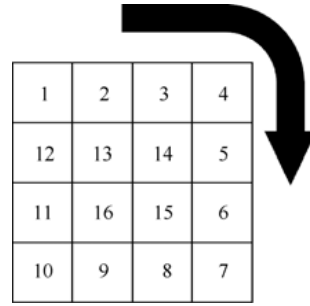


Fig.2 Traverse each pixel in the image clockwise

图 2 顺时针遍历图像中的每个像素

2.2 提取节点局部描述子

将人脸图像一般化为传统的图结构之后,接下来需要解决的问题是如何描述图结构中每个节点.考虑到人脸图像具有局部一致性,本文认为低分辨率人脸图像中的每个节点可以用其周围的 k 个最近邻描述(包括该节点本身),下面针对提取节点局部描述子,对如下 3 个问题进行讨论:(1) 特征提取;(2) 图结构中边权重的计算;(3) k 值的选择节点.

对于特征提取,在本文只利用每个节点的 k 个最近邻的 R,G,B 信息来对节点特征进行描述,并把图结构中一个节点 i 的特征表示为

$$F_i = [\alpha_{i1}f_1, \alpha_{i2}f_2, \dots, \alpha_{ik}f_k]^T \quad (3)$$

f_1, f_2, \dots, f_k 为节点 i 的 k 个近邻 v_1, v_2, \dots, v_k 的 R,G,B 信息,节点 i 到 k 个近邻的边 e_{ij} 的权重满足 $w_{i1} < w_{i2} < \dots < w_{ik}$.而 α_{ij} 表示为

$$\alpha_{ij} = 1 - \frac{j}{k}, j = 1, 2, \dots, k \quad (4)$$

接下来讨论图结构中边的权重的计算以及 k 值的选择.在将低分辨率人脸图像转化为图结构之后,本文认为图结构中的每个节点到其他所有节点都存在边(包括节点本身),对于每条边的权重,有多种不同的计算方法,每种计算方法都将直接影响 k 近邻的选择.在基于规则矩形网格的超分辨率方法以及卷积神经网络中,通常强利用中心像素周围的一个矩形网格区域内的所有像素对中心像素进行描述,而忽略掉纹理信息的相似性.这种方法表示在图结构中就是:对于图 G 中的一条连接节点 $v_i=[index_i, x_i, f_i]$ 和 $v_j=[index_j, x_j, f_j]^T$ 的边 e_{ij} ,其权重 w_{ij} 为

$$w_{ij} = \|x_i - x_j\|_2 \quad (3)$$

这样节点的 k 近邻表示的是在图像坐标系中与该节点距离最近的 k 个节点,而节点的纹理信息可能相似也可能差距很大.本文把节点中的所有信息看作一个整体,即 $x'=[x,\delta]^T$,这里, δ 用来确定 R,G,B 信息所占比重,这样边 e_{ij} 的权重 w_{ij} 表示为

$$w_{ij} = \|x'_i - x'_j\|_2 \quad (4)$$

此时, w_{ij} 表示两个节点的相似程度, k 近邻表示 k 个与当前节点最相似的节点,它不仅要求这两个节点所对应的像素在图像坐标系中距离近,同时还要求在像素值上也相似.考虑到对 HR 图像下采样的倍数 D 越大,生成的 LR 图像中相邻像素在 HR 图像中的实际位置相差越大,为了保证对每个节点计算得到的 k 近邻尽量都分布在该节点附近,本文取 $\delta = \frac{1}{D}$.

如图 3(a)所示为传统的利用网格内所有像素描信息描述中心像素的方法,图 3(b)所示为利用图结构中节点间相似性为度量计算 k 近邻来描述节点的方法.从图中可以看出,这两种方法的主要区别在于对图像中靠近边缘区域的描述.在图 3(a)所示传统的基于网格的方法中,在对图像中人脸上一个像素点进行描述时,由于只考虑距离的相似性,这个像素点的描述信息中同时包含人脸以及背景的像素信息;而在图 3(b)所示方法中,由于同时考虑了距离以及纹理的相似性,在对一个像素点的描述信息中,更多的是人脸上与该像素相似的像素信息.针对不同的需要,这两种方法都有各自的优点,但当试图用回归方法解决人脸超分辨率问题时,我们更希望在对人脸上的一个像素点的描述过程中,应尽量使用该像素周围同样属于人脸的有效信息来描述这个像素,而不是掺杂进背景等其他无效的信息.这也是本文采用的图结构来描述人脸图像的主要原因.



(a) 基于规则网格方法对一个像素局部信息的描述

(b) 本文方法对一个像素局部信息的描述

Fig.3

图 3

对于 k 值的选择,在利用网格的方法中,当想通过放大网格尺寸来获得对像素点更多描述信息时,有效信息增加的同时,无效信息也会增加,这往往导致当网格内有效信息还不足以更好地描述当前像素点时,过多的无效信息就已经拉低了整体的描述能力.在利用图结构描述图像的方法中,当增大 k 值时,能够很好地抑制无效信息的增加,并且获得更多的有效信息,从而达到更好的局部信息的描述能力.同样,考虑到对 HR 图像下采样的倍数 D 越大,生成的 LR 图像中相邻像素在 HR 图像中的实际位置相差越大,为了保证每个节点的局部描述子尽可能少地包含无效信息,本文取 $k = \left\lfloor \frac{100}{D} \right\rfloor$.

2.3 训练神经网络回归模型

BP 神经网络作为最常用的一种非线性多目标回归工具,利用反向传播算法,能够直接从输入的大量训练数据中学习输入样本与目标值之间存在的复杂映射关系,而不需要人工设计变量之间存在的映射关系,避免了人工干预的影响,对非线性多目标预测问题有着更好的拟合能力.提高神经网络性能的方法往往包括增加神经元的个数以及网络的深度.同时,按照集成学习的思想,增加网络的基数同样可以提高网络的性能^[8].本文充分考虑到人脸图像之间的局部一致性,采取为图结构中的每个节点训练一个网络的方法进行超分辨率回归.但是,与集成学习不同,本文将整张人脸图像的超分辨率问题以像素为单位将其划分为多个针对人脸局部进行超分辨率的子问题,并且每个网络只解决一个局部的超分辨率问题.

本文将由低分辨率人脸图像中每个节点的特征 F^{LR} 预测对应的高分辨率节点 R, G, B 值 f^{SR} 的过程看作一个多目标非线性回归问题, 并采用 BP 神经网络进行预测. 网络的回归目标可以表示为

$$R_j = \arg \min \sum_{i=1}^n \| F_{ij}^{LR} R_j - f_{ij}^{SR} \|_2^2 \quad (7)$$

式中, R_j 表示第 j 个节点的回归子, F_{ij}^{LR} 表示第 i 个低分辨率图像中第 j 个节点的特征, f_{ij}^{SR} 表示第 i 个高分率图像中第 j 个节点的 R, G, B 值.

通过对图结构中的每一个节点逐地进行训练, 将人脸超分辨率问题转化成相对简单的多个子问题, 为了减小总体的参数规模, 本文尽可能地简化了每个节点的网络结构. 首先, 每个节点的网络结构相同; 其次, 对于网络的深度, 实验中发现, 在有限数量的训练样本情况下, 增加网络的深度非但不能提高回归的准确度, 反而因为训练样本过少以及梯度弥散等问题使得结果更差, 所以最终只使用包含一个隐藏层的网络结构; 再次, 对于隐藏层神经元的数量, 后续实验结果表明, 隐藏层神经元的数量不宜过多, 综合考虑到训练速度、参数规模以及回归的准确率, 本文认为网络中隐藏层神经元数量为 50 比较合适, 并且随着训练数据的增加, 适当增加神经元的数量往往能取得更好的结果.

3 实验结果与分析

本节首先介绍实验比较采用的两种质量度量指标以及实验所使用的两组数据集. 而后对第 2 节提出的算法的各个部分进行实验验证. 最后, 在 FEI^[34] 和本文作者所在实验室采集的 DUT Multi-Views 数据集上, 将提出的基于图结构的超分辨率回归方法与现有的针对人脸以及自然图像的最新超分辨率方法进行了对比.

3.1 度量标准与数据集

超分辨率问题的评价标准包括峰值信噪比(peak signal-to-noise rate, 简称 PSNR)和结构相似度量(structural similarity, 简称 SSIM)两种, 其中, PSNR 的计算公式如下:

$$PSNR(I_i, I_j) = 10 \cdot \lg \frac{P^2}{\frac{1}{WH} \sum_{x,y} (I_i(u,v) - I_j(u,v))^2} \quad (8)$$

其中, I_i, I_j 分别表示需要比较的两幅图像, W, H 为图像的高度和宽度, P 为最大灰度值, 在 8 位图中取值 255. 而 SSIM 计算公式如下:

$$SSIM(I_i, I_j) = \frac{1}{WH} \sum_{x,y} \frac{(2\mu_i\mu_j + \beta_i^2 P^2)(2\sigma_{ij} + \beta_j^2 P^2)}{(\mu_i^2 + \mu_j^2 + \beta_i^2 P^2)(\sigma_i^2 + \sigma_j^2 + \beta_j^2 P^2)} \quad (9)$$

其中, $\mu_i(x,y)$ 和 $\mu_j(x,y)$ 分别表示在两幅图像中以 (u,v) 为中心的像素块的均值(像素块大小通常为 8×8); $\sigma_i(x,y)$ 和 $\sigma_j(x,y)$ 分别表示在两幅图像中以 (u,v) 为中心的像素块的方差; $\sigma_{ij}(x,y)$ 表示两幅图像中像素块的协方差; β_i 和 β_j 为两个常数, 通常取 $\beta_i=0.01, \beta_j=0.03$. 实验运用的数据集包括 DUT Multi-View 和 FEI, 具体信息如下.

(1) DUT Multi-View 数据集: 由大连理工大学软件学院几何计算与数字媒体技术实验室采集. 由 2D 多视角人脸图像组成. 其中, 2D 多视角人脸图像共包含来自 547 个中国人的 8 205 个样本, 年龄分布为 18 岁~25 岁, 其中, 每个目标包含 15 个同步采集的不同视角的样本. 15 个不同视角的样本分别由 15 台分布在距离目标相同半径(1.5m)的球面上的佳能单反相机(DSLR)采集.

(2) FEI 数据集^[34]: FEI 数据集中共包含来自 200 个人(100 名男性, 100 名女性)的 400 张正脸人脸图像, 每人包含自然和微笑两种表情. 图像的原始分辨率为 360×260 .

3.2 验证实验

在第 2.1 节中, 为了加快训练的速度以及提高回归精度, 提出了利用相邻的上一个节点的神经网络参数初始化当前节点神经网络参数的方法. 为了验证提出方法的有效性, 本文与随机初始化的方法进行了对比. 实验中, 在完全相同的单核单线程的硬件环境下, 使用 DUT Multi-View 数据集中的 400 张多视角人脸图像进行训

练,100 张人脸图像进行测试,HR 图像分辨率为 64×64 ,LR 图像分辨率为 32×32 .分别计算了 100 张测试图像的平均 PSNR 和平均 SSIM.从表 1 中可以看出,通过近邻初始化网络权重神经网络的训练时间减少为原来的 $\frac{1}{3}$,同时网络得到更好的收敛,在 100 个测试图像中 PSNR 有了 4.8dB 的提高,而 SSIM 也提高了 0.6.

Table 1 Comparison between initial network weight randomly and initial it by neighbor

表 1 随机初始化与近邻初始化结果对比

	Initial randomly	Initial by neighbor
Train time (h)	2.101 3	0.635 4
PSNR	30.266 0	35.006 9
SSIM	0.904 7	0.961 6

在第 2.2 节中,本文就为什么选用图结构来描述人脸图像进行了详细的讨论,分析了传统的基于网格的特征描述方法与基于图结构特征描述方法的主要区别,为了证明提出方法的有效性,同样在 DUT Multi-View 数据集中进行了实验.实验中为了模拟基于网格的特征描述方法,使用公式(3)计算每条边的权重,即只考虑坐标的相似性而忽略纹理的相似性.选取 400 张多角度人脸图像进行训练,100 张多角度人脸图像进行测试,其中 HR 图像分辨率为 64×64 ,LR 图像分辨率为 32×32 .神经网络隐藏层神经元个数为 50.分别计算两种方法取不同 k 值时,在测试集上的平均 PSNR 以及平均 SSIM,得到如图 4 所示结果.图中,PSNR1 和 SSIM1 表示用公式(3)计算权重得到的结果,PSNR2 和 SSIM2 为由公式(4)计算权重得到的实验结果,横坐标为 k 的值,左边的纵坐标表示 PSNR 值,右边纵坐标表示 SSIM 值.从图中可以看出,基于图结构的局部特征描述方法相对于只考虑坐标相似的基于网格的方法取得了更好的结果.这说明,在人脸超分辨率问题中,在对每个像素的局部信息进行描述时,同时考虑坐标相似性与纹理相似性会更有效地描述当前像素的特征.从图中还可以看出,在基于网格的特征描述方法中,随着 k 值的增大,无效信息和有效信息同时增加,并且很快无效信息占据主导地位,当 $k=15$ 时,算法取得最大精度,之后随着 k 值的增加,精度便开始下降.而在本文的方法中,由于尽可能避免加入无效信息,所以在 k 值较小时,便取得了较网格方法更好的精度(这里对于 $k=1$ 时,两种方法结果不同的原因,本文认为主要是由于每个节点的描述信息过少,神经网络收敛不稳定导致).同时,随着 k 值的增大,当基于网格的方法的精度已经不能继续提高时,本文方法的精度却随着 k 值的增大而进一步提高.值得注意的是,本文在相同的训练集以及测试集上对最新的人脸超分辨率算法 SRLSPS^[26]进行了实验,得到的平均 PSNR 和平均 SSIM 值分别为 33.413 5dB 和 0.952 7,而当 $k=15$ 时,基于网格的特征描述方法的平均 PSNR 和平均 SSIM 值分别为 33.275 9dB 和 0.947 5,这说明,对基于网格的特征描述方法的模拟几乎可以达到最新人脸超分辨率算法的精度,也进一步证明了本文方法的有效性.

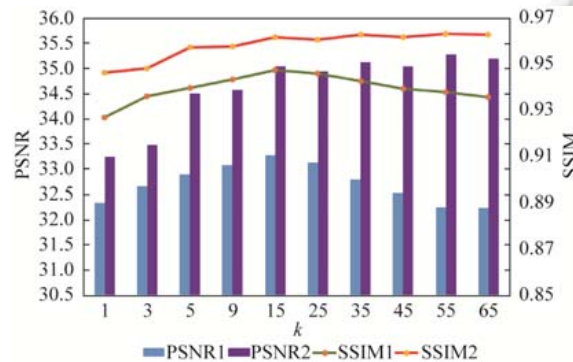


Fig.4 Comparison between grid based method and our graph based method

图 4 基于网格方法与本文的基于图结构方法比较

对于隐藏层神经元的数量,依旧选用上述训练集和测试集,本文针对取不同隐藏层神经元数量时对算法的影响进行了评估,得到如图 5 所示结果.图中横坐标表示隐藏层神经元数量,可以看出当隐藏层神经元数量在 10~100 之间时,实验结果基本保持不变.而当隐藏层神经元数量进一步增加到 150~500 后,在有限数量的训练样

本的情况下,网络出现过拟合,准确率下降.为了在保证精度的前提下,尽可能地简化每个子网络的结构,本文在拥有少量训练样本的情况下取隐藏层神经元数量为 50,当训练样本数量增加时,隐藏层神经元数量也会适当增加.

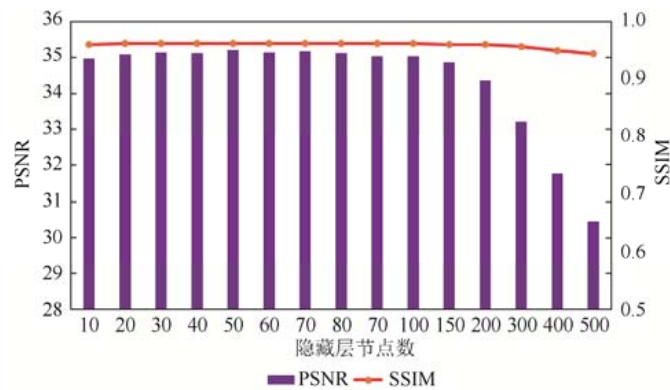


Fig.5 The influence of the number of neur for regression precision

图 5 不同隐藏层神经元数对回归精度的影响

3.3 对比实验

首先本文在 FEI^[39]数据集上将基于图结构的人脸超分辨率回归方法与最新的人脸超分辨率方法行了对比.实验中设定 HR 图像分辨率为 180×130 ,对 HR 图像进行 2、4、8 倍降采样得到的 LR 图像分辨率分别为 90×65 、 45×33 、 23×18 .对比的人脸超分辨率算法包括:LCR^[40]、LINE^[41]、SRLSPS^[38].其中,LCR^[40]在邻域嵌入过程中,通过自适应地选择最相关的像素块,避免了传统邻域嵌入算法中使用固定数量近邻导致的欠拟或过拟合.而 LINE^[41]采取由粗到细的方法,迭代地更新每个 LR 像素块所对应的相关像素块,从而解决图像降采样过程中 LR 图像与 HR 图像在流形空间局部一致性假设不成立的问题.对于每一种方法都使用随机选取出的 300 张人脸图像进行训练,剩余的 100 张人脸图像进行测试.本文的方法中,神经网络隐层神经元数为 50.

Table 2 Comparison of performance of our method against state-of-the-art method on FEI dataset

表 2 本文方法与最新人脸超分辨率方法在 FEI 数据集上的对比

SRF	Our		LCR		LINE		SRLSPS	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
2	40.361 7	0.976 1	36.677 7	0.951 5	36.629 7	0.940 0	37.484 2	0.967 3
3	34.310 8	0.916 3	33.499 7	0.908 6	33.607 3	0.909 3	30.125 8	0.847 9
4	28.889 6	0.793 3	25.611 5	0.725 9	25.820 4	0.745 5	24.750 5	0.639 4

实验结果见表 2.可以看出,在 2 倍、4 倍以及 8 倍降采样的 LR 测试图像上,相比其他 3 种方法,本文方法都取得了最好的实验结果,特别是在 2 倍和 8 倍降采样测试图像上,PSNR 值相比其他 3 种方法提高了 3dB~4dB.这充分说明,相比利用规则矩形网格描述图像局部的方法,本文通过将 LR 人脸图像转化为传统图结构,并在对每个像素的描述过程中,同时考虑到了几何和纹理的影响,很好地表达了每个像素的局部信息;再加以分治的思想,针对人脸局部结构相似的特点,为每一个像素所形成的局部区域训练一个简单的神经网络,充分学习到了每个 LR 图像像素与 HR 图像像素之间的映射关系,最终使得本文的方法在 2 倍、4 倍、8 倍降采样的测试 LR 图像上都取得了最好结果.值得说明的是,上述与本文对比的 3 种方法都只是在灰度通道对人脸图像进行超分辨率操作,而本文的方法利用神经网络同时回归 RGB 三通道的像素值,但是为了方便比较,本文将得到的三通道结果图像转为灰度图像,并在灰度图像上与 3 种方法进行了对比.图 6 为本文方法和其他 3 种方法在 FEI 数据集上的部分结果,可以看出,对于 8 倍降采样的 LR 人脸图像,本文方法获得了更多的细节表现,更好地重建出人脸图像中关键的五官信息.

随后,在 DUT Multi-View 多视角人脸数据集上,本文与人脸超分辨率算法 SRLSPS^[38],以及 SRCNN^[42]、SelfExSR^[43]、VDSR^[44]这 3 种最新的针对自然图像的超分辨率算法进行了对比.其中, SRCNN^[42]与 VDSR^[44]采

用深度卷积神经网络来实现超分辨率重建,而SelfExSR^[43]直接根据输入LR图像中存在的自相似性进行超分辨率重建.实验过程中,设定HR图像分辨率为128×128,对HR图像进行3、4、8倍降采样得到的LR图像分辨率分别为43×43、32×32和16×16.对于SRLSPS^[38]、SRCNN^[42]、VDSR^[44]和本文方法,选取DUT Multi-View多视角人脸数据集中2000张图像进行训练.在本文的方法中,设定每个子神经网络隐藏层神经元数为100.而SelfExSR^[43]通过输入图像的自相似性进行超分辨率重建,不需要训练集.实验中测试样本集为DUT Multi-View多视角人脸数据集中的500张不同视角的人脸图像,且训练样本集与测试样本集不存在交叉.

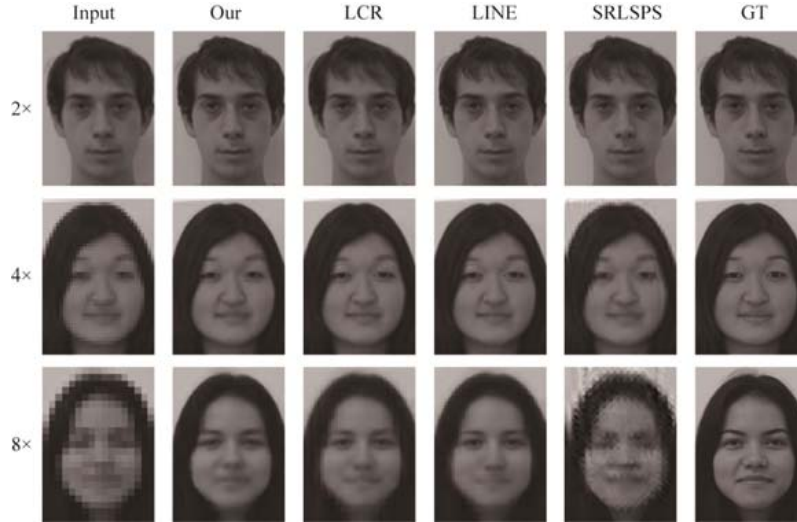


Fig.6 Some of results on FEI dataset

图6 FEI数据集上的部分实验结果

实验结果见表3.对于3倍降采样的LR人脸图像,本文方法得到的结果与基于图像自相似性的SelfExSR^[38]相当,超过了基于卷积神经网络的SRCNN^[42]以及基于邻域嵌入的SRLSPS^[38],可以看出,在3倍降采样人脸图像上,基于图结构的神经网络回归方法超越了浅层卷积神经网络.在4倍降采样人脸图像上,本文方法达到了与利用20层卷积层的VDSR相当的结果.进一步地,在8倍降采样人脸图像上,基于图结构的神经网络回归方法超过所有其他方法,在PSNR和SSIM上均取得了最好结果.说明本文方法在处理高倍降采样的人脸图像上,相比卷积神经网络等其他方法更具优势.这也充分证明第2.2节中所得出的结论,即对于人脸超分辨率问题,在提取每个像素局部特征时,将图像转化为传统图结构,同时考虑节点间几何与纹理的相似性,相比直接不加区分地利用规则网矩形格中所有信息的方法具有更好的描述能力.图7中给出了实验中5种方法得到的部分实验结果.从图7中可以看出,对于4倍降采样的LR人脸图像,本文方法在PSNR和SSIM上虽然没有VDSR出色,但在双眼、鼻子、嘴巴附近,特别是脸颊边缘,获得了更多的细节表现.进一步地,对于8倍降采样的LR人脸图像,SelfExSR^[43]和VDSR^[44]的结果已经很难辨认出五官信息,出现了严重的涂抹情况,SRLSPS^[38]结果中的边缘也严重抖动.而此时,通过为每一个像素训练一个简单神经网络,并将整张人脸图像的超分辨率问题分解成对每个像素的超分辨率回归问题,使得本文方法在高降采样倍数的情况下,在纹理边缘位置获得了出众的细节表现,很大程度上避免了涂抹情况的产生.

Table 3 Comparison of performance of our method against state-of-the-art method on DUT Multi-View

表3 本文方法与最新人脸超分辨率方法在DUT Multi-View数据集上的对比

SRF	Our		SRLSPS		SRCNN		SelfExSR		VDSR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
3	36.215 8	0.953 5	34.234 4	0.938 1	35.076 7	0.945 3	36.259 9	0.958 0	36.875 4	0.960 4
4	33.921 5	0.924 4	31.314 6	0.891 8	32.567 7	0.908 5	33.434 7	0.925 3	34.307 5	0.933 7
8	28.956 9	0.819 9	26.966 2	0.753 8	26.734 2	0.747 9	26.866 8	0.756 7	28.319 2	0.809 0

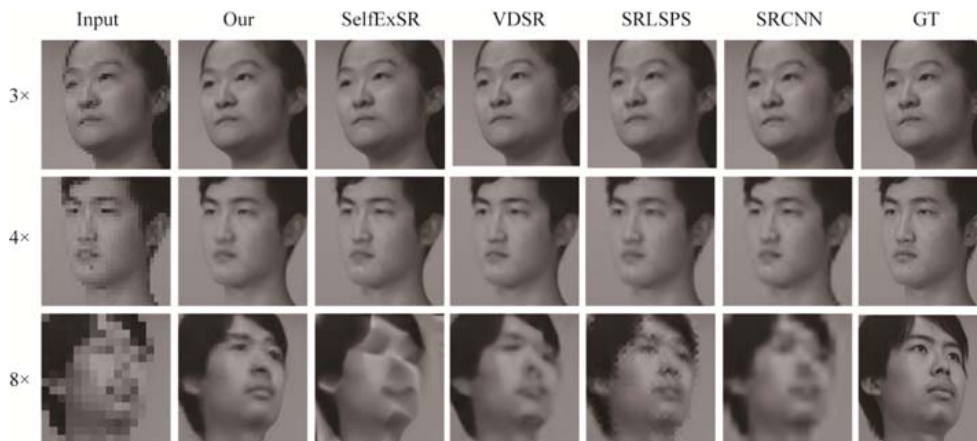


Fig.7 Some of results on DUT Multi-View dataset

图7 DUT Multi-View 数据集上的部分实验结果

4 结论与总结

本文针对人脸图像具有局部一致性的特点,提出了一种基于图结构的人脸超分辨率神经网络回归方法.通过将 LR 人脸图像转换为图结构,在局部特征描述时,同时考虑节点的纹理与几何的相关性,相比传统的基于网格的方法更好地表达了每个像素周围的局部特征.同时,通过为每一个像素训练一个简单的子神经网络,将整张人脸图像上的超分辨率问题分解成对每个像素的超分辨率回归问题,使得本文方法在高降采样倍数的情况下,在纹理边缘位置获得了出色的细节表现,很大程度上避免了涂抹情况的产生.在子网络训练过程中,充分利用相邻像素较为相似的特点,本文提出一种神经网络参数传递方法,通过用已经收敛的相邻像素的网络参数来初始化下一个网络的参数,极大地提高了神经网络的收敛速度以及预测精度.在与 3 种最新人脸超分辨率算法的比较中,本文方法取得了最好的结果;在与最新的基于深度卷积神经网络等最新的自然图像超分辨率算法的比较中,本文方法在高降采样倍数的 LR 人脸图像上表现出极大的优势.本文提出的图神经网络(GNN)并不局限于解决二维人脸图像的超分辨率问题,它还可以应用于诸如 3D 网格等各种具有不规则拓扑结构的数据,解决不同的问题.

References:

- [1] Shin HC, Roth HR, Gao M, Lu L, Xu, Z, Nogueis I, Summers RM. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. on Medical Imaging*, 2016,35(5):1285–1298. [doi: 10.1109/TMI.2016.2528162]
- [2] Rastegari M, Ordonez V, Redmon J, Farhadi A. XNOR-Net: ImageNet classification using binary convolutional neural networks. In: Bastian L, Jiri M, Nicu S, Max W, eds. *Proc. of the European Conf. on Computer Vision*. Cham: Springer Int'l Publishing, 2016. 525–542. https://link.springer.com/chapter/10.1007/978-3-319-46493-0_32
- [3] Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2016. 2414–2423. [doi: 10.1109/CVPR.2016.265]
- [4] Wang CF, Su L, Zhang WG, Huang QM. No reference video quality assessment based on 3D convolutional neural network. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(Suppl.(2)):103–112 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/16025.htm>
- [5] Liu N, Li CH. Single image super-resolution reconstruction via deep convolutional neural network. *China Sciencepaper*, 2015, 10(2):201–206 (in Chinese with English abstract).
- [6] Hu W, Cheung G, Kazui M. Graph-Based dequantization of block-compressed piecewise smooth images. *IEEE Signal Processing Letters*, 2016,23(2):242–246. [doi: 10.1109/LSP.2015.2510379]

- [7] Pang J, Cheung G. Graph Laplacian regularization for image denoising: Analysis in the continuous domain. *IEEE Trans. on Image Processing*, 2017,26(4):1770–1785. [doi: 10.1109/TIP.2017.2651400]
- [8] Niepert M, Ahmed M, Kutzkov K. Learning convolutional neural networks for graphs. In: Balcan MF, Weinberger KQ, eds. *Proc. of the 33rd Annual Int'l Conf. on Machine Learning*. New York: ACM, 2016. <http://proceedings.mlr.press/v48/niepert16.html>
- [9] Puy G, Kitić S, Pérez P. Unifying local and non-local signal processing with graph CNNs. *arXiv:1702.07759*, 2017.
- [10] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R, eds. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2016. 3837–3845.
- [11] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. *arXiv: 1611.05431*, 2016.
- [12] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2016. 770–778. [doi: 10.1109/CVPR.2016.90]
- [13] Gao X, Zhang K, Tao D, Li X. Image super-resolution with sparse neighbor embedding. *IEEE Trans. on Image Processing*, 2012,21(7):3194–3205. [doi: 10.1109/TIP.2012.2190080]
- [14] Dai S, Han M, Xu W, Wu Y, Gong Y. Soft edge smoothness prior for alpha channel super resolution. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2007. 1–8. [doi: 10.1109/CVPR.2007.383028]
- [15] Sun J, Xu Z, Shum HY. Image super-resolution using gradient profile prior. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2008. 1–8. [doi: 10.1109/CVPR.2008.4587659]
- [16] Ji H, Fermüller C. Robust wavelet-based super-resolution reconstruction: Theory and algorithm. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(4):649–660. [doi: 10.1109/TPAMI.2008.103]
- [17] Belekos SP, Galatsanos NP, Katsaggelos AK. Maximum a posteriori video super-resolution using a new multichannel image prior. *IEEE Trans. on Image Processing*, 2010,19(6):1451–1464. [doi: 10.1109/TIP.2010.2042115]
- [18] Zhu Y, Li K, Jiang J. Video super-resolution based on automatic key-frame selection and feature-guided variational optical flow. *Signal Processing, Image Communication*, 2014,29(8):875–886. [doi: 10.1016/j.image.2014.06.005]
- [19] Li K, Zhu Y, Yang J, Jiang J. Video super-resolution using an adaptive superpixel-guided auto-regressive model. *Pattern Recognition*, 2016,51,59–71. [doi: 10.1016/j.patcog.2015.08.008]
- [20] Su H, Zhou J, Zhang ZH. Survey of super-resolution image reconstruction methods. *Acta Automatica Sinica*, 2013,39(8):1202–1213 (in Chinese with English abstract).
- [21] Yang J, Wright J, Huang TS, Ma Y. Image super-resolution via sparse representation. *IEEE Trans. on Image Processing*, 2010,19(11):2861–2873. [doi: 10.1109/TIP.2010.2050625]
- [22] Zhang Y, Liu J, Yang W, Guo Z. Image super-resolution based on structure-modulated sparse representation. *IEEE Trans. on Image Processing*, 2015,24(9):2797–2810. [doi: 10.1109/TIP.2015.2431435]
- [23] Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Shi W. Photo-Realistic single image super-resolution using a generative adversarial network. *arXiv:1609.04802*, 2016. [doi: 10.1109/CVPR.2017.19]
- [24] Wang Z, Liu D, Yang J, Han W, Huang T. Deep networks for image super-resolution with sparse prior. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2015. 370–378. [doi: 10.1109/ICCV.2015.50]
- [25] Chakrabarti A, Rajagopalan AN, Chellappa R. Super-Resolution of face images using kernel pca-based prior. *IEEE Trans. on Multimedia*, 2007,9(4):888–892. [doi: 10.1109/TMM.2007.893346]
- [26] Wang X, Tang X. Hallucinating face by eigentransformation. *IEEE Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2005,35(3):425–434. [doi: 10.1109/TSMCC.2005.848171]
- [27] An L, Bhanu B. Face image super-resolution using 2D CCA. *Signal Processing*, 2014,103:184–194. [doi: 10.1016/j.sigpro.2013.10.004]
- [28] Wu W, Liu Z, He X. Learning-Based super resolution using kernel partial least squares. *Image and Vision Computing*, 2011,29(6):394–406. [doi: 10.1016/j.imavis.2011.02.001]
- [29] Park JS, Lee SW. An example-based face hallucination method for single-frame, low-resolution facial images. *IEEE Trans. on Image Processing*, 2008,17(10):1806–1816. [doi: 10.1109/TIP.2008.2001394]
- [30] Capel D, Zisserman A. Super-Resolution from multiple views using learnt image models. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Vol.2. 2001. II. [doi: 10.1109/CVPR.2001.991022]
- [31] Yang J, Wright J, Huang TS, Ma Y. Image super-resolution via sparse representation. *IEEE Trans. on Image Processing*, 2010, 19(11):2861–2873. [doi: 10.1109/TIP.2010.2050625]

- [32] Li M, Cheng J, Le X, Luo HM. Super-Resolution based on sparse dictionary coding. Ruan Jian Xue Bao/Journal of Software, 2012,23(5):1315–1324 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3989.htm> [doi: 10.3724/SP.J.1001.2012.03989]
- [33] Chang H, Yeung DY, Xiong Y. Super-Resolution through neighbor embedding. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Vol.1. 2004. I. [doi: 10.1109/CVPR.2004.1315043]
- [34] Huang H, Wu N. Fast facial image super-resolution via local linear transformations for resource-limited applications. IEEE Trans. on Circuits and Systems for Video Technology, 2011,21(10):1363–1377. [doi: 10.1109/TCSVT.2011.2163461]
- [35] Huang H, He H, Fan X, Zhang J. Super-Resolution of human face image using canonical correlation analysis. Pattern Recognition, 2010,43(7):2532–2543. [doi: 10.1016/j.patcog.2010.02.007]
- [36] Gao X, Zhang K, Tao D, Li X. Joint learning for single-image super-resolution via a coupled constraint. IEEE Trans. on Image Processing, 2012,21(2):469–480. [doi: 10.1109/TIP.2011.2161482]
- [37] He H, Siu WC. Single image super-resolution using Gaussian process regression. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2011. 449–456. [doi: 10.1109/CVPR.2011.5995713]
- [38] Jiang J, Chen C, Ma J, Wang Z, Wang Z, Hu R. SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior. IEEE Trans. on Multimedia, 2017,19(1):27–40. [doi: 10.1109/TMM.2016.2601020]
- [39] Thomaz CE, Giraldi GA. A new ranking method for principal components analysis and its application to face image analysis. Image and Vision Computing, 2010,28(6):902–913. [doi: 10.1016/j.imavis.2009.11.005]
- [40] Jiang J, Hu R, Wang Z, Han Z. Noise robust face hallucination via locality-constrained representation. IEEE Trans. on Multimedia, 2014,16(5):1268–81. [doi: 10.1109/TMM.2014.2311320]
- [41] Jiang J, Hu R, Wang Z, Han Z. Face super-resolution via multilayer locality-constrained iterative neighbor embedding and intermediate dictionary learning. IEEE Trans. on Image Processing, 2014,23(10):4220–4231. [doi: 10.1109/TIP.2014.2347201]
- [42] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2016,38(2):295–307. [doi: 10.1109/TPAMI.2015.2439281]
- [43] Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 5197–5206. [doi: 10.1109/CVPR.2015.7299156]
- [44] Kim J, Kwon LJ, Mu LK. Accurate image super-resolution using very deep convolutional networks. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 1646–1654. [doi: 10.1109/CVPR.2016.182]

附中文参考文献:

- [4] 王春峰,苏荔,张维刚,黄庆明.基于3D卷积神经网络的无参考视频质量评价.软件学报,2016,27(增刊(2)):103–112. <http://www.jos.org.cn/1000-9825/16025.htm>
- [5] 刘娜,李翠华.基于多层卷积神经网络学习的单帧图像超分辨率重建方法.中国科技论文,2015,10(2):201–206.
- [20] 苏衡,周杰,张志浩.超分辨率图像重建方法综述.自动化学报,2013,39(8):1202–1213.
- [32] 李民,陈建,乐翔,罗环敏.稀疏字典编码的超分辨率重建.软件学报,2012,23(5):1315–1324. <http://www.jos.org.cn/1000-9825/3989.htm> [doi: 10.3724/SP.J.1001.2012.03989]



呼延康(1992—),男,内蒙古太仆寺旗人,硕士生,主要研究领域为图像处理,3D人脸重建.



余乐天(1995—),男,学士,主要研究领域为图像处理,图像合成.



樊鑫(1977—),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为计算机视觉与图像处理,医学影像分析.



罗钟铉(1966—),男,博士,教授,博士生导师,CCF专业会员,主要研究领域为计算几何,图像处理,机器视觉.