



中国科技期刊卓越行动计划项目入选期刊

控制与决策

CONTROL AND DECISION



基于多级隐空间信息约束的噪声人脸超分辨率算法

滕韬, 于晓升, 吴成东

引用本文:

滕韬,于晓升,吴成东. 基于多级隐空间信息约束的噪声人脸超分辨率算法[J]. *控制与决策*, 2024, 39(5): 1469–1477.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1835>

您可能感兴趣的其他文章

Articles you may be interested in

人脸性别约束下的深度随机森林表情识别

Facial expression recognition using deep random forest under gender constraints

控制与决策. 2021, 36(7): 1693–1698 <https://doi.org/10.13195/j.kzyjc.2019.1703>

基于自注意力生成对抗网络的图像超分辨率重建

Image super-resolution reconstruction based on self-attention GAN

控制与决策. 2021, 36(6): 1324–1332 <https://doi.org/10.13195/j.kzyjc.2019.1290>

复杂背景下全景视频运动小目标检测算法

Panoramic video motion small target detection algorithm in complex background

控制与决策. 2021, 36(1): 249–256 <https://doi.org/10.13195/j.kzyjc.2019.0686>

基于混合模型驱动的红外与可见光图像融合

Infrared and visible image fusion based on hybrid model driving

控制与决策. 2021, 36(9): 2143–2151 <https://doi.org/10.13195/j.kzyjc.2019.1749>

基于条件对抗生成孪生网络的目标跟踪

Conditional generative adversarial siamese networks for object tracking

控制与决策. 2021, 36(5): 1110–1118 <https://doi.org/10.13195/j.kzyjc.2019.1215>

基于多级隐空间信息约束的噪声人脸超分辨率算法

滕 轺, 于晓升, 吴成东[†]

(东北大学 机器人科学与工程学院, 沈阳 110169)

摘要: 为了实现强噪声和模糊干扰下的低清人脸图像重建, 提出一种基于多级隐空间信息约束的噪声人脸超分辨率算法. 首先设计一个用于人脸有效信息提取的特征蒸馏网络, 并通过统计性抗干扰模型和隐空间特征对比算法移除噪声等无效信息, 构建一个具有高噪声鲁棒性的人脸信息提取模型; 然后, 设计人脸重建网络, 该网络利用提取的人脸特征重建高清人脸图像; 最后, 通过人脸身份嵌入模型和离散小波变换模型, 分别从超球面身份度量空间和小波域进一步对重建人脸的身份信息和空间结构进行约束. 实验结果表明, 所提出的算法不仅能够有效去除高噪声环境下的人脸噪声, 而且还能有效提升人脸图像分辨率, 获得更高的峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity index, SSIM), 具有较好的实用性.

关键词: 特征蒸馏; 隐空间信息约束; 图像超分辨率; 图像去噪; 深度神经网络; 超球面度量空间; 小波变换

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1835

引用格式: 滕轺, 于晓升, 吴成东. 基于多级隐空间信息约束的噪声人脸超分辨率算法[J]. 控制与决策, 2024, 39(5): 1469-1477.

Noisy face super-resolution method based on multi-level latent space information constraints

TENG Zi, YU Xiao-sheng, WU Cheng-dong[†]

(Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110169, China)

Abstract: To super-resolve low-resolution (LR) face image suffering from strong noise and fuzzy interference, we propose a noisy face super-resolution (FSR) method based on multi-level latent space information constraints. Firstly, we design a feature distillation network to extract effective facial information, which exploits a statistical anti-interference model and a latent contrast algorithm to remove invalid information such as noise. Then, we design a face reconstruction network, which utilizes the extracted face features to reconstruct high-resolution (HR) face images. Finally, we deploy a face identity embedding model and a discrete wavelet transform model to further supervise the reconstruction of identity information and spatial structure from the hypersphere identity metric space and wavelet domain respectively. The experimental results show that the proposed method not only removes the noise from face in the high noise environment, but also improves the resolution of the face image effectively, which obtains higher PSNR, SSIM, and good practicability.

Keywords: feature distillation; latent information constraint; image super-resolution; image denoising; deep neural network; hypersphere metric space; wavelet transform

0 引言

人脸图像超分辨率重建 (face super-resolution, FSR) 技术是计算机视觉领域很重要的一个研究分支, 主要是从同一场景下捕获的一幅或多幅低分辨率人脸图像 (low-resolution, LR) 中重建出高分辨率人脸图像 (high-resolution, HR). 该技术对于提高公安机

关破案效率、维持社会治安稳定具有重要意义^[1]. 许多学者已经开展了一系列面向人脸图像的超分辨技术的研究^[2], 早期的研究工作, 如插值方法^[3]、邻域嵌入^[4]和稀疏表示^[5]等, 虽然这些算法能够取得一定效果, 但重建图像中往往会出现振铃、模糊和虚化等问题, 严重限制了重建图像的质量.

收稿日期: 2022-10-24; 录用日期: 2023-01-29.

基金项目: 国家自然科学基金项目 (U20A20197, 61973063); 辽宁省重点研发计划项目 (2020JH2/10100040); 辽宁省自然科学基金项目 (2021-KF-12-01); 中国科学院光电信息处理重点实验室开放基金项目 (OEIP-O-202005).

[†]通讯作者. E-mail: wuchengdong@mail.neu.edu.cn.

随着深度学习方法逐渐成为计算机视觉领域主流的研究方向,神经网络在特征提取和非线性映射方面的成功应用,为解决人脸图像超分辨率问题提供了新的方案.现有基于深度学习的人脸超分辨率方法通过神经网络直接建立低分辨率图像到高分辨率图像之间的非线性映射,凭借其强大的信息表征能力、灵活的网络结构以及端到端高效且简洁的优化方式,在图像重建的质量和速度上均超越了经典的基于学习的方法.例如, Yu等^[6]利用生成对抗网络进行人脸超分辨率处理,提出一种对抗损失来判别生成人脸的真实性,并提出一种空间变换网络用于反卷积网络的补偿环节. Zhu等^[7]提出使用任务交替级联框架对未对齐的低分辨率人脸进行超分辨率处理,该网络有两个分支,其中一支卷积神经网络对面部元素进行定位,另一支负责恢复所对应位置的面部高频信息. Song等^[8]提出了一种两阶段的方法,首先由卷积神经网络生成面部组件,然后通过组件增强方法合成细粒度的面部结构. Cao等^[9]提出了基于注意感知机制的人脸SR算法,使用深度强化学习的方法找到参与人脸重建的面部补丁,通过充分利用图像上下文的相互依赖关系进行面部超分辨率处理. Ma等^[10]提出一种迭代式协作卷积神经网络,利用两个循环神经网络分别进行人脸图像重建和关键点预测,经过多次迭代提取更加准确的先验信息,从而逐步提高重建人脸图像质量. Wang等^[11]提出将生成式人脸先验应用于人脸超分辨率问题,并提出一种通道拆分式空间特征转换模型,对人脸重建特征与先验特征进行融合,从而得到具有高清细节和纹理的人脸图像.但由于基于生成式先验的方法只能恢复具有固定风格(FHQ数据集风格^[12])的人脸图像,为了提升视觉感受指标而生成与原人脸无关的细节特征,使生成结果无法较好地保留身份信息.此外,现有大多数的人脸超分辨率方法的研究均是在输入图像是无噪声的假设下进行的.在实际应用场景中,被噪声污染的输入图像会导致模型的性能急剧下降,重建的人脸图像会存在明显的身份信息混淆,无法满足实际应用需求.

针对现有方法无法解决具有强噪声、模糊的极低分辨率图像超分辨率问题,本文提出一种基于多级隐空间信息约束的噪声人脸超分辨率算法,其贡献可以归纳为3个方面.

1) 为解决现有方法无法有效移除强噪声、模糊等降质因子对人脸重建过程的干扰的问题,本文提出一种用于人脸有效信息提取的特征蒸馏网络.为了从输入图像中提取关键信息的同时忽略图像中的无

效细节(如噪声、模糊等信息),设计一种隐空间蒸馏损失对特征提取过程进行监督,使网络学习如何只提取与高清人脸图像特征相似的重要特征.为了使蒸馏网络在提取到核心特征的同时避免噪声和模糊信息的干扰,还设计一种统计性抗干扰模块,对提取特征添加随机高斯噪声,对特征蒸馏网络增加扰动以提高网络的抗干扰能力,使网络关注更为关键的信息,从而得到表征能力更强的隐层特征.

2) 为解决现有方法无法准确估计低清噪声人脸图像中的语义先验而导致身份信息丢失的问题,本文提出一种身份识别约束模型来减小重建人脸图像与高清人脸图像的身份差异.利用预训练好的身份信息嵌入模型对身份相关特征进行提取,并在超球面度量空间下对重建人脸身份向量从角度和幅值两个方面进行精确约束,从而重建出纹理细节真实且身份信息不变的高分辨率人脸图像.

3) 为解决现有方法无法在高噪声环境下有效重建人脸高频细节的问题,本文提出一种基于小波变换的高频特征约束模型.该方法可以有效捕捉图像信号中的空间信息和频率信息,通过将重建人脸图像和高清重建人脸解耦成不同频率的子带,并计算对应高频带的损失,对人脸高频细节的重建过程进行约束,从而重建出具有高质量高频细节纹理的人脸图像.定量对比结果表明,本文算法的重建结果在PSNR和SSIM方面的性能均显著优于目前的技术水平.

1 本文算法

1.1 模型框架

如图1所示,本文模型主要由4个部分组成,即特征蒸馏网络、人脸重建网络、身份信息嵌入模块和离散小波变换模块.

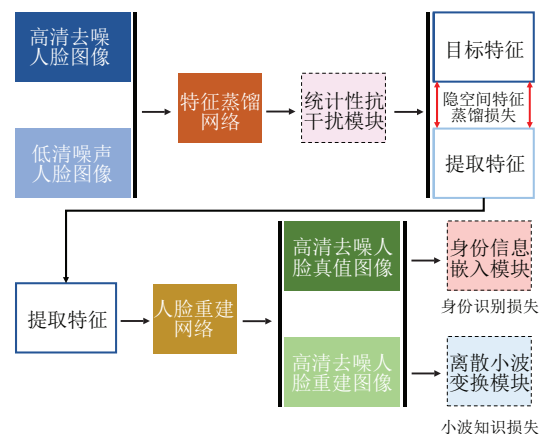


图1 本文算法

将低清噪声人脸图像的重建过程分为3个层次: 隐空间特征级重建、语义级重建和像素级重建。现有的方法大多只关注于像素级的人脸信息重建, 然而, 在低清噪声人脸图像中, 像素级人脸信息的准确重建十分艰难, 仅通过像素级信息约束容易导致重建人脸图像的语义失真。而本文提出的隐空间特征级重建、语义级重建和像素级重建三层次递进策略则是一种更加完备的人脸图像重建算法(如图2所示)。首先, 特征蒸馏网络通过设计统计性抗干扰模型和隐空间特征对比算法来对人脸图像中的噪声等无效信息进行移除, 从而实现噪声影响具有较高鲁棒性的隐空间特征级重建。然后, 身份信息嵌入模块进行身份识别约束来减小重建人脸图像与高清人脸图像的身份差异, 实现人脸图像的语义级重建。其中身份识别约束可以视为一种控制特征, 这些控制特征可用来指导基本面部特征的重组, 以重建出纹理细节真实且身份信息不变的高分辨率人脸图像。最后, 离散小波变换模块通过捕捉人脸图像信号中的空间信息和频率信息,

将重建人脸图像和高清真值人脸解耦成不同频率的子带, 并计算对应高频带的损失以实现人脸高频细节的像素级重建。

与现有方法直接学习低分辨率人脸到目标分辨率人脸的映射过程不同, 本文模型整体采用图像分辨率渐进提升的重建策略。每个阶段中, 均依次从隐空间特征级重建、语义级重建和像素级重建3个层次对低清噪声人脸图像进行复原, 实现当前阶段下的最优重建效果后, 送入到更高分辨率映射网络进一步重建人脸关键信息, 从而得到具有更加精细化纹理的人脸图像。

通过以上步骤, 本文模型可以实现完备的、由粗到精的人脸去噪和超分辨率重建过程, 从而显著地减少人脸重建图像中的噪声和伪影, 有效提升低清噪声人脸图像重建质量。在单一阶段网络中, 特征蒸馏网络与人脸重建网络由多个残差卷积块^[13]和残差注意力块^[14]构成(如图3所示), 两个子网络的对应层采用跨层连接方式, 从而最大化卷积层之间的信息流。

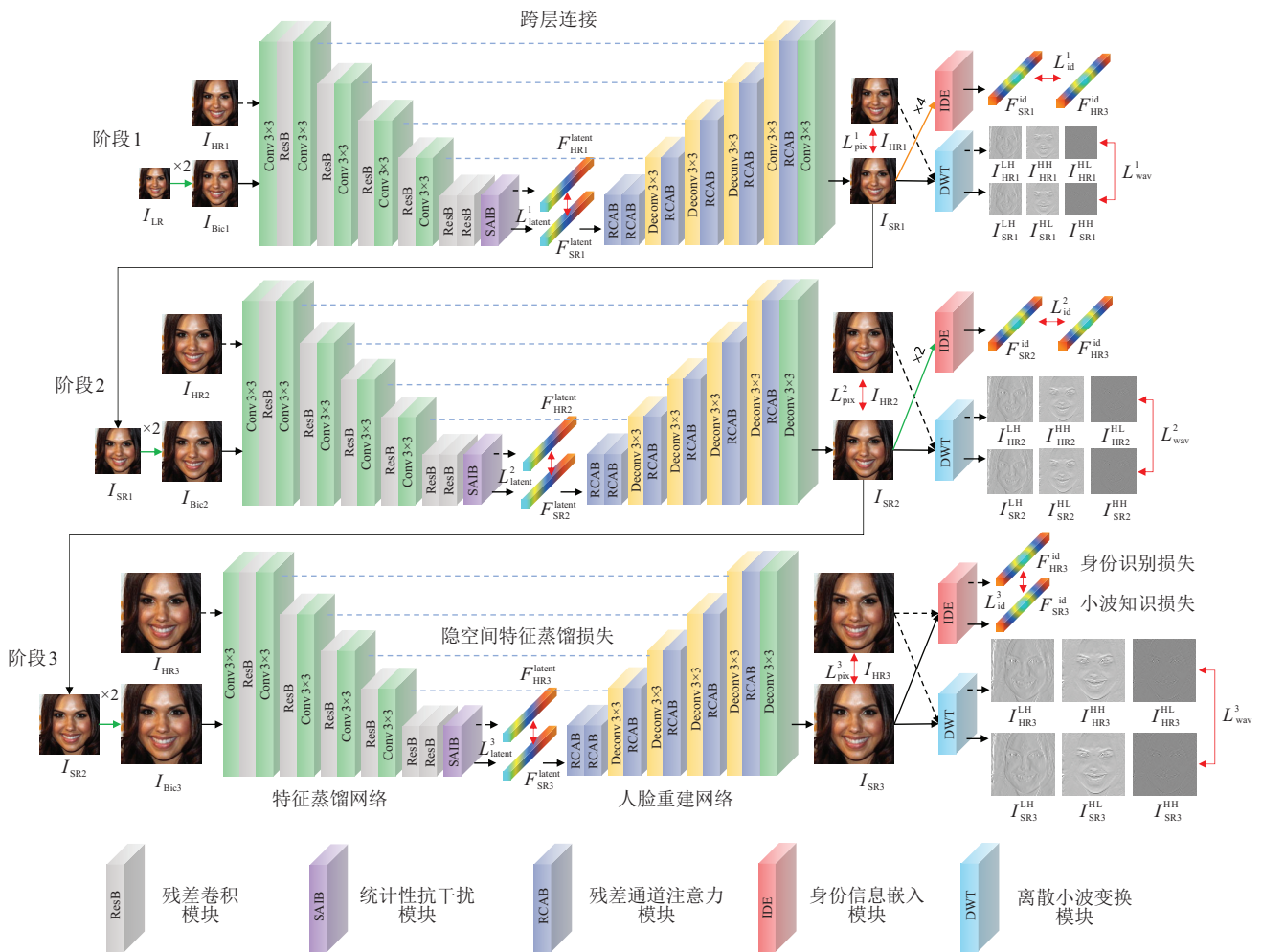


图2 基于多级隐空间信息约束的噪声人脸超分辨率算法框架示意

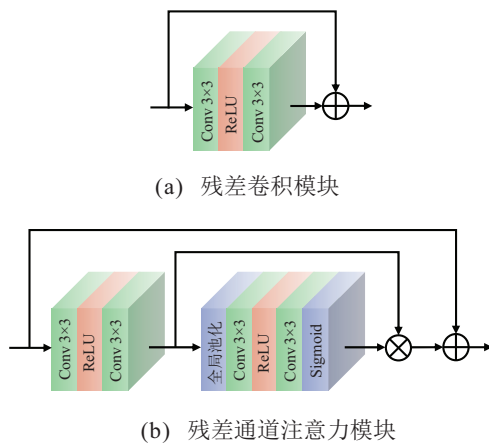


图3 残差单元结构

1.2 隐空间特征蒸馏

现有方法主要从像素层面对人脸信息的重建过程进行约束,然而,在低清噪声人脸图像中,像素级人脸信息的准确重建十分艰难,仅通过像素级信息约束容易导致重建人脸图像的语义失真.为解决这一问题,本文提出对低清人脸图像进行隐空间特征级约束.隐层特征约束是指在隐空间下将低清人脸图像的隐层特征与对应的高清人脸图像进行对齐.输入图像经过编码器进行多层网络特征提取精炼而来的隐空间特征,以远小于输入图像的特征维度,对面部轮廓、纹理、年龄、性别、肤色等特征进行编码,并作为控制特征在重建阶段对基本面部特征的恢复与重组进行指导.然而,现有方法通常仅在重建网络末端计算像素级损失,并利用反向传播对隐层特征重建进行约束,效率较低,导致训练过程耗时长,且噪声和模糊等降噪因子会进一步提高训练难度.与其不同,本文方法直接在隐空间计算损失并对隐层特征进行约束,有效提高了隐空间特征的重建精度和网络训练效率.与此同时,利用隐层特征损失在网络中间部分进行反向传播,可以更好地引导编码器网络的训练过程,提高网络特征提取能力.重建的隐层特征也可以更好地在解码器网络中表达,高效地指导解码器对面部特征进行恢复与重组.为了使蒸馏网络在提取到核心特征的同时避免噪声和模糊信息的干扰,网络中采用统计性抗干扰模块对提取特征添加均值为0、方差为1的随机高斯噪声 δ ,并通过两层卷积层 $g(\cdot)$ 处理,对特征蒸馏网络增加扰动以提高网络的抗干扰能力,使网络关注更为关键的信息,从而得到表征能力更强的隐层特征.基于统计性抗干扰模块的特征蒸馏损失 L_{latent}^i 定义如下:

$$F_t^{\text{latent}} = \text{SAIB}(\text{Dis}(\theta_1, I_t)) = g(\text{Dis}(\theta_1, I_t) + \delta), \quad (1)$$

$$L_{\text{latent}}^i = \|F_{\text{HR}i}^{\text{latent}} - F_{\text{SR}i}^{\text{latent}}\|, \quad i \in (1, 2, 3). \quad (2)$$

其中: θ_1 为蒸馏网络 $\text{Dis}(\cdot)$ 的参数, F_t^{latent} 表示蒸馏特征 $\text{Dis}(\theta_1, I_t)$ 经统计性抗干扰模块 $\text{SAIB}(\cdot)$ 运算得到的隐层特征, t 表示输入图像 I 的类别为高清人脸或低清噪声人脸,表示当前网络所处阶段.

现有的特征蒸馏算法通常是将模型知识从一个结构复杂、参数量大的模型迁移到另一个具有简单结构的模型,从而实现模型的轻量化处理^[15].不同的是,本文提出的隐空间特征蒸馏方法则是利用单一网络学习如何只从输入图像中提取重要信息,通过隐层特征约束的方式发掘网络自身的特征表达和提取能力,使网络关注更为关键的信息,从而得到表征能力更强的隐层特征,而不依赖具有较高参数量的复杂模型的监督,从而加速模型训练过程,降低模型训练难度,实现了在保持较小参数量和模型复杂度的同时提升模型的特征提取能力和图像重建质量.

1.3 人脸身份嵌入

人脸超分辨率作为图像超分辨率领域中的特定应用问题,通常可采用人脸先验信息对重建过程进行指导,例如面部组件热力图^[16]、带有关键点和解析图的面部形状^[17]、人脸小波系数^[18]以及人脸局部结构先验信息^[19],并取得了较好的超分辨率重建效果.但在高噪声环境中,低清噪声人脸图像中的语义先验信息难以被准确估计,因此,这类方法的重建性能也会出现明显下降.与直接使用人脸语义先验信息的方法不同,本文利用人脸的身份先验信息来指导人脸图像超分辨率过程,以重建出纹理细节真实且身份信息不变的高分辨率人脸图像.

为了准确恢复身份相关的人脸细节,本文提出一种身份识别约束模型来减小重建人脸图像与高清人脸图像的身份差异.首先,由于超球面空间在人脸身份表示方面的性能优异,本文将其作为身份度量空间,并利用预训练好的身份信息嵌入模型 LightCNN_v9 ^[20]对身份相关特征进行提取.然后,采用欧几里得正则化操作,将身份特征映射到超球面空间用于计算身份损失.如图4所示,在超球面身份度量空间中,低分辨率噪声人脸特征与高清去噪人脸特征在角度和幅值两方面存在着明显差异;并且伴随着分辨率的降低,低-高清人脸特征对之间的平均角度差异和幅值差异会逐渐加大.这说明角度和幅值这两个指标可以反映深度表征特征的退化程度,并以此来预测人脸重建图像的质量.因此,本文提出一种基于特征解构的身份识别损失,通过计算超球面度量空间下重建人脸身份向量与高清人脸身份向量在角度和幅值两个方面的损失 (L_a^i, L_m^i) ,求和后得到身份

识别损失 L_{id}^i 如下:

$$L_a^i = 1 - \frac{(F_{SRi}^{id})^T (F_{HRi}^{id})}{\|F_{SRi}^{id}\|_2 \|F_{HRi}^{id}\|_2}, \quad i \in (1, 2, 3); \quad (3)$$

$$L_m^i = \left| \|F_{SRi}^{id}\|_2 - \|F_{HRi}^{id}\|_2 \right|; \quad (4)$$

$$L_{id}^i = L_a^i + L_m^i. \quad (5)$$

其中: F_{SRi}^{id} 和 F_{HRi}^{id} 表示重建人脸 I_{SRi} 和高清人脸 F_{HRi} 经身份信息嵌入模型 LightCNN_v9 运算输出的身份信息特征向量, $\|\cdot\|_2$ 为 L_2 正则化操作, i 表示当前网络所处阶段.

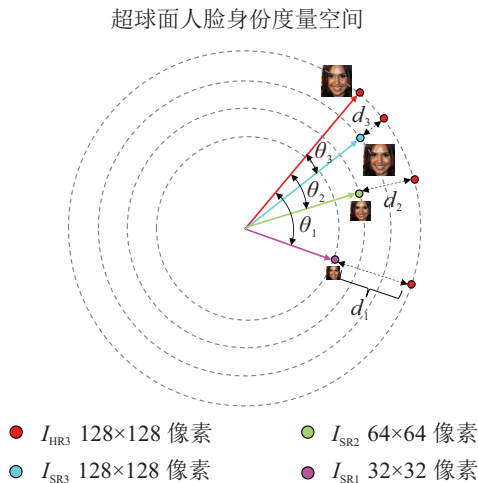


图4 超球面度量空间示意

通过在超球面度量空间下对重建人脸身份向量从角度和幅值两个方面进行精确约束, 本文模型可重建出纹理细节真实且身份信息不变的高分辨率人脸图像.

1.4 小波知识蒸馏

现有超分辨率模型通常可以准确重建人脸图像的低频细节, 但往往无法有效重建人脸区域的高频细节, 且这一现象在高噪声环境下表现得更为明显. 为解决这一问题, 本文提出利用离散小波变换方法将重建人脸图像和高清重建人脸解耦成不同频率的子带, 并计算对应高频带的 L_1 损失, 对人脸高频细节的重建过程进行约束.

离散小波变换方法作为一种数学工具常用于金字塔图像解耦. 在本文中, 重建人脸图像和高清人脸图像被解耦成4个子频带, 分别是 LL、LH、HL、HH. 其中: LL 表示低频子带, 其余为高频子带. 如图5所示, 将离散小波变换定义为 $\Psi(\cdot)$, 则图像 I 的高低子频带图像可分别表示为 $\Psi^H(I)$ 和 $\Psi^L(I)$. $\Psi^H(I) = I^{LH}, I^{LH}, I^{HH}$, 则小波知识损失 L_{wav}^i 定义为

$$L_{wav}^i = \|\Psi^H(I_{HRi}) - \Psi^H(I_{SRi})\|_1, \quad i \in (1, 2, 3). \quad (6)$$

其中: I_{SRi} 和 F_{HRi} 分别表示重建人脸和高清人脸图

像, i 表示当前网络所处阶段. 相较于其他频率分析方法 (例如傅里叶变换), 小波变换方法能更有效地捕捉图像信号中的空间信息和频率信息^[21], 从而使得本文模型能够重建具有高质量高频细节纹理的人脸图像.

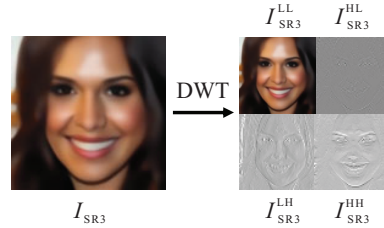


图5 离散小波变换示意

2 实验研究

2.1 实验准备

本文使用 CelebA 数据集^[22], 共选取 40 000 幅人脸图像进行训练, 并使用接下来的 1 000 幅人脸图像作为测试集. 同时, 对裁剪后的人脸区域不进行任何预对准操作, 裁剪后图像大小为 128×128 . 为验证本文算法在噪声和模糊干扰下仍能重建清晰人脸图像, 实验采用 3 种退化模型来模拟低分辨率输入图像. 在下采样部分: 本文采用双三次下采样算法 (Bicubic) 制作比例因子为 8 的低分辨率图像 (简称 Bic); 然后, 在 8 倍下采样图像中加入噪声等级为 15 的高斯噪声, 得到带噪声的低清人脸图像 (简称 BicN); 最后, 为制作同时受噪声和模糊因素影响的降质人脸图像, 实验先对高清人脸进行标准差为 1.5、大小为 7×7 的高斯核模糊, 再对图像进行比例系数为 8 的双三次下采样, 并加入噪声级别为 30 的高斯噪声 (简称 BBicN).

在模型权重初始化方面, 考虑到固定方差的高斯分布初始化导致深层次网络训练不稳定和收敛难的问题, 采 He 等^[23] 的方法对卷积层进行初始化. 在训练参数的设置方面, 使用 Adam (adaptive moment estimation) 算法作为损失函数的优化器, 它利用梯度的一阶矩估计和二阶矩估计动态调整每个参数的学习率, 实验中, 梯度一、二阶矩估计的系数分别为 0.9 和 0.999. 与其他自适应学习率优化算法相比, 其收敛速度更快, 有效避免了高方差参数导致的波动, 使训练过程更加平稳. 批量数据大小为 16, 模型在 Titan X GPU 上进行训练.

为了评估 SR 结果的质量, 本文采用两个客观的图像质量评估指标: 峰值信噪比 (peak signal to noise ratio, PSNR) 和结构相似度 (structural similarity, SSIM). 所有的度量都在超分辨图像的 y 通道 (YCbCr 颜色空间) 上执行.

2.2 消融实验

为了表明本文提出的隐空间特征蒸馏、人脸身份嵌入和小波知识蒸馏3个部分是如何提高网络重建性能的,进一步,对上述3个模块对噪声人脸超分辨过程的影响进行消融研究.实验结果如图6所示,进行4组实验来重建人脸图像.

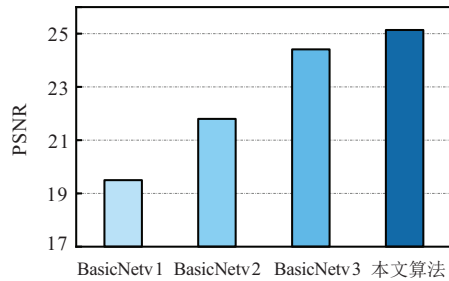


图6 消融实验结果

首先,在本文提出的标准模型中删除隐空间特征蒸馏、人脸身份嵌入和小波知识蒸馏模块,其余部分构成第1个验证网络“BasicNetv1”;然后,在验证网络“BasicNetv1”中加入隐空间特征蒸馏模块,得到网络“BasicNetv2”,用于观测隐层特征约束对模型性能的加持;同样,通过在网络“BasicNetv2”中加入人脸身份嵌入模块,可以得到网络“BasicNetv3”,进一步观测人脸身份嵌入是否能够有效提高模型重建性能;最后,在网络“BasicNetv3”中加入小波知识蒸馏模块,得到本文提出的完备的基于多级隐空间信息约束的噪声人脸超分辨率算法.图6显示了上述关于各模块的消融实验结果.结果表明:1)与无隐层特征约

束的“BasicNetv1”模型相比,隐空间蒸馏损失对人脸特征提取过程的监督可有效提升“BasicNetv2”模型性能,这说明隐空间特征蒸馏方法通过隐层特征约束的方式发掘网络自身的特征表达和提取能力,使网络关注更为关键的人脸信息,对人脸超分辨过程有积极的影响;2)与仅有隐层特征约束的“BasicNetv2”模型相比,利用人脸的身份先验信息来指导人脸图像超分辨率过程的网络“BasicNetv3”可以在超球面度量空间下对重建人脸身份向量从角度和幅值两个方面进行精确约束,并作为语义信号指导基本面部特征的重组,更加显著地提高人脸图像超分辨率性能;3)将小波知识蒸馏与隐空间特征蒸馏、人脸身份嵌入相结合能够取得最优效果,说明小波变换方法可以有效捕捉图像信号中的空间信息和频率信息,提升超分辨率模型性能.消融实验中,4个模型性能的逐步提升也说明,从隐空间特征级重建、语义级重建和像素级重建3个层次对低清噪声人脸图像进行复原,可以实现更加完备的人脸去噪和超分辨率重建过程,有效提升低清噪声人脸图像重建质量.

2.3 实验结果分析

2.3.1 定量对比结果

为了评估本文算法的性能,选取目前最先进的人脸超分辨率算法PaniniNet^[24]、DIC^[10]、GPEN^[25]、GFPGAN^[11]与本文算法对不同降质过程下的CelebA测试集进行对比测试.表1展示了本文算法与现有方法在CelebA数据集上的定量评估结果.

表1 本文算法与现有方法在不同退化模型下的客观评价对比结果

method	Bic		BicN		BBicN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Panini-Net	23.56	0.6248	22.67	0.6013	21.70	0.5820
DIC	26.23	0.7441	24.15	0.6889	22.70	0.6033
GPEN	25.88	0.7339	24.49	0.6925	23.14	0.6162
GFPGAN	25.91	0.7390	24.85	0.7047	23.51	0.6483
本文算法	26.87	0.7786	25.14	0.7194	24.29	0.6856

得益于本文提出的多级隐空间信息约束算法,所设计的统计性抗干扰模型和隐空间特征对比算法成功地移除了人脸图像中的噪声和模糊等无效信息,从而有效地构建了高噪声鲁棒性人脸信息提取模型.同时,通过人脸身份嵌入模型和离散小波变换模型,分别从超球面身份度量空间和小波域进一步对重建人脸的身份信息和空间结构进行约束,使本文模型能够重建出纹理细节真实且身份信息不变的高分辨率人脸图像.定量对比结果表明,本文算法的重建结果在PSNR和SSIM方面的性能均显著优于目前的技术水平.

2.3.2 定性对比结果

为了进一步评估本文算法重建的视觉效果,实验在不同降质过程下将本文算法与现有方法进行了重建人脸图像的主观视觉对比.如图7所示,在无噪声模糊影响下,现有方法均可重建出较为理想的人脸细节,并在一定程度上保留了身份信息.其中,由于DIC具有迭代式协作的网络结构,利用两个循环神经网络分别进行人脸图像重建和关键点预测,经过多次迭代提取更加准确的先验信息,从而取得了较好的人脸重建质量.

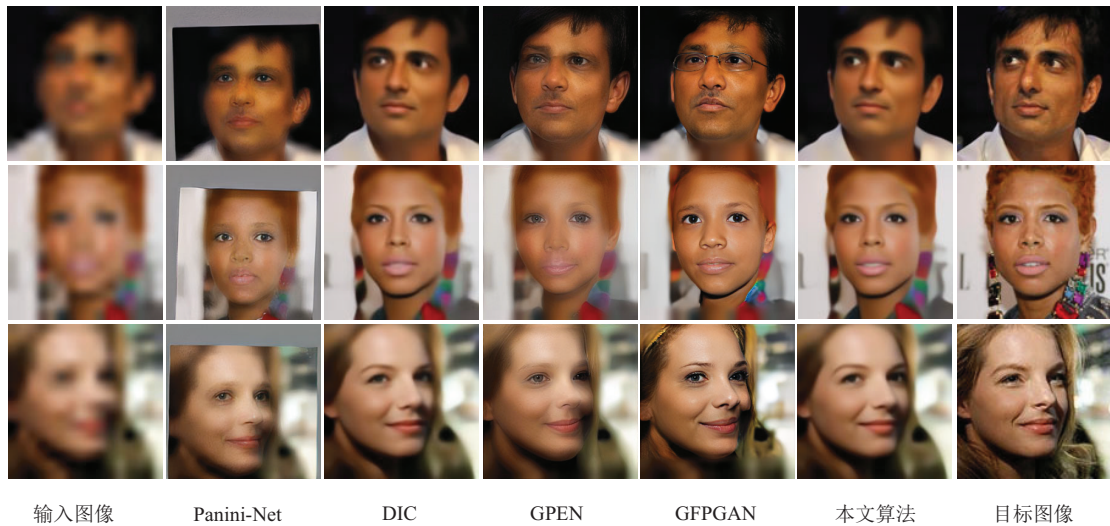


图7 本文算法与现有方法在Bic退化模型上的主观对比结果

然而在噪声影响下(BicN), 现有方法的性能均出现了不同程度的下降, 而本文算法仍然可以重建出较清晰的人脸图像. 如图8所示, 本文算法重建的人脸图像有着更为清晰的面部纹理细节, 并较好地保留了身份信息. 虽然GFPGAN通过将生成式人脸先验与人脸重建特征进行融合, 重建出具有高频细节纹理的

人脸图像, 但由于基于生成式先验的方法只能恢复具有固定风格的人脸图像, 为了提升视觉感受指标而生成与原人脸无关的细节特征, 使得生成结果无法保留身份信息. 因此可以明显看出, 尽管降质过程逐渐复杂, 本文算法重建的人脸图像的视觉质量并没有出现明显的降低, 具有较好的实用性.

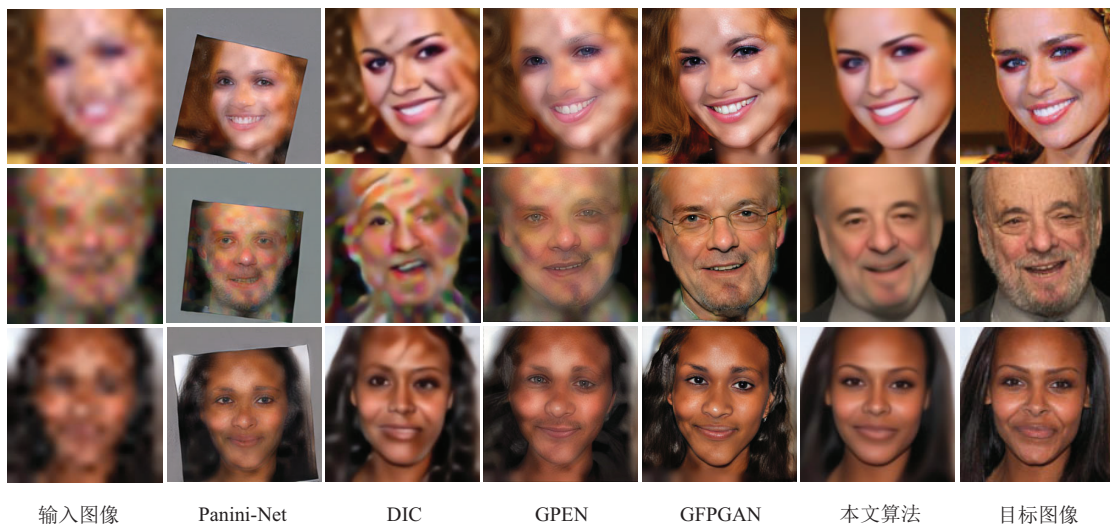


图8 本文算法与现有方法在BicN退化模型上的主观对比结果

如图9所示, 在引入模糊降质因子后(BBicN), 被污染的低清人脸输入图像会导致各模型的性能急剧下降, 重建的人脸图像会存在明显的身份信息混淆, 无法满足实际应用需求. 而本文模型通过由粗到精的人脸去噪和超分辨率重建, 利用隐空间蒸馏损失对特征提取过程进行监督, 使网络学习如何只提取与高清人脸图像特征相似的核心特征, 避免了噪声和模糊信息的干扰, 显著地减少了人脸重建图像中的噪声和伪影. 同时, 由于本文模型在超球面度量空间下对重

建人脸身份向量从角度和幅值两个方面进行精确约束, 成功重建了纹理细节真实且身份信息不变的高分辨率人脸图像. 此外, 本文模型中的离散小波变换方法有效地捕捉到图像信号中的空间信息和频率信息, 从而使本文模型可以重建具有高质量高频细节纹理的人脸图像. 可以明显看出, 尽管降质过程逐渐复杂, 本文算法重建的人脸图像的视觉质量并没有出现明显的降低, 因此具有较好的实用性.

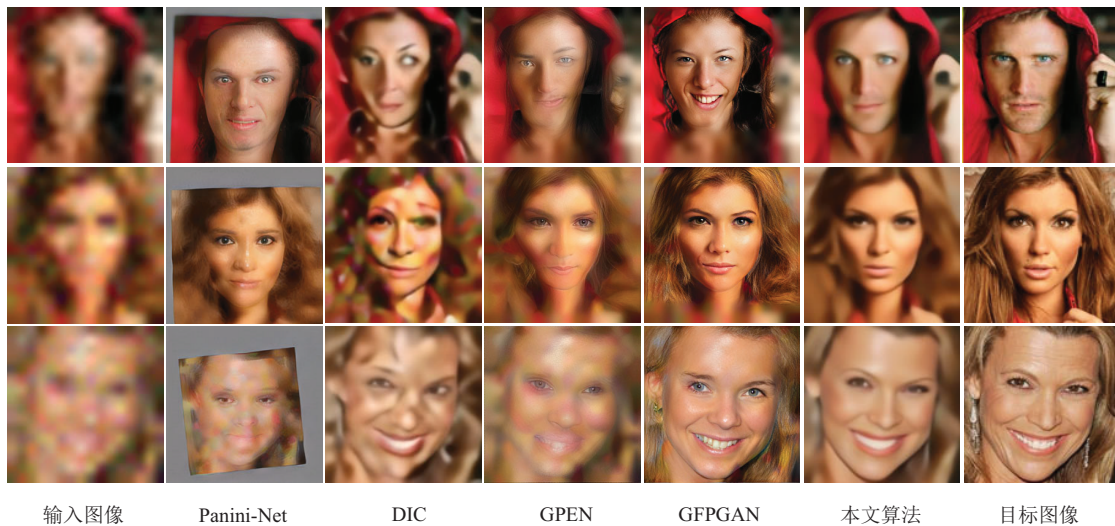


图9 本文算法与现有方法在BBicN退化模型上的主观对比结果

2.3.3 计算复杂度分析

如表2所示,在NVIDIA RTX 2080Ti GPU上,将本文算法与另外4种超分辨率网络(Panini-Net、DIC、GPEN和GFPGAN)的模型参数和计算速度进行比较.因本文模型仅在训练阶段从隐空间特征级重建、语义级重建和像素级重建3个层次对模型的学习过程进行约束,而在模型的部署测试阶段,隐空间特征蒸馏、人脸身份嵌入和小波知识蒸馏模块均不必调用,仅利用编码-解码器主干网络对低清噪声人脸图像进行复原,故在实现减少模型参数的同时,提高了推理速度,能够更好地应用于现实场景.计算结果显示,本文模型能以更低的模型参数和更快的运行速度获得更好的性能.

表2 本文算法与现有方法在Bic退化模型上的时效性对比

method	model size/m	running time/ms	PSNR/dB
Panini-Net	34.76	102.55	23.56
DIC	21.77	83.61	26.23
GPEN	28.10	98.45	25.88
GFPGAN	31.39	85.09	25.91
本文算法	18.43	69.82	26.87

3 结论

本文提出了一种基于多级隐空间信息约束的噪声人脸超分辨率算法.通过特征蒸馏网络对人脸有效信息进行提取,并通过统计性抗干扰模型和隐空间特征对比算法移除噪声等无效信息,构建了一个具有高噪声鲁棒性的人脸信息提取模型.随后,人脸重建网络利用提取到的人脸特征进行高清人脸图像重建.同时,通过人脸身份嵌入模型和离散小波变换模

型,分别从超球面身份度量空间和小波域进一步对重建人脸的身份信息和空间结构进行约束,使本文模型可以重建出纹理细节真实且身份信息不变的高分辨率人脸图像.在公开人脸数据集上进行对比测试的结果表明,本文算法相比于主流算法,在评价性能指标和视觉效果上都获得了更好的效果.

参考文献(References)

- [1] 王雪松,晁杰,程玉虎.基于自注意力生成对抗网络的图像超分辨率重建[J].控制与决策,2021,36(6):1324-1332.
(Wang X S, Chao J, Cheng Y H. Image super-resolution reconstruction based on self-attention GAN[J]. Control and Decision, 2021, 36(6): 1324-1332.)
- [2] Liu A R, Liu Y H, Gu J J, et al. Blind image super-resolution: A survey and beyond[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 5461-5480.
- [3] Fritsch F N, Carlson R E. Monotone piecewise cubic interpolation[J]. SIAM Journal on Numerical Analysis, 1980, 17(2): 238-246.
- [4] 吴成东,卢紫微,于晓升.基于加权随机森林的图像超分辨率算法研究[J].控制与决策,2019,34(10):2243-2248.
(Wu C D, Lu Z W, Yu X S. Image super resolution reconstruction algorithm based on weighted random forest[J]. Control and Decision, 2019, 34(10): 2243-2248.)
- [5] Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.

- [6] Yu X, Porikli F. Ultra-resolving face images by discriminative generative networks[C]. European Conference on Computer Vision. Cham: Springer, 2016: 318-333.
- [7] Zhu S Z, Liu S F, Loy C C, et al. Deep cascaded Bi-network for face hallucination[C]. European Conference on Computer Vision. Cham: Springer, 2016: 614-630.
- [8] Song Y B, Zhang J W, He S F, et al. Learning to hallucinate face images via component generation and enhancement[J/OL]. 2017, arXiv: 1708.00223.
- [9] Cao Q X, Lin L, Shi Y K, et al. Attention-aware face hallucination via deep reinforcement learning[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 1656-1664.
- [10] Ma C, Jiang Z Y, Rao Y M, et al. Deep face super-resolution with iterative collaboration between attentive recovery and landmark estimation[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 5568-5577.
- [11] Wang X T, Li Y, Zhang H L, et al. Towards real-world blind face restoration with generative facial prior[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, 2021: 9164-9174.
- [12] arnas T, Laine S, Aila T M. A style-based generator architecture for generative adversarial networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 4396-4405.
- [13] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 770-778.
- [14] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]. European Conference on Computer Vision. Cham: Springer, 2018: 294-310.
- [15] Gou J P, Yu B S, Maybank S J, et al. Knowledge distillation: A survey[J]. International Journal of Computer Vision, 2021, 129(6): 1789-1819.
- [16] Yu X, Fernando B, Ghanem B, et al. Face super-resolution guided by facial component heatmaps[C]. European Conference on Computer Vision. Cham: Springer, 2018: 219-235.
- [17] Chen Y, Tai Y, Liu X M, et al. FSRNet: End-to-end learning face super-resolution with facial priors[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 2492-2501.
- [18] Huang H B, He R, Sun Z N, et al. Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution[C]. 2017 IEEE International Conference on Computer Vision. Venice, 2017: 1698-1706.
- [19] Jiang J J, Chen C, Ma J Y, et al. SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior[J]. IEEE Transactions on Multimedia, 2017, 19(1): 27-40.
- [20] Wu X, He R, Sun Z N, et al. A light CNN for deep face representation with noisy labels[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(11): 2884-2896.
- [21] Mallat S. Wavelet bases[C]. A Wavelet Tour of Signal Processing. Amsterdam: Elsevier, 1999: 220-320.
- [22] Liu Z W, Luo P, Wang X G, et al. Deep learning face attributes in the wild[C]. IEEE International Conference on Computer Vision. Santiago, 2016: 3730-3738.
- [23] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification[C]. IEEE International Conference on Computer Vision. Santiago, 2016: 1026-1034.
- [24] Wang Y H, Hu Y J, Zhang J. Panini-net: GAN prior based degradation-aware feature interpolation for face restoration[J/OL]. 2022, arXiv: 2203.08444.
- [25] Yang T, Ren P R, Xie X S, et al. GAN prior embedded network for blind face restoration in the wild[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, 2021: 672-681.

作者简介

滕轺(1987-), 男, 博士生, 从事图像处理、图像超分辨率重建等研究, E-mail: tengzi@stu.neu.edu.cn;

于晓升(1984-), 男, 副教授, 博士, 从事人工智能与模式识别等研究, E-mail: yuxiaosheng@mail.neu.edu.cn;

吴成东(1960-), 男, 教授, 博士生导师, 从事图像智能处理、机器人控制等研究, E-mail: wuchengdong@mail.neu.edu.cn.