

分类号_____

密 级_____

UDC _____

编 号_____

武 汉 大 学
博 士 学 位 论 文

基于一致流形学习的人脸超分辨率算法
研究

研 究 生 姓 名： 江俊君

指导教师姓名、职称： 胡瑞敏、教授

学 科 、 专 业 名 称： 通信与信息系统

研 究 方 向： 图像处理与分析

二零一四年十二月

A Dissertation Submitted to
Academic Degrees Evaluation Committee of Wuhan University for
The Degree of Doctor of Philosophy in Engineering

Face Super-resolution via Consistent Manifold Learning

Ph.D Candidate: Jiang Junjun

Supervisor : Hu Ruimin

Major : Communication and
Information System and

Wuhan University

Wuhan, Hubei, P.R.China

Dec 2014

论文创新点

论文针对基于学习的人脸超分辨率方法在实际监控应用中面临的三大挑战：有限高维数据样本构成的高维空间不具有局部线性性质、流形空间易受噪声影响和高低分辨率图像空间流形几何结构不一致，开展了人脸图像库的表达理论、人脸图像的视觉表示理论、高低分辨率的流形保持学习理论的研究，创新性地解决了基于学习的人脸超分辨率技术中的多个关键问题，其主要创新点如下：

(1) 提出基于最近特征线流形学习的人脸超分辨率方法，实现从基于样本点的图像表示到基于特征线的图像扩充表示的转变。本文针对实际人脸库样本不足导致样本空间局部线性假设不成立的问题，引入最近特征线思想，提出基于最近特征线投影样本点的近邻查询与保持方法，保持相对传统流形学习方法更小局部内的线性关系，更加体现了流形学习的本质。

(2) 提出基于图像块局部约束表示的噪声鲁棒人脸超分辨率方法，实现人脸图像块从稀疏表示到局部约束表示的转变。本文针对前沿稀疏表示人脸超分辨率方法对噪声图像块表示不稳定的问题，分析了噪声对人脸图像块表示的影响，建立了图像块局部约束表示模型，赋予相似图像块更大的权重系数，使得图像块的表示对噪声更加鲁棒。

(3) 提出基于高分辨率流形空间约束的人脸超分辨率方法，实现从保持低分辨率流形结构到高分辨率流形空间约束的一致流形学习的转变。本文针对现有方法只考虑低分辨率流形结构而忽略原始高分辨率流形结构的问题，建立高分辨率流形正则约束的高低分辨率空间映射模型，使重构后的高分辨率流形结构与原始高分辨率流形结构一致。

(4) 提出基于局部迭代嵌入和字典渐进更新的人脸超分辨率方法，实现从单一层次流形保持学习到多层次流形保持学习的转变。本文针对基于学习的人脸超分辨率方法在复杂降质环境下性能急剧下降的问题，首次将低分辨率人脸字典更新的思想引入到超分辨率重构中，建立局部迭代嵌入模型和低分辨率人脸字典渐进更新模型，在渐近更新后的高低分辨率空间学习它们之间的关系，从而进行更加准确地邻域嵌入学习与预测。

摘要

在刑侦破案中，犯罪嫌疑人的人脸图像是刑侦人员最关注的目标之一。但是在实际监控中，由于嫌疑人距离摄像头较远且拍摄设备本身的分辨率有限，导致拍摄得到的人脸图像尺寸很小，难以辨识。如何提高所获人脸图像的空间分辨率，改善人脸图像质量、提升人脸图像清晰度和辨识价值是极具挑战性的课题。人脸超分辨率技术是解决上述问题较经济且有效的手段之一，它可以根据观察到的低质量、低分辨率人脸图像，结合由人脸图像样本库学习得到的先验知识重构出清晰的高分辨率人脸图像。

近年来，学者们提出了很多基于学习的人脸超分辨率算法，其基本假设是高分辨率图像（块）构成的流形空间与低分辨率图像（块）构成的流形空间具有局部几何结构一致性，学习低分辨率图像与高分辨率图像的关系，并利用此关系预测输入低分辨率人脸图像对应的高分辨率人脸图像。但是，上述基于流形一致性假设的人脸超分辨率方法在实际监控应用中仍然面临以下三大挑战：（1）实际应用中的人脸库规模有限，有限样本点只能构成高维人脸流形空间的一个稀疏子空间，即使最相邻的样本点也难以称之为局部，局部线性假设不成立；（2）大案、要案多发生在夜间，嫌疑目标人脸图像受噪声干扰严重，现有人脸图像块表示方法在噪声干扰下的表示不稳定；（3）图像降质带来的信息丢失导致高低分辨率图像之间存在“多对一”的不确定性关系，高低分辨率图像空间流形几何结构不一致。为此，本文开展基于一致流形学习的人脸超分辨率研究，在人脸图像库表达理论、人脸图像的视觉表示理论、高低分辨率的一致流形学习理论等方面取得了创新性成果：

（1）基于最近特征线流形学习的样本表达能力扩充技术

本文针对实际人脸库样本不足导致样本空间局部线性假设不成立的问题，引入最近特征线思想，连接任意两个样本图像（块）形成特征线，将人脸图像库中的任意两个样本图像（块）扩充为特征线上的无数多个样本，通过保持相对传统基于局部线性嵌入流形学习方法更小局部内的线性关系，来揭示高低分辨率人脸流形空间的非线性关系，更加体现了流形学习的本质。实验结果表明，提出方法相较于基于流形学习的前沿人脸超分辨率算法^[3]在CAS-PEAL-R1人脸库^[4]上的PSNR值提升了1个dB以上。

（2）基于局部约束表示的噪声图像块鲁棒表示技术

本文针对基于稀疏表示的前沿人脸超分辨率方法对噪声图像块表示不稳定的问题，分析了噪声对图像块表示的影响，建立了图像块局部约束表示模型，在图像块样本空间中自适应地选取与输入图像块距离近的样本图像块，惩罚距离远的样本图像块，利用局部性来约束图像块的表示系数，使得图像块的表示对噪声更加鲁棒。在FEI人脸库^[5]上的实验结果表明，提出方法相对于稀疏表示方法^[2]在PSNR和SSIM^[6]值上分别提升了1.02 dB和0.0168，重构

人脸图像的主观结果也更优。

(3) 基于高分辨率流形空间约束的高低分辨率图像空间映射技术

本文针对现有基于流形学习的人脸超分辨率方法只考虑低分辨率流形结构而忽略原始高分辨率流形结构的问题, 建立高分辨率流形正则约束的高低分辨率空间映射模型, 利用不受下采样和噪声影响的原始高分辨率流形结构作为先验约束, 约束重构后的高分辨率流形空间, 使重构后的高分辨率流形结构与原始高分辨率流形结构一致。相较于基于低分辨率图像流形约束的方法, 提出方法在 CAS-PEAL-R1 人脸库^[4]和 AR 人脸库^[7]上的 PSNR 值分别提升了 0.34 dB 和 0.27 dB, SSIM 值分别提升了 0.0044 和 0.0049。

(4) 基于局部迭代嵌入和字典渐进更新的图像观测模型学习与逼近技术

本文针对基于学习的人脸超分辨率方法在复杂降质环境下性能急剧下降的问题, 首次将低分辨率人脸字典更新的思想引入到人脸超分辨率重构中, 建立局部迭代嵌入模型和低分辨率人脸字典渐进更新模型, 渐近地学习得到多层中间低分辨率人脸字典, 逐渐减小高低分辨率图像空间的维度差, 在渐近更新后的高低分辨率空间学习它们之间的关系, 从而进行更加准确地邻域嵌入学习与预测。在 CAS-PEAL-R1 人脸库^[7]的实验结果表明, 提出方法相对于前沿稀疏表示方法^[8]在 PSNR 值和 SSIM 值上分别提升了 1.66 dB 和 0.0303。

综上所述, 本文通过借鉴人类视觉感知和认知的机理来完成人脸图像库的扩充表达、人脸图像的噪声鲁棒表示、高低分辨率一致流形关系的学习。在基础理论和关键技术方面的研究成果可为解决实际监控人脸超分辨率重构问题提供了新的途径。

关键字: 视频监控, 人脸超分辨率, 流形学习, 近邻嵌入, 字典学习

Abstract

In the detection of the criminal investigation, the face image of the suspect is the most concerned goals for investigators. But in the actual monitoring, the suspect is far from the camera and the resolution of the surveillance camera is so limited, which lead to the size of the captured face image is small. Thus, it is difficult to identify the suspect. How to improve the spatial resolution and quality of the captured face, and enhance image clarity and discriminability of face value is a challenging field. Face super-resolution is the technology aiming at solving the problem of inducing one high-quality and high-resolution (HR) face image from the given low-quality and low-resolution (LR) face images by combining with complementary information from consequence image frames or learning the prior knowledge from the given image database.

In recent years, a number of learning based face image super-resolution approaches has been proposed. Based on the assumption that the LR and HR manifold spaces share the similar local structure, these learning based face image super-resolution approaches try to learn the relationship between the LR and HR training sets, and then use the learned model to infer the HR face image of the input LR image. But face image super-resolution technology still faces the following three challenges especially in the surveillance conditions: (i) the high-dimensional space consisted by the high-dimensional data is not linear. The size of current face database is limited; (ii) the manifold space is susceptible to noise. In serious case, the face image of suspect often severely affected by noise; (iii) the local structure of the low and high resolution image manifold spaces is inconsistent. In this thesis, we study the problem of enhancing the representation ability of the given database, designing a robust image patch representation, learning the neighbor embedding model with consistent coupled manifolds, and the main contribution of the paper is highlighted in the following:

(1) Face database representation expansion through nearest feature line manifold learning.

To solve the inconsistency between the LR and HR manifold spaces caused by the limited sample in the face database, we introduce the concept of nearest feature line, and it can extend every two sample points of the training set to numberless points on the feature line by connecting the two points, which improves the expression ability of the original samples. Our algorithm maintains linear relationship in a smaller local space than traditional manifold learning based methods, and reveals the non-linear relationship between HR and LR face manifold space. This reflects the nature of manifold learning theory. Experimental results show that the gain of the proposed face super-resolution algorithm over [3] is 1dB in term of PSNR in the CAS-PEAL-R1 face database [4].

(2) Noise robust patch representation via Locality-Constrained Representation.

The noise will reduce the accuracy and stability of sparse representation based face

super-resolution method. To this end, we propose a noise robust face super-resolution method via locality-constrained representation. It analysis the influence of noise on the representation of image patch, and establish a locality-constrained representation model, which uses the locality prior to constraint the representation. This method can project each image patch into its neighborhoods in the training set adaptively, leading to a robust image patch representation, which solves the not unique problem of least square representation [1] and the unstable problem of sparse representation [2]. The average PSNR and SSIM improvements of locality-constrained representation method over the sparse representation method [2] are 1.02 dB and 0.0168 respectively in the FEI face database [5]. The subjective results of reconstructed face images are also better.

(3) Projection between the LR and HR spaces through HR manifold constraint.

The current manifold learning based super-resolution approaches mainly focus on the structure of the LR manifold and neglect the HR one. To this end, we propose to establish a HR manifold constrained mapping model between the LR and HR images. In particular, we use the original HR manifold structure to regularize the reconstructed HR manifold space, making the local geometric structure of the reconstructed HR image patch manifold and the original HR image patch manifold in consistency. When compared with the LR manifold constrained face Super-resolution method, the gain of the proposed method is 0.34 dB and 0.0044 in terms of PSNR and SSIM in CAS-PEAL-R1 face database [4], 0.27 dB and 0.0049 in terms of PSNR and SSIM in AR face database [7].

(4) Learning and fitting of the image degeneration model via Multilayer Locality-Constrained Iterative Neighbor Embedding and Intermediate Dictionary Learning.

The performance of learning based face super-resolution methods will drop sharply when the degeneration process is very complex. To this end, we introduce the LR dictionary learning to the face image super-resolution task for the first time, and establish a multilayer locality-constrained iterative neighbor embedding model and an intermediate dictionary learning model. Then we propose to update the LR training set by constructing intermediate dictionaries and modeling the relationship in much more consistent LR and HR spaces. Experimental results show that the gain of the proposed face super-resolution algorithm over [2] is 1.66 dB and 0.0303 in terms of PSNR and SSIM in the CAS-PEAL-R1 face database [4].

In summary, this thesis completes the express expansion of face image database, robust face image representation and learning of the relationship between high and low resolution images by referring the mechanism of human visual perception and cognition. It provides a new way to the actual monitor face image super-resolution reconstruction on the theory and key technologies.

Key Words: Video Surveillance, Face Super-resolution, Manifold Learning, Neighbor Embedding, Dictionary Learning

目录

摘要	I
ABSTRACT	III
第一章 绪论	1
1.1 研究背景与意义	1
1.2 人脸超分辨率面临的问题与研究现状	2
1.2.1 监控视频人脸图像的特点与面临的挑战	2
1.2.2 研究现状与存在的问题	3
1.3 本文的研究内容	14
1.4 本文的组织结构	15
第二章 基于最近特征线流形学习的人脸超分辨率算法	18
2.1 引言	18
2.2 流形学习算法	20
2.3 基于邻域嵌入的图像超分辨率算法	22
2.4 基于最近特征线流形学习的人脸超分辨率算法	22
2.4.1 符号定义	23
2.4.2 基于最近特征线流形学习的人脸超分辨率算法	23
2.4.3 最近特征线人脸超分辨率方法的改进	26
2.4.4 人脸图像小块划分	27
2.5 实验及结果分析	30
2.5.1 人脸库介绍	30
2.5.2 算法参数设置	30
2.5.3 不同算法主客观对比结果	30
2.5.4 训练库大小对算法结果的影响	32
2.6 总结与展望	33
第三章 基于图像块局部约束表示的噪声鲁棒人脸超分辨率算法	34
3.1 引言	34
3.2 图像块表示方法	37
3.2.1 最小二乘表示法	37
3.2.2 稀疏表示法	38
3.3 局部约束鲁棒人脸超分辨率算法	38
3.3.1 图像块局部约束表示	38
3.3.2 目标函数优化	39

3.3.3 局部约束鲁棒人脸超分辨率算法.....	40
3.4 稀疏性和局部性分析.....	42
3.4.1 稀疏性分析.....	42
3.4.2 局部性分析.....	44
3.5 实验及结果分析.....	46
3.5.1 人脸库介绍.....	47
3.5.2 FEI 人脸库上的结果.....	48
3.5.3 参数分析.....	50
3.5.4 算法噪声鲁棒性测试.....	53
3.5.5 真实图像的超分辨率重构结果.....	55
3.6 总结与展望.....	57
第四章 基于高分辨率流形空间约束的人脸超分辨率算法	59
4.1 引言.....	59
4.2 基于高分辨率流形结构保持的嵌入学习.....	62
4.2.1 符号定义及问题提出.....	62
4.2.2 局部诱导的支撑回归方法.....	63
4.3 实验结果分析.....	68
4.3.1 对比算法与参数设置.....	68
4.3.2 CAS-PEAL-R1 人脸库上的对比结果.....	69
4.3.3 AR 人脸库上的对比结果.....	73
4.3.4 算法参数对重构结果的影响.....	74
4.3.5 真实世界人脸图像上的重建结果.....	79
4.3.6 讨论.....	81
4.4 总结与未来工作展望.....	82
第五章 基于局部迭代嵌入和字典渐进更新的人脸超分辨率算法.....	84
5.1 引言.....	84
5.2 相关工作.....	87
5.2.1 符号定义及问题提出.....	87
5.2.2 图像块表示方法.....	88
5.3 提出的“由粗到精”人脸超分辨率方法.....	91
5.3.1 局部约束迭代邻域嵌入方法.....	92
5.3.2 多层局部约束迭代邻域嵌入方法.....	94
5.3.3 计算复杂度.....	96
5.4 实验结果.....	96
5.4.1 人脸库描述.....	97
5.4.2 参数分析.....	97
5.4.3 与前沿算法的对比.....	101

5.4.4 实际人脸超分辨率重构结果.....	103
5.4.5 讨论.....	104
5.5 结论.....	106
第六章 总结与展望	107
6.1 本文工作总结	107
6.2 未来工作展望	108
附录 部分推导与数学证明.....	110
附录 A 图像噪声强度与正则化参数的关系.....	110
参考文献	111
已获得研究成果情况	119
致 谢	125

第一章 绪论

1.1 研究背景与意义

进入二十一世纪以后，我国的社会经济发展已经进入了新的发展阶段，取得了令人瞩目的成就。但是发展中不平衡、不协调、不可持续的问题依然突出。近年来不仅藏独、疆独活动频繁，社会仇富、仇官问题也日渐突出，尤其是无直接利益关系的群体事件、武汉建行爆炸案和周克华枪击案等恶性案件呈急剧上升趋势，公共安全形势不容乐观。中国正处于工业化、城镇化快速发展时期，各种利益关系错综复杂，各种传统的和非传统的、自然的和社会的风险、矛盾交织并存，维护社会稳定的任务艰巨，国家安全面临的形势更加复杂严峻。2005 年以来中国已投入 3200 亿建成了全世界最大的城市安防监控网络。视频监控技术在防控打击违法犯罪方面发挥着越来越重要的作用^[9]。

公安部物证鉴定中心对 5000 起常规案件物证影像资料统计表明^[10]（如表 1-1 所示）：犯罪嫌疑人的人脸图像是刑侦人员最关注的目标，但是在实际监控应用中，由于摄像头和感兴趣目标距离通常较远、监控系统的带宽和存储资源有限以及环境噪声和器件噪声等因素的影响，白天获得的监控图像质量差的比例为 60%，而夜间这一比例则高达 95%。图 1 为两个典型案例中采集到的人脸图像，它们都存在分辨率低、噪声大的问题，感兴趣人脸图像无法被人或者机器所辨识。监控视频嫌疑目标人脸图像模糊不清的现象十分普遍，因此在原始的低质量感兴趣目标图像基础上重建高质量可辨识的人脸图像成为视频侦查的核心技术需求。

人脸超分辨率（Face Super-Resolution）是一种根据观测到的低质量、低分辨率（Low-Resolution）人脸图像，结合由高低分辨率人脸图像训练对学习到的先验知识重建清晰高分辨率（High-Resolution）人脸图像的技术。它可以有效增强低质量人脸图像的分辨率，恢复人脸的特征细节信息，这对于提高人脸图像清晰度，增加人脸辨识准确性，进而提高公安机关破案率具有重要意义^[11, 12]。

表 1-1 对 5000 起案件物证影像资料统计^[10]。

影像时间	白天	晚上
图像质量差比	60%	95%
主要原因	低分辨率、模糊	噪声、低分辨率、模糊
关注目标	人脸图像	其它
所占比例	75%	25%

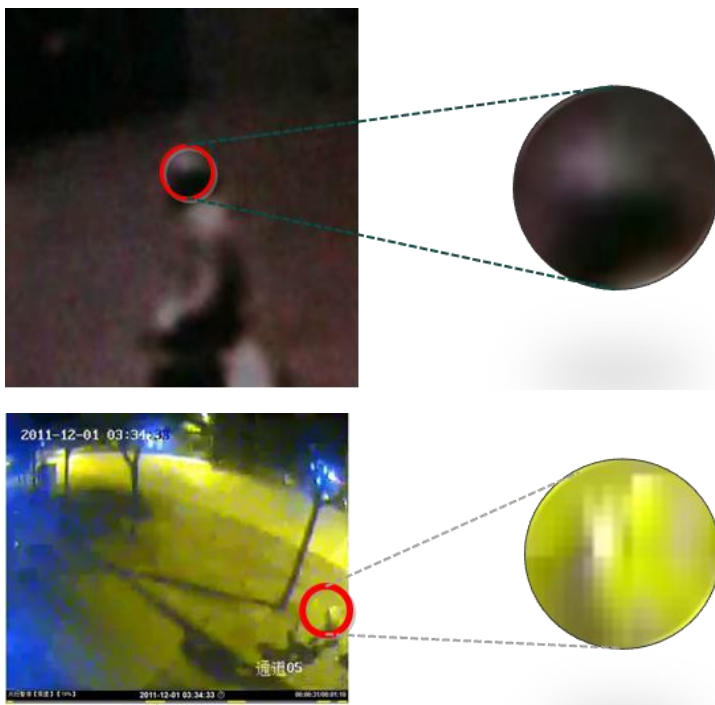


图 1-1 典型案件监控人脸图像质量示例。周克华重庆 3·19 枪杀哨兵枪案（左）和王海剑武汉 12·1 银行爆炸案（右）。

1.2 人脸超分辨率面临的问题与研究现状

1.2.1 监控视频人脸图像的特点与面临的挑战

为了解决实际监控环境的人脸超分辨率问题，本课题组在公安部重点科技攻关计划《刑事侦察图像比对系统》（项目号：2008ZDXMHBST011）、国家自然科学基金面上项目《面向极低质量监控图像的鲁棒性人脸超分辨率算法研究》（项目号：61172173）和国家科技支撑计划课题《警用视频图像信息应用关键技术研究及示范系统》（项目号：2013AA014602）等项目的资助下对人脸超分辨率技术和通用图像增强（包括图像去噪、图像去模糊、图像去雾、图像光照变换和图像几何变换）等刑事侦查图像比对分析领域的关键技术展开了充分的调研与研究。在此过程中发现，现有文献报道的人脸超分辨率方法主要针对高分辨率图像加模糊、下采样后得到的仿真低分辨率人脸图像，而且可以获得较满意的结果。但是，现有方法在以实际监控低质量人脸图像作为输入时，仍然无法满足刑侦实际应用的需求。经分析，主要可以归结为以下几点原因：

第一，因为人类对人脸图像的微小变化极其敏感，因此在对人脸图像进行表示时，需要更多的训练样本对其进行精细表示。而实际应用中构建的人脸图像样本库的规模非常有限，训练样本过少会导致无法准确地对人脸图像的特征细节进行表示；

第二，在视频监控的实际应用中，受环境因素、摄像机因素、成像因素等多方面因素的影响，实际监控摄像头得到的低分辨率人脸图像受噪声影响严重。而现有人脸超分辨率方法对实际监控噪声敏感，输入图像中所包含的微小噪声扰动会对人脸图像块的表达造成很大影响，现有方法对噪声图像块的表示不稳定；

第三，现有图像超分辨率方法认为高低分辨率流形空间具有局部一致性，然而，从真实场景中拍摄获取低分辨率人脸图像是一种图像降质过程，在此过程中会有信息的丢失（如人脸图像上用来区别不同个体的高频细节特征），进而导致在实际情况下存在多张低分辨率人脸图像（块）可能对应同一张高分辨率人脸图像（块）的情形，而这种低分辨率图像与高分辨率图像之间的“多对一”对应关系将使得现有算法的流形假设不成立；

第四，在实际监控中，低分辨率图像获取过程极其复杂，例如：模糊类型和模糊强度不清、下采样倍数大、输入人脸图像存在大量噪声。然而，现有方法假设高低分辨率流形空间具有局部一致性，试图通过单一映射学习得到低分辨率图像到高分辨率图像的对应关系，实际监控成像过程的复杂性进一步加大了这种单一映射学习的难度，学习得到的映射关系的准确性无法保证。

1.2.2 研究现状与存在的问题

相对于图像降质过程（通过对原始高分辨率场景进行采样得到低分辨率图像），图像超分辨率问题是一个反问题¹，即利用已知的观测得到的一张或者多张低分辨率图像来推断恢复出原始未知的高分辨率图像，其目的就是补偿低分辨率图像在图像降质过程中丢失的高频细节信息。20世纪60年代Harris^[13]和 Goodman^[14]首次提出“图像超分辨率”这一概念，他们通过定义线性插值或样条函数插值函数来增加输入图像的空间分辨率。然而，超分辨率重建这一思想并未得到广泛的认可，直到1984年Tsai和Huang^[15]提出一种根据连续多帧具有相对运动的低分辨图像重建一张高分辨率图像的方法之后，图像超分辨率重建技术才得到学者们广泛的关注和深入的研究。目前，该技术已经成为图像处理领域的一个非常重要的研究方向^[16, 17]。

重建图像的高频细节信息不可能无中生有，为了获得图像降质过程中丢失的高频细节^[18]。一方面，可以利用视频序列中的连续多帧图像之间提供的互补信息来超分辨率重建原始高分辨率图像，其核心思想就是用时间带宽（获取同一场景的多帧图像序列）换取空间分辨率，实现时间分辨率向空间分辨率的转换，即**基于多帧重建的图像超分辨率算法**；另一方面，可以利用图像的先验知识（自然图像统计规律或者样本高低分辨率图像之间的对应关系），来指导或约束图像超分辨率重建过程，以获得符合人类视觉先验的高分辨率图像，即**基于学习的图像超分辨率算法**。除了上述根据重构高分辨率图像的信息来源进行分类外，也可以根据超

¹反问题是根据事物的演化结果，由可观测的现象来探求事物的内部规律或所受的外部影响，由表及里，索隐探秘，起着倒果求因的作用。

分辨率重构的任务进行分类。比如可以分为特定领域的图像超分辨率方法（包括本文所关注的人脸超分辨率^[19]、特定场景超分辨率^[20]和图形作品的超分辨率重构^[21]等等）和通用图像超分辨率方法。

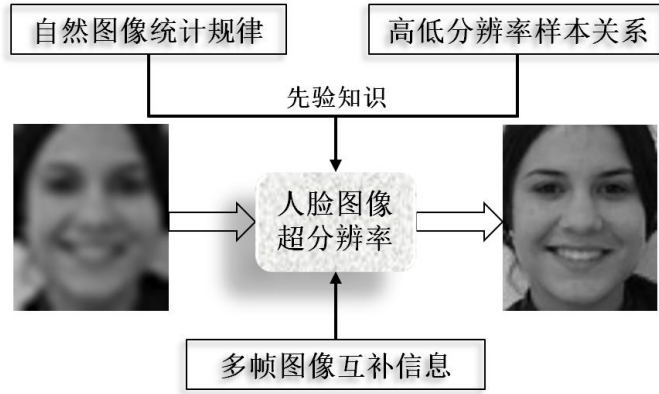


图 1-2 人脸超分辨率额外高频细节的来源。

基于多帧重建的图像超分辨率方法需要从同一场景中获取连续多帧低分辨率图像并进行配准，然后将多帧低分辨率图像信息映射到高分辨率图像的网格中，从而实现对低分辨率图像的分辨率增强^[22-27]。然而，多帧序列图像超分辨率重建在下面两种条件下经常难以满足要求，一是多帧图像序列之间存在的互补信息并不充分（即帧间互补信息量不足），二是随着超分辨率重建要求提高倍数的增加序列中的互补信息会显示得相对不足。此时，试图通过增加输入低分辨率图像的数量来得到更多的高频细节的方法仍然难以达到超分辨率复原的目的。Lin等人^[28]已经证明即使在准确配准且输入低分辨率图像没有噪声的条件下，基于多帧序列图像重建的超分辨率算法有效的放大倍数也仅仅只有1.6，放大倍数有限这一缺陷极大的限制了这类算法的应用前景。仅靠增加图像的数量并不一定能产生新的高频细节，在这种情况下，图像本身的先验知识显得非常重要。基于学习的图像超分辨率技术作为一个前沿的研究领域应运而生^[29-31]。

根据先验知识在获取方式上的不同，基于学习的图像超分辨率方法又可以分为基于自然图像统计的方法^[32-36]和基于训练样本学习^[30, 37-40]的方法两大类。自然图像统计先验一般是指图像通过滤波器之后的响应的统计分布规律。通过对图像滤波响应的统计分布进行建模，求得模型参数并以此来约束接下来的超分辨率重建过程。然而用解析化的数学表达式来定义图像的统计分布（如概率密度）的方式很难刻画自然图像的丰富特性。事实也是如此，对于超分辨率重建这一病态程度极高的反问题，利用自然图像统计先验的方法很难给出令人满意的结果。而Elad等人^[31, 41]也指出，直接从样本图像中学习得到的先验知识将更加有利于图像超分辨率重建。

由于仅仅依靠自然图像本身的统计先验知识不足以获取令人满意的超分辨率重建性能，

Freeman等人^[29]最早提出基于样本学习的超分辨率方法。他们认为，在图像块这一级别上的高低分辨率图像之间的对应关系是可再生的。也就是说，如果两个低分辨率样本图像块在视觉上存在相似关系，那么它们对应的两个高分辨率图像块在视觉上也是相似的。因此，通过提前收集足够数量的高低分辨率样本图像块对，就可以通过机器学习的方式来分块预测观测低分辨率图像中丢失的高频细节信息。相比于解析式的先验知识（显示先验知识），通过样本所引入的额外信息（隐式先验知识）进行超分辨率重建，可带来突破性的性能提升^[42]。尤其是对于人脸这类具有高度结构化的对象，如五官几何分布具有对称性和一致性，不同器官也具有一致的方向性特点，因此，人脸图像训练样本可以提供更多有利于超分辨率过程的先验知识^[1, 18, 43]。

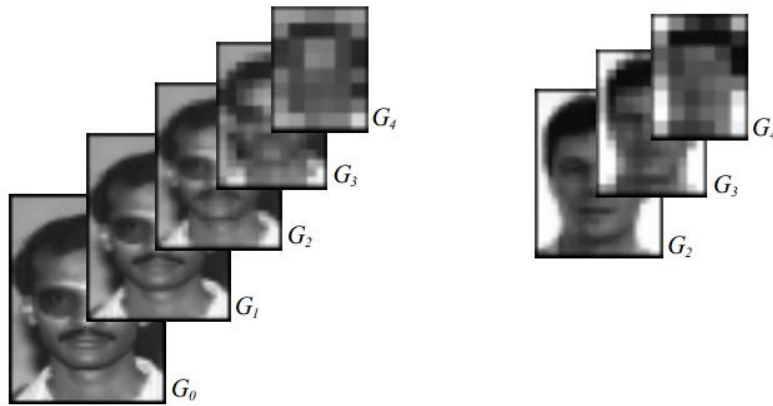


图1-3 左图为经过多次模糊和下采样所得到的高分辨率高斯金字塔、右图为输入低分辨率图像包含的低分辨率高斯金字塔。图像超分辨率的目标就是根据输入的低分辨率高斯金字塔来估计低层高分辨率高斯金字塔的过程^[19]。

2000年, Baker和Kanade^[19]专门针对人脸图像, 首次提出一种“人脸幻构(face hallucination)”的思想。他们认为简单的平滑约束很容易使重建的高分辨率图像由于高频信息的丢失而过于平滑, 因此通过对特定的目标、场景或图像集合进行学习, 将高低分辨率图像空间像素匹配关系加入到人脸图像重构的过程中, 以此来增加可用的信息, 其原理如图1-3所示。随后, Liu等人^[40, 44]提出人脸重构的两步法, 利用参数全局模型主成分分析 (Principal Component Analysis^[45, 46], 简称PCA) 和非参数局部模型马尔可夫随机场 (Markov Random Field^[47], 简称MRF) 分别合成人脸的全局信息和局部信息, 如图1-4所示。继上述两种代表性的算法之后, 出现了基于学习的人脸超分辨率重建算法的研究热潮, 学者们针对人脸超分辨率问题提出很多变型。这些方法大致可以分为两大类^[48]: 基于全局脸的人脸超分辨率方法^[3, 44, 49-58]和基于局部脸的人脸超分辨率方法^[1, 2, 43, 51, 59-72]。前者将每幅人脸图像看作一个整体, 通过利用人脸图像的全局结构信息保证超分辨率出来的人脸图像在结构上与原始真实高分辨率人脸图像具有一致性; 为了获得更多的人脸特征细节信息, 后者则是将人脸图像按照位置划分成多个很小的图像块, 再对每个图像块分别进行超分辨率重建, 最后融合所有重建得到的高分辨率人脸图

像。以下将对这两类人脸超分辨率方法进行综述。

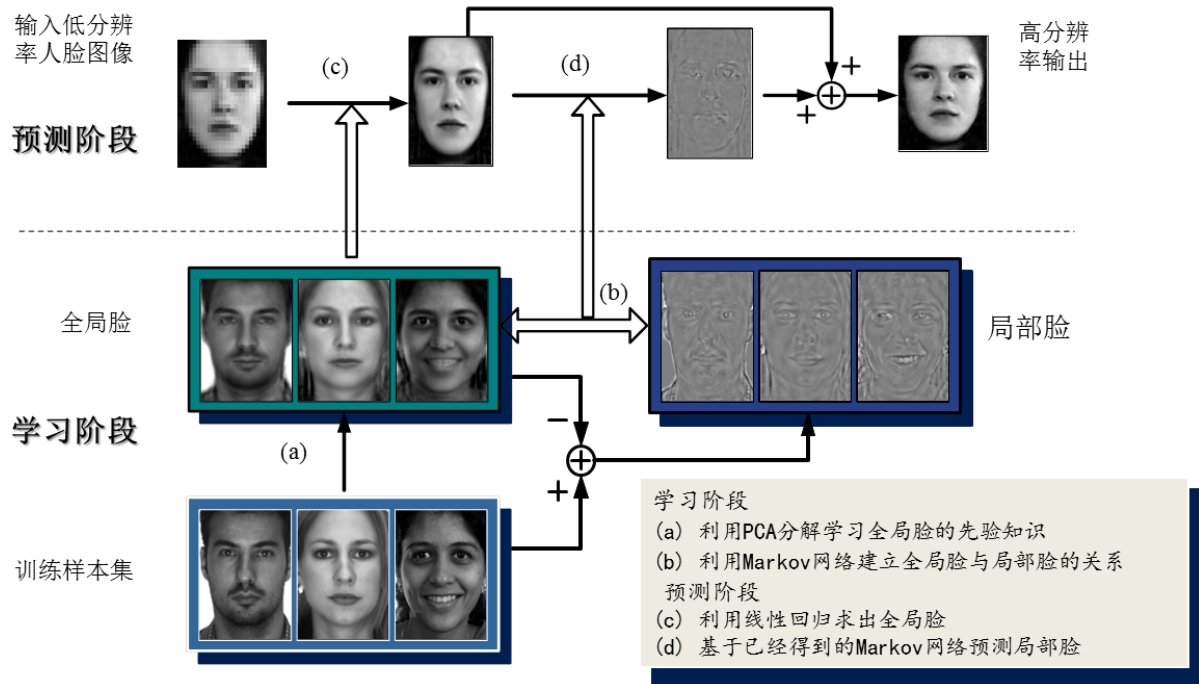


图1-4 从左到右依次是原始高分辨率人脸图像、重构全局脸图像和残差人脸图像^[40,44]。

1.2.2.1 基于全局脸的人脸超分辨率方法

基于全局脸的人脸超分辨率方法的基本思想是通过训练样本进行子空间分解(包括主成分分析^[45,46]、局部保持投影^[73]、非负矩阵分解^[74,75]、典型相关分析^[76]和偏最小二乘分解^[77]等),得到一组能够表征人脸图像特征的基向量,然后将输入图像投影到该子空间,在此子空间进行超分辨率重构。

例如,2001年,牛津大学的Capel 和 Zisserman^[55]首次提出利用训练样本库中的所有人脸图像进行PCA分解,通过选取适量特征脸得到一个主成分特征空间,并以此作为超分辨率重构的先验约束,结合最大后验概率(Maximum a Posteriori^[33],简称MAP)估计从主成分特征空间中超分辨率重构出一张高分辨率人脸图像。与利用MRF模型从高低分辨率图像中学习图像先验的方法相比,该方法利用训练样本库中的所有人脸图像构建主成分特征空间,并以此作为超分辨率重构的先验约束,由于PCA统计模型可以通过选取对人脸表达有用的重要成分,因此提出的方法相较于传统基于MRF模型的人脸超分辨率具有更强的表达能力,大大提升了人脸超分辨率重构的性能。

鉴于PCA模型对人脸有着非常强的刻画能力,2005年,香港中文大学的Wang和Tang^[54]提出一种基于PCA特征变换法的全局人脸超分辨率算法。如图1-5所示,该方法首先利用主成分分析对训练样本库中的所有人脸图像进行PCA分解,将输入的低分辨率人脸图像投影到低分辨

率训练样本所张成的特征空间上得到投影系数，然后将该投影系数通过特征转换得到原始低分辨率空间的投影系数，最后把新得到的转换投影系数直接映射到对应的原始高分辨率图像上，线性加权得到最终重构的高分辨率人脸图像。由于增强了低分辨率人脸图像中的特征信息，重建后的高分辨率图像去除了大量噪声并保留大部分的人脸特征。在上述方法的基础上，印度理工学院的Chakrabarti等人^[53]进一步提出了一种基于核化PCA的人脸超分辨率算法。

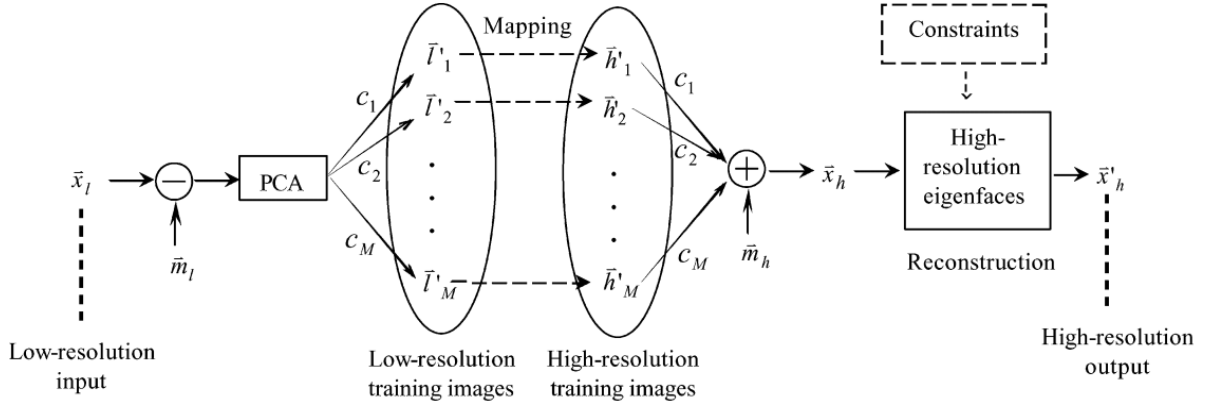


图1-5 基于特征转换的人脸超分辨率方法^[54]。

2008年，韩国全南国立大学的Jeong-Seon Park等人^[52]进一步提出将人脸图像分解成形状分量和形状无关的纹理分量，并利用PCA分解将输入人脸图像的形状和纹理分别投影到对应的主成分特征空间，并利用迭代误差后向投影由输入低分辨率人脸图像重构高分辨率人脸图像。这种根据重建误差更新重构出来的高分辨率图像的思想可以逐步优化、补偿估计得到的高分辨率人脸图像，实验结果表明，该方法与传统PCA分解方法相比有所提高。

2007年，浙江大学的Zhuang等人^[51]提出一种基于局部保持投影（Locality Preserving Projections^[73]，简称LPP）人脸超分辨率算法（Locality Preserving Hallucination，简称LPH）。该方法提出两点假设：第一，相同分辨率的人脸图像来自同一的潜在低分辨率流形空间；第二，高低分辨率训练图像对分别嵌入在各自的潜在低维空间，而且上述两个低分辨率空间具有一致性。基于这两点假设，LPH方法利用局部保持投影策略将高低分辨率人脸图像投影到具有一致流形结构的低维空间，然后再在此低分辨率空间进行邻域嵌入学习。与Wang和Tang提出的全局人脸超分辨率算法^[54]相比，该算法增加了更多的细节信息，同时也带来更多的噪声，尤其是在人脸轮廓、鼻子、嘴巴等部位。

2010年，西安交通大学Huang等人^[3]提出利用典型相关分析法（Canonical Correlation Analysis^[76]，简称CCA）提取高低分辨率人脸图像的相关子空间，从而增强高分辨率人脸图像和低分辨率人脸图像流形空间拓扑结构的一致性。该方法将典型相关分析同时运用于全局脸重构的残差脸的补偿中，最终获得了很好的重构效果。An Le等人^[50]在上述方法的基础上，提出一种基于2DCCA的人脸超分辨率重构方法，并获得了很好的重构效果。同年，伊利诺伊大学香槟分校的Yang等人^[8]提出利用非负矩阵分解（Nonnegative Matrix Factorization^[74, 75]，简称

NMF) 来重构全局高分辨率人脸图像。与PCA不同的是, NMF的线性加权系数必须为非负。PCA分解得到的特征脸是一张具有完整脸部特征的基图像, 而NMF得到的则是“局部”脸。这种基于局部基图像组合的表示形式具有很直观的语义解释, 即对整张人脸的感知由对组成整张人脸的各个器官(局部脸)的感知构成。

2011年, 四川大学的Wu等人^[49]提出一种基于核偏最小二乘分解(Kernel Partial Least Squares^[77], 简称KPLS)的人脸超分辨率算法, 利用KPLS建立高低分辨率人脸图像之间的非线性映射关系, 利用该映射关系就可以预测输入低分辨率人脸图像对应的高分辨率人脸图像。该方法将原图像特征从线性空间推广到了非线性空间, 因此可以表达出图像的高阶统计特性, 取得了较好的重建效果。

近年来, 国内文献对基于全局脸的人脸超分辨率方法也进行了相关报道。

例如, 西安交通大学黄华等人^[78]在2006年提出基于粒子滤波的人脸超分辨率重构方法。2008年四川大学的吴炜等人^[79]提出基于多分辨率的人脸幻构算法, 采用Kirsch算子提取了高频特征, 使得重构的人脸更加精细。2010年, 中科院自动化所张雪松等人^[80]提出一种基于特征子空间规整化的人脸超分辨率重建方法。2011年, 武汉大学兰诚栋等人^[81]提出一种基于语义约束的全局人脸超分辨率算法。

全局脸超分辨率算法将人脸图像看作一个整体, 对其进行处理与学习, 利用人脸这一具有高度结构化对象的全局特性重建出输入低分辨率人脸图像对应的高分辨率人脸图像, 实现人脸图像的超分辨率复原。这类方法能够在高放大倍数条件下较好地保持人脸图像的全局信息。但是由于人脸图像是一个非常高维的信息(人脸的特征表示动辄上万维), 少量的训练人脸样本图像(目前最大样本库的对象数也不超过2000)表示能力有限, 无法对此高维的数据进行准确的表示, 进而导致重建出的人脸图像缺乏关键的特征细节信息。因此, 为了增强给定训练样本人脸图像的表达能力, 近年来大量文献提出对整张输入人脸图像进行分块的方式进行表示, 这种分块表示合成的方式大大提升了原有样本集合的表达能力, 受到了人们广泛的青睐。下面我们将对这类基于局部图像块的人脸超分辨率方法进行综述。

1.2.2.2 基于局部脸的人脸超分辨率方法

基于局部脸的人脸超分辨率方法的基本思想是将输入人脸图像和所有人脸样本图像利用同样的方式进行分块, 学习低分辨率图像样本块与高分辨率图像样本块之间的对应关系, 以此学习得到的关系推测输入低分辨率图像每一个图像块的高频细节信息, 融合所有超分辨率重建的图像块便可得到最后的高分辨率人脸图像。

例如, 2000年, 三菱电子研究院的Freeman等人^[29, 38]提出基于分块的图像超分辨率方法。他们认为, 在图像块这一级别上的高低分辨率图像之间的对应关系是可再生的。也就是说, 如果两个低分辨率样本图像块在视觉上存在相似关系, 那么它们对应的两个高分辨率图像块

在视觉上也是相似的。如图 1-6 所示，该方法以马尔可夫网络为基础体系，对图像的空间关系进行建模，把观察图像分成一些 5×5 或 7×7 的图像小块，通过学习获得表示高分辨率图像块之间的转移概率矩阵和表示高分辨率图像块和低分辨率图像块之间的转移概率矩阵。通过大量的样本图像块学习得到 MRF 网络参数，利用信任传播机制为输入低分辨率图像块找到一个后验概率的局部最大值，从而估计出输入低分辨率图像块所对应的高分辨率块。与传统基于多帧重建的图像超分辨率算法相比，这种基于学习的方法可以从训练样本库中获取丰富的高频信息，在放大倍数较大（如 4 倍）时，仍能获得较高的超分辨率重构质量。然而，该方法对训练样本的选择要求比较高，当输入待重构的低分辨率图像与训练样本图像差异较大时，超分辨率重构的结果较差。

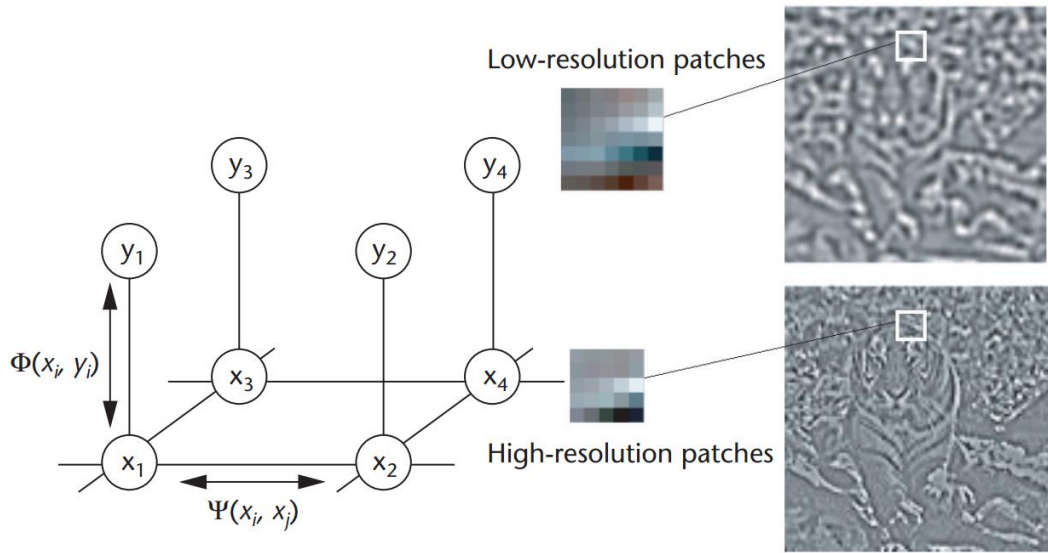


图 1-6 基于 MRF 网络模型的图像超分辨率方法。 $\Phi(x_i, y_i)$ 和 $\Psi(x_i, y_i)$ 分别代表重建约束项和局部兼容项^[38]。

2003 年，西安交通大学的 Sun Jian 等人^[82]提出了一种基于图像基元（Primal Sketch）学习的图像超分辨率方法。作者认为，自然图像基元（如边缘、纹理和拐点等）的潜在维度要远小于原始图像块空间，因此可以用少量的基元图像便可以有效地表达丰富的自然图像。而且，通过构建图像基元空间，在一定程度上可以减少原始高低分辨率空间“多对一”的不确定性问题，从而提高了学习的准确性。

2004 年，香港科技大学的 Hong Chang 等人^[83]首次将流形学习的思想引入到图像超分辨率重建中。提出一种基于局部线性嵌入（Locally Linear Embedding^[84]，简称 LLE）流形学习的单张图像超分辨率。他们认为高低分辨率图像块流形空间局部几何结构具有一致性，对于给定待重建的低分辨率图像块，先在低分辨率样本图像块空间中找到它的 K 个近邻样本图像块，并构建它们之间的近邻关系（局部线性重建系数），基于上述一致性假设，将该近邻关系直接保持到高分辨率图像块空间中，与对应的 K 个高分辨率近邻样本图像块进行线性加权

便可以得到最终的超分辨率重建图像。2007 年, 香港科技大学的 Wei Fan 等人受文献^[82]和文献^[83]的启发, 提出一种基于图像基元邻域嵌入的图像超分辨率算法^[85]。首先, 利用大量训练样本块学习得到图像基元, 然后在图像基元空间进行邻域嵌入学习, 重构高分辨率图像。基于近邻保持的图像超分辨率原理示意图如图 1-7 所示。随后, 印度理工学院的 Kumar 等人^[86]提出一种基于正交局部保持投影的人脸超分辨率方法。他们利用正交局部保持投影学习得到一组高低分辨率基图像, 然后在此空间进行局部保持学习。Zhang 等人^[68]将上述基于邻近嵌入保持的方法从像素域拓展到 DCT 变换域, 进而提出一种基于 DCT 域的邻近嵌入保持人脸超分辨率方法。近年来, 基于这种局部流形一致性假设, 人们大量的相关邻域嵌入图像超分辨率算法^[87-93]。然而, 这些方法中的近邻数 K 是固定的, 在进行局部线性重建的时, 既定数量的近邻样本的组合会造成对观测图像块的过拟合或者拟合不当的问题, 进而导致合成出的高分辨率图像存在模糊, 或者合成出原始高分辨率图像中本身没的“细节信息”。

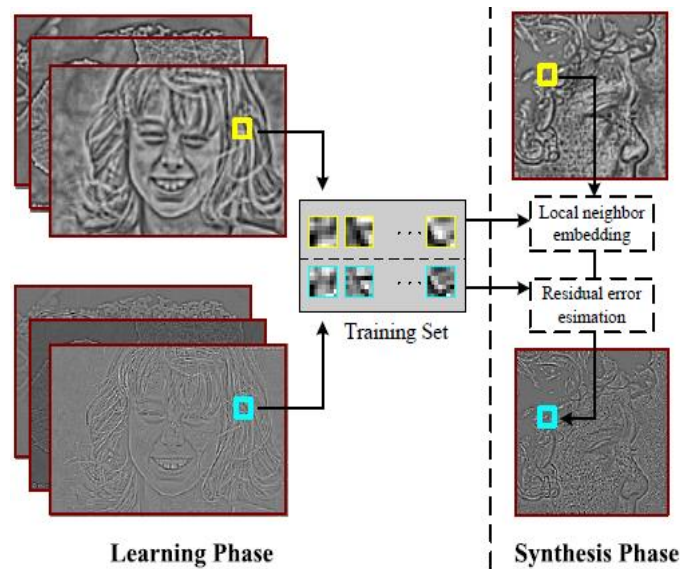


图1-7 基于近邻保持的图像超分辨率原理示意图^[85]。

针对上述问题, 美国伊利诺伊大学香槟分校的杨建超等人在 2008 年^[94]和 2010 年^[8]相继提出基于稀疏表示的图像超分辨率方法和基于耦合稀疏字典学习的图像超分辨率方法。他们认为观察信号在过完备字典上的表示系数是稀疏的, 即表示系数中非零元素的个数非常少 (图 1-8 为稀疏表示模型示意)。在图像超分辨率重构中, 图像块能够被一个合适的过完备字典稀疏地线性表示。具体地, 对输入低分辨率图像块在低分辨率字典上进行稀疏编码 (表示), 基于高低分辨率图像块流形空间具有一致性 (低分辨率空间获得的表示系数与真实高分辨率空间一样) 假设, 直接将上述系数映射到对应高分辨率图像块字典上, 最终加权得到高分辨率图像块输出。通过引入稀疏先验, 可以提升原始高图像块空间的表达能力。主观实验结果表明, 在无噪声或者噪声强度很小的情况下, 该方法可以获得比传统插值放大方法更优

的效果,但是当输入低分辨率图像中存在强噪声时,相较于传统插值放大算法的 RMSE 增益迅速下降,主观效果与传统插值放大算法相当。

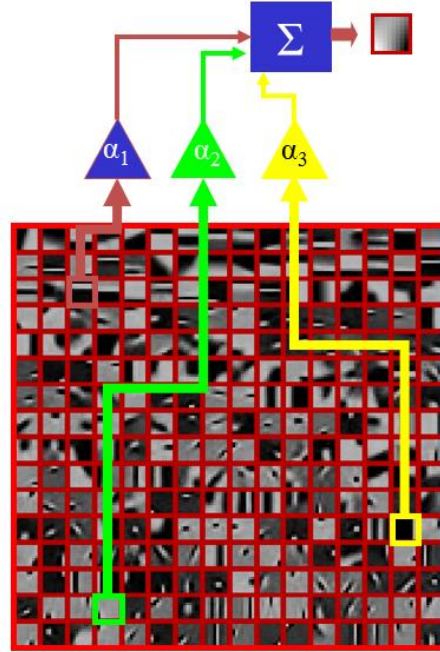


图 1-8 稀疏表示模型示意: 每一个图像块均可以由原子库中少数几个原子来线性表示。

由于人脸是一类高度结构化的物体,人脸图像上每个图像块所在的位置信息对于人脸图像的合成与分析具有非常重要的作用。受此思想的启发, Ma 等人^[1]提出一种基于位置图像块的人脸超分辨率方法。该方法使用训练集中所有与输入图像块相同位置的人脸图像块重建高分辨率人脸图像,避免流形学习或者特征提取等步骤,提高了效率,同时也提升了合成图像的质量。然而,由于该方法采用最小二乘法进行求解,当训练样本中图像的个数比图像块的维数大时,图像块的表示系数并不唯一。因此,2011 年西安电子科技大学 Jung 等人^[2]提出一种基于凸优化的位置图像块人脸超分辨率方法,将凸优化理论加入到图像块求解表示中,可以解决方程的解不唯一的问题,为了使输入图像块的表示尽量稀疏,该方法在合成输入的图像块时可能选取一些与输入的图像块差异很大的图像块进行线性重建,没有考虑到局部性这一特征,因此重建效果并不理想。2013 年,武汉大学王中元等人^[63, 64]在基于稀疏表示方法的基础上提出一种基于自适应范数的人脸超分辨率方法,但是如何根据特定输入图像选择合适的范数是非常困难的。

上述基于图像块表示的方法均是基于高分辨率图像块流形和低分辨率图像块流形拥有相似的局部几何结构这一基本假设,保持低分辨率空间的合成权重系数不变,并直接将低分辨率训练样本替换成高分辨率训练样本,从而重构出高分辨率图像。然而,在实际情况下,高分辨率图像与低分辨率图像是“多对一”的映射关系^[69, 87, 95, 96],高低分辨率流形具有一致性局部几何关系的假设并不成立。为了进行更好的邻域嵌入学习,2009 年,哈尔滨工业大学

的 Bo Li 等人^[59,97]提出通过流形对齐的思想来解决这一问题。他们认为高分辨率图像块流形和低分辨率图像块流形具有一个公共的潜在低维流形空间,通过学习两个投影矩阵便可以把原始高分辨率图像块流形和低分辨率图像块流形同时投影到这个潜在的公共流形空间中,然后在此公共流形空间中进行基于邻域嵌入的图像超分辨率重构。

2011 年,西安交通大学的 Huang 等人^[65]针对现有算法需要高存储和高计算复杂度的问题,提出一种基于局部线性转换的人脸超分辨率方法,该方法利用多个局部线性转换来估计低分辨率图像与高分辨率图像之间在像素域上的关系。对人脸图像上的每一个图像块,以上线性转换关系相对固定。因此,可以通过离线学习的方式得到,避免的复杂的学习与求解过程,而且也不需要预先存储好高低分辨率训练样本图像。所以该方法不仅高效,而且还节省内存。2013 年,他们在上述方法的基础上进一步提出一种基于局部非线性转换的人脸超分辨率方法^[98]。

2012 年,长安大学的 Ma 等人^[71]结合人脸图像块的位置先验信息和稀疏表示理论,提出一种改进位置块的稀疏表示人脸超分辨率算法。2013 年,西安交通大学的齐春等人^[60,91]利用 PCA 分析法为每个位置块的训练图像块集合构建高低分辨率字典对,在此基于上进行回归预测学习,获得了较好的重建结果。

2013 年,电子科技大学 Li 等人^[62]提出一种利用四类相似性作为约束进行人脸图像超分辨率重构。这四类相似性约束包括:输入低分辨率图像与样本低分辨率图像的相似性约束、输入低分辨率图像与样本高分辨率图像的局部结构相似性约束、重构之后的近邻图像块平滑约束和空间相似性约束。

2014 年,清华大学 Huang 等人^[99]提出广义自适应联合核回归和自适应字典学习的方法。利用局部结构约束和非局部自相似性约束,提出的联合回归先验模型可以用规整化病态超分辨率重建问题。作者将核回归先验模型推广到多尺度和多方向上,并结合了自适应字典学习和基于字典稀疏的先验知识。该方法对通用图像和人脸图像均获得了较好的重构效果。

除了上述一些具有代表性的局部脸超分辨率算法外,学者们还提出大量基于局部图像块的图像超分辨率方法来解决图像块的表示^[34,100-106]、高低分辨率字典学习^[107-111]和高分辨率图像回归预测^[33,70,99,108,112-116]基于图像分解的超分辨率^[56,117,118]等问题。最近,也有部分学者们开始关注实际环境(包括多姿势^[119-123]、变化光照等因素)中的人脸超分辨率问题,这些算法的好坏依赖于对齐前处理、特征点检测,人脸的语义理解等技术。此外,他们一般要求训练样本库中存在大量同一对象且相近模态(姿势、表情和光照等)的样本^[72,122,124,125],而对于监控条件下的人脸超分辨率问题,这些条件都很难达到。

国内文献中也有许多局部脸方法的报道。

例如,2004 年,浙江大学的黄丽等人^[126]提出基于多尺度和多方向特征的人脸超分辨率算法。2008 年,中科院自动化所的张雪松等人^[127]针对超分辨率重构与学习过程中没有充分

利用人脸图像块流形空间局部相似性的问题,提出一种基于局部保持投影的自适应流形学习方法。2011年,西安交通大学的马祥和齐春针对传统基于学习的人脸超分辨率算法重构高分辨率人脸图像缺乏足够多高频特征细节的问题,提出一种全局重建和位置块残差补偿相结合的人脸超分辨率新算法^[128]。四川大学的吴炜等人^[129]提出基于偏最小二乘(Partial Least Squares, 简称 PLS)的人脸超分辨率技术。2013年,福州大学的兰诚栋等人^[130,131]采用后验信息构建稀疏原子库的超分辨率人脸重建,通过事后采集现场图像,训练只对输入图像的清晰内容稀疏而对噪声内容不稀疏的低维原子集和与之相对应的高维原子集,计算低维空间的稀疏系数并映射到高维空间以合成出重建人脸图像。对于实际拍摄的监控图像输入,重建结果比传统方法更加有效。

上述基于局部脸的人脸超分辨率算法通过对整张人脸进行小块划分,大大增强了原来训练样本的表达能力,提升了超分辨率重构的性能。然而,由于高分辨率图像与低分辨率图像是“多对一”的映射关系。尤其是在图像块级别上,低分辨率样本空间的近邻关系未必在高分辨率空间中保持。此外,与全局脸方法相对,基于图像块的局部脸方法对输入图像中的噪声更加敏感,小的噪声扰动可能会导致图像块表示系数的巨大差异。

1.2.2.4 现有方法存在的问题

近年来,虽然国内外学者已经对人脸超分辨率技术进行了大量的研究,对于仿真条件下四倍以下超分辨率重建的结果已经非常接近原始输入高分辨率人脸图像,但仍无法满足刑侦实际应用的需求。因为基于学习的人脸超分辨率方法的本质就是基于高低分辨率人脸样本所构成的高低分辨率流形空间具有相似局部几何结构这一假设,以保持局部几何结构为基础,学习高低分辨率图像块的对应关系。对于实际监控中的成像过程极其复杂,噪声、模糊不确定,且要求较高的图像超分辨率放大倍数,局部几何结构一致性的假设不再成立。为了进行鲁棒地基于学习人脸超分辨率重建,以下技术难点亟待攻克:

(1) 样本表达能力的扩充

基于学习的超分辨率方法成败在于样本库的选择,样本库的丰富性是这类算法获得良好效果的基础。当输入待重建人脸图像与样本库差异很大时,重建效果往往很差。所以,必须通过扩大样本规模来解决流形空间采样不够稠密导致局部流形线性假设不成立的问题。然而,人脸库的建立是一个复杂而又耗时的过程,此外,即使能获得大量样本,算法在进行求解计算时也需要巨大的存储空间,且带来很高的运算复杂度。因此,通过简单地增加样本数量来改进传统方法在实际中并不可行。如何扩充有限样本的人脸图像库的表达能力是人脸超分辨率的一个技术难点;

(2) 噪声图像块的鲁棒表达

当输入低分辨率图像存在噪声，尤其是非高斯分布的实际噪声时，现在人脸超分辨率方法可能对输入图像进行过拟合，即把输入图像中存在的噪声也重建出来，而再将这组过拟合的重建系数直接映射到高分辨率图像流形空间，势必导致超分辨率重建结果的不确实性。由于超分辨率问题的病态性，即超分辨率重建的结果不唯一。因此，引入何种视觉先验知识并设计一种对噪声鲁棒的图像块表示方法是人脸超分辨率的第二个技术难点；

(3) 高低分辨率图像空间的一致性学习

当超分辨率重建倍数增大或者低分辨率人脸图像存在较大模糊时，高分辨率图像和低分辨率图像之间的“多对一”关系将进一步凸显，高低分辨率流形的局部几何关系不一致性变得更加明显。现有基于学习的图像超分辨率方法基本思想均是通过将低分辨率图像流形的几何结构保持到重建高分辨率图像流形空间进行超分辨率重建。然而，这将导致重建高分辨率流形与原始高分辨率流形空间的不一致性。因此，学习高低分辨率样本空间的映射关系，同时保持原始高分辨率空间的局部几何结构（使重建后的高分辨率流形空间与原始高分辨率流形空间具有一致性）是人脸超分辨率的第三个技术难点；

(4) 图像成像过程的层次化建模

实际待重建低分辨率图像由原始高分辨率图像经过模糊、下采样再加入少量噪声而得到，降质过程会导致信息的丢失，当图像模糊很强或者下采样倍数很大时，低分辨率图像丢失信息就越多，其中包含的信息量就越少。现有人脸超分辨率算法试图通过直接构建低分辨率图像与原始高分辨率图像之间的关系，并以此关系来预测重建高分辨率图像。然而由于低分辨率图像中过多信息的丢失，使得直接构建低分辨率图像与原始高分辨率图像之间的关系变得非常困难。因此，如何对成像模型进行层次化建模，经过多次映射学习得到低分辨率图像与原始高分辨率图像之间的关系是人脸超分辨率的第四个技术难点。

1.3 本文的研究内容

基于人脸图像超分辨率的理论框架，本项目将重点研究以下三个关键科学问题：第一、高低分辨率图像字典对扩充表示；第二、高低分辨率图像的视觉先验表示；第三、高低分辨率图像的一致保持学习。为攻克上述人脸图像超分辨率中的三大科学问题，本论文着重围绕以下四个方面进行研究（图 1-9 所示为本文四大研究内容与三大科学问题之间的关系）：

(1) 基于最近特征线流形学习的人脸图像库表达能力扩充

研究专门针对人脸图像的人脸图像库表达能力扩充方法，提高原始人脸图像库的表示能力与效能，具体研究内容包括：研究成像因素（光照、距离、模糊等）对样本图像产生的影响；给定样本图像，研究扩充样本图像表达能力的方法；研究扩充后样本空间的字典设计与优化

方法。

(2) 基于局部约束表示的噪声图像块鲁棒表示

研究基于局部约束先验的噪声鲁棒人脸图像表示,提高人脸图像的表示准确性和稳定性,具体研究内容包括:图像局部约束先验的统计获取;图像局部约束先验的形式化描述;基于图像局部约束先验模型的代价函数构建与求解方法。

(3) 基于高分辨率流形空间约束的高低分辨率图像空间映射

研究基于图约束的高低分辨率图像一致保持学习方法,约束重建后高分辨率图像流形空间,使重建后高分辨率图像流形空间与原始高分辨率图像流形空间具有一致性。具体研究内容包括:低分辨率图像空间到高分辨率图像空间的映射模型;高分辨率图像空间流形结构保持学习方法;高低分辨率图像空间双向保持的迭代优化算法求解。

(4) 基于局部迭代嵌入和字典渐进更新的图像观测模型学习与逼近

研究高低分辨率图像空间层次化观测模型的构建方法,揭示人脸图像降质模型内在规律,提高人脸超分辨率重建结果的精度。具体研究内容包括:低分辨率样本人脸图像的超分辨率与更新机制;测试人脸图像的多层次超分辨率优化方法。

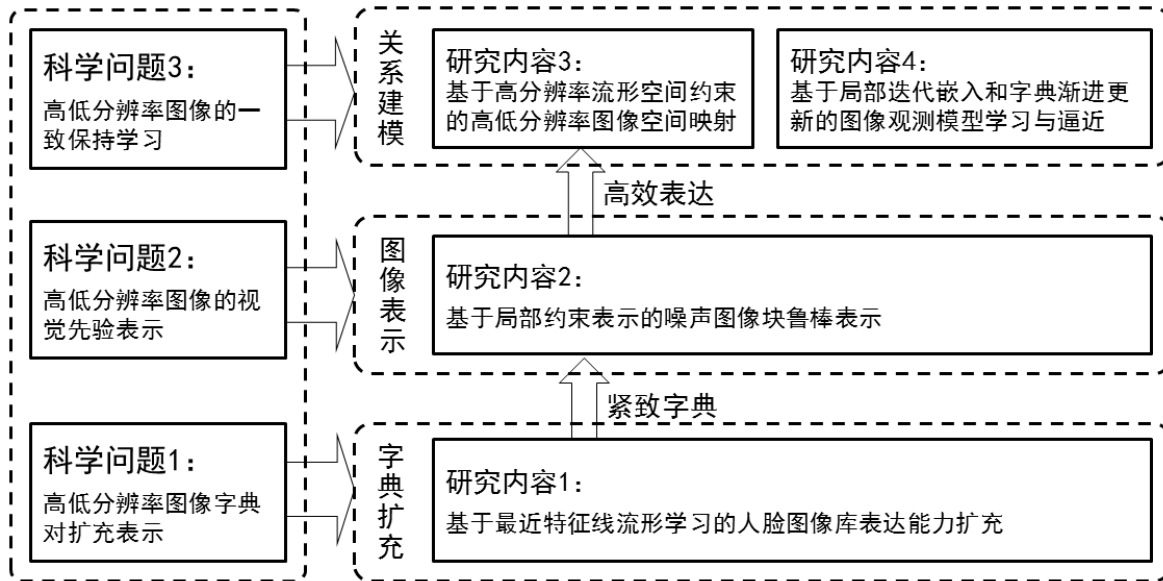


图 1-9 本文四大研究内容与三大科学问题之间的关系。

1.4 本文的组织结构

全文共分 6 章,组织结构如图 1-10 所示。

第一章为绪论。提出了本文的研究背景和意义,回顾了人脸超分辨率技术发展的历程,分析了当前实际监控人脸超分辨率研究面临的关键问题,最后给出了本文要研究的四个工作

内容。

第二章为基于最近特征线流形学习的人脸超分辨率算法。详细讨论了传统基于流形学习的人脸超分辨率算法通过学习高低分辨人脸图像块流形之间相似的局部几何结构关系,获得低分辨率块的局部表示系数,并将该系数直接嵌入(投影)到对应高分辨率图像块上,融合所有块得到高分辨率人脸。在欠采样条件下(即小样本问题,目前,最大样本库的样本数也不过 2000),样本数据只能构成人脸流形空间的一个稀疏空间,即使最相邻的点也难以称其为局部。因此,对欠采样的样本数据,由局部形成全局的流形嵌入思想从本质上并不适合。利用最近特征线可以把样本空间中任意两个特征点扩充为特征线上的无数多个特征样本点这一特性,不用新增样本数(避免增加样本导致计算复杂度和所需存储空间的增大),但是可以有效扩充原始样本的表达能力。本算法通过保持相对传统基于流形学习算法更小局部内的线性关系来揭示高分辨率和低分辨率人脸流形空间的非线性关系,从而更加体现了流形学习的本质。

第三章为基于图像块局部约束表示的噪声鲁棒人脸超分辨率算法。主要针对当前人脸超分辨率过程表示图像块不唯一和不稳定的问题,提出一种基于局部约束表示的人脸超分辨率重建方法。将输入的低分辨率人脸图像、高低分辨率训练集内人脸图像分别划分为相互交叠的图像块;对输入的低分辨率人脸图像的每一个图像块,根据图像块的表示具有局部性这一先验,计算由低分辨率训练集内每个图像相应位置图像块进行线性重建时的最优权重系数;将低分辨率训练集内每个图像相应位置图像块,用一一对应的高分辨率训练集内每个图像相应位置图像块进行替换,加权合成高分辨率图像块;按照人脸上的位置,融合成高分辨率人脸图像。本文提出局部约束表示模型,自适应地选择训练集内样本图像块空间中与输入图像块近邻的图像块来线性重建输入图像块,得到最优权重系数,合成出高质量的高分辨率图像。

第四章为基于高分辨率流形空间约束的人脸超分辨率算法。将目前基于学习的人脸超分辨率方法大体分为基于编码表示和基于回归预测两大类,由于高低分辨率图像的“多对一”关系,前者基于高低分辨率图像块具有“一致表示系数”的假设在实际情况下并不成立;基于回归预测的方法通过学习整个(全局)高低分辨率样本空间的映射关系,并用该关系推测高分辨率图像。然而它无法根据低分辨率输入(观察值)对映射关系进行自适应调整。针对以上问题,我们提出一种局部诱导支撑集回归的人脸超分辨率方法。将高低分辨率图像块具有“一致表示系数”的假设放松为高低分辨率图像块的表示系数具有一致的支撑集的假设(更加灵活);在输入低分辨率图像块支撑集上学习高分辨率图像的映射关系(自适应性);在不受图像降质过程影响的高分辨率图像空间,设计一种局部诱导的支撑集生成方法(流形一致);最后,提出一种迭代算法,逐次更新超分辨率图像合成块和支撑集,从而得到最优解。

第五章为基于局部迭代嵌入和字典渐进更新的人脸超分辨率算法。传统基于学习的人脸超分辨率方法均是通过挖掘低分辨率人脸图像空间的几何、重建关系(合成系数),并将这

种关系保持到高分辨率空间，从而得到一张高分辨率人脸图像。当下采样倍数较大或者低分辨率空间受模糊、噪声等因素影响时，高低分辨率图像空间的维度差将非常大（表现为高分辨率流形不一致），对重建效果产生严重的影响；针对上述问题，提出一种“由粗到精”字典渐进更新的人脸超分辨率方法。不仅根据上一步的高分辨率重构结果更新输入低分辨率测试人脸图像，而且对低分辨率人脸训练样本进行更新，减小高低分辨率训练空间的维度差，从而可以更加准确地进行学习与预测。该方法首次将低分辨率人脸字典更新的思想引入到人脸超分辨率重构中，在渐近更新后的低分辨率空间与高分辨率空间学习高低分辨率图像之间的关系。

第六章为总结与展望。总结了本文的研究工作，介绍了本文的主要研究成果和创新点，最后对未来工作进行了研究展望。

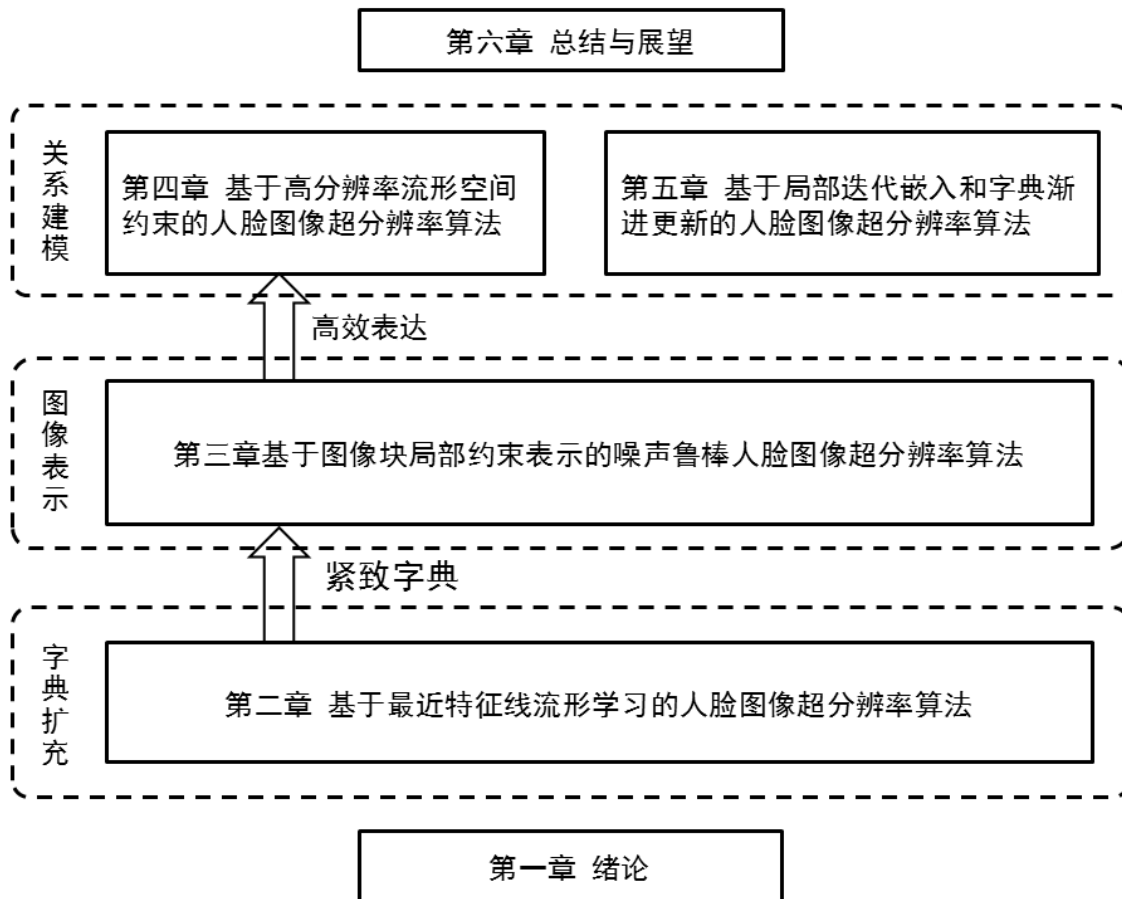


图 1-10 全文组织结构。

第二章 基于最近特征线流形学习的人脸超分辨率算法

2.1 引言

与基于多帧重建的超分辨率算法相比,基于学习的方法只需要输入一张低分辨率图像便可获得更高的放大倍数和更好的效果,而且避免了基于多帧重建方法由于图像配准不精确导致重建效果不理想的问题,因而获得了更多的关注。基于学习的图像超分辨率方法可以根据高低分辨率图像组成的训练库,学习低分辨率图像到高分辨率图像之间的映射关系。输入一张低分辨率图像,就可以利用学习得到的映射关系预测出一张高分辨率图像^[19,29,30,40]。上述映射关系的获取方式多种多样,既可以通过线性或者非线性回归技术学习低分辨率图像和高分辨率图像之间的显示回归关系,又可以通过马尔可夫随机场模型、金字塔模型和近年来备受关系的流形学习技术来学习低分辨率图像到高分辨率图像之间隐式的对应关系。本章主要关注利用流形学习技术进行高低分辨率图像隐式对应关系构建的图像超分辨率方法。

基于流形学习的图像超分辨率方法的基本前提假设是低分辨率样本图像块流形空间与高分辨率样本图像块流形空间具有局部几何结构一致性。给定一张低分辨率图像,他们首先获得输入低分辨率图像块在训练样本图像块空间中的局部几何结构,然后将低分辨率图像块局部几何结构保持到重构后的高分辨率图像块空间中,从而重构出高分辨率图像。近年来许多学者提出了大量基于流形学习的图像超分辨率方法。

例如,Chang 等人^[83]首次将局部线性嵌入流形学习算法引入到图像超分辨率重构中,提出一种邻域嵌入的图像超分辨率重构算法。Wang 和 Tang^[54]提出一种特征变换法,利用主成分分析将输入的低分辨率人脸图像投影到低分辨率训练样本所张成的特征子空间中,再把投影系数直接映射到高分辨率图像上,得到高分辨率图像。文献^[44]利用流形学习理论提出一种人脸超分辨率重构的新两步法,先根据局部保持投影和径向基函数回归得到全局脸,再由基于局部重建的方法补偿人脸特征细节信息。文献^[51]提出一种基于局部保持投影的自适应流形学习人脸超分辨率方法,从人脸图像构成的局部子流形分析人脸的内在特性,有效地重构出低分辨率人脸由于降质丢失的高频细节成分。Huang 等人^[3]提出利用典型相关分析法提取高低分辨率图像的相关子空间,从而增强高低分辨率图像流形空间拓扑结构的一致性。该方法将典型相关分析同时运用于全局脸重构的残差脸的补偿中,最终获得了很好的重构效果。

上述基于流形学习的图像超分辨率方法都是基于高分辨率样本图像块和低分辨率样本图像块分别构成的流形空间具有相似局部几何结构这一假设,从训练样本空间的局部线性几何结构中发掘样本空间的全局非线性结构特征,算法以保持样本空间的局部线性几何结构为基础。然而,在实际情况中,该局部几何结构一致性假设的成立必须满足以下两个前提^[84]:第一,样本数据稠密地采样于潜在的流形空间;第二,样本数据不受噪声的干扰。本章只考虑前者对流形假设的影响,实验样本不受噪声干扰。

在现有人脸超分辨率算法中，样本数据一般均为欠采样（即小样本问题，目前，最大样本库的样本数也不过 2000）。因此，它们只能构成高维人脸流形空间的一个稀疏空间，即使最相邻的点也难以称其为局部。因此，对欠采样的训练样本数据，通过局部几何结构线性保持形成全局非线性结构保持的流形学习思想从本质上并不适合于欠采样样本数据的图像超分辨率重构。

因此，基于流形学习的人脸超分辨率算法要想获得好的学习效果，必须通过扩大训练样本的规模来解决流形空间采样不够稠密的问题。然而，人脸图像库的建立是一个复杂而又耗时的过程。此外，即使能获得大量样本，算法在进行求解计算时也需要巨大的存储空间，且带来非常高的运算复杂度。因此，通过简单地增加样本数量来改进传统基于流形学习的方法在实际应用中并不可行。

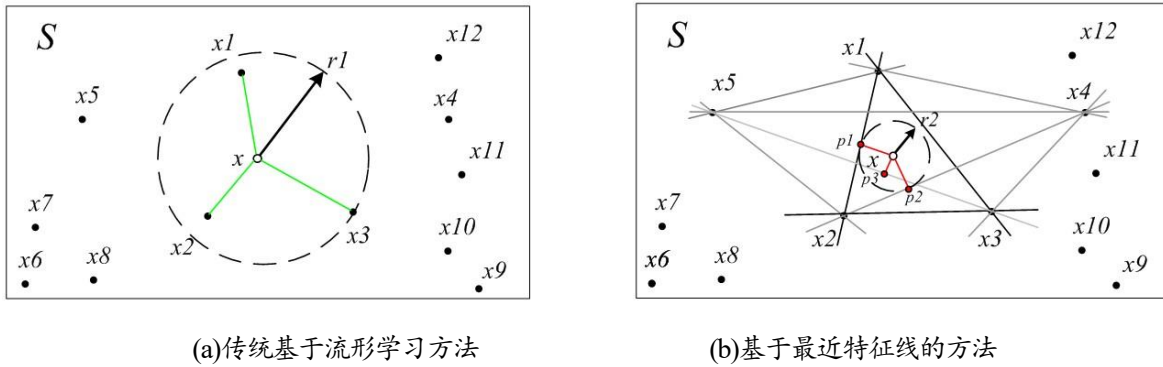


图 2-1 基于流形学习方法和基于最近特征线方法查询点局部保持方式对比。

为此，本文通过引入最近特征线（Nearest Feature Line，简称 NFL）思想^[132, 133]，提出了一种基于最近特征线流形学习的人脸超分辨率算法。图 2-1 为传统流形学习方法和基于最近特征线方法查询点局部保持方式对比示意图。其中，空心圆点表示查询点 x ，实心圆点 $x_1 \sim x_{12}$ 为样本空间 S 的 $N(N=12)$ 个样本点。在图 2-1(a)中， x_1 、 x_2 和 x_3 为查询点 x 的 $K(K=3)$ 个最近邻样本点；在图 2-1(b)中， $x_1 \sim x_5$ 为样本空间 S 中筛选出的 $K'(K'=5)$ 个与 x 近邻的样本点， p_1 、 p_2 和 p_3 为 x 的 K 个最近投影样本点。传统流形学习方法和最近特征线方法分别通过保持查询点 x 与其 K 个近邻样本点（图 2-1 (a) 中半径为 r_1 的圆内的样本点）和 K 个近邻投影样本点（图 2-1 (b) 中半径为 r_2 的圆内的投影样本点）的线性特征来近似高低分辨率人脸流形空间的非线性关系。从图中可以看出， $r_2 < r_1$ ，当样本数 N 增大时，可以选择较大的 K' 值，使 x 的 K 个近邻投影样本点落在一个非常小的区域内 ($r_2 \ll r_1$)。因此，基于最近特征线方法是以保持更小局部的几何结构特征，来揭示高低分辨率人脸流形空间的非线性关系，它更加体现了流形学习局部邻域结构保持的本质。

算法通过保持相对传统基于流形学习算法更小局部内的线性关系来揭示高分辨率样本

图像块和低分辨率样本图像块人脸流形空间的非线性关系，体现了流形学习的本质。提出的方法具有如下几个方面的特点：

(1) 首次将最近特征线思想引入到基于流形学习的人脸超分辨率重构中，它可以把样本空间中任意两个样本点扩充为特征线上的无数多个样本点，从而有效扩充人脸图像样本空间的表达能力；

(2) 算法利用最近特征线准则来定义样本空间中样本点之间的邻域关系，通过保持高低分辨率图像块样本空间的这种邻域关系，从而正确揭示高分辨率和低分辨率人脸流形之间的局部相似结构特征；

(3) 通过对原始样本空间的筛选，再利用特征线方法对其分析，增强了样本对输入查询点的表达能力，并降低算法的运算复杂度。

本章接下来的安排如下：第 2.2 节简要介绍流形学习算法；第 2.3 节介绍了传统基于邻域嵌入的图像超分辨率算法；第 2.4 节详细介绍提出的基于最近特征线流形学习人脸超分辨率算法的核心环节，包括算法所用到的符号的预定义，最近特征线流形学习方法，并针对算法的复杂度问题进行分析，给出一种加速策略；第 2.5 节通过在 CAS-PEAL-R1 人脸库上的一系列实验验证算法的有效性，并与前沿流形学习人脸超分辨率方法进行主客观对比；第 2.6 节对算法进行总结，特别是讨论了算法可能的改进方向。

2.2. 流形学习算法

流形 (Manifold)，是局部具有欧氏空间性质的空间，为嵌入在高维空间中的低维数据集。假设观测数据是均匀采样于高维欧氏空间的低维流形，那么流形学习就是从观察到的高维数据中恢复出潜在的低维流形结构，即找到高维数据空间中的潜在低维流形，并求出相应的嵌入映射，以实现数据维数的约简或者数据的可视化。它是从观测到的现象中去寻找事物的本质，找到产生数据的内在规律。流形学习方法是模式识别和机器学习领域中的基本方法，分为线性流形学习算法和非线性流形学习算法，线性方法包括传统的方法如主成分分析和线性判别分析等，典型的非线性流形学习算法有局部线性嵌入，等距映射和拉普拉斯特征映射等。局部线性嵌入是流形学习中一种最主要的算法，其基本思想被广泛应用于图像超分辨率重构中。

局部线性嵌入算法相关工作最早发表在 Science (2000) 上^[84]，是非线性降维技术的里程碑。它可以通过非监督的方式将高维空间非线性映射到低维空间中，该技术被广泛地应用于图像分类与聚类、多维数据的可视化等领域。局部线性嵌入算法的基本思想为保持原嵌入流形中局部邻域间的几何结构（相互重构关系），将高维观测数据映射到潜在低维流形中，具体算法实现分以下三步：

(1) 邻域点选择。

设 D 维空间中有 N 个数据样本 $X = \{x_i\}_{i=1}^N$, $x_i \in \mathbb{R}^D$ 。如果这些数据样本足够稠密, 可以认为数据样本空间中的任意一个样本点可以用它的 K 个近邻样本或者固定半径 ε 的球状邻域内的近邻样本线性表示。

(2) 计算重构权重 W 。

对于每一个样本点 x_i , 计算能最好地重构它的权值 w_{ij} , 使重构误差最小。且重构误差定义为:

$$\varepsilon(W) = \sum_i \left\| x_i - \sum_j w_{ij} x_j \right\|_2^2, \quad (2-1)$$

并且权值要满足两个约束条件: 一方面, 每一个数据点 x_i 都只能由它的邻近点来表示, 若 x_j 不是近邻点, 则 $w_{ij} = 0$; 另一方面, 权值矩阵的每一行的和为 1, 即 $\sum_j w_{ij} = 1$ 。因此, 求解 W 的过程其实就是求解带约束的最小二乘问题

(3) 获得低维嵌入空间 \mathbb{R}^d 。

保持权值 W 不变, 优化 d 维坐标下低维嵌入 $y_i \in \mathbb{R}^d$, 映射嵌入满足如下条件:

$$\varepsilon(Y) = \sum_i \left\| y_i - \sum_j w_{ij} y_j \right\|_2^2. \quad (2-2)$$

与步骤 (2) 中的误差函数类似, 不过在这里是固定重建权值 W 优化 Y 。由文献[84]可知, 对 Y 的求解也就是一定约束条件下求解稀疏矩阵的特征向量问题。

以上三步可以用图 2-2 进行说明。

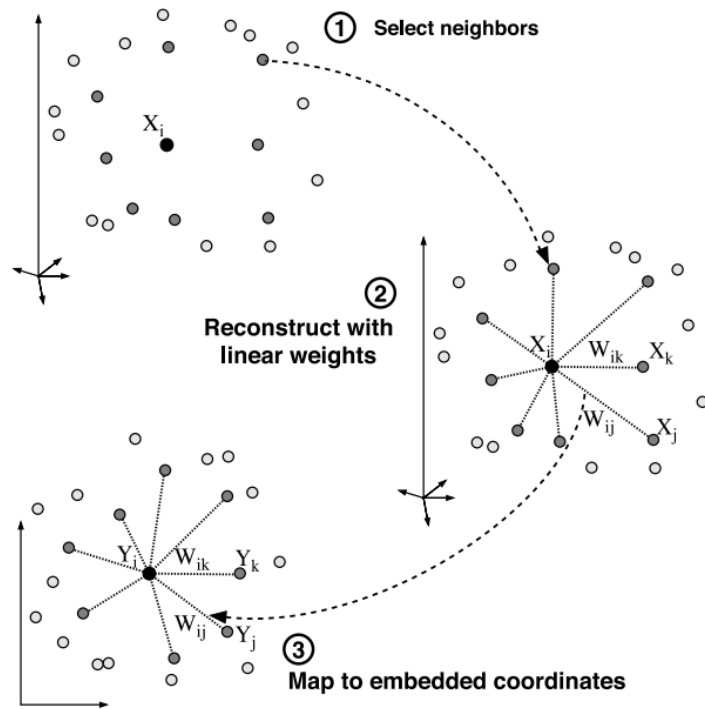


图 2-2 局部线性嵌入流形学习算法流程图[84]。

2.3 基于邻域嵌入的图像超分辨率算法

基于低分辨率样本图像块流形空间与对应的高分辨率样本图像块流形空间在局部几何结构上具有一致性，邻域嵌入的图像超分辨率方法使用局部线性嵌入流形学习的思想实现高分辨图像的重构。如图 1-7 是基于邻域嵌入的图像超分辨率算法的基本思想。在训练阶段，提取低分辨率样本图像块（特征²）和对应高分辨率样本图像块（特征），构成高低分辨率图像块字典对。在合成阶段，输入一张低分辨率图像，对于每个低分辨率图像块（特征），首次在低分辨率图像块字典中找到 K 个最相似的低分辨率图像块，然后用这 K 个图像块对输入低分辨率图像块进行合成并得到重构权重，再利用 K 个低分辨率图像块对应的 K 个高分辨率图像块和上述得到的重构权重，线性加权得到目标高分辨率图像块，最后，融合所有重构的高分辨率图像块便可得到高分辨率重构图像。算法 2-1 给出了传统基于邻域嵌入算法的超分辨率重建过程。

算法2-1 基于邻域嵌入的图像超分辨率算法

1. **输入：** 高低分辨率图像块字典对 $X = \{x_i\}_{i=1}^N$ and $Y = \{y_i\}_{i=1}^N$ ，低分辨率图像 $X' = \{x'_i\}_i^M$ ，近邻个数 K ；
2. **输出：** 高分辨率图像 Y' ；
3. 对于输入低分辨率图像上的每一个图像块 x'_i ；
4. 从 $X = \{x_i\}_{i=1}^N$ 中找到 x'_i 的 K 个近邻图像块，下标记为 $C_K(x'_i)$ ；
5. 计算重构权重：

$$w^* = \arg \min_w \left\{ \left\| x'_i - \sum_{k \in C_K(x'_i)} x_k w_k \right\|_2^2 \right\};$$

6. 线性加权 K 个低分辨率图像块对应的 K 个高分辨率图像块和上述重构权重，获得的高分辨估计：

$$y'_i = \sum_{k \in C_K(x'_i)} y_k w_k^*$$

7. 重复步骤3-步骤5，直到处理完所有低分辨率图像块；
 8. 按位置融合所有合成得到的高分辨率图像块，对于交叠位置直接加权求平均，最终得到输入低分辨率图像 X' 的高分辨率估计图像 Y' 。
-

2.4 基于最近特征线流形学习的人脸超分辨率算法

本节将提出一种基于最近特征线样本扩展的局部流形保持人脸图像超分辨率方法。2.4.1

² 一阶梯度或者二阶梯度信息

小节介绍一些符号的定义, 2.4.2 节详细介绍了基于最近特征线流形学习的人脸超分辨率算法, 2.4.3 小节针对算法复杂度过度的问题, 提出一种样本图像块选择策略。

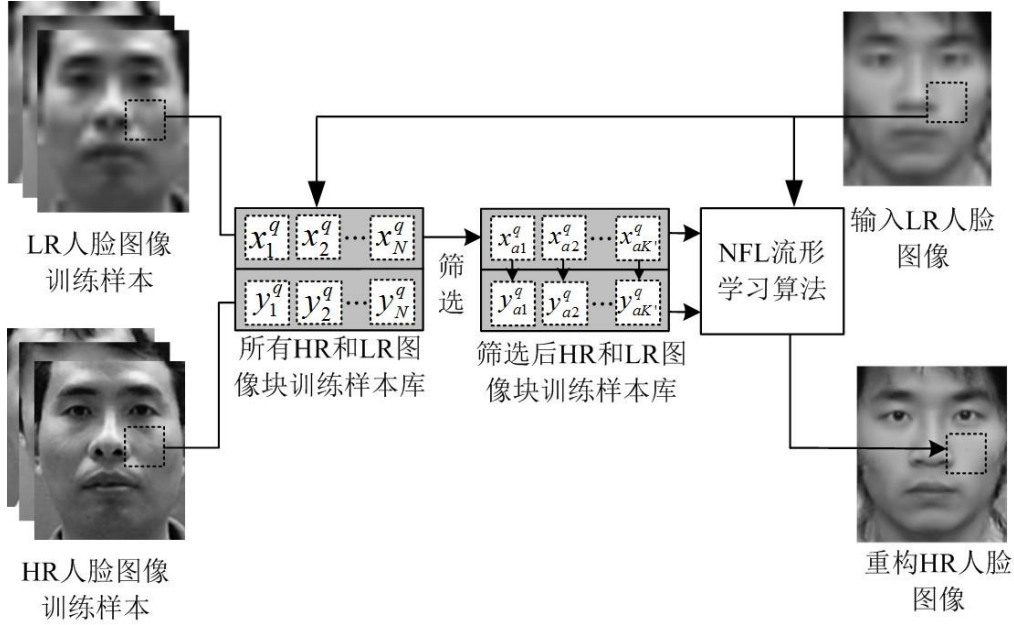


图 2-3 基于最近特征线的人脸超分辨率流程图。

图 2-3 为本文算法流程框图。对于某一个位置块, 高分辨率样本图像块和低分辨率样本图像块组成训练库, 首先根据输入低分辨率图像块, 在低分辨率图像块样本空间中筛选出与其最匹配的低分辨率图像块, 利用这些低分辨率图像块和对应的高分辨率图像块组成新的训练库, 然后通过最近特征线流形学习算法合成出高分辨率图像块, 最后融合所有得到的高分辨率图像块, 便可重构出高分辨率人脸图像。

2.4.1 符号定义

给定一组由高分辨率人脸图像样本和低分辨率人脸图像样本对组成的训练样本库 $\langle X, Y \rangle = \{\langle x_i, y_i \rangle | 1 \leq i \leq N\}$, 其中 N 为样本的个数, 低分辨率人脸样本 x_i 由高分辨率人脸样本 y_i 下采样得到 (同时用双三次插值算法将输入低分辨率人脸图像上采样到与高分辨率人脸样本尺寸相同大小)。对于基于位置块^[1, 43]的人脸超分辨率重构方法, 先将高分辨率人脸样本和低分辨率人脸样本划分为 M 个交叠的块 $\langle X, Y \rangle = \{\langle x_i^j, y_i^j \rangle | 1 \leq i \leq N, 1 \leq j \leq M\}$ 。输入一张低分辨率人脸图像 x_i , 用上述方法将其划分为 M 个块, $x_i = \{x_i^j | 1 \leq j \leq M\}$, 分块重构高分辨率人脸图像块 y_i^j , 融合所有重构得到的高分辨率人脸图像块便可以得到最终的高分辨率人脸图像 y_i 。

2.4.2 基于最近特征线流形学习的人脸超分辨率算法

计算机低维视觉研究表明, 人脸是一类高度结构化的物体, 具有很强的结构相似性, 对

齐后的人脸在同一位置上的纹理变化是平缓的。对于两张不同的人脸，可将同一位置的两个块图像（图像块按行或者列扫描得到的一列像素值）看成高维空间中的两个点，它们的连线（特征线）表示某一种因素（如光照、纹理等）引起的平缓变化过程，该特征线就可以近似这个位置块特征样本的轨迹。如图 2-4 所示，利用特征连接两种不同光照的人脸图像，便可生成多种很多其它光照的人脸图像，从而将原来的两个样本点扩充为了特征线上的无数多个样本点。

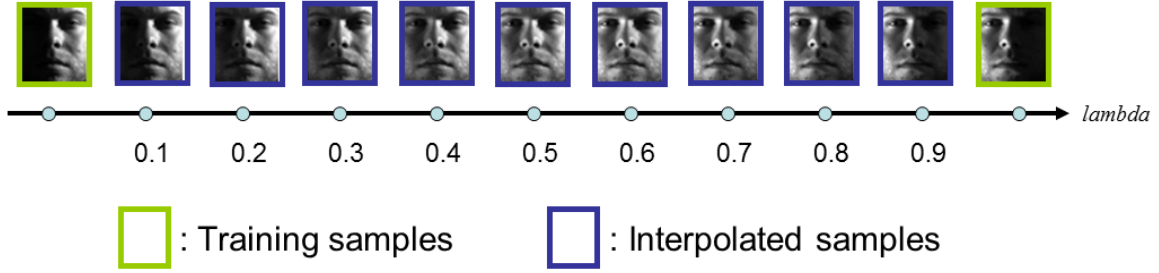


图 2-4 利用特征线扩展样本点示意。绿色框人脸为两个观察样本，蓝色框人脸为特征线插值得到的人脸图像。

利用特征线的这种扩展特性，我们将其引入到图像超分辨率重构中，提出一种基于最近特征线流形学习的人脸超分辨率方法。把人脸图像块样本空间中的样本点两两相连形成特征线，将已有样本点扩充为特征线上的无数多个样本点，从而增强人脸图像块样本空间的表达能力；在扩充之后的样本空间中，利用最近特征线准则来定义样本空间中样本点之间的近邻关系，通过保持高低分辨率图像块样本空间的这种邻域关系，从而正确揭示高低分辨率人脸图像块流形之间的局部相似结构特征；通过保持这种局部相似结构特征，重构输入的低分辨率人脸图像块对应的高分辨率块，融合所有高分辨率块得到高分辨率人脸图像。下面以合成第 q 个低分辨率块 x_i^q 为例，根据该位置块的样本 $\langle X^q, Y^q \rangle = \{\langle x_i^q, y_i^q \rangle | 1 \leq i \leq N\}$ 提供的先验信息，分以下六步来说明提出的算法中重构高分辨率块 y_i^q 的过程：

(1) 将第 q 个低分辨率训练样本块所组成的集合 $\{x_i^q | 1 \leq i \leq N\}$ 中的任意两点 $x_{I_1}^q$ 和 $x_{I_2}^q$ ($1 \leq I_1 < I_2 \leq N$) 连成直线形成 $N(N-1)/2$ 条特征线 $\overline{x_{I_1}^q x_{I_2}^q}$ 。如图 2-5 所示， $x_{I_1}^q$ 和 $x_{I_2}^q$ 为低分辨率样本空间上的两个特征点；

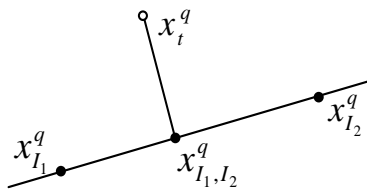


图 2-5 特征点、投影点和特征线。

(2) 计算 x_t^q 在特征线 $\overline{x_{I_1}^q x_{I_2}^q}$ 上的投影 x_{I_1, I_2}^q (如图 2-5):

$$x_{I_1, I_2}^q = x_{I_1}^q + u_{I_1, I_2}^q (x_{I_1}^q - x_{I_2}^q), \quad (2-3)$$

其中, $u_{I_1, I_2}^q = \frac{(x_t^q - x_{I_1}^q)(x_{I_2}^q - x_{I_1}^q)}{(x_{I_2}^q - x_{I_1}^q)(x_{I_2}^q - x_{I_1}^q)}$ 。那么, x_t^q 到特征线 $\overline{x_{I_1}^q x_{I_2}^q}$ 的距离便可定义为:

$$d(x_t^q, \overline{x_{I_1}^q x_{I_2}^q}) = \|x_t^q - x_{I_1, I_2}^q\|, \quad (2-4)$$

这里, $\|\bullet\|$ 为欧氏距离;

(3) 计算 x_t^q 的 K 个最近邻投影样本点。因为每条特征线表征由某一因素引起的纹理上的平缓变化, 所以可以假定 K 个最近邻投影样本点位于不同的特征线上。因而, 计算 x_t^q 的 K 个最近邻投影样本点就是根据第(2)步所计算出的距离, 找出 x_t^q 与所有特征线的 $N(N-1)/2$ 个投影点中距离最近的 K 个, 记作 $\{x_{a,b}^q\}_{(a,b) \in A(q)}$, 其中, (a,b) 为最近邻投影样本点所在特征线的下标号对, $A(q)$ 为 K 对下标号对组成的集合;

(4) 计算重构权重 W^q 。用以上 K 个最近邻投影样本点 $\{x_{a,b}^q\}_{(a,b) \in A(q)}$ 线性合成 x_t^q , 最小化目标函数 \mathcal{E}^q :

$$\mathcal{E}^q = \|x_t^q - \sum_{(a,b) \in A(q)} w_{a,b}^q x_{a,b}^q\|, \quad (2-5)$$

且 $\sum_{(a,b) \in A(q)} w_{a,b}^q = 1$ 。

为了求解 W^q , 对 x_t^q 引入局部 Gram 矩阵 G_q :

$$G_q = (x_t^q \mathbf{1}^T - M_P^q)^T (x_t^q \mathbf{1}^T - M_P^q), \quad (2-6)$$

其中, $\mathbf{1}$ 是 K 维的全 1 向量, M_P^q 为 $|x_t^q| \times K$ 的矩阵 ($|x_t^q|$ 为块图像 x_t^q 中像素的个数), M_P^q 的每一列元素由 x_t^q 的最近邻投影样本点构成。因此, 最小化公式 (3) 有如下形式闭解:

$$W^q = \frac{G_q^{-1} \mathbf{1}}{\mathbf{1}^T G_q^{-1} \mathbf{1}}. \quad (2-7)$$

(5) 计算与 x_t^q 的 K 个低分辨率最近邻投影样本点对应的 K 个高分辨率最近邻投影样本点

$\{y_{a,b}^q\}_{(a,b) \in A(q)}$:

$$y_{a,b}^q = y_a^q + u_{a,b}^q (y_a^q - y_b^q), \quad (2-8)$$

$u_{a,b}^q$ 为第②步所计算出来的 u_{I_1, I_2}^q 在 $I_1 = a, I_2 = b$ 时的取值;

(6) 用上述得到的 K 个高分辨率最近邻投影样本点 $\{y_{a,b}^q\}_{(a,b) \in A(q)}$ 线性合成高分辨图像块 y_i^q ，合成系数为 w^q ：

$$y_i^q = \sum_{(a,b) \in A(q)} w_{a,b}^q y_{a,b}^q \quad (2-9)$$

对于输入低分辨率图像上的所有块，均利用上述六步来合成其对应的高分辨率块，将所有合成得到的高分辨率块按位置块叠加起来，然后除以每个像素位置交叠的次数，便可重构出高分辨率人脸图像。

2.4.3 最近特征线人脸超分辨率方法的改进

特征线思想的引入，增强了样本库的表达能力，更好地保持了高分辨率和低分辨率人脸流形空间的局部几何结构。但同时也带来了两个严重的问题：第一，计算复杂度急剧增加。当特征线上所有的点变成样本点之后，计算低分辨率输入查询点 x_i^q 的 K 个最近邻投影样本点转化为计算 x_i^q 与所有 $N(N-1)/2$ 条特征线上的投影的距离，而传统方法只需要计算 x_i^q 与 N 个样本点的距离；第二，当距离 x_i^q 很远的两点所形成的特征线位于 x_i^q 附近时，特征线方法将失效。如图 2-6， $x_{1,2}^q$ 和 $x_{3,4}^q$ 分别是 x_i^q 在特征线 $\overline{x_1^q x_2^q}$ 和 $\overline{x_3^q x_4^q}$ 上的投影， x_i^q 到特征线 $\overline{x_3^q x_4^q}$ 的距离更短，但它离样本点 x_1^q 和 x_2^q 更近。在此情况下， x_i^q 与 $x_{1,2}^q$ 有更相似的特征（纹理），但是，最近特征线方法将优先选择 $x_{3,4}^q$ 来表示 x_i^q ，这与通常情况并不符合，这也是最近特征线方法用在人脸识别和语音分类等问题上的最大障碍。

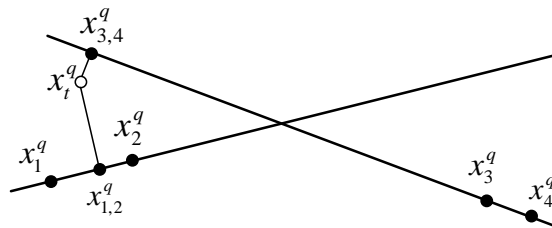


图 2-6 最近特征线法失效的情形。

针对以上两个问题，本章提出了一种统一的解决方案：对于低分辨率输入块 x_i^q ，首先对 N 个样本点进行筛选，从 N 个样本中选出 K' 个与 x_i^q 最相邻的样本 $\{x_{a1}^q, x_{a2}^q, \dots, x_{aK'}^q\}$ ，再利用 2.3.2 节提出的最近特征线方法来分析这 K' 个样本点形成的子空间的局部几何结构。通过选择恰当的 K' 值，一方面，可以减小算法的运算复杂度（运算复杂度从 $O(N^2)$ 降低到 $O(N + K'^2)$ ， $K' \ll N$ ），另一方面，由于所选 K' 个样本本身为 x_i^q 的近邻，它们所形成的特

征线对 x_i^q 具有更好的表达能力。因此，通过对样本空间进行筛选，不仅可以使运算复杂度大大降低，而且新得到的样本对低分辨率输入块具有更好的表达能力。

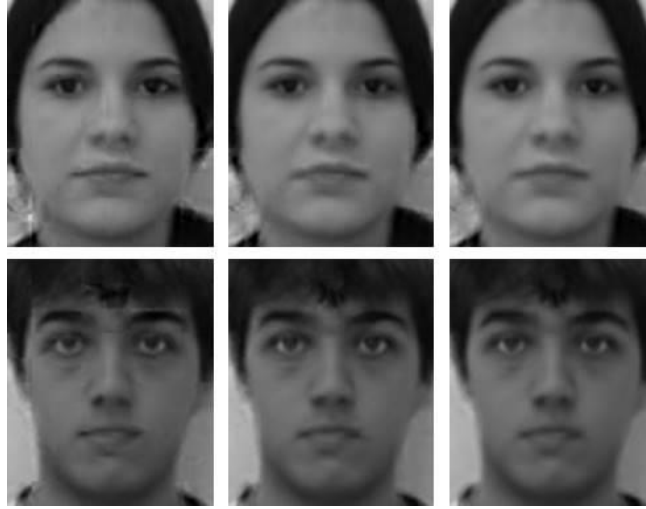


图 2-7 图像块交叠对重构结果的影响。输入低分辨率人脸图像块大小为 3×3 ，第一列为相邻图像块之间交叠像素分别是 0 个像素的结果；第二列为相邻图像块之间交叠像素分别是 1 个像素的结果；第三列为相邻图像块之间交叠像素分别是 2 个像素的结果。

2.4.4 人脸图像小块划分

基于图像块的人脸超分辨率方法中，相邻图像块的重建是分别进行的，如不考虑他们的兼容性，重构的高分辨率人脸图像将具有非常明显的块效应。如图 2-7 所示为不同程度的图像块交叠对重构结果的影响。可以看出如果相邻图像块之间没有交叠，重构人脸将存在明显的块效应，交叠越多重构结果越好。图 2-8 为人脸图像小块划分时边缘图像块的“回退”策略示意。具体地，按照从左到右、从上到下的顺序进行划分图像块，当划分图像块至图像边缘时，如果剩余尺寸比预先设置的图像块的尺寸小则以原图像的边缘为基准进行回退划分，包括当横向划分到图像右边边缘处时，向左回退并以右边边缘为基准进行分块；当纵向划分到图像底边边缘处时，向上回退并以底边边缘为基准进行分块。

算法 2-2 给出了提出的基于最近特征线流形学习算法的人脸超分辨率重建过程。

算法2-2 基于最近特征线流形学习的人脸图像超分辨率算法

1. **输入:** 高分辨率人脸图像训练样本集合 $\{y_i\}_{i=1}^N$ ，低分辨率人脸图像训练样本集合 $\{x_i\}_{i=1}^N$ ，低分辨率图像 x_i 近邻个数 K ;
 2. **输出:** 高分辨率图像 y' ;
-

3. 按照人脸图像的位置信息，对高低分辨率人脸图像训练样集合对和输入低分辨率人脸图像进行分块： $\{y_i^j | 1 \leq i \leq N, 1 \leq j \leq M\}$ ， $\{x_i^j | 1 \leq i \leq N, 1 \leq j \leq M\}$ 和 $\{x_i^j | 1 \leq j \leq M\}$ ；
4. 对于输入低分辨率人脸图像上的每一个图像块 x_t^q ；
5. 分别提取高低分辨率训练集中第 q 个图像块，形成高低分辨率训练图像块集合 $H^q = \{y_i^q | 1 \leq i \leq N\}$ 和 $L^q = \{x_i^q | 1 \leq i \leq N\}$ ；
6. 从低分辨率训练图像块集合 L^q 中选出与图像块 x_t^q 欧氏距离最近的 K' 个图像块，形成筛选后的低分辨率训练图像块集合 $L_K^q = \{x_i^q | x_i^q \in C(x_t^q), |C(x_i^q)| = K'\}$ ， $C(x_i^q)$ 表示图像块 x_i^q 的邻域集合， $|C(x_i^q)|$ 表示邻域集合 $C(x_i^q)$ 中图像块的个数；
7. 将筛选后的低分辨率训练图像块集合 L_K^q 中所有图像块作为样本点两两相连，形成 $C_K^2 = \frac{K'(K'-1)}{2}$ 条特征线，构成的特征线集合表示成 $L_L^q = \{\overline{x_{i_1}^q x_{i_2}^q} | 1 \leq i_1 < i_2 \leq N\}$ ，其中 $x_{i_1}^q$ 和 $x_{i_2}^q$ 为低分辨率训练图像块集合 L_K^q 中的两个样本点， $\overline{x_{i_1}^q x_{i_2}^q}$ 为连接样本点 $x_{i_1}^q$ 和 $x_{i_2}^q$ 生成的特征线；
8. 计算图像块 x_t^q 在特征线集合 L_L^q 中所有特征线上的投影点，构成投影集合 $L_P^q = \{x_{i_1, i_2}^q | 1 \leq i_1 < i_2 \leq N\}$ ， $x_{i_1, i_2}^q = x_{i_1}^q + u_{i_1, i_2}^q (x_{i_1}^q - x_{i_2}^q)$ ，其中， x_{i_1, i_2}^q 表示图像块 x_t^q 在特征线 $\overline{x_{i_1}^q x_{i_2}^q}$ 上的投影点， u_{i_1, i_2}^q 为位置参数， $u_{i_1, i_2}^q = \frac{(x_t^q - x_{i_1}^q)(x_{i_2}^q - x_{i_1}^q)}{(x_{i_2}^q - x_{i_1}^q)(x_{i_2}^q - x_{i_1}^q)}$ ；
9. 计算图像块 x_t^q 到特征线集合 L_L^q 中所有特征线的距离，通过求取图像块 x_t^q 与投影点 x_{i_1, i_2}^q 的距离实现，距离 $d(x_t^q, \overline{x_{i_1}^q x_{i_2}^q}) = \|x_t^q - x_{i_1, i_2}^q\|$ ，其中， $\|x_t^q - x_{i_1, i_2}^q\|$ 表示图像块 x_t^q 到投影点 x_{i_1, i_2}^q 的欧氏距离；
10. 根据上一步得到的 $d(x_t^q, \overline{x_{i_1}^q x_{i_2}^q})$ ，查找 K 个距离最小的投影点 x_{i_1, i_2}^q ，构成集合 $L_K^q = \{x_{a, b}^q\}_{(a, b) \in A(q)}$ ，其中， $A(q)$ 为 K 个距离最小的投影点 x_{i_1, i_2}^q 的下标号 i_1, i_2 所构成的集合，集合 L_K^q 为图像块 x_t^q 在低分辨率人脸样本块空间上的 K 个最近的投影点；
11. 利用低分辨率人脸样本块空间上的 K 个最近的投影点对对于输入低分辨率人脸图像块线性重构，得到线性重构权重；
12. 计算在高分辨率人脸样本块空间 H^q 对应的 K 个高分辨率最近投影点，

$y_{a,b}^q = y_a^q + u_{a,b}^q(y_a^q - y_b^q)$, $u_{a,b}^q$ 为步骤9所得 u_{i_1,i_2}^q 在 $i_1 = a, i_2 = b$ 时的取值;

13. 将步骤10所得低分辨率人脸样本块空间上的 K 个最近的投影点, 替换为步骤12所得高分辨率人脸样本块空间上的 K 个样本点, 使用步骤11所得权重系数, 加权重构出高分辨率的图像块;
14. 重复步骤5-步骤13, 直到处理完所有低分辨率图像块;
15. 将所有加权重构出高分辨率的图像块按照位置叠加, 然后除以每个像素位置交叠的次数, 重构出高分辨率人脸图像。

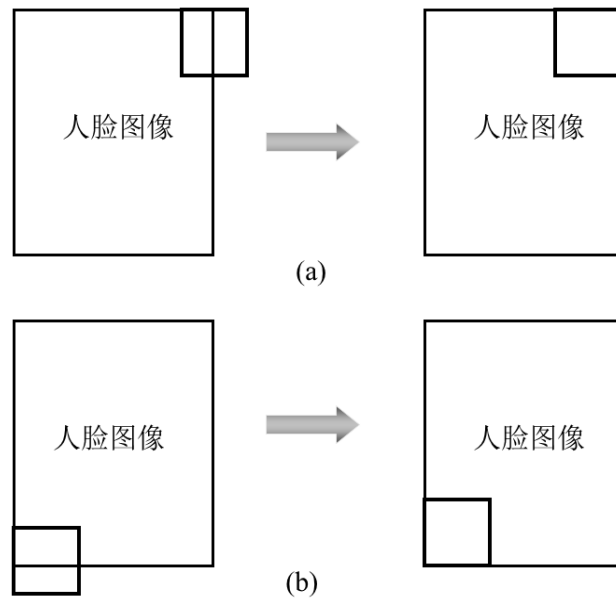


图 2-8 人脸图像小块划分时边缘图像块的“回退”策略示意。

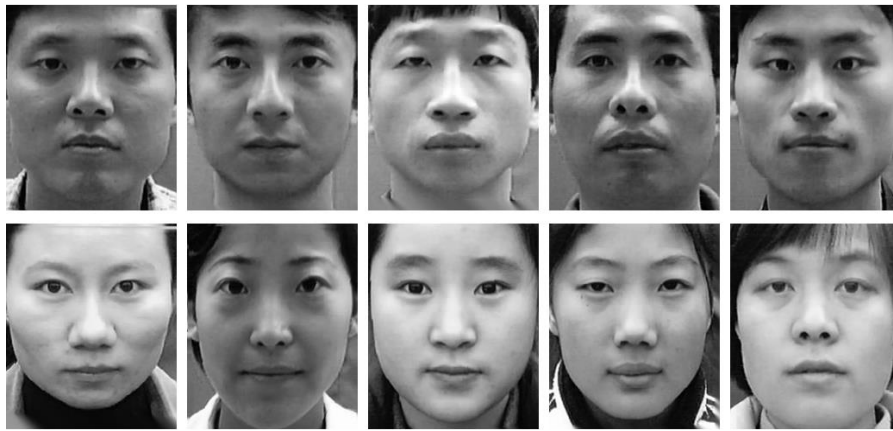


图 2-9 CAS-PEAL-R1 人脸库示例。

2.5 实验及结果分析

2.5.1 人脸库介绍

本章采用 CAS-PEAL-R1 人脸库^[4]进行实验, 选用所有 1040 个个体的中性表情、正常光照下的正面人脸图像。抠取人脸区域并将其裁剪成 112×100 像素, 再手工标定人脸上的五个特征点(两眼中心、鼻尖和两个嘴角)并进行仿射变换对齐, 得到原始的高分辨率人脸图像(如图 2-9)。低分辨率人脸图像由高分辨率人脸图像 4 倍下采样后再 4 倍 Bicubic 上采样得到。随机选择 1000 张作为训练样本, 将剩余 40 张作为测试图像。

2.5.2 算法参数设置

本算法涉及四个参数, 即最近邻投影点数 K 和预筛选块个数 K' , 以及高分辨率和低分辨率块大小和块之间交叠的像素。实验表明, 当 K 取 3 时, 重构效果最好。由 2.2 小节的讨论, 在确定 K' 的值时, 既要考虑所选出样本的表达能力, 又要考虑算法的复杂度, K' 的取值不能太大。实验发现, 当 $K'=25 \sim 35$ 时, 可以获得较好的重构效果。参考文献^[83]和文献^[85]中参数设置的方法, 我们也将高分辨率和低分辨率块的大小分别设为 7×7 , 交叠 4 个像素。

2.5.3 不同算法主客观对比结果

图 2-10 为经典的流形学习算法^[54]、目前文献报道中最好的流形学习算法^[3]和本章提出算法的重构结果对比(受篇幅所限, 只列出其中 7 组结果)。实验中, 调节对比算法的参数, 使其效果达到最佳。文献^[54]算法(图 2-10 (b))重构的结果有很明显的人工痕迹, 而且在轮廓部分存在不同程度的鬼影。这主要是因为文献^[54]算法在主成分空间中寻求高分辨率和低分辨率图像样本的关系来合成全局脸, 从而丢失了过多的细节信息。进一步观察人脸图像的细节就会发现, 文献^[3]算法重构的人脸(图 2-10 (c))比较平滑且在轮廓部分不够清晰, 本算法重构结果具有更加清晰的人脸边缘轮廓和更多的局部特征细节(图 2-10 (d)), 原因在于, 文献^[3]虽然在其算法第二步运用流形学习方法进行了残差脸补偿, 但它只是在一个更大的“局部”来保持高分辨率和低分辨率图像流形的几何特征, 在合成高分辨率块时会产生很大误差, 主要表现为局部边缘的重影和毛刺现象。

另外, 从所有 40 幅测试人脸重构效果的客观峰值信噪比(Peak Signal to Noise Ratio, 简称 PSNR)盒图(图 2-11)可以看出, 与 Bicubic 插值算法相比, 两种对比算法和本章算法的重构结果客观 PSNR 值均有不同程度的提升。另外, 本章算法重构结果的平均 PSNR 值比文献^[54]和文献^[3]的算法分别高出 2.6dB 和 1.4dB。



图 2-10 不同算法主观重构效果对比: (a) 输入低分辨率图像 (b) 双三次插值结果 (c) Wang 全局特征转制方法 (d) Huang 的 CCA 方法 (e) 本章算法的效果 (f) 原始高分辨率图像。(在电子版上放大观看, 不同算法的对比效果将更加明显)。

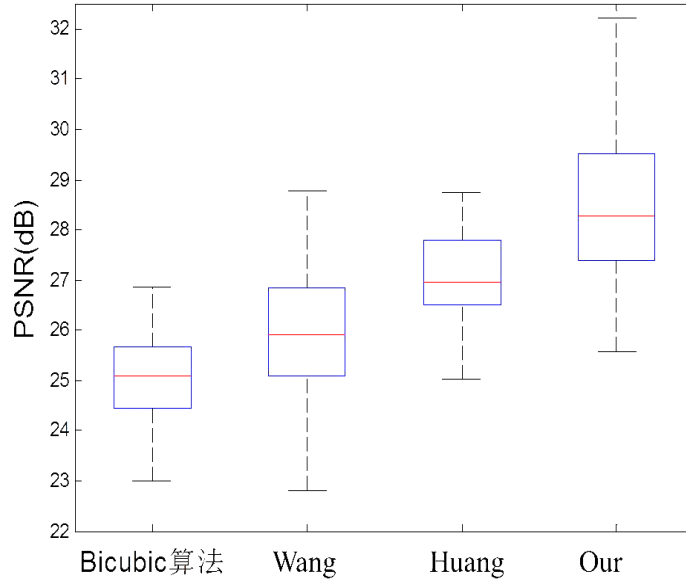


图 2-11 不同算法 PSNR 值对比(四种算法的平均 PSNR 值分别为 25.03 dB, 25.96 dB, 27.10 dB, 28.51dB)。

2.5.4 训练库大小对算法结果的影响

本小节将测试在欠采样条件下提出算法和传统基于邻域嵌入方法在性能上的差别。为了对比公平起见，我们为上述两种方法选择同样的位置块高低分辨率图像块训练字典。这两种方法唯一的差别就是 K 近邻样本获取方式。在提出方法中，我们选取了 K 个最近邻的投影特征样本，而传统基于邻域嵌入方法直接选择了 K 个最近邻的样本。表 2-1 给出了两种方法在不同训练库大小下的参数设置情况。

图2-12给出了提出方法在不同训练库大小下相当于传统基于邻域嵌入方法的客观 PSNR 值提升结果。从该图可以看出，提出的方法在绝大多数图像上的重建结果均好于对比算法。而且，随着训练样本库的减小，提出方法相对于对比算法的增益更加明显。例如，当训练库为 1000 时，提出方法相对于对比算法的 PSNR 增益为 0.55 dB。当训练库样本数据减少到 400 时，而这一增益值变成了 1.0 dB。说明了提出方法在欠采样条件下（训练样本数据比较小）仍然可以获取很好的效果。这主要归因于提出方法可以有效地扩充原始样本点的表达能力。

表 2-1 两种不同方法的近邻个数。

训练库大小	传统邻域嵌入方法	提出方法	
	K	K	K'
400	5	3	20
700	5	3	25
1000	7	3	30

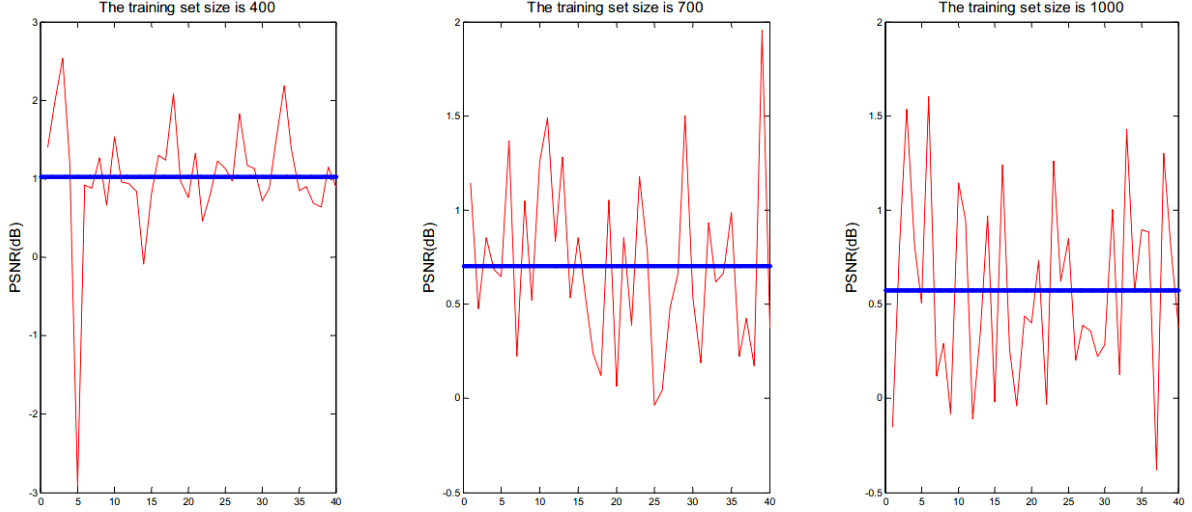


图 2-12 提出方法在不同训练库大小下相较于传统基于邻域嵌入方法的客观 PSNR 增益。横坐标为测试图像序号，红线为提出方法相对于传统基于邻域嵌入方法在不同测试图像上的 PSNR 增益，蓝线为所有 40 幅测试图像在某一训练集大小的 PSNR 增益。

2.6 总结与展望

针对在欠采样（训练样本数有限）的条件下，传统基于流形学习的人脸超分辨率方法无法正确揭示高分辨率和低分辨率人脸流形空间之间的相似局部几何结构的问题，本章首次将最近特征线思想引入到人脸超分辨率的流形学习中，提出一种基于最近特征线的人脸图像超分辨算法。算法利用最近特征线来扩充原始人脸图像样本的表达能力和定义样本点之间的邻域关系，以保持人脸图像块的流形结构于很小的局部为目标，揭示高分辨率和低分辨率人脸流形之间的局部相似结构特征，体现了流形学习局部邻域结构保持的本质。此外，对于输入查询点，算法先对原始样本空间进行筛选，再利用特征线方法对其分析，不仅减小了算法的复杂度，而且增强了样本的表达能力。算法所重构的高分辨率人脸图像具有更丰富的局部细节特征，验证了算法的有效性。

在本章的实验中，只用到了图像的像素值作为每个块的特征，而像素值特征容易受到光照、噪声等因素的影响。因此，在下一步的工作中，将会考虑利用更能反应图像块内容的特征，比如图像块的直方图特征、方向梯度特征、特征子空间系数特征等等，相信可以取得更好的重构效果。

第三章 基于图像块局部约束表示的噪声鲁棒人脸超分辨率算法

3.1 引言

根据流形学习理论，2004年香港科技大学的Chang等人^[83]基于高分辨率样本空间和低分辨率样本空间具有相似局部几何特征这一假设，提出一种基于邻域嵌入（Neighbor Embedding，简称NE）的图像超分辨率重构方法，通过线性加权近邻样本点重构目标高分辨率图像块，该方法利用少量训练样本便可获得很好的重构效果。由于NE方法^[83]对样本的处理具有较好的灵活性和对观察信号（图像块）具有较强的表达性能，因此受到了超分辨率研究学者们的广泛关注，近年来学者们相继提出了大量基于流形学习的图像超分辨率方法。例如，Fan等人^[85]提出一种图像基元和邻域嵌入结合的方法，能得到较传统邻域嵌入方法更好的重构结果。李波等人从流形对齐的角度出发，提出一种公共流形（Common Manifold）的思想，对Chang等人的NE方法^[83]进行了有效的扩展。他们将高低分辨率样本空间映射到一个公共的流形空间，在此公共流形空间进行邻域嵌入学习。

但是，由于上述方法所选择的近邻个数 K 是固定的，在对输入图像块进行表示的过程中会出现过拟合或拟合不当的问题。为此，2010年Yang等人提出首次将稀疏编码（Sparse Coding）理论引入到图像超分辨率重构过程中，提出基于稀疏表示的图像超分辨率方法。为了对观察得到的图像块进行表示，该方法在图像块表示的目标函数中施加稀疏正则约束项。一方面，要求观察图像块能够被准确表达，另一方面，要求用来重构的样本图像块比较少。利用上述稀疏编码策略，可以达到选取尽量少的样本图像块准确地表示观察图像块。与固定邻域个数的邻域嵌入学习方法相比，Yang等人提出的稀疏表示方法可以根据观察图像块自适应地确定所选样本图像块的个数，解决了基于邻域嵌入图像超分辨率方法中图像块表示过拟合或拟合不当的问题。

然而，上述方法没有充分利用人脸图像的结构先验信息，超分辨率重构效果不佳。众所周知，人脸图像是一类高度结构化的物体，所有人脸均是由两只眼睛、鼻子和嘴巴等相对位置稳定的器官构成，该位置先验信息在人脸的合成与重构过程中具有非常重要的作用。基于上述认识，近年来学者们提出许多基于位置图像块的人脸超分辨率算法。图3-1是这类算法的基本原理框架图，首先，输入低分辨率人脸图像和高低分辨率样本人脸图像样本都按照同样的划分方式被划分成多个图像小块；其次，对于输入的每一个低分辨率图像块，用所有低分辨率人脸样本相同位置的图像块对其进行表示得到表示系数；然后，把低分辨率样本图像块替换成对应的高分辨率样本图像块，同时保持表示系数不变，便可线性加权得到重构的高分辨率图像块；最后，重复上述过程，得到所有位置的重构高分辨率人脸图像块，按照重构出来

的人脸图像块在人脸图像上的位置，对重构得到的高分辨率人脸图像块进行融合，最终得到输入低分辨率人脸图像的目标高分辨率人脸图像。由上述重构步骤可知，基于位置图像块的人脸超分辨率方法的关键就是如何获得输入低分辨率人脸图像块在低分辨率样本人脸图像空间的表示系数。

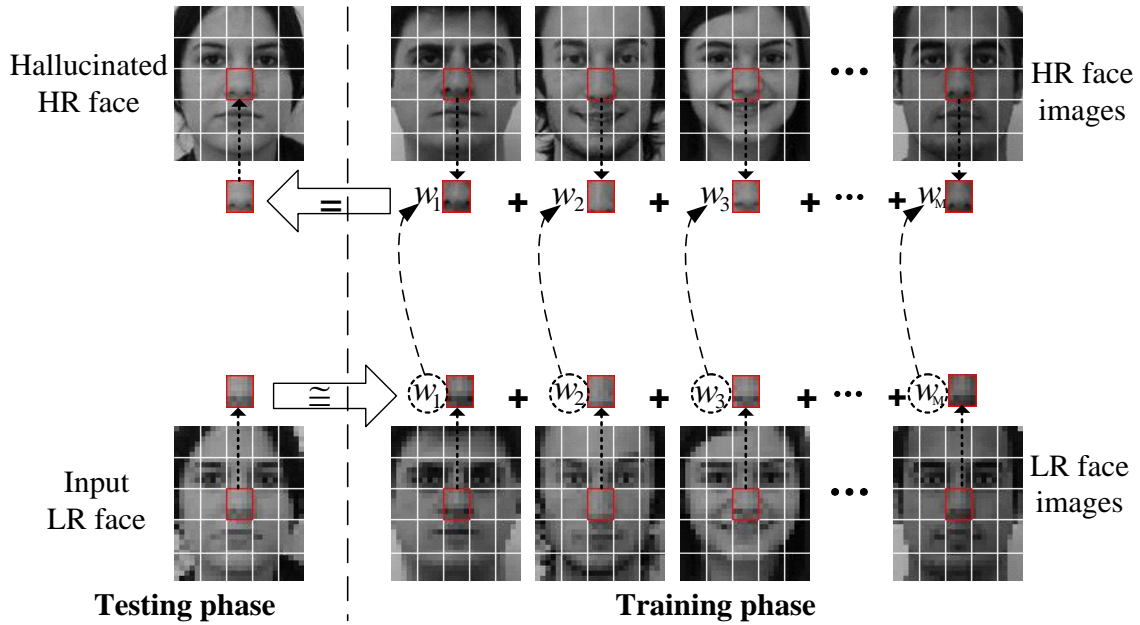


图 3-1 基于位置图像块的人脸超分辨率算法的框架图。

例如，西安交通大学Ma等人^[1]为了对输入人脸图像块进行表示，首次提出基于最小二乘表示（Least Squares Representation，简称LSR）的位置图像块人脸超分辨率算法。该方法利用低分辨率训练集中和输入低分辨率图像块具有相同位置的所有图像块对输入低分辨率图像块进行最小二乘表示，再利用该系数与对应训练集高分辨率图像块进行线性加权，最终重构出高分辨率人脸图像块。然而，当训练样本中图像的个数比图像块的特征表示维数大时，基于最小二乘的图像块表示方法会导致图像块的表示系数不唯一。为此，2011年Jung等人^[2]提出一种基于凸优化的位置图像块人脸超分辨率方法，将稀疏约束加入到图像块表示中，提出一种基于稀疏表示（Sparse Representation，简称SR）的位置图像块人脸超分辨率算法，使得最小二乘法具有唯一解。该方法结合了Yang等人的稀疏约束表示方法和Ma等人的基于LSR^[1]的位置图像块先验方法的优点，其性能好于上述两种方法。然而，SR方法^[2]存在以下两点不足之处：一方面，为了使输入图像块的表示尽可能的稀疏，它在合成输入的图像块时可能会选取一些与输入图像块差异很大的样本图像块进行线性重构，进而合成出一些原始高分辨率人脸图像中不存在的虚假高频细节信息；另一方面，由于稀疏表示对强噪声非常敏感，观察图像块的噪声会造成表示系数的不稳定（相似图像块在受到噪声干扰后，表示系数差异巨大）。

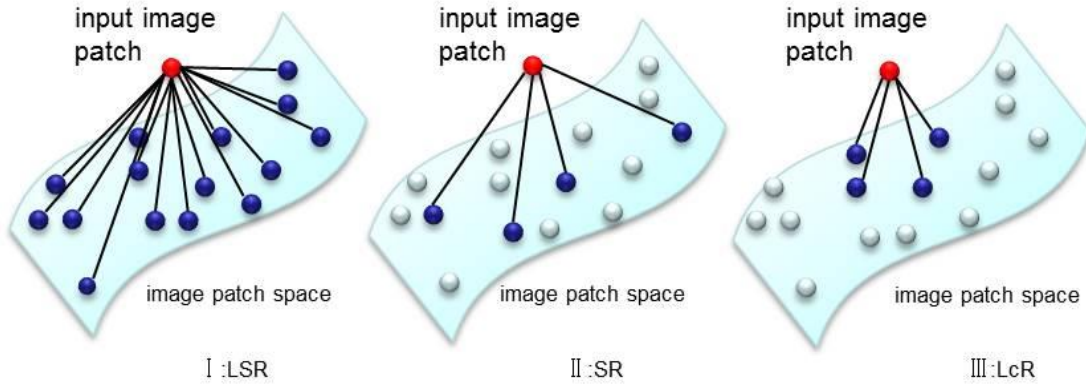


图 3-2 三种不同的图像块表示方法原理示意。

近年来,计算视觉相关研究成果表明:局部性先验对于揭示数据空间潜在的非线性流形结构非常重要。这一先验知识被广泛应用于图像聚类、图像分类、图像标记和图像识别等相关领域,并获得了比传统方法更好的性能。受这些研究成果的启发,本章将局部性先验知识引入到图像超分辨率重构过程中,提出一种基于局部约束表示(Locality-constrained Representation,简称LcR)的人脸超分辨率算法。图3-2是三种不同图像块表示方法的原理示意。在对输入图像块进行表示时,利用局部性来约束图像块的表示系数,在图像块样本空间中自适应地选取与输入图像块距离近的样本图像块,惩罚距离远的样本图像块,实现对图像块的鲁棒表示。特别地,当输入噪声图像块时,通过选择少量的相似图像块进行表示,可以达到抗噪的目的。而在Ma的方法中,他们选取了所有样本图像块对噪声图像块进行重构,不仅合成出了图像像素信息,也合成出了不需要的噪声信息,因此Ma的方法对噪声不鲁棒。提出方法具有以下几个优点:

(1) 首次将局部约束先验引入到人脸超分辨率重构中,提出一种基于局部约束表示的人脸超分辨率算法,该方法利用样本空间的相似样本点进行线性表示,可很好地揭示人脸图像块空间的非线性流形结构;

(2) 利用局部性来约束图像块的表示系数,选取与输入图像块距离近的样本图像块,惩罚距离远的样本图像块,使得基于LSR方法中图像块表示不唯一以及基于稀疏表示方法中图像块表示不稳定的问题得到有效缓解;

(3) 通过调节局部约束项的权重,自适应选择与输入噪声图像块相似的样本图像块,提出方法对噪声人脸图像、尤其是实际监控环境下获得的噪声人脸图像的超分辨率重构十分鲁棒。

本章接下来的安排如下:第3.2节介绍两种现有图像块表示方法:LSR^[1]和SR^[2];第3.3节详细介绍提出方法的思路,包括图像块局部约束表示目标函数的构建、优化和基于局部约束鲁棒人脸超分辨率算法的详细实现细节;第3.4节分别定性和定量地分析了LSR^[1]、SR^[2]和LcR三种表示方法的稀疏性和局部性特点;第3.5节通过在FEI人脸库^[5]、CUM+MIT人脸图像库^[134]和实际监控人脸图像库上的主客观对比实验验证算法的有效性以及对噪声的鲁棒性,并分析

了局部约束正则参数、图像块大小与图像块交叠、对齐误差等因素对算法性能的影响；第3.6节对算法进行总结，特别是讨论了算法在下一阶段可能的改进方向。

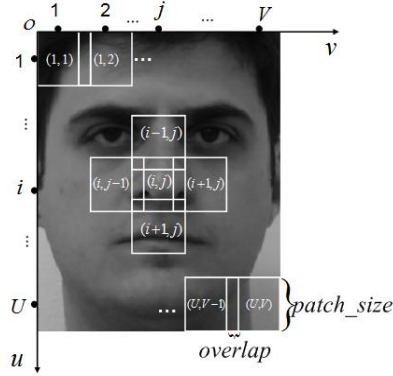


图 3-3 人脸图按位置被分成图像块示意。

3.2 图像块表示方法

设 Y^n ($n=1, \dots, N$) 为训练集人脸图像， N 是训练集大小。每幅人脸图像被分成 M 个交叠的图像块，记这 M 个图像块组成的集合为 $\{Y^n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$ ， $M = UV$ 。 U 为人脸图像每一列包括的图像块数， V 为人脸图像每一行包括的图像块数， (i, j) 表示图像块在人脸图像位置块坐标系 $o-uv$ 下的位置坐标。如图3-3表示， $patch_size$ 和 $overlap$ 分别表示图像块的大小和相邻图像块之间交叠的像素个数。

对于观察到的一幅人脸图像 X ，对其用上述同样的图像块划分方式进行分块 $\{X(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$ ，对于位于位置 (i, j) 上的图像块，它可以用对应位置的所有 N 个训练样本图像块进行表示，得到一个 N 维的表示权重向量 $w(i, j) = [w_1(i, j), w_2(i, j), \dots, w_N(i, j)]^T$ 。

接下来的3.2.1节和3.2.2节将分别介绍本章引言部分提到的两种图像块表示方法：Ma等人提出的LSR法和Jung等人提出的SR法^[2]。

3.2.1 最小二乘表示法

通过引入图像块先验，Ma等人利用样本库中的所有相同位置（位置同为 (i, j) ）的图像块对观察图像块 $X(i, j)$ 进行联合表示：

$$X(i, j) = \sum_{n=1}^N Y^n(i, j) w_n(i, j) + e, \quad (3-1)$$

其中， e 为观察图像块的重构误差。

上述表示权重向量可以转化为下面带约束的最小二乘拟合问题：

$$w^*(i, j) = \arg \min_{w(i, j)} \| X(i, j) - \sum_{n=1}^N Y^n(i, j) w_n(i, j) \|_2^2 \quad \text{s.t.} \quad \sum_{n=1}^N w_n(i, j) = 1. \quad (3-2)$$

它可以利用^[135]中提出的方法进行求解。

3.2.2 稀疏表示法

在实际情况下，由于训练集中样本个数比图像块的特征表示维数大，式（3-2）的解不唯一。为了解决这类解空间过大的问题，通常的做法是在目标函数中引入正则约束项，通过正则化技术得到满足某种条件的唯一解。例如，Jung等人引入稀疏表示理论，利用训练样本的一个子集对输入图像块 $X(i, j)$ 进行表示，而不是对其进行联合表示（每个样本图像块均参与）。SR方法^[2]将式（3-2）中的优化问题转化为标准的稀疏表示问题：

$$\min_w \|w(i, j)\|_0, \text{ s.t. } \|X(i, j) - \sum_{n=1}^N Y^n(i, j)w_n(i, j)\|_2^2 \leq \varepsilon, \quad (3-3)$$

其中 $\| \cdot \|_0$ 代表 ℓ_0 范数，表示一个向量中非零元素的总个数。式（3-3）是一个组合优化问题，直接求解将非常困难。近来备受关注的稀疏理论表明，倘若式（3-3）的解足够稀疏（ w 中非零元素个数非常少），那么它就可以用带 ℓ^1 范数约束的最优化问题来逼近：

$$\min_w \|w(i, j)\|_1, \text{ s.t. } \|X(i, j) - \sum_{n=1}^N Y^n(i, j)w_n(i, j)\|_2^2 \leq \varepsilon, \quad (3-4)$$

其中 $\| \cdot \|_1$ 代表 ℓ^1 范数，表示一个向量中所有元素绝对值的总和， $\varepsilon \geq 0$ 为允许的重构误差阈值。

通过引入稀疏约束正则项，SR方法^[2]一方面可以使原来不适定的问题（式（3-2））具有唯一的解，另一方面也可使得对观察图像块的重构误差尽量小。

3.3 局部约束鲁棒人脸超分辨率算法

3.3.1 图像块局部约束表示

SR方法^[2]通过引入稀疏正则项，使得LSR方法^[1]中图像块表示具有唯一稳定解，然而SR方法^[2]过分强调稀疏性，忽略了图像块流形空间的局部性质。因为SR方法^[2]选取出来的样本图像块（字典原子）可能与输入待表示图像块有较大的差异，而通过施加局部约束可以解决这一问题。近来在机器学习领域的研究成果也表明，局部性先验好于稀疏性先验，特别是在流形空间非线性结构的发现与低维嵌入学习中。局部性会导致稀疏性，反之则不然。

近来，Yu等人提出的局部坐标编码^[136]（Local Coordinate Coding，简称LCC）方法认为图像的编码应该满足局部性。同时，他们也从理论上证明了在一些特定的假设下，局部性先验比稀疏性先验更加重要，对于一些非线性学习问题，基于局部性先验的方法往往能够获得比稀疏性先验方法更好的结果。受到这一思想的启发，我们将局部约束先验加入到式（3-2）中，进而提出一种基于局部约束的图像块表示方法，目标函数如下：

$$\min \|d(i, j) \circ w(i, j)\|_2^2, \quad \text{s.t.} \quad \|X(i, j) - \sum_{n=1}^N Y^n(i, j)w_n(i, j)\|_2^2 \leq \varepsilon \quad \sum_{n=1}^N w_n(i, j) = 1, \quad (3-5)$$

其中，“ \circ ”表示两个向量之间的点积， $d(i, j)$ 是一个用来惩罚重构权重 w 的 N 维向量，它可以通过 $X(i, j)$ 与训练样本的欧式距离获得：

$$d_n(i, j) = \|X(i, j) - Y^n(i, j)\|_2, \quad 1 \leq n \leq N. \quad (3-6)$$

可以将式 (3-5) 写成如下拉格朗日形式：

$$\begin{aligned} w^*(i, j) = \arg \min_{w(i, j)} & \left\{ \|X(i, j) - \sum_{n=1}^N Y^n(i, j)w_n(i, j)\|_2^2 + \tau \sum_{n=1}^N [d_n(i, j)w_n(i, j)]^2 \right\} \\ \text{s.t.} & \quad \sum_{n=1}^N w_n(i, j) = 1. \end{aligned} \quad (3-7)$$

式 (3-7) 包括两项：第一项为重建误差项，衡量重构图像块与输入图像块之间的差异；第二项为局部约束保持项。 τ 为正则化参数，用来平衡重构误差项和解的局部性。当 $\tau = 0$ 时，提出的 LcR 方法就退化成了 LSR 方法^[1]。我们在实验部分将详细讨论 τ 对算法的影响。在提出的 LcR 方法中，局部约束项具有两个作用：首先，它可以使原 LSR 方法^[1]解空间过大的问题变得解空间唯一；其次，它在某种程度上可以对输入图像块进行稀疏表示（尽管这种“稀疏”性不是在 ℓ^0 意义下的稀疏，在 3.4.1 节我们也将讨论这一点）。

3.3.2 目标函数优化

目标函数 (3-7) 可以写作如下矩阵形式：

$$w^*(i, j) = \arg \min_{w(i, j)} \left\{ \|X(i, j) - Y(i, j)w(i, j)\|_2^2 + \tau \|Dw(i, j)\|_2^2 \right\}, \quad \text{s.t.} \quad \sum_{n=1}^N w_n(i, j) = 1. \quad (3-8)$$

$Y(i, j)$ 是位置为 (i, j) 的所有样本图像块组成的矩阵，它的每一列即为样本图像块 $Y^m(i, j)$ 按行扫描拉直形成的列向量。 D 是一个 $N \times N$ 的对角矩阵，其对角线上的每个元素为：

$$D_{nn} = d_n(i, j), \quad 1 \leq n \leq N. \quad (3-9)$$

令 $J(w) = \|X(i, j) - Y(i, j)w(i, j)\|_2^2 + \tau \|Dw(i, j)\|_2^2$ ，由 $\sum_{n=1}^N w_n(i, j) = 1$ 得

$$\begin{aligned} J(w) &= \|X - \sum_{n=1}^N Y^n w_n\|_2^2 + \tau \|Dw\|_2^2 \\ &= \|X \sum_{n=1}^N w_n - \sum_{n=1}^N Y^n w_n\|_2^2 + \tau \|Dw\|_2^2 \\ &= \left\| \sum_{n=1}^N (X - Y^n) w_n \right\|_2^2 + \tau \|Dw\|_2^2 \\ &= \|[X - Y^1, X - Y^2, \dots, X - Y^N]w\|_2^2 + \tau \|Dw\|_2^2. \end{aligned} \quad (3-10)$$

假设 $\text{ones}(N, 1)$ 是一个 N 维元素全为 1 的列向量，并令 $C = X(i, j)\text{ones}(N, 1)^T - Y(i, j)$ ， $G(i, j) = C^T C$ ， $Q = G + \tau D^2$ ，式 (3-10) 可以进一步简化为：

$$\begin{aligned}
 J(w) &= \| (X \cdot \text{ones}(N,1)^T - Y)w \|_2^2 + \tau \| Dw \|_2^2 \\
 &= w^T (X \cdot \text{ones}(N,1)^T - Y)^T (X \cdot \text{ones}(N,1)^T - Y)w + \tau w^T D^T Dw \\
 &= w^T C^T Cw + \tau w^T D^2 w \\
 &= w^T Gw + \tau w^T D^2 w \\
 &= w^T Qw.
 \end{aligned} \tag{3-11}$$

在 $\sum_{n=1}^N w_n(i, j) = 1$ 这个约束条件下, 最小化 $J(w)$, 可由拉格朗日乘数法进行求解。令:

$$J(w, \lambda) = w^T Qw + \lambda(\text{ones}(N,1)^T w - 1), \tag{3-12}$$

目标函数 $J(w, \lambda)$ 对 w 求导有:

$$\frac{\partial J(w, \lambda)}{\partial w} = 2Qw + \lambda \text{ones}(N,1) = 0, \tag{3-13}$$

由上面等式可求得 w :

$$w = -\frac{\lambda}{2} Q^{-1} \text{ones}(N,1). \tag{3-14}$$

目标函数 $J(w, \lambda)$ 对 λ 求导有:

$$\frac{\partial J(w, \lambda)}{\partial \lambda} = \text{ones}(N,1)^T w - 1 = 0, \tag{3-15}$$

将 (3-14) 代 (3-15), 有

$$\text{ones}(N,1)^T w = \text{ones}(N,1)^T \left(-\frac{\lambda}{2} Q^{-1} \text{ones}(N,1)\right) = 1, \tag{3-16}$$

进而有,

$$-\frac{\lambda}{2} = (\text{ones}(N,1)^T Q^{-1} \text{ones}(N,1))^{-1}. \tag{3-17}$$

再将 (3-17) 代入到 (3-14) 中便可以得到最终的解向量:

$$w = \frac{Q^{-1} \text{ones}(N,1)}{(\text{ones}(N,1)^T Q^{-1} \text{ones}(N,1))}. \tag{3-18}$$

上述解的形式正是参考文献^[83]的结果。如文献^[83]所述, 为了避免对矩阵 Q 求逆, 一个更有效的求解 w 的方法是直接求解线性方程 $Qw = \text{ones}(N,1)$, 然后再将求得的结果进行缩放使得

$\sum_{n=1}^N w_n(i, j) = 1$ 。那么, 最优重构权重就可以通过下式得到:

$$w = Q \setminus \text{ones}(N,1). \tag{3-19}$$

式 (3-19) 中的符号 “ \setminus ” 代表矩阵除法运算。

3.3.3 局部约束鲁棒人脸超分辨率算法

在人脸超分辨率算法中, 训练集由多张高分辨率人脸图像和对应的多张低分辨率人脸图像

组成。高分辨率人脸图像集合记作 $\{Y_H^n\}_{n=1}^N$ ，相应地，低分辨率人脸图像集合记作 $\{Y_L^n\}_{n=1}^N$ 。人脸超分辨率的任务就是：给定一张观察得到的分辨率人脸图像 X_L ，通过训练高低分辨率人脸图像样本对，重构潜在高分辨率人脸图像 X_H 。

首先，我们利用同样的划分策略将所有高低分辨率人脸图像样本对和输入的低分辨率人脸图像划分成图像小块。然后，对于输入的每一个低分辨率图像小块，可以利用我们提出的表示方法对其进行线性拟合表示，得到在低分辨率样本图像块集合上的表示系数。由于低分辨率图像块流形与高分辨率图像块流形具有相似的局部几何结构，那么目标高分辨率人脸图像相应位置上的高分辨率图像块可以利用以上得到的表示系数和对应高分辨率样本图像块集合线性加权得到。最后，根据重构得到的高分辨率图像块所在的位置，融合所有高分辨率图像块便可以得到最终的高分辨率人脸图像。

为了进一步提升基于局部约束表示的人脸超分辨率算法的性能，通过实验测试发现：运用合适的图像统计先验知识，对输入人脸图像进行简单的前处理就可以获得较大的性能提升。正如文献^[29, 95]所提到的，在给定图像中频带信息 I_{Medium} 条件下，假设图像的高频信息 I_{High} 条件独立于图像的低频带信息 I_{Low} 。上述思路可形式化为如下数学表达式：

$$p(I_{\text{High}} | I_{\text{Low}}) = p(I_{\text{High}} | I_{\text{Medium}}). \quad (3-20)$$

基于上述假设，图像高频细节信息重构只需要利用输入图像的中频带信息，而无需对图像高频带细节信息的预测没有影响的图像低频带信息。文献^[95]指出，为了获得图像的中频带成分，我们只需要对输入的低分辨率图像进行简单地插值操作。在本章中，我们直接利用双三次插值对输入人脸图像进行前处理，将输入人脸图像插值放大到目标高分辨率图像尺寸。提出的人脸超分辨率过程如算法3-1所示。

算法3-1 基于局部约束表示的人脸超分辨率

1. **输入：** 高分辨率人脸图像训练集 $\{Y_H^n\}_{n=1}^N$ ，低分辨率人脸图像训练集合 $\{Y_L^n\}_{n=1}^N$ ，输入低分辨率人脸图像 X_L ，图像块大小，近邻图像块之间的交叠像素，局部正则化参数；
 2. **输出：** 重构高分辨率人脸图像 X_H ；
 3. 利用双三次插值方法对输入低分辨率人脸图像进行上采样到目标高分辨率人脸图像大小；
 4. 将高分辨率人脸图像训练集、低分辨率人脸图像训练集合和插值放大之后的低分
-

分辨率人脸图像按照同样的划分策略分成图像小块;

5. 对于输入低分辨率人脸图像上的每一个小块 $X_L(i, j)$;
6. 计算输入低分辨率图像小块 $X_L(i, j)$ 和训练集中所有 N 个低分辨率图像小块 $\{Y_L^n(i, j)\}_{n=1}^N$ 之间的欧式距离:

$$d_n(i, j) = \|X_L(i, j) - Y_L^n(i, j)\|_2, \quad 1 \leq n \leq N;$$

7. 在低分辨率样本图像块集 $\{Y_L^n(i, j)\}$ 上为 $X_L(i, j)$ 计算最优表示权重 $w^*(i, j)$:

$$w^*(i, j) = \arg \min_{w(i, j)} \left\{ \|X_L(i, j) - Y_L w(i, j)\|_2^2 + \tau \|Dw(i, j)\|_2^2 \right\}, \text{ s.t. } \sum_{n=1}^N w_n(i, j) = 1;$$

8. 通过下式计算目标高分辨率图像小块 $X_H(i, j)$:

$$X_H(i, j) = \sum_{n=1}^N Y_H^n(i, j) w_n^*(i, j);$$

9. 重复步骤6-步骤8直至所有图像块被处理完毕;
10. 将所有加权重构出的高分辨率的图像块按照位置叠加, 然后除以每个像素位置交叠的次数, 重构出高分辨率人脸图像 X_H 。

3.4 稀疏性和局部性分析

近年来, 在数据表示与分析相关研究领域, 关于稀疏性先验和局部性先验的优劣问题的讨论受到了学者们的大量关注。受人类视觉系统信息处理过程的启发, 大量学者主张信号的稀疏性是一个重要的先验知识, 它对于解决计算机视觉问题是至关重要的。与此同时, 在一般的模式识别问题(如数据表示和维数约简)中, 局部性先验同样被证明是非常有用的。而且在一些特定的假设下, 局部性先验比稀疏性先验更加重要。

在本小节接下来的部分, 我们将分别从定性分析和定量分析两个角度来验证两种对比方法(LSR^[1]方法和SR方法^[2])和提出的LcR方法的稀疏性和局部性。本节所有实验结果均是基于FEI人脸图像数据库的, 更多关于这个数据库的介绍, 请参见第3.5节实验部分内容。

3.4.1 稀疏性分析

为了测试LSR方法^[1]、稀疏表示和LcR方法不同, 我们将所有测试图像上所有图像小块的每个表示权重向量首尾拼接起来, 得到一个长度为 L^3 的拼接表示权重向量 \bar{w} 。画出三种不同表

³ L =测试图像的张数 \times 单个图像块表示的维度 \times 每幅人脸垂直方向图像块的个数 $U \times$ 每幅人脸水平方向图像块的个数 V =

$40 \times 360 \times 15 \times 12 = 2592000 = 2.592 \times 10^6$. 其中 $U = \text{ceil} \left\{ \frac{\text{imrow} - \text{overlap}}{\text{patch_size} - \text{overlap}} \right\}$,

示方法得到的拼接表示权重向量，如图3-4所示。由第一行右半部分的结果可知，LSR方法^[1]并不稀疏，因为大多数的表示权值都不接近于零值。第三行为提出LcR方法的结果，其表示权值在某种程度是稀疏的，尽管还不如SR方法^[2]的结果稀疏。与些同时，实验中也观察到了一个有趣的现象：LSR方法^[1]和SR方法^[2]所得到的权值比较均匀地分布在零值的上下两侧，而提出的LcR方法所获得的权值可以近似认为是“非负”的（绝大部分权值都大于零）。上述分析表明：在对输入图像块进行表示的过程中，提出的LcR方法通过选择少量相似样本图像块进行“加性”表示，这也更具直观的语义解释，反映了人类思维中“局部构成整体”的概念。

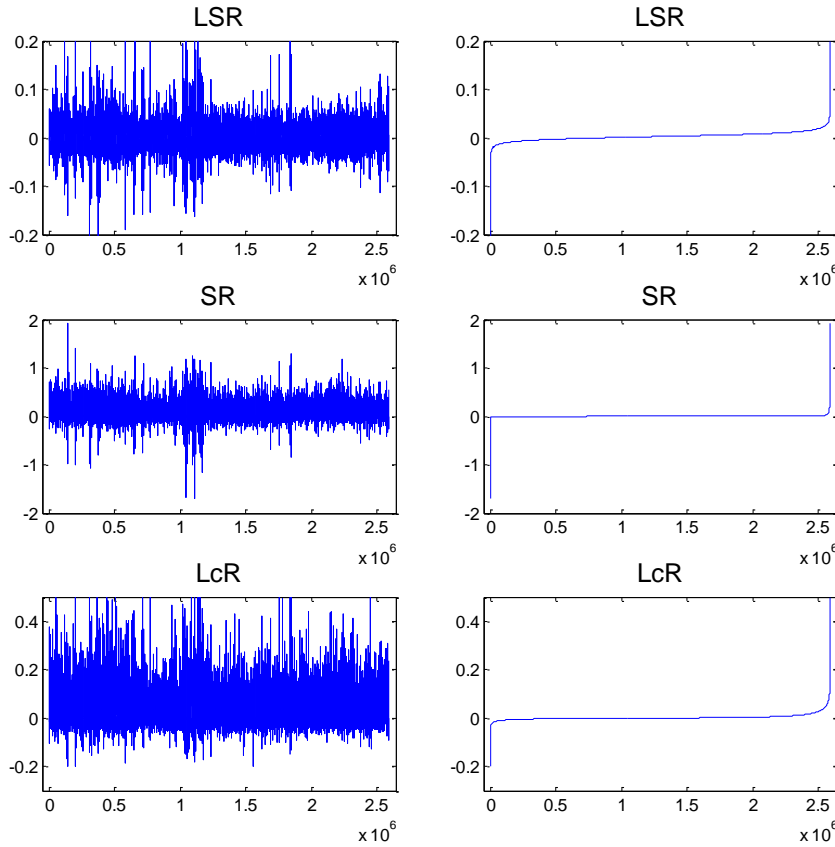


图3-4 不同表示方法的稀疏性测试：第一行为LSR方法^[1]，第二行为SR方法，第三行为提出的LcR方法。左半部分为拼接表示权重向量画出来的结果，右半部分为拼接表示权重向量按从小到大排列之后画出来的结果。排序之后的拼接表示权重向量越平滑且水平接近零值，则说明该方法表示系数越稀疏。

接下来，我们通过一种稀疏性度量指标——Gini index (GI)，来定量地度量不同表示方法的稀疏性。具体地，GI度量拼接化表示权重向量 $\tilde{w} = [w_1, w_2, \dots, w_L]$ 的稀疏性公式如下：

$$V = \text{ceil} \left\{ \frac{\text{imcol} - \text{overlap}}{\text{patch_size} - \text{overlap}} \right\}, \text{imrow} = 120, \text{imcol} = 100, \text{patch_size} = 12, \text{overlap} = 4. \text{ceil}(x) \text{ 返回大于或者等于指定表达式的最大整数的函数。}$$

$$GI(\vec{w}) = 1 - 2 \sum_{i=1}^L \frac{|w_{[i]}|}{\|\vec{w}\|_1} \left(\frac{L-i+1/2}{L} \right), \quad (3-21)$$

假设 $\vec{w}_{GI} = [w_{[1]}, w_{[2]}, \dots, w_{[L]}]$ ($|w_{[1]}| \leq |w_{[2]}|, \dots, |w_{[L]}|$) 是 $\vec{w} = [w_1, w_2, \dots, w_L]$ 从小到大排序后的结果, 那么 $w_{[i]}$ 为 \vec{w}_{GI} 的第 i 个元素, $[1], [2], \dots, [L]$ 为排序后的下标号。

表3-1给出了三种不同方法的GI。由该表可以看出, 尽管SR方法^[2]的稀疏性是最好的, 但是提出方法的稀疏性相较于LSR方法^[1]的提升也是非常显剧的。从表3-1可知, 提出方法的稀疏性比SR方法^[2]差24%, 但是相较于LSR方法^[1]提高了85%。

表 3-1 三种不同方法的 GINI 指标。

不同方法	LSR方法 ^[1]	SR方法 ^[2]	LcR方法
GI	0.2181	0.5325	0.4040

3.4.2 局部性分析

文献^[135, 136]认为局部性比稀疏性更加重要。文献^[84]也指出: 稠密采样的流形空间中的每个数据点均可以被周围少量近邻样本进行线性表示。在本小节, 我们将测试了不同算法的局部性。与3.4.1小节中拼接所有表示权重向量的方式类似, 我们拼接得到一个长度为 L 的距离向量 \vec{d} , 并对其进行从大到小排序, 得到

$$\text{sort}(\vec{d}) = [\vec{d}_{[1]}, \vec{d}_{[2]}, \dots, \vec{d}_{[L]}], \vec{d}_{[1]} \leq \vec{d}_{[2]} \leq \dots \leq \vec{d}_{[L]}. \quad (3-22)$$

根据上述排序标号, 我们也对拼接表示权重向量 \vec{w} 进行重新排序, 得到依距离排序的表示权重向量 \vec{w}_d :

$$\vec{w}_d = [w_{[1]}, w_{[2]}, \dots, w_{[L]}]. \quad (3-23)$$

以从小到大排序后的距离 $\text{sort}(\vec{d})$ 为横坐标, 以拼接依距离排序的表示权重向量 \vec{w}_d 为纵坐标作图 (如图3-5所示), 画出了三种不同方法重排后的结果。由结果可见, 随着样本图像块与输入待表示图像块的距离的增大, 表示权值在减小。与两种对比方法相对, 提出方法的这种趋势最明显 (大的表示权重基本上都集中在小的距离上)。也就是说, 提出的方法更多地选择相似的样本图像块进行表示。我们也发现, LSR方法^[1]和SR方法^[2]可能出现不稳定的尖峰, 而我们的方法可以很好的抑制这种尖峰, 因而表示权值更加稳定。

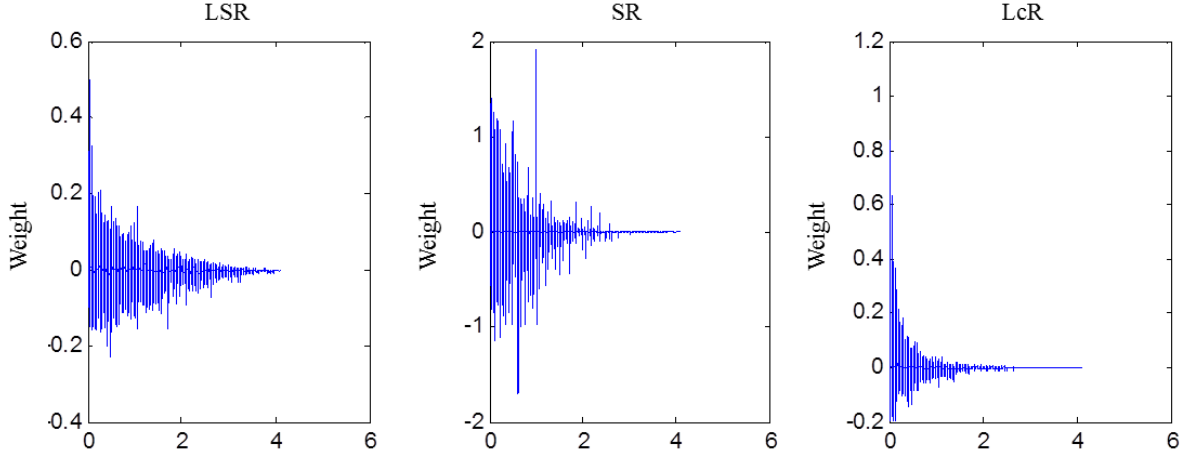


图 3-5 三种不同表示方法的拼接权重向量随距离变化示意。

与稀疏性分析类似，本小节也另外从定量分析的角度测试了三种不同表示方法的局部性。为此，我们定义了一种局部性度量指标—— K 均值距离（ K -Mean Distance，简称 K -MD）因而， K -MD定义如下：

$$K\text{-MD} = \left\{ \frac{\sum_{C_K(X(i,j))} \|X(i,j) - Y^n(i,j)\|_2^2}{K} \right\}^{\frac{1}{2}}. \quad (3-24)$$

其中， $X(i,j)$ 和 $\{Y^n(i,j) | m \in C_K(X(i,j))\}$ 分别表示输入图像块和与之对应的 K 个最重要的训练样本图像块。其中， $C_K(X(i,j))$ 表示 K 个表示权重最大的训练样本图像块的下标。表示权重越大，那么说明这个样本图像块越重要。根据 K -MD的定义，一种表示方法的局部性可以通过计算输入图像块与 K 个表示权重最大的训练样本图像块的平均距离来度量。因此， K -MD值越小，则说明这种方法的局部性越好。

图3-6是根据三种不同表示方法的 K -MD值所画的直方图。可以看出，在 K 取不同值（ $K=5, K=20, K=50, K=100$ ）的情况下，提出的LcR方法的 K -MD值都集中在较小的范围内，而且均值都是最小的。这表明与其它两种对比的表示方法相比，提出方法用作表达的样本图像块与输入图像块是最相似的。与此同时，我们也注意到，提出方法的 K -MD值随着 K 的增大而增大，而两种对比方法在不同 K 值下的 K -MD值却相对稳定。这说明了对比方法不关注样本图像块与输入图像块的相似性，他们比较随机地（在不考虑相似性的意义下）选择样本图像块。而对于提出的LcR方法，训练样本集中相似的图像块更容易选中，如果 K 值增大，必然会选择那些不相似的样本图像块，从而导致 K -MD值的增大。上述结果均表明，提出的方法更有可能选择近邻的样本图像块，可以很好的揭示样本图像块流形空间的局部几何结构。

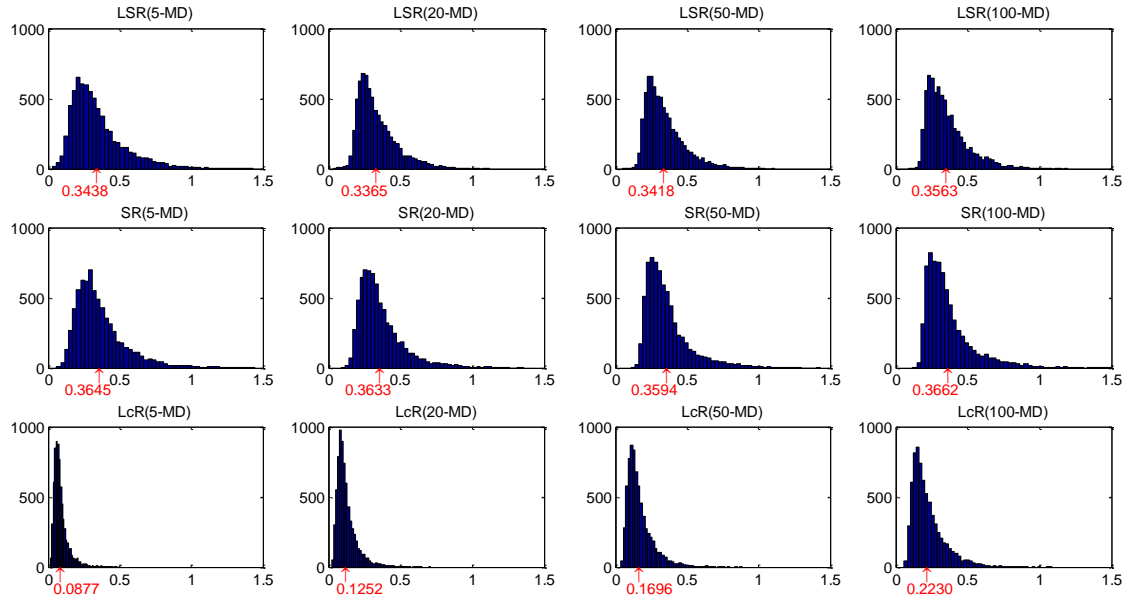


图3-6 三种不同表示方法的K-MD值直方图: 第一行是LSR方法^[1], 第二行是SR方法^[2], 第三行是LcR方法. 红色箭头表示均值。

3.5 实验及结果分析

为了验证提出算法的有效性, 此节将对比几个不同的人脸超分辨率算法, 包括Wang等人提出的特征转换全局脸方法^[54], Chang等人提出的NE方法^[83], Ma等人提出的LSR方法^[1], Jung等人提出的SR方法^[2]和Zhang等人提出的DCT域邻域嵌入方法^[68]。3.5节所有实验结果均在FEI人脸库^[5] (关于它的介绍请参见接下来的3.5.1小节) 进行。3.5.2小节将给出主观人脸重构结果和客观指标, 如峰值信噪比 (Peak Signal to Noise Ratio, 简称PSNR) 和结构相似度^[6] (Structural SIMilarity, 简称SSIM)。同时, 3.5.3小节也测试了不同参数对算法结果的影响, 包括局部约束正则参数、图像块大小与图像块交叠和对齐误差。3.5.4小节将验证提出方法对噪声的鲁棒性。此外, 为了进一步证明提出方法的优越性, 3.5.5小节也报告了提出方法和对比算法在一些实际低分辨率人脸图像 (包括监控人脸图像) 上的重构结果。



图 3-7 FEI 人脸库示例。

3.5.1 人脸库介绍

本章采用FEI人脸库^[5]进行实验,该数据库包括了200个个体(包括100位男士和100位女士)的400幅人脸图像,每个人有一张正面中性表情和一张正面笑脸图像,他们年龄从19岁到40岁,他们或带不同的配饰、或有不同的发型、或有不同的装扮。图3-7是该数据库的一些示例人脸图像。库中所有人脸图像均被裁剪成 120×100 像素大小。在实验中,我们随机选择180个人(360张图像)张作为训练样本,将剩余20个人(40张图像)作为测试图像。注意:所有测试图像均不在训练集中。低分辨率图像由高分辨率图像加模糊(窗口大小为 4×4 的平均模糊)并下采样(4倍双三次插值下采样)得到。因此,低分辨率人脸图像大小为 30×25 像素。

为了公平起见,我们为不同的对比算法尝试各种参数设置,使其达到最好性能。具体地,对于Wang等人提出的特征转换全局局方法^[54],提取主成分信息能量占原始信号能量的百分比为99%(约100幅本征脸(Eigenface));对于Chang等人提出的NE方法^[83],其近邻个数设为50;对于Jung等人提出的SR方法^[2],图像块重构误差容忍值设为1.0;对于Zhang等人提出的DCT域邻域嵌入法^[68],近邻个数设为10。在所有基于图像块的方法中,高分辨率图像块大小被设为 12×12 像素,相邻图像块之间交叠像素为4个;对应的低分辨率图像块大小被设为 4×4 像素,相邻图像块之间交叠像素为1个(更多关于不同图像块大小和相邻图像块交叠程度的分析,请参见3.5.3.2节)。对于提出的LcR方法,只有局部约束正则参数 τ 需要调整。在3.5.3.1小节,我们将测试不同局部约束正则参数设置对提出LcR算法的影响。在接下来的实验中,如果没有特别的说明,对于没有前处理的LcR方法, τ 值设为0.04;对于有前处理的LcR方法, τ 值设为0.02。



图3-8 不同算法人脸超分辨率重构结果。从左到右依次为:输入低分辨率人脸图像,Wang方法的重构结果^[54],NE方法的重构结果,LSR方法^[1]的重构结果,SR方法^[2]的重构结果,DCT邻域嵌入方法^[68]的重构结果,LcR方法的重构结果,PreLcR方法的重构结果,原始高分辨率人脸图像。

3.5.2 FEI 人脸库上的结果

图3-8给出了8组不同方法超分辨率重构结果。第一列为输入的低分辨率人脸图像，最后一列为原始高分辨率人脸图像，第二列到第八列为7种不同算法的人脸超分辨率重构结果。由该重构结果图可以得出以下几点结论：

(1) 基于图像块的方法好于基于全局脸的方法。Wang等人提出的特征转换全局脸超分辨率方法^[54]重构的人脸图像过于平滑且与原始高分辨率人脸图像相似性不够，此外，在重构人脸的轮廓会引入“鬼影”效应（如图3-8第二列）。相较而言，基于图像块的方法特征细节合成与重构人脸相似度方面展现出了他们的优越性，增强了边缘和纹理等细节信息。这主要可解释为Wang等人的全局脸方法^[54]是一种基于统计模型的方法，在训练样本规模不够的情况下无法很好地揭示样本分布的本质规律。例如，在我们的实验中，训练样本只有360个，而每个样本数据的特征表示维度却高达 $30 \times 25 = 750$ 维。基于统计模型的方法在训练样本数不是远远大于特征表示维度的情况下是无法揭示数据间本质规律的。尤其是在输入低分辨率人脸图像与训练样本库中的人脸图像差异较大时（图八第二行结果），这种基于全局脸的人脸超分辨率重构方法无法重构训练集中不存在的新的结构信息。

(2) 基于位置图像块的方法（LSR方法^[1]、SR方法^[2]、DCT邻域嵌入方法^[68]和提出的LcR方法）好于不用位置图像块先验的方法（Chang等人的NE方法^[83]）。从图3-8第三行结果可以看出，Chang等人的NE方法^[83]重构的结果比较模糊且有明显的人工痕迹。通过引入位置块这一人脸图像独有的先验知识，基于位置图像块的方法可以很好得保持这种潜在的语义结构信息。

(3) 通过进一步的观察与对比可以看出，与另外三种基于位置图像块的方法（LSR方法^[1]，SR方法^[2]和 DCT方法^[68]）相对，提出的LcR方法可以重构出更多的高频细节信息，重构结果没有明显的人工痕迹更加自然，这主要归因于局部约束这一先验的引入。此外，我们也给出了增加前处理的LcR方法（图3-8倒数第二列）的主观重构结果。通过该前处理，重构的高分辨率人脸图像在视觉上更加令人舒服。这也说明了合适先验知识在图像超分辨率重构过程中的重要性。

(4) 表3-2第二列给出了不同算法在超分辨率放大四倍情况下的PSNR和SSIM两种客观指标。双三次插值方法被视为基准算法进行比较。同样，我们可以看到，基于图像块方法进行比较Wang等人提出的全局脸方法^[54]和双三次插值方法好得多。基于DCT的方法^[68]不如其他一些基于图像块的方法，这是由于在高低分辨率局部几何结构一致性假设在DCT空间失效（我们的实验表明：高低分辨率图像像素空间之间的流形一致性比在DCT空间更好）。提出方法获得了最高的PSNR和SSIM值，相较于第二好的PSNR和SSIM增益分别为0.65 dB和0.0097。加入前处理之后，上述增益更是达到1.02 dB和0.0168。这足见探索新的视觉先验约束图像表示模型的重要性。

表 3-2 在不同下采样倍数的条件下, 不同算法的 PSNR (dB) 和 SSIM 对比。

下采样倍数	4		8		16	
不同对比方法	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	27.49	0.8417	22.94	0.6617	19.01	0.5322
Wang ^[54]	27.75	0.7582	26.05	0.7297	24.16	0.6954
NE ^[83]	31.23	0.8975	27.75	0.8088	24.60	0.7283
LSR ^[1]	31.90	0.9032	26.88	0.7813	23.94	0.7063
SR ^[2]	32.11	0.9048	26.88	0.7814	23.90	0.6993
DCT ^[68]	30.55	0.9011	26.55	0.7881	21.79	0.6327
LcR	32.76	0.9145	27.86	0.8102	24.60	0.7320
PreLcR	33.13	0.9216	28.64	0.8345	25.07	0.7429
相对增益	1.02	0.0168	0.89	0.0257	0.47	0.0146



图3-9 不同算法在8倍下采样和16倍下采样下的人脸超分辨率重构结果。从左到右依次为: 输入低分辨率人脸图像, Wang方法的重构结果^[54], NE方法^[83]的重构结果, LSR方法^[1]的重构结果, SR方法^[2]的重构结果, DCT邻域嵌入方法^[68]的重构结果, LcR方法的重构结果, PreLcR方法的重构结果, 原始高分辨率人脸图像。第一行和第二行为下采样8倍的结果, 第三行和第四行是下采样16倍的结果。

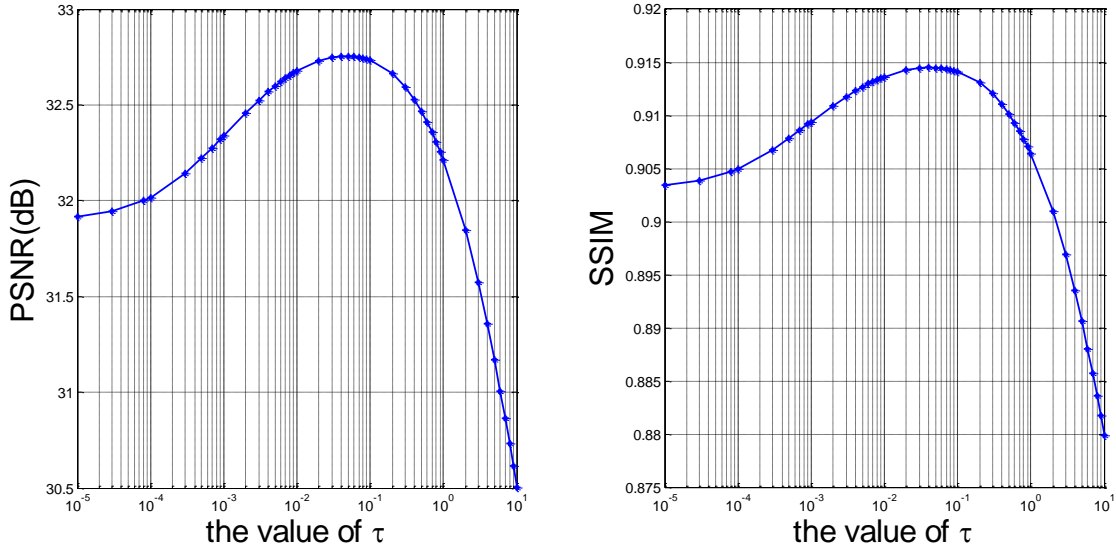


图 3-10 提出 LcR 方法在不同 τ 值下的 PSNR 和 SSIM 客观结果。

接下来将测试不同算法对于输入极低分辨率人脸图像的重构结果。表3-2的第三和第四列分别给出了低分辨率测试人脸图像的超分辨率放大8倍和16倍两种情况下的结果。可以看出，与对比方法相比，提出方法的增益仍然是十分明显的。实验结果也表明，当输入非常低分辨率的人脸图像时，基于全局脸的人脸超分辨率算法性能要好于基于局部脸的方法。此外，我们在图3-9中也给出了低分辨率测试人脸图像的下采样倍数扩大为8倍和16倍两种情况下对应的主观超分辨率重构结果。总的来说，随着下采样倍数的增大，超分辨率重构的结果会变得越来越差。尽管如果，提出算法的重构结果仍然是最好的，而对比算法在人脸的眼睛、嘴巴和轮廓部分会产生一些明显的人工效应。结合表3-2和图3-9的结果可以看出，引入合适先验的位置图像块方法是人脸超分辨率重构的一个很好的方向。

3.5.3 参数分析

3.5.3.1 不同 τ 值下的结果

为了验证局部性这一先验知识的有效性，我们测试了不同局部约束正则化参数 τ 对提出算法的影响。正则化参数 τ 可以控制局部性约束在整个代价函数中所占的比重。如图3-10所示，我们给出了PSNR和SSIM值随正则化参数 τ 的变化曲线。当 $\tau=0$ 时，提出的方法就退化为LSR方法^[1]，算法的性能是受限的。随着 τ 值的增加，客观指标不断提升。这表明了局部性这一先验的有效性。然而，我们同时也必须看到，正则化参数 τ 的取值也不宜设置过大。因为在 τ 值设置很大情况下，重建误差项在整个目标函数中所占的比例就会很小，导致重构结果与输入图像有较大差异。因此，一个合适的正则化参数 τ 是提出算法获得很好性能的关键。

3.5.3.2 不同块大小的结果

对于基于图像块的超分辨率方法，图像块大小的设置是非常重要的。一方面，如果图像块设置过小，它只能为图像的重构提供非常有限的信息，无法保持人脸结构特征且重构的高分辨率人脸图像容易受噪声的干扰；另一方面，如果图像块设置太大，重构的高分辨率人脸图像将非常平滑、缺乏特征细节信息。此外，随着图像块的增大，图像块的特征表示维度也随之增加，这就需要训练样本库包含更多的样本人脸图像（否则无法对高维图像块进行准确表示），尤其是在输入待重构人脸图像与样本库人脸图像差异很大的情况下。表3-3给出了不同方法在不同图像块大小和相邻图像块具有不同程度交叠条件下的PSNR和SSIM值对比。可以看出，在相同图像块大小的条件下，相邻图像块的交叠像素越多，人脸超分辨率重构的效果就越好。然而，更多的交叠像素也就意味着更高的计算复杂度。不论块大小（4~24 pixels）和块交叠（0~20 pixels）怎么变化，提出的LcR方法的性能始终优于所有对比基于图像块的方法。为了权衡算法的重构性能和计算复杂度，我们推荐高分辨率图像块大小为 12×12 像素，相邻图像块之间交叠为4个像素。

表3-3 不同方法在不同图像块大小和不同交叠条件下的PSNR和SSIM值对比。

Patch Size	Overlap	NE ^[83]		LSR ^[1]		SR ^[2]		LcR		PreLcR	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
4	0	28.80	0.8140	28.28	0.8042	27.91	0.7991	28.81	0.8149	32.63	0.9105
8	0	31.36	0.8838	30.58	0.8715	30.33	0.8661	31.51	0.8875	32.92	0.9163
	4	32.17	0.9044	31.18	0.8918	31.23	0.8912	32.19	0.9049	33.17	0.9216
12	0	30.03	0.8556	31.40	0.8897	31.45	0.8883	32.22	0.9030	32.87	0.9164
	4	31.23	0.8975	31.90	0.9032	32.11	0.9048	32.76	0.9145	33.11	0.9210
	8	31.47	0.9015	32.04	0.9083	32.50	0.9145	32.90	0.9186	33.16	0.9224
16	0	30.66	0.8906	31.79	0.8973	31.76	0.8934	32.40	0.9074	32.73	0.9142
	4	32.17	0.9037	32.15	0.9067	32.28	0.9061	32.81	0.9163	32.92	0.9185
	8	32.43	0.9099	32.27	0.9100	32.58	0.9138	32.90	0.9185	32.97	0.9196
	12	32.60	0.9146	32.36	0.9135	32.77	0.9191	32.99	0.9213	33.01	0.9208
24	0	31.18	0.8835	31.88	0.8970	31.50	0.8865	32.29	0.9056	32.41	0.9077
	8	31.58	0.8951	32.31	0.9086	32.05	0.9030	32.66	0.9146	32.64	0.9138
	16	31.73	0.8999	32.41	0.9122	32.46	0.9132	32.71	0.9167	32.70	0.9156
	20	31.78	0.9020	32.41	0.9134	32.52	0.9157	32.72	0.9177	32.71	0.9160

3.5.3.3 对齐误差对人超分辨率重构的影响

输入人脸图像的准确对齐对于基于位置图像块的人脸超分辨率重构方法而言是至关重要的，因为对齐误差的存在将使人脸图像位置先验知识在人脸超分辨率重构的过程中所起的作用大打折扣。在以上所有实验中，输入人脸图像与训练库中的样本图像都是通过手工或者自动算法进行准确对齐的。为了说明对齐误差对最终人脸超分辨率重构结果的影响，我们在水平方向和垂直方向平移低分辨率图像，人工生成一些不对齐的低分辨率人脸图像。水平方向平移和垂直方向平移从-3像素到3像素不等。与3.5.2小节类似，我们利用来自FEI人脸库中的所有40张测试人脸图像进行测试。图3-11给出了40张测试人脸图像在不同对齐误差下的平均PSNR和SSIM值。图3-12给出了一张不同对齐误差下人脸超分辨率重构结果示例。以上主客观结果均表明：对齐误差对重构结果影响巨大，对齐误差越大，重构结果越差。

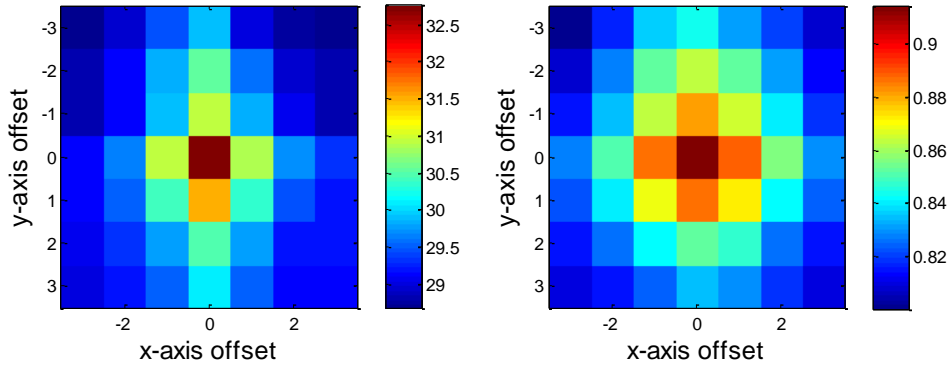


图 3-11 输入人脸图像在不同对齐误差下的 PSNR 和 SSIM 值变化。

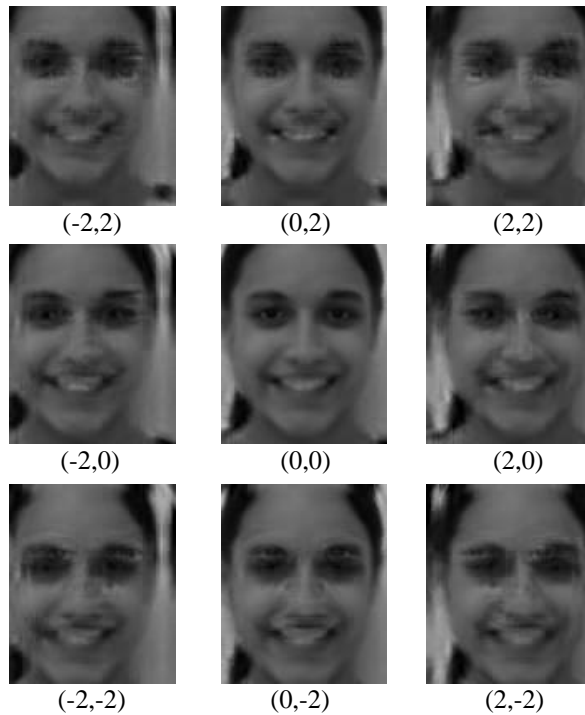


图3-12 不同对齐误差下人脸超分辨率重构结果示例。(0,0) 表示输入低分辨率人脸图像是严格对齐的。

3.5.4 算法噪声鲁棒性测试

当前大多图像超分辨率方法都是简单地假设输入低分辨率图像是干净的,没有考虑噪声对算法性能的影响。然而,在实际情况下,获得的待重构的图像不可避免的会受到一些拍摄环境噪声、器件噪声等因素的影响。为了处理带噪声的输入图像,他们一般都是先通过对输入图像进行去噪,然后再进行超分辨率重构。然而,由去噪声算法引起的人工效应极有可能在接下来的超分辨率重构过程中保留或者进一步被放大。在此小节中,将验证提出LcR方法的抗噪声性能。

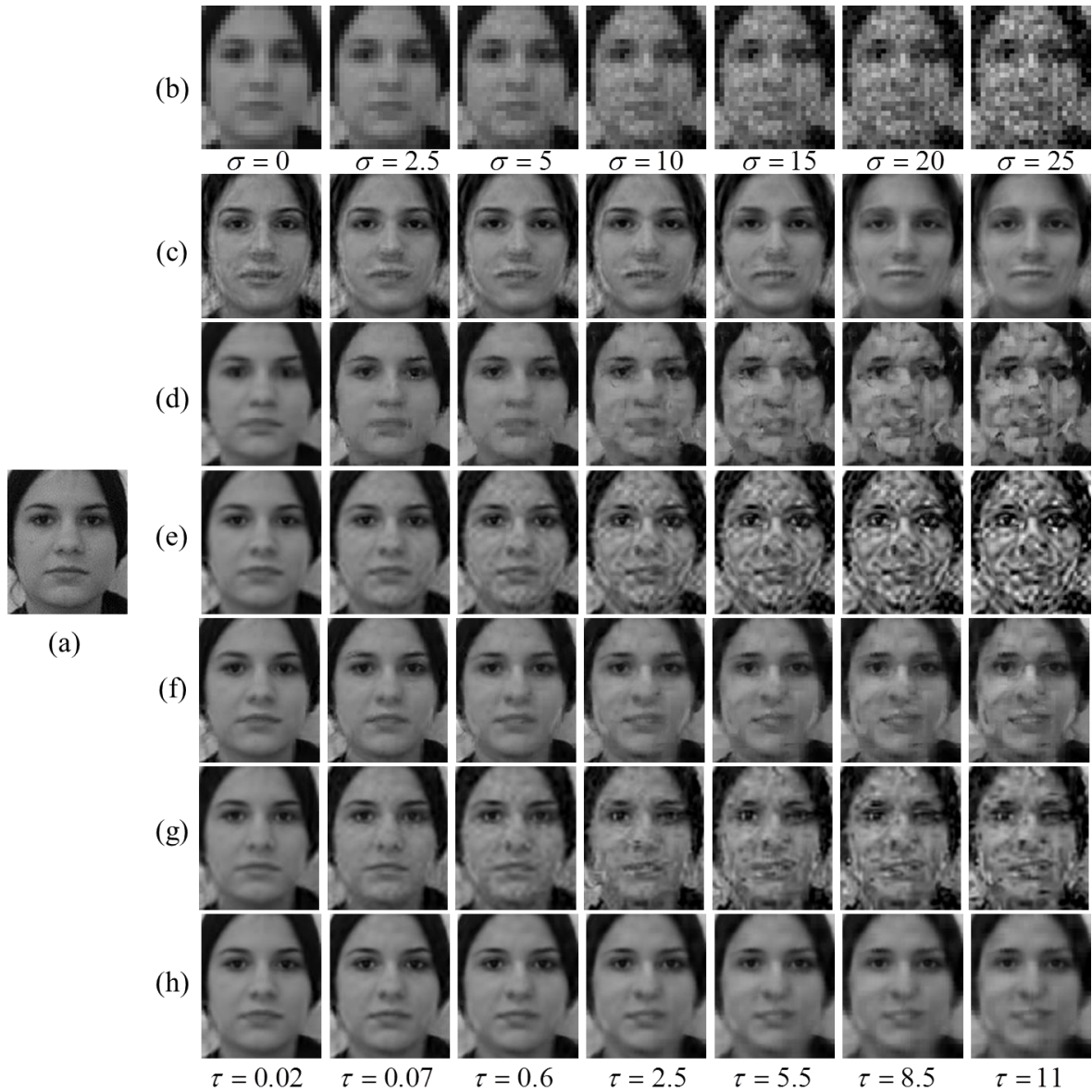


图3-13 不同程度噪声输入下不同算法的超分辨率重构结果。(a) 原始高分辨率人脸图像, (b) 输入低分辨率人脸图像, (c) Wang方法的重构结果^[54], (d) NE方法的重构结果, (e) LSR方法^[1]的重构结果, (f) SR方法的重构结果, (g) DCT邻域嵌入方法^[68]的重构结果, (h) 提出方法的重构结果。注意: 最后一行下面的 τ 值表示不同噪声条件下提出方法获得最好性能的局部约束正则化参数。

为了模拟噪声，在实验中，我们对输入低分辨率人脸图像加入不同方差水平(σ)的高斯噪声。对于对比算法，我们通过调节对应算法的参数（NE算法^[83]的近邻个数 K ，SR算法^[2]的误差容忍阈值 τ ），使在噪声条件下达到最好的重构结果。虽然通过增大图像块和增加交叠像素个数会产生更好的性能，但是为了公平起见，我们对不同的基于图像块的方法采样和3.5.2节设置一样的图像块大小和相邻图像块交叠像素。如图3-13所示，随着噪声的增大，NE方法^[83]，LSR方法^[1]和DCT邻域嵌入方法^[68]的性能急剧下降。当 $\sigma > 10$ 时，这些算法基本无法去除输入图像中的噪声，不能得到一个满意的结果。然而，SR方法^[2]和提出的方法能够较好的去除输入人脸图像中的大部分噪声；当噪声非常大时（ $\sigma > 20$ ），SR方法^[2]的重构结果有一些扭曲，而提出的方法却可以很好地保持人脸的主要特征（图3-13（h）），尽管提出算法的重构结果也会有些平滑且有一些块效应。SR方法^[2]在强噪声下的性能不佳问题是可以解释的，因为在稀疏表示理论中提到，稀疏表示只对较小的噪声很鲁棒，对于强噪声则无法恢复原始信号。换言之，一般无法仅仅通过稀疏先验来实现对噪声的鲁棒。图3-13结果还表明Wang等人提出的基于全局脸的方法^[54]可以比较好的保持全局人脸结构信息，同时也去除了大部分的噪声。然而，这些重构高分辨率人脸图像不是非常干净且与原始高分辨率人脸图像的差距较远。

从图3-13（h）可以看出，正则化参数 τ 对提出算法而言是非常重要的， τ 值的设置与输入图像的噪声大小紧密相关。输入图像含噪声越大，那么就需要给予局部约束更大的权重，即 τ 值需要设置得更大一些，反之亦然。更多关于输入图像噪声与局部约束正则化参数 τ 的选取问题，请参见附录A。为了更加直观地理论局部性约束在噪声鲁棒的超分辨率重构过程中起的作用，我们可以设想这样一种情况：假设输入的低分辨率人脸图像受噪声干扰非常严重，对其直接进行表示不仅不能排除噪声的干扰，还可能合成甚至放大噪声。提出方法通过引入局部约束先验，调节一个合适的（较大的） τ 值，便可以选择训练样本集中最相似的一些样本图像块去噪声图像块进行拟合，而不是对强噪声进行重构。这里必须强调的是：所选取的图像块个数必须尽量少（满足稀疏性），与此同时，所选取的图像块必须与输入图像块尽量相似（满足局部性）。正如3-4节所讨论的，同时达到稀疏性和局部性正是我们提出的基于局部约束表示的人脸超分辨率算法所具有的特点。

在本章中，我们把提出算法的去噪性能主要归因于自适应地选择近邻样本图像块。特别地，通过选取距离近的样本图像块，惩罚距离远的样本图像块，提出的方法能选择最相关的图像，从而避免过拟合或者拟合不当的问题。与使用固定近邻个数 K 的NE方法^[83]相比，提出的方法对噪声更加鲁棒。接下来，我们将通过一组实验验证这一点。两个对比算法分别是：我们提出的具有自适应选择近邻样本特性的LcR方法和基于位置块的NE方法^[83]（固定近邻个数）。与3.5.2节类似，我们选择360幅图像作为训练，剩余40幅图像作为测试集。为了达到最优性能，我们已经为两种方法调节相应的参数（局部约束正则项参数 τ 和近邻个数 K ）。图3-14不同噪声水平下两种不同方法的客观PSNR和SSIM指标。从图3-14可以看出，提出的方法的性能在所

有噪声水平下均好于邻域嵌入法。随着噪声水平的增大,这种优势更加明显。当噪声水平从0增加到25时,提出方法相对于NE方法^[83]的增益依次为0.30 dB, 0.73 dB, 0.94 dB, 1.09 dB, 1.36 dB。当 $\sigma \geq 10$ 时,NE方法^[83]的最优参数 K 非常小($K=1$ 或者 $K=3$),这说明了局部约束先验对噪声特别是强噪声的去除的重要性。这一点也是直观的,因为选择太多的无关的训练样本图像块不仅不能去除噪声,而且还可能合成图像信号自身之外的噪声信息。

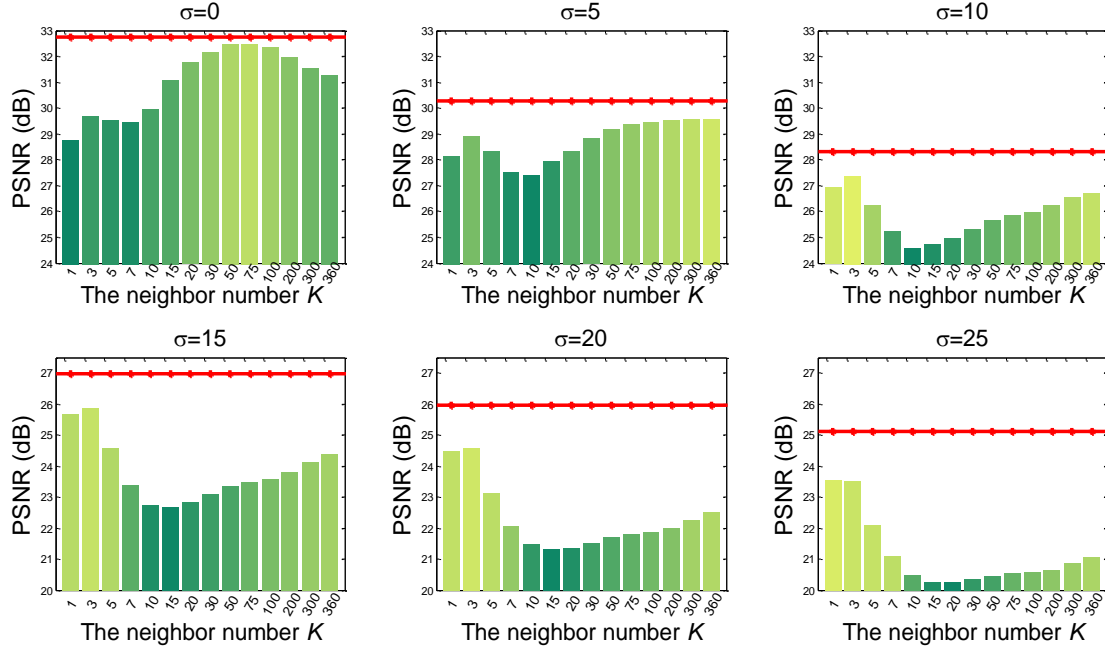


图3-14 不同噪声水平下两种不同的近邻选择策略的PSNR指标对比。红色代表提出方法的性能,直方图是NE方法^[83]在不同 K 值下的结果,在 $K=10$ 左右性能的下是NE方法^[83]对输入图像过拟合引起的。

3.5.5 真实图像的超分辨率重构结果

在上述实验中,所有低分辨率人脸图像均是通过平滑模糊和下采样原始高分辨率人脸图像得到。然而这种平滑模糊后加下采样的降质方式无法模拟实际低分辨率图像的获取过程。因此,为了验证提出算法对实际条件下获得低分辨率人脸图像的有效性,本小节将做如下两组测试:(1)CMU+MIT人脸库^[134]重构实验;(2)模拟实际监控环境人脸超分辨率重构实验。

(1) CMU+MIT人脸库实验。首先,我们人工地抠取感兴趣人脸,并根据人脸的两眼位置将其对齐到训练集人脸图像中,所有抠取得到的人脸图像的长宽大约为25到45个像素。其次,我们利用双三次插值方法将抠取且对齐了的人脸图像放大到 120×100 像素,以插值后得到的“高分辨率”人脸图像作为输入低分辨率人脸图像。最后,我们调节不同人脸超分辨率方法的算法参数,对输入图像进行超分辨率重构。提出的方法只需要根据输入图像来调整局部约束正则参数,而其它参数均保持不变,图3-15为提出方法在CMU+MIT一些典型测试图像上的人脸超分辨率重构结果,我们总共给出了五组对比实验结果。可以看出,尽管输入的低

分辨率人脸图像与训练样本库差异较大且可能含有较强的噪声，提出方法的结果仍然比较理想，且好于所有对比方法。

(2) **模拟实际监控环境人脸超分辨率重构实验。**图3-16为监控摄像机拍摄得到的输入图像。第一行为监控摄像机在低光照且拍摄对象距离摄像机较远位置下拍摄得到的低分辨率人脸图像，第二行为监控摄像机在正常光照且人脸靠近拍摄机下拍摄得到的高分辨率人脸图像。第二行抠取得到的高分辨率人脸图像可以用作“原始”高分辨率人脸图像（作为对比人脸图像）。与前述（1）中的处理方式类似，我们通过手工抠取人脸、对齐和上采样等步骤便可得到对齐之后的低分辨率人脸彩色图像。然后，对其进行彩色转灰度，再进行色阶调整，便可得到算法的输入低分辨率人脸图像（图3-17的第一行）。



图3-15 不同算法在 CMU+MIT 人脸库上的重构结果。左到右依次为：输入低分辨率人脸图像，Wang方法的重构结果^[54]，NE方法^[83]的重构结果，LSR方法^[1]的重构结果，SR方法的重构结果，DCT邻域嵌入方法的重构结果，LcR方法的重构结果。

如图3-17所示，其第一列是从图3-16的第一行中得到的输入低分辨率人脸图像。在这组实验中，我们同样调整了局部约束正则化参数，使提出的方法达到最优效果。从图3-17可以看出，当输入低分辨率监控人脸图像质量很差时（不仅有噪声，而且受到不同程度的模糊），提出的方法均可以获得一个比较好的结果。对于输入低分辨率监控人脸图像中存在的非高斯噪声，Wang等人提出的全局脸方法无法^[54]获得比较好的结果；LSR方法^[1]不仅没有去除噪声，而且放大了输入图像中的噪声；SR方法^[2]的结果不够干净和平滑。NE方法^[83]和DCT邻域嵌入方法^[68]可以部分限制输入低分辨率图像中的噪声，我们将这一优点归因于局部性先验知识的引入。然而，由于NE方法^[83]和DCT邻域嵌入方法^[68]固定了近邻个数 K ，因此，在对输入图像块进行表示时，会有过拟合或者拟合不当的问题，这也是这两种方法超分辨率重构出来的结果出现块效应的原因。相比较而言，提出的方法达到了最好的重构高分辨率人脸图像，去除了输入低分辨率人脸中的大部分噪声而且重构的人脸与原始高分辨率人脸非常相似。



图 3-16 监控摄像机拍摄得到的人脸图像。上面一行为监控摄像机在低光照下拍摄得到的低分辨率人脸图像，下面一行为监控摄像机在正常光照下拍摄得到的高分辨率人脸图像。

3.6 总结与展望

在本章中，我们提出一种对噪声鲁棒的人脸超分辨率方法——基于局部约束表示的人脸像超分辨率方法。受流形空间局部线性性这一思想的启发，我们考虑引入局部约束先验，在对观察图像块进行表示时，通过选取距离近的样本图像块，惩罚距离远的图像块，达到稳定、鲁棒表示图像块的目的。提出的方法可以使原来不适定的最小二乘问题具有稳定且唯一的解。此外，在提出的方法中，每个输入低分辨率图像块被少量通过自适应选择得到的相似样本图像块线性表示，表示系数可以同时达到稀疏性（少量）和局部性（相似）。在仿真低分辨率人脸图像和实际拍摄得到的低分辨率人脸图像上的实验结果均说明了提出算法的优越性。



图 3-17 不同算法在模拟监控环境下获得的低分辨率人脸图像的重构结果。左到右依次为：输入低分辨率人脸图像，Wang 方法的重构结果^[54]，NE 方法^[83]的重构结果，LSR 方法^[1]的重构结果，SR 方法^[2]的重构结果，DCT 邻域嵌入方法^[68]的重构结果，LcR 方法的重构结果，原始高分辨率人脸图像。

未来我们将从图像块相似性度量、图像块字典构建和算法计算复杂优化等三个方面对已提出的LcR方法进行改进：

(1) 在提出的局部约束表示模型中，我们利用输入低分辨率图像块和低分辨率训练样本图像块的欧氏距离来定义局部性，无法发现数据空间内在的几何结构信息。因此，设计一种新的距离测量算法，它可以利用训练样本图像块的内在流形结构（准确的先验知识），将是我们的下一步的第一个工作。

(2) 本章实验结果已表明，局部性对于基于图像块的方法而言是至关重要的。然而，我们只是简单地使用原始训练样本图像块进行局部性约束表示。给定原始训练样本图像块，如何学习一个紧凑和具有表达能力的训练字典，以更好地发挥局部性和稀疏性将是我们的下一步的第二个工作。

(3) 在所有基于图像块的人脸超分辨率算法中，相互交叠图像块的表示和重构是非常耗时的，这将导致我们的方法在难以在实际中应用（例如，人脸识别和实时三维人脸合成这些对实时性很高的应用）。由于每个目标的高分辨率人脸图像块的重构是相互独立的，因此我们可以通过并行处理的方式进行加速算法。

第四章 基于高分辨率流形空间约束的人脸超分辨率算法

4.1 引言

图像的降质过程是不可逆、病态的，由于降质过程会导致信息的损失，高分辨率图像与低分辨率图像之间的对应关系是“多对一”的关系^[95]。即对于观察得到的一张低分辨率图像，可能存在多张高分辨率图像与之对应。高低分辨率流形空间具有一致局部几何结构的假设在实际情况下并不成立，低分辨率图像空间的近邻关系并不能真实地反映原始高分辨率图像空间的近邻关系，如图4-1所示。

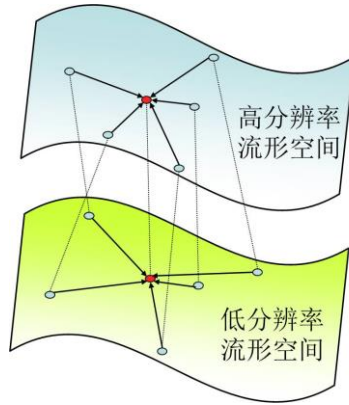


图4-1 高分辨率流形空间与低分辨率流形空间的局部近邻关系不一致示意。

现有基于学习的图像超分辨率方法大致可以分为两大类：基于编码表示方法和基于回归预测方法。

基于编码表示的图像超分辨率方法假设由高分辨率图像块构成的流形空间和由低分辨率图像块构成的流形空间具有相似的局部几何结构特性^[83]。这类方法先是得到低分辨率图像空间的几何关系（样本间的重构关系），然后将得到的几何关系保持到重构后的高分辨率图像空间，进而推测出目标高分辨率图像。如图4-2所示，首先，对于输入的低分辨率图像块利用低分辨率训练集图像块进行表示，得到表示系数。然后，目标高分辨率图像块便可用同样的表示系数和对应的高分辨率训练集图像块加权得到。最典型的基于编码表示的图像超分辨率是由Chang等人^[83]提出的近邻保持方法（Neighbor Embedding, 简称NE），对于输入低分辨率图像块，该方法通过选择低分辨率图像块流形空间中最相似的 K 个样本对其进行线性表示，再把表示系数直接映射到对应的 K 个高分辨率样本图像块上，得到目标高分辨率图像块。随后Yang等人^[8]提出一种基于稀疏表示（Sparse Representation, 简称SR）的图像超分辨率方法，该方法通过引入稀疏性先验，可以避免NE方法^[83]由于选择固定 K 个样本导致的对输入低分辨率图像块过拟合或者拟合不当的问题。我们之前提出的局部约束表示方法^[137, 138]（Locality-constrained Representation, 简称LcR）在图像块的表示过程中引入局部性约束，通过

选取距离近的样本图像块，惩罚距离远的样本图像块，表示系数同时达到了稀疏性和局部性，也获得了很好的人脸超分辨率重构性能。

然而，由于高低分辨率图像之间“多对一”的对应关系，基于编码表示的人脸超分辨率方法的流形一致性的前提假设太过理想，在实际情况下并不成立。尽管近年来提出了很多改进方法来降低这种不一致性带来的影响，但是它们的本质还是将受降质过程影响的低分辨率流形空间的几何关系保持到重构后的高分辨率流形空间中，无法正确揭示训练样本构成流形空间的几何结构关系。

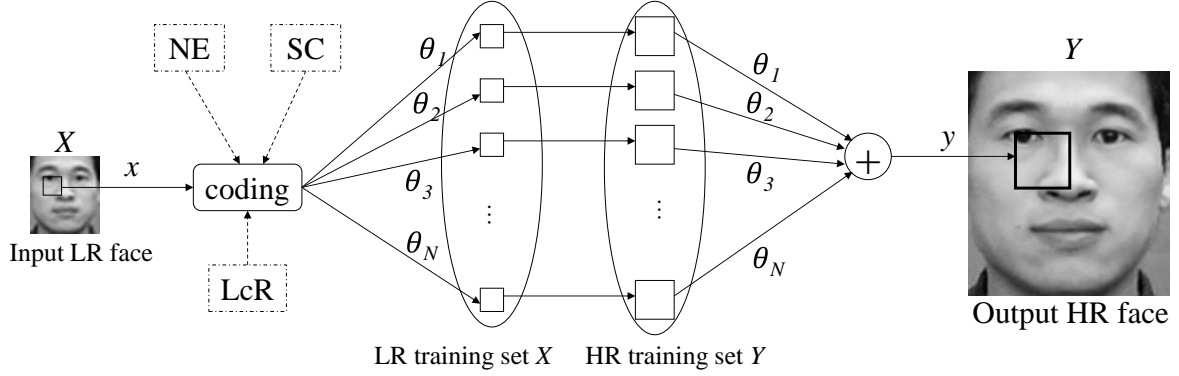


图4-2 基于编码表示的人脸超分辨率方法原理示意图。

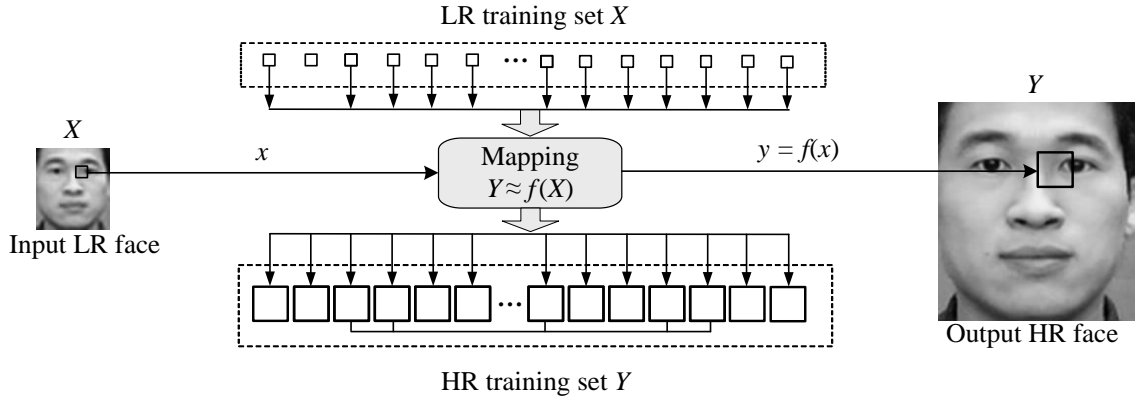


图4-3 基于回归预测的人脸超分辨率方法原理示意图。

另外一类基于学习的图像超分辨率方法是基于回归预测的方法。这类方法将图像超分辨率看作一种回归问题，直接通过学习一系列低分辨率图像空间到高分辨率图像空间的映射函数来预测目标高分辨率图像块，它们不需要任何额外的假设。图4-3是这类方法的原理示意图。Tappen等人^[139]将图像超分辨率重构中的高频细节估计问题看作一个聚类样本库上的多线性回归问题。受此回归思想的启发，Ni和Nguyen^[140]提出一种支撑向量回归图像超分辨率方法。Kim和Kwon^[113]在此基础上提出一种核岭回归（Kernel Ridge Regression, 简称KRR）方法。为了去除由回归预测导致的重构图像边缘模糊和振铃效应，这些方法往往考虑采用连续性先验

来约束重构后的高分辨率图像。近来, Tang等人^[114]试图在稀疏编码空间进行线性回归学习, 并提出一种基于贪婪稀疏编码回归 (Greedy Sparse Coding Regression, 简称GSCR) 的图像超分辨率方法。通过引入人脸图像的位置块先验, Huang和Wu^[65]提出一种针对人脸图像的局部线性转换 (Local Linear Transformations, 简称LLT) 图像超分辨率算法。该方法为每一个位置上的人脸图像块学习一个由低分辨率图像空间到高分辨率图像空间的线性转换函数, 每次预测只需要简单的矩阵相乘运算, 因此该算法非常有效且节省资源。LLT方法^[65]本质上是一种全局回归模型, 提前计算得到的线性转换函数无法根据观察到的低分辨率图像进行自适应的调整, 因此它的推广能力非常有限。

从以上的讨论, 我们可以看出: 高分辨率图像流形空间和低分辨率图像流形空间之间局部几何结构的不一致性是基于编码表示的图像超分辨率方法的最大障碍。而基于回归预测的图像超分辨率方法不依赖于这种强的“相同的表示权重”假设, 学习方式更加灵活可靠。为此, 在给定低分辨率图像的条件下, 我们直接学习低分辨率图像空间到高分辨率图像空间的映射函数来预测目标高分辨率图像, 而不需要基于编码表示的图像超分辨率方法中“相同的表示权重”这一强制性假设。为了充分利用不受降质过程影响的高分辨率图像空间的几何结构特征, 我们进一步提出了一种局部诱导支撑回归框架: 在观察低分辨率图像块的支撑集空间学习低分辨率样本图像子集到高分辨率样本图像子集的映射关系。通过引入流形结构的局部性约束, 本章提出两种支撑集的生成方式, 第一种支撑集生成方法是由输入低分辨率图像块和低分辨率训练样本之间 K 个最近邻居的下标构成, 也就是所谓的低分辨率图像块支撑诱导的生成策略; 第二种支撑集生成方法是由目标高分辨率图像块和高分辨率训练样本之间 K 个最近邻居的下标构成, 即高分辨率图像块支撑诱导的生成策略。对于高分辨率图像块支撑诱导的生成策略, 由于无法提前获得输入低分辨率图像块对应的目标高分辨率图像块, 为此我们设计一种迭代方案来更新支撑集, 从上一步估计得到的高低分辨率支撑集样本对中学习新的映射关系。

与以往的人脸超分辨率方法相比, 本章提出方法的贡献可以概括为以下几点:

(1) 传统基于编码表示的方法假设对应高低分辨率图像块具有“相同的表示权重”, 然而这一假设在实际情况下过于苛刻。本章提出的方法将对应高低分辨率图像块具有“相同的表示权重”这一假设条件放松为对应高低分辨率图像块具有“同样支撑集”。因而, 提出的方法对图像的重构将更加灵活;

(2) 与基于回归预测方法中学习一个全局的映射函数不同 (即所有观察图像块共用一个映射模型), 本章提出的方法为每一个观察图像块设计一个对应的高低分辨率图像支撑集, 并在此特定支撑集上学习高低分辨率图像之间的对应关系。因而, 提出的方法相对于输入图像块而言具有很好的自适应性;

(3) 与那些完全不考虑原始高分辨率图像块空间几何结构的回归方法相比, 本章提出的方

法利用原始高分辨率图像块空间的几何近邻关系来定义支撑集，并使用该几何结构来约束高低分辨率样本图像块之间的映射关系。经过多次迭代，提出的方法能够重构出更多的人脸特征细节。

本章接下来的安排如下：第4.2节详细介绍提出的人脸超分辨率方法。在这一小节中，我们先介绍了本章所用到的符号，接着介绍两类学习模型（基于编码表的方法和基于回归预测的方法）的形式化表示，最后给出了提出局部诱导方法的两种形式（低分辨率图像块指导下局部诱导的支撑集回归方法（Low-Resolution patch guided Locality-induced Support Regression, 简称LR-LiSR）和高分辨率图像块指导下局部诱导的支撑集回归方法（High-Resolution patch guided Locality-induced Support Regression, 简称HR-LiSR））以及它们的详细实现细节；第4.3节通过在CAS-PEAL-R1人脸库^[4]、带遮挡的AR人脸库^[7]和两种实际图像数据库（包括CUM+MIT人脸图像库^[134]和实际监控人脸图像库）上的主客观对比实验验证算法的有效性以及对遮挡人脸的鲁棒性，并分析了迭代次数、支撑集大小即近邻个数、支撑集产生方式、训练集大小、全局约束等因素对算法性能的影响；第4.4节对算法进行讨论与分析，分析了全局约束和局部约束的优劣、高分辨率流形结构保持和低分辨率流形结构保持的优劣、一揽子方案和迭代方法的优劣；第4.5节对本章进行了总结，讨论了非正面人脸图像的超分辨率问题以及在下一阶段可能的改进方向。

4.2 基于高分辨率流形结构保持的嵌入学习

4.2.1 符号定义及问题提出

在本节中，我们给出了提出方法的具体细节。首先介绍本章用到的一些术语及对应的数学符号。如上所述，人脸超分辨率问题可以转化为在给定高分辨率人脸图像训练样本 $Y = \{y_n\}_{n=1}^N$ 和低分辨率人脸图像训练样本 $X = \{x_n\}_{n=1}^N$ 的条件下，由输入低分辨率人脸图像 x_i 预测目标高分辨率人脸图像 y_i 的过程。其中， N 表示训练集的大小。在我们的方法中，高分辨率训练样本人脸图像、低分辨率训练样本人脸图像和输入低分辨率人脸图像都被划分成 M 个相互交叠的图像块， $\{y_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$ ， $\{x_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$ ，和 $\{x_i(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$ 。其中， $M=UV$ ， U 代表一张人脸图像上每一列的图像块数， V 代表一张人脸图像上每一行的图像块数， (i, j) 表示图像块在图像块坐标系下的坐标。基于图像块的人脸超分辨率方法就是将输入低分辨率人脸图像块的每一个低分辨率块 $x_i(i, j)$ 超分辨率重构成高分辨率图像块 $y_i(i, j)$ 。按照图像块的位置，拼接、融合所有超分辨率重构得到的高分辨率图像块 $\{y_i(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$ ，就可以得到一张目标高分辨率人脸图像 y_i 。下面介绍两类基于学习的人脸超分辨率方法（基于编码表示的方法和基于回归预测的方法）的形式化描述：

(1) **基于编码的人脸超分辨率方法。**利用与输入低分辨率图像块相同位置的所有低分辨率

样本图像块对输入低分辨率图像块进行表示，由此获得输入低分辨率图像块的编码系数：

$$w^* = \arg \min \{ \|x_i(i, j) - X(i, j)w\|_2^2 + \lambda E(w) \}, \quad (4-1)$$

其中， $X(i, j)$ 可以看作样本图像块字典库，其每一列是由位置为 (i, j) 的所有低分辨率样本图像块 $x_m(i, j)$ 拉直而构成，因此有 $X(i, j) = [x_1(i, j), x_2(i, j), \dots, x_N(i, j)]$ 。 $E(w)$ 是关于编码系数 w 的先验知识，这种先验知识可以使得重构后的人脸图像具有某些性质，例如平滑性或稀疏性， λ 是一个正则化参数因子，用来控制低分辨率图像块的重构误差和关于编码系数 w 的先验信息之间的平衡。一旦获得了最优编码系数 w^* ，这些基于编码表示的方法直接将上述得到的编码系数直接映射到对应同样位置的高分辨率训练样本图像块， $Y(i, j) = [y_1(i, j), y_2(i, j), \dots, y_N(i, j)]$ ，从而获得高分辨率图像块：

$$y_i(i, j) = Y(i, j)w^*. \quad (4-2)$$

基于编码表示的人脸超分辨率方法都假设每对高低分辨率图像块具有相同的表示系数。然而，由于受到降质过程的影响，原始高分辨率图像与降质之后的低分辨率图像之间是“多对一”的映射关系，这种强制性的“相同的表示权重”假设通常在实际情况下并不成立。

(2) 基于回归预测的人脸超分辨率方法。与基于编码表示的人脸超分辨率方法不同，此类方法没有严格要求对应高低分辨率图像块具有相同的表示权重。他们直接构建高分辨率训练图像空间与低分辨率训练图像空间的映射关系， $f: x \rightarrow y$ ，然后根据输入低分辨率人脸图像块 $x_i(i, j)$ ，利用学习得到的关系模型来预测目标高分辨率图像块：

$$y_i(i, j) = f(x_i(i, j)). \quad (4-3)$$

基于编码表示的方法要求高低分辨率图像之间具有“相同的表示权重”这一强约束条件，而基于回归预测的方法无需这种强的约束条件，直接学习高低分辨率图像的对应关系，因而基于回归预测的模型更加灵活（由于同一位置上的人脸图像块具有很强结构相似性，通过回归模型来预测目标图像块是有可能的）。

4.2.2 局部诱导的支撑回归方法

受上述基于编码表示和基于回归预测两类方法的启发，在本章中，我们提出一种新颖的局部线性回归方法。通过直接构建高低分辨率支撑训练样本（训练样本的标号属于支撑集的所有样本）之间的对应关系，预测目标高分辨率人脸图像。特别地，上述支撑集信息可以通过以下两种方式获得：第一种是由输入低分辨率图像块和低分辨率训练样本之间 K 个最近邻居的下标构成，也就是所谓的低分辨率图像块支撑诱导的生成策略；第二种是由目标高分辨率图像块和高分辨率训练样本之间 K 个最近邻居的下标构成，也就是所谓的高分辨率图像块支撑诱导的生成策略。这也就是提出局部线性回归方法的两个变种：低分辨率图像块指导下局部诱导的支撑集回归方法（LR-LiSR）和高分辨率图像块指导下局部诱导的支撑集回归方法（HR-LiSR）。在接下来的两个小节，我们将详细介绍这两种方法。

4.2.2.1 低分辨率图像块指导下局部诱导的支撑集回归方法

相较于基于编码表示的人脸超分辨率方法^[1, 2, 8, 43, 54, 71, 83]，在我们提出的方法中，高低分辨率图像块具有“相同的表示权重”这一强约束条件被放松成“相同支撑集”，即只要求用来表示高低分辨率对的训练样本图像块是一致的，而无需满足表示系数的数值也一致的强约束。因此，提出的方法是在高低分辨率支撑训练样本空间来构建高低分辨率图像块之间的关系，而不是在整个高低分辨率训练样本空间。

假设给定了某一位置 (i, j) 上高低分辨率样本图像块对：

$$\Omega = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}, \quad (4-4)$$

为了表达方便，上式省掉了位置指示项 (i, j) 。对于一个输入低分辨率图像块 x_i ，我们尝试学习得到一个从支撑低分辨率样本图像块空间到支撑高分辨率样本图像块空间的映射函数， $f(x, P) = Px$ 。通过最小化以下目标函数，便可以得到要求的映射矩阵：

$$O_{LR-LiSR} = \sum_{n \in S_L} (Px_n - y_n)^2 + \alpha \|P\|_F^2, \quad (4-5)$$

其中， P 是要求解的映射函数， $\|P\|_F^2$ 是平滑项， α 是一个防止矩阵不可逆的微小正则因子（例如， $\alpha = 10^{-5}$ ）， S_L 是输入低分辨率图像块在训练样本空间的支撑集。那么，接下来的任务就是如何为输入低分辨率图像块定义（估计）其支持集。

近年来，流形学习理论表明^[84, 141]，虽然我们观察到的数据均被表示成高维的特征向量，但其实它们的潜在维度比较低，也就是说，这些高维观察数据可能嵌入在一个相对较低维的潜在空间。流形学习的核心思想就是局部几何结构的度量和保持。这种数据局部性被广泛地应用于许多计算机视觉应用中，如聚类^[142]、维数约减^[84, 141]和图像分类^[135, 136, 143]等问题中。在本章节，我们提出利用流形的这种局部结构来定义高低分辨率图像块流形的回归关系。其核心思想是：局部地（样本图像块空间小区域内）学习高低分辨率图像块空间的映射关系。特别地，图像块的近邻信息被用来指导支撑集的定义。那么，支撑集就可以定义成如下形式：

$$S_L = \text{supp}(dist_L|_K), \quad (4-6)$$

其中， $dist_L|_K$ 表示 $dist_L$ 中 K 个最小的值， $\text{supp}(dist_L|_K)$ 就 $dist_L$ 中 K 个最小的值的下标号组成的集合。 $dist_L$ 是输入低分辨率图像块与低分辨率训练样本图像块之间的平方欧式距离：

$$dist_{Ln} = \|x_i - x_n\|_2^2, \quad 1 \leq n \leq N, \quad (4-7)$$

其中， x_n 是低分辨率训练样本图像块集合中的第 n 个低分辨率图像块。

假设 $X_{S_L} = \{x_n | x_n \in X, n \in S_L\}$ 和 $Y_{S_L} = \{y_n | y_n \in Y, n \in S_L\}$ ，那么目标函数（4-5）可以写作如下矩阵形式：

$$O_{LR-LiSR} = \|PX_{S_L} - Y_{S_L}\|_F^2 + \alpha \|P\|_F^2. \quad (4-8)$$

基于 LR-LiSR 人脸超分辨率方法的伪代码如算法 4-1 所示。

算法4-1 低分辨率图像块指导下局部诱导的支撑集回归方法

1. **输入:** 高分辨率人脸图像训练集 $X = \{x_n\}_{n=1}^N$, 低分辨率人脸图像训练集 $Y = \{y_n\}_{n=1}^N$, 输入低分辨率人脸图像 x_i , 图像块大小, 近邻图像块之间的交叠像素, 支撑集大小 (近邻个数 K), 最大迭代次数 $maxIter$;
2. **输出:** 重构后的高分辨率人脸图像 y_i ;
3. 将高分辨率人脸图像训练集、低分辨率人脸图像训练集和插值放大之后的低分辨率人脸图像按照同样的划分策略分别分成图像块, $\{y_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$, $\{x_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$, $\{x_i(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$;
4. 对于输入低分辨率人脸图像上的每一个图像块 $x_i(i, j)$;
5. 计算输入低分辨率图像块 $x_i(i, j)$ 和训练集中所有 N 个低分辨率样本图像块 $\{x_n(i, j)\}_{n=1}^N$ 之间的平方欧式距离:

$$d_{L_n}(i, j) = \|x_i(i, j) - x_n(i, j)\|_2^2, \quad 1 \leq n \leq N ;$$

6. 获得低分辨率图像诱导的支撑集:

$$S_L = \text{supp}(\text{dist}_L |_K) ;$$

7. 获得低分辨率支撑样本集合和对应的高分辨率支撑样本集合:

$$X_{S_L} = \{x_n | x_n \in X, n \in S_L\} \text{ 和 } Y_{S_L} = \{y_n | y_n \in Y, n \in S_L\} ;$$

8. 求解优化目标函数 (4-8) 得到低分辨率支撑样本集合到高分辨率支撑样本集合的映射矩阵:

$$P = Y_{S_L} X_{S_L}^T (X_{S_L} X_{S_L}^T + \alpha I)^{-1} ;$$

9. 通过下式估计高分辨率图像块 y_i :

$$y_i = P x_i ;$$

10. 重复步骤5-步骤9直至所有图像块按照同样的重构方法被处理完毕;
 11. 将所有加权重构出的高分辨率图像块按照其在人脸图像上的位置进行叠加, 然后除以每个像素位置被交叠的次数, 最终重构出高分辨率人脸图像 y_i 。
-

4.2.2.2 高分辨率图像块指导下局部诱导的支撑集回归方法

在以上提出的 LR-LiSR 方法中, 支撑集的产生和目标高分辨率图像块的重构是相互独立的。实际上 LR-LiSR 方法只考虑了低分辨率流形空间的几何结构, 而忽略了样本近邻关系更加可信的原始高分辨率图像块流形空间的结构信息。在接下来, 我们将从两个方面进一步改进提出的 LR-LiSR 方法: 第一, 与 LR-LiSR 方法估计支撑集的方式不同, 提出通过利用高分辨率

流形结构来获取支撑集信息；第二，设计了一种迭代策略来更新支撑集和映射函数。通过以上两个改进，利用映射函数估计所得的高分辨率图像块作为指导信息，将进一步改善支持集的精确性，进而可以得到更好的高分辨率图像块预测。

基于以上讨论，我们给出HR-LiSR的目标函数：

$$O_{HR-LiSR} = \sum_{n \in S_H} (Px_n - y_n)^2 + \alpha \|P\|_F^2 \quad (4-9)$$

给定高低分辨率图像训练样本 X 和 Y ，我们的目标就是同时获得支撑集 S_H 和高分辨率图像之间的映射函数 P 。此处的支撑集是通过估计得到的高分辨率图像块诱导产生的，由估计得到的高分辨率图像块的 K 个最相似的高分辨率样本图像块的下标组成：

$$S_H = \text{supp}(dist_H|_K), \quad (4-10)$$

$dist_H$ 是输入高分辨率图像块与高分辨率训练样本图像块之间的平方欧式距离：

$$dist_{Hn} = \|y_t - y_n\|_2^2, \quad 1 \leq n \leq N, \quad (4-11)$$

其中， y_m 是高分辨率训练样本图像块集合中的第 m 个高分辨率图像块， $dist_H|_K$ 表示 $dist_H$ 中 K 个最小的值， $\text{supp}(dist_H|_K)$ 就 $dist_H$ 中 K 个最小的值的下标号组成的集。

假设 $X_{S_H} = \{x_n | x_n \in X, n \in S_H\}$ 和 $Y_{S_H} = \{y_n | y_n \in Y, n \in S_H\}$ 目标函数可以简化为如下形式：

$$O_{HR-LiSR} = \|PX_{S_H} - Y_{S_H}\|_F^2 + \alpha \|P\|_F^2. \quad (4-12)$$

该目标函数就是要最小化低分辨率图像支撑样本到高分辨率图像支撑样本的映射误差，给定了支撑集 S_H ，上式是一个关于 P 的二次型，它的全局最优解可以通过二次规划算法进行求解。在本章节，我们直接对式（4-12）进行求导，并令导数为零来找到全局最优解。

利用矩阵性质 $\text{tr}(AB) = \text{tr}(BA)$ ， $\|A\|^2 = \text{tr}(AA^T)$ ，and $\text{tr}(A) = \text{tr}(A^T)$ ，我们有：

$$\begin{aligned} O_{HR-LiSR} &= \|PX_{S_H} - Y_{S_H}\|_F^2 + \alpha \|P\|_F^2 \\ &= \text{tr}\{(PX_{S_H} - Y_{S_H})(PX_{S_H} - Y_{S_H})^T\} + \alpha \text{tr}(PP^T). \end{aligned} \quad (4-13)$$

为了最小化式（4-13），我们直接对它求关于 P 的导数并置零，于是我们得到映射矩阵 P ：

$$\begin{aligned} 2PX_{S_H}X_{S_H}^T - 2Y_{S_H}X_{S_H}^T + 2\alpha P &= 0 \\ \Rightarrow P(X_{S_H}X_{S_H}^T + \alpha I) &= Y_{S_H}X_{S_H}^T \\ \Rightarrow P &= Y_{S_H}X_{S_H}^T(X_{S_H}X_{S_H}^T + \alpha I)^{-1}. \end{aligned} \quad (4-14)$$

本章提出的基于回归预测的人脸超分辨率算法根据每个位置块上学习得到的回归关系对输入低分辨率图像 x_i 进行预测，最终超分辨率重构得到高分辨率人脸图像 y_i 。我们首次根据位置将输入的整张低分辨率人脸图像划分为多个低分辨率图像小块，按照光栅扫描顺序（从左到右，从上到下）对所有图像块进行处理。然后，我们利用提出的LR-LiSR方法和HR-LiSR方法对每个

输入低分辨率图像块进行预测，估计得到高分辨率图像块。根据文献^[83]和文献^[8]，我们对近邻图像块施加兼容性约束（近邻图像块加权求平均）。基于HR-LiSR人脸超分辨率方法的伪代码如算法4-2所示。

算法4-2 高分辨率图像块指导下局部诱导的支撑集回归方法

1. **输入：** 高分辨率人脸图像训练集合 $Y = \{y_n\}_{n=1}^N$ ，低分辨率人脸图像训练集合 $X = \{x_n\}_{n=1}^N$ ，输入低分辨率人脸图像 x_i ，图像块大小，近邻图像块之间的交叠像素，支撑集大小（近邻个数 K ），最大迭代次数 $maxIter$ ；
2. **输出：** 重构高分辨率人脸图像 y_i ；
3. 将高分辨率人脸图像训练集、低分辨率人脸图像训练集按照同样的划分策略分成图像块， $\{y_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$ 和 $\{x_n(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}_{n=1}^N$ ；
4. 初始化： $r = 0$ ， $y_i^0 = Bicubic(x_i)$ ；
5. 将上采样之后的输入低分辨率人脸图像 y_i^0 按照与步骤3同样的方式划分为图像块， $\{y_i(i, j) | 1 \leq i \leq U, 1 \leq j \leq V\}$ ；
6. 对于上一步估计得到高分辨率人脸图像上的每一个图像块 y_i^r ；
7. 计算 y_i^r 和训练集中所有 N 个高分辨率样本图像块 $\{y_n(i, j)\}_{n=1}^N$ 之间的平方欧式距离：

$$dist_{Hn}^r = \|y_i^r - y_n\|_2^2, \quad 1 \leq n \leq N;$$

8. 获得高分辨率图像诱导的支撑集：

$$S_H^r = \text{support}(dist_H^r |_K);$$

9. 获得低分辨率支撑样本集合和对应的高分辨率支撑样本集合：

$$X_{S_H}^r = \{x_n | x_n \in X, n \in S_H\} \quad \text{和} \quad Y_{S_H}^r = \{y_n | y_n \in Y, n \in S_H\};$$

10. 固定支撑集 S_H^r ，求解优化目标函数（4-12）得到低分辨率支撑样本集到高分辨率支撑样本集的映射矩阵：

$$P^r = Y_{S_H}^r (X_{S_H}^r)^T (X_{S_H}^r (X_{S_H}^r)^T + \alpha I)^{-1};$$

11. 通过下式计算目标高分辨率图像块：

$$y_i^{r+1} = P^r x_i;$$

12. 迭代次数加1： $r = r + 1$ ；
 13. 重复步骤5步骤11直至 $r > maxIter$ ；
 14. 重复步骤5-步骤12直至所有图像块被处理完毕；
 15. 将所有加权重构出的高分辨率图像块按照其在人脸图像上的位置进行叠加，然后除以每个像素位置被交叠的次数，最终重构出高分辨率人脸图像 y_i 。
-

4.3 实验结果分析

在本章实验中，我们详细介绍了实验细节，通过与前沿人脸超分辨率算法的对比，验证了提出人脸超分辨率算法的有效性。本章实验的目的是为了回答以下四个问题：

- 与前沿人脸超分辨率算法的对比结果如何，提出方法对于实际情况拍摄得到的低分辨率人脸图像的重构结果如何？
- 在高分辨率样本图像块空间进行迭代式 K 近邻搜索（生成支撑集），并从支撑集高低分辨率样本图像块对中学习映射函数的迭代策略，这种迭代策略是否必要、是否收敛？
- 流形空间局部性先验约束的引入，是否有利于人脸超分辨率方法成功地重构出一张很好的高分辨率人脸图像？
- 迭代次数、支撑集大小即近邻个数、支撑集产生方式、训练集大小和全局约束等因素如何影响超分辨率方法的结果？

4.3.1 对比算法与参数设置

本章实验在CAS-PEAL-R1人脸库^[4]、AR人脸库^[7]和CMU+MIT人脸库^[134]进行。关于这几个数据库的详细介绍将在给出对应实验结果的时候进行说明。我们对比了提出方法（LR-LiSR和HR-LiSR）和目前几个最主流方法的主客观结果。这些对比算法主要包括：Wang等人的特征转换全局脸方法^[54]、三种基于编码的人脸超分辨率方法（Chang的NE方法^[83]，Yang的SR方法^[8]和我们之前提出的LcR方法^[137, 138]）和三种基于回归的人脸超分辨率方法（Kim等人提出的KRR方法^[113]，Tang等人提出的GSCR方法^[114]和Huang等人提出的LLT方法^[65]）。尽管目前有许多有影响力的人脸超分辨率工作，这里只给出了一些代表的工作。

为了得到最好的重构结果，我们为不同的对比算法尝试各种参数设置，使其达到最好性能。具体地，对于Wang等人提出的特征转换全局脸方法^[54]，在CAS-PEAL-R1人脸库和AR人脸库的所有实验中，提取主成分信息能量占原始信号能量的99%。对于基于图像块的方法，高分辨率图像块大小设为 12×12 像素，相邻图像块之间交叠4个像素；对应的低分辨率图像块大小设为 4×4 像素，相邻图像块之间交叠1个像素。很显然地，在一定的图像块大小下，相邻图像块之间的交叠像素越多，人脸超分辨率重构的结果将越好；对于Chang等人提出的NE方法^[83]，在CAS-PEAL-R1人脸库的实验中支撑集大小即近邻个数设为50，在AR人脸库的实验中支撑集大小即近邻个数设为100；对于Yang等人提出的SR方法^[8]和Kim等人提出的KRR方法^[113]，我们直接利用论文作者在其主页上公布的源代码进行测试；对我们之前提出的LcR方法^[137, 138]，在CAS-PEAL-R1人脸库和AR人脸库的所有实验中局部约束参数均设置成0.04；对于Tang等人提出的GSCR方法^[114]，我们随机地提取了50,000组高低分辨率图像块对，并训练得到1024个高低分辨率图像块原子字典对；对于我们提出的LR-LiSR方法，只有一个参数（支撑集大小即近

邻个数 K)需要调节。在CAS-PEAL-R1人脸库和AR人脸库的所有实验中,支撑集大小即近邻个数 K 都被设为100;对于我们提出的HR-LiSR方法,它只有两个参数需要调节,支撑集大小即近邻个数 K 和最大迭代次数 $maxIter$ 。在CAS-PEAL-R1人脸库和AR人脸库的所有实验中,支撑集大小即近邻个数 K 同样地都被设为100,最大迭代次数 $maxIter$ 则设成3(在接下来的小节,我们将测试这两个参数对提出算法的影响)。

4.3.2 CAS-PEAL-R1 人脸库上的对比结果

在本小节中,我们将在公开人脸库——CAS-PEAL-R1人脸库上进行算法性能测试,它包括1040个对象的30871幅人脸图像。我们只利用每个对象的中性表情和正常光照人脸图像进行实验。在所有的1040张正面人脸图像中,我们随机选择1000张人脸图像进行训练,利用剩下其它40张人脸图片作为测试图像。因此,所有的测试对象都不在训练人脸图像中出现。所有的人脸图像均是由两个眼睛进行对齐并裁剪为 112×100 像素(如图4-4所示)。低分辨率人脸图像则通过平滑(尺寸 4×4 的一个平均滤波器)和下采样(4倍双三次插值,因而低分辨率人脸图像的大小是 28×25 像素)对应的高分辨率人脸图像的方式得到。



图4-4 CAS-PEAL-R1人脸库示例。

为了定量地评价各方法的性能,我们采用峰值信噪比(Peak Signal to Noise Ratio, 简称PSNR)和结构相似度^[6](Structural SIMilarity, 简称SSIM)两种指标。众所周知,由于PSNR评价指标很难真实反映人眼的主观感受,因此,我们采用另外一种更加符合人眼主观感受的评价指标SSIM作为辅助来评估各方法的性能。表4-1为不同方法在CAS-PEAL-R1人脸库上的客观结果。可以看到,提出的方法在PSNR和SSIM两种指标上均优于最好的对比方法(LcR方法^[137, 138])。从表4-1可以看出,结合了人脸位置先验信息的图像块超分辨率方法往往可以获得更好的性能,例如SR方法^[8], LcR方法^[137, 138]和我们提出的LiSR方法。这主要是因为样本位置块具有足够强的表达能力,可以对整个人脸进行准确表示和重构。全局脸方法对输入人脸图像比较敏感,特别是在输入的低分辨率人脸图像与训练样本不相似时,全局脸人脸超分辨率方法的重构结果将非常差。作为一种通用图像超分辨率方法, KRR方法^[113]并不适用于人脸图像,因为它所获得的先验知识是由一般自然场景图像所产生的,而没有考虑人脸图像的结构先验信息。GSCR方法^[114]从高低分辨率人脸图像训练对中学习得到一个字典对,其性能好于KRR方法^[113],但是不如基于位置块的方法。虽然LLT方法^[65]是一个基于图像块的方法,而且

结合了人脸的位置信息，但其性能明显不如其它基于位置的图像块方法。这主要是因为该方法在训练阶段只能为每一个位置提前学习得到一个固定的映射函数，因而不能自适应地根据给定的低分辨率图像块进行调整。我们先前提出的LcR法^[137, 138]和本章提出的LiSR方法的性能比其它对比较方法均有显著的提升，这是因为在它们成功地引入了局部性约束这一先验知识。而本章的方法相对于第二好的LcR方法在PSNR和SSIM上分别提升了0.46 dB 和0.0054。

表4-1 不同方法在CAS-PEAL-R1人脸库上的PSNR (dB)和SSIM性能对比。

方法	PSNR	SSIM
Wang ^[54]	26.62	0.8254
NE ^[83]	27.91	0.8868
SR ^[8]	28.27	0.8968
LcR ^[137, 138]	28.84	0.9083
KRR ^[113]	26.91	0.8597
GSCR ^[114]	28.18	0.8971
LLT ^[65]	27.64	0.8823
LR-LiSR	28.96	0.9093
HR-LiSR	29.30	0.9137
Improvement	0.46	0.0 054

图4-5为我们提出的方法以及几种代表性的人脸超分辨率方法重建的结果。可以看出，Wang等人提出的特征转换全局脸方法^[54]重建的人脸有很明显的人工效应和锯齿状的效应（如眼睛，嘴和鼻子等部位），该方法由于只考虑了全局结构因而不能有效地重建特征细节信息。图4-5(c)为NE方法^[83]的结果。这种方法是具有代表性的基于流形学习的人脸超分辨率方法，与提出方法的本质差别在于其用来保持相似局部几何结构的方式不同：在NE方法中^[83]，输入低分辨率图像块的局部几何结构是在低分辨率图像块构成的流形空间获取到的；相比之下，我们的方法利用上一步估计得到的高分辨率图像块，在高分辨率图像块流形空间计算局部流形结构。换句话说，我们的方法在给出估计得到的高分辨率图像块下，挖掘高分辨率图像块流形结构信息，充分利用了低分辨率输入人脸图像和高低分辨率两个流形结构，从而实现更加鲁棒和准确的一致流形保持学习。本章提出的方法更多的考虑了不受降质过程（噪声和高放大倍数）影响的高分辨率图像块流形空间的几何结构信息，这也解释了为什么NE方法^[83]产生模糊的结果和不合理的面部细节。



图4-5 不同算法在CAS-PEAL-R1人脸库上的超分辨率重构结果对比: (a) 输入低分辨率人脸图像; (b) Wang等人提出全局脸方法的重构结果^[54]; (c) NE方法的重构结果^[83]; (d) SR方法的重构结果^[8]; (e) LcR方法的重构结果^[137, 138]; (f) KRR方法的重构结果^[113]; (g) GSCR方法的重构结果^[114]; (h) LLT方法的重构结果^[65]; (i) LR-LiSR方法的重构结果; (j) HR-LiSR方法的重构结果; (k) 原始高分辨率人脸图像。

图4-5(d)和4-5(e)是由两种基于图像块的人脸超分辨率方法所得到的结果: SR方法^[8]和LcR方法^[137, 138]。SR方法^[8]利用稀疏约束进行超分辨率重建, 而LcR方法^[137, 138]利用局部性约束对目标函数进行正则以获得一个稳定的解。SR方法^[8]和LcR方法^[137, 138]可以获得令人满意的效果。然后, 由于这些方法都是基于高低分辨率图像块流形空间具有相似局部几何结构这一假设的, 而这一假设在实际情况下未必成立, 因此它们无法有效地重构人脸图像上的特征细节信息。KRR方法^[113]是一种通用图像超分辨率方法, 学习得到的通用图像先验模型无法应用于人脸图像。相似的, 从通用图像中学习得到的梯度先验也没有考虑人脸这一具用特殊物体的结构性, 因而KRR方法^[113]同样也不适用于人脸图像。此外, KRR方法^[113]作为一种核回归方法, KRR方法^[113]的核函数是在低分辨率图像空间定义的, 而忽略了更加可信且更具判别性的高分辨率流形空间几何结构信息。因此, KRR方法^[113]超分辨率重构出来的人脸图像非常平滑、缺乏高频细节信息。GCSR方法^[114]随机地在训练集中采样高低分辨率图像块对, 其超分辨率重构出来的结果有明显的人工效应。图4-5 (h)是LLT方法^[65]的重构建结果, 该方法利用全局块回归模型来建立低分辨率图像训练样本到高分辨率图像训练样本之间的映射关系。然后利用构

建得到的回归模型来预测输入低分辨率图像块对应的目标高分辨率图像块。与我们提出方法的结果相比，LLT方法^[65]的重构结果过于平滑且缺乏人脸特征细节信息。这主要是因为LLT方法^[65]回归模型是通过整个训练集得到，它是一种全局回归模型，不能根据输入低分辨率图像块自适应地调整。正是考虑到这一缺陷，我们提出的方法根据输入低分辨率图像块先得到与之对应的高低分辨率图像块支撑集，然后在此高低分辨率样本图像块支撑集合上学习高低分辨率图像块之间的对应关系。为了更好的估计支撑集，我们提出了一种迭代映射策略，渐进地优化重构的高分辨率图像块。如图4-5 (i)和4-5 (j)所示，HR-LiSR方法超分辨率重构出来的人脸与原始高分辨率人脸图像非常相似，且拥有比对比算法更多的特征细节信息和更清晰的边缘轮廓。



图4-6 AR人脸库示例。

表4-2 不同方法在AR人脸库上的PSNR (dB)和SSIM性能对比。

方法	PSNR	SSIM
Wang ^[54]	25.13	0.7030
NE ^[83]	28.16	0.8572
SR ^[8]	28.80	0.8603
LcR ^[137, 138]	29.35	0.8738
KRR ^[113]	25.72	0.8213
GSCR ^[114]	28.10	0.8424
LLT ^[65]	28.10	0.8424
LR-LiSR	29.16	0.8695
HR-LiSR	29.43	0.8745
Improvement	0.08	0.0007

4.3.3 AR 人脸库上的对比结果

在本小节中，我们在AR人脸库上进行实验，验证提出方法对遮挡的鲁棒性。在本实验中，遮挡类型主要包括戴墨镜或者围巾，而这种遮挡方式往往也是实际应用中经常出现的两种情况。本章选用了AR人脸库的一个子集，该子集包括100个对象（含50名男性和50名女性）的600张正面人脸图像，也就是说，每个人都有6张正面人脸图像，其中两张是戴太阳镜的，两张是带围巾的，另外两张是没有遮挡的，如图4-6所示。同样地，对于100个对象的全部600张正面人脸图像，我们随机选择90个对象（540图像）作为训练人脸图像，剩余10个对象（60张）作为测试。所有的测试图像都没有出现在训练样本中。所有的高分辨率人脸图像是由两个眼睛对准并裁剪为 112×88 像素大小而得到。低分辨率人脸图像则通过平滑（尺寸 4×4 的一个平均滤波器）和下采样（4倍双三次插值，因而低分辨率人脸图像的大小是 28×22 像素）对应的高分辨率人脸图像的方式得到，因而低分辨率人脸图像的尺寸为 28×22 像素。

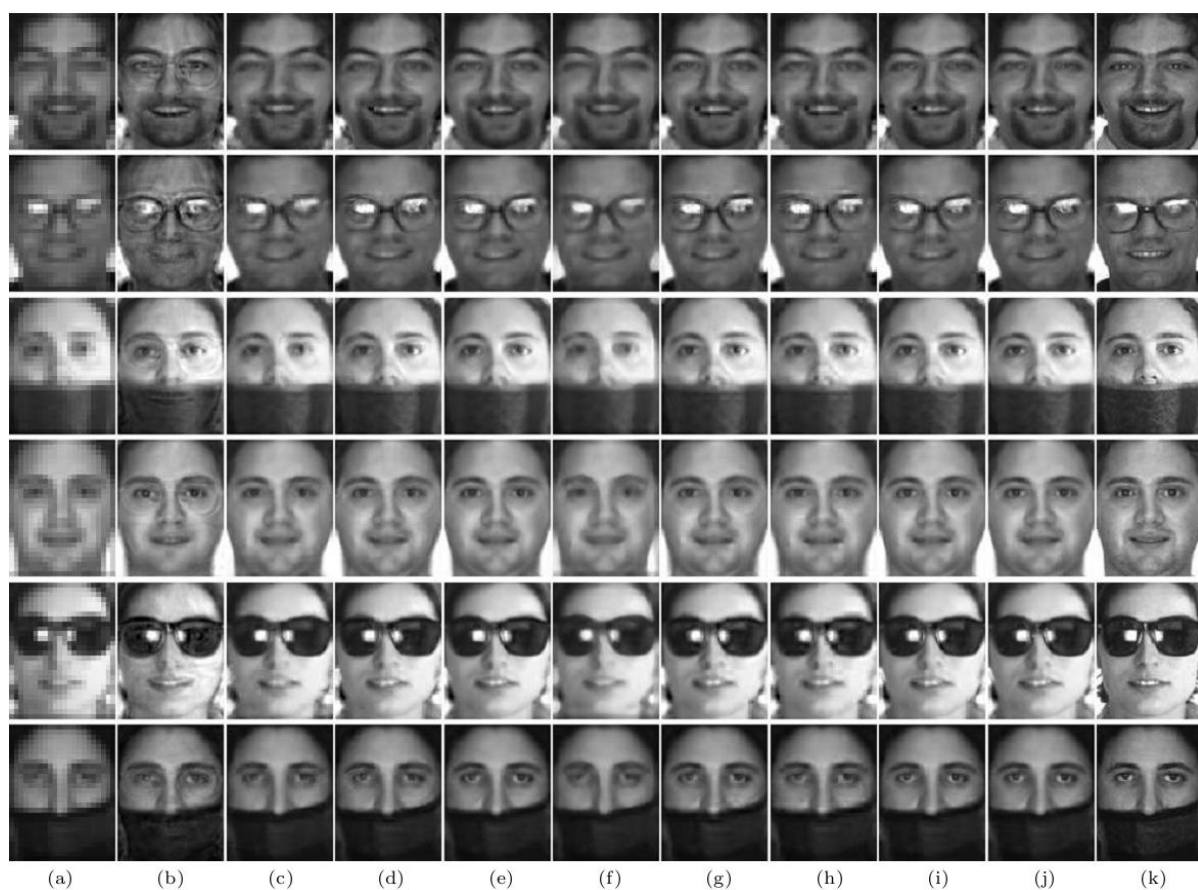


图4-7 不同算法在AR人脸库上的超分辨率重构结果对比: (a) 输入低分辨率人脸图像; (b) Wang等人提出全局脸方法的重构结果^[54]; (c) NE方法的重构结果^[83]; (d) SR方法的重构结果^[8]; (e) LcR方法的重构结果^[137, 138]; (f) KRR方法的重构结果^[113]; (g) GSCR方法的重构结果^[114]; (h) LLT方法的重构结果^[65]; (i) LR-LiSR方法的重构结果; (j) HR-LiSR方法的重构结果; (k)原始高分辨率人脸图像。

表4-2给出了不同方法在所有60张测试图像上的平均PSNR和SSIM客观指标对比结果。可以看出，提出的方法优于所有对比方法。此外，我们也在图4-7给出了所有方法的主观对比结果。通过这组实验可以看出，我们的方法在存在遮挡的情况下，仍然获得了最好的性能。Wang等人提出的特征转换全局脸方法^[54]在人脸的边缘、眼睛和嘴巴等部位有明显的“鬼影”效应。NE方法^[83]则可能重构一些原始高分辨率人脸图像中没有的特征细节，或者丢掉一些重要的信息。而KRR方法^[113]和LLT方法^[65]两种方法重构的高分辨率人脸图像比较模糊，且在人脸的轮廓部位有一些振铃效应。

4.3.4 算法参数对重构结果的影响

在接下来的小节，我们讨论了算法参数（包括迭代次数，支撑集大小即近邻个数 K ，支撑集产生方式，训练集大小和全局约束）对提出方法重构结果的影响。这里我只给出在CAS-PEAL-R1人脸库上的结果，在AR人脸库上可以得到类似的结论。

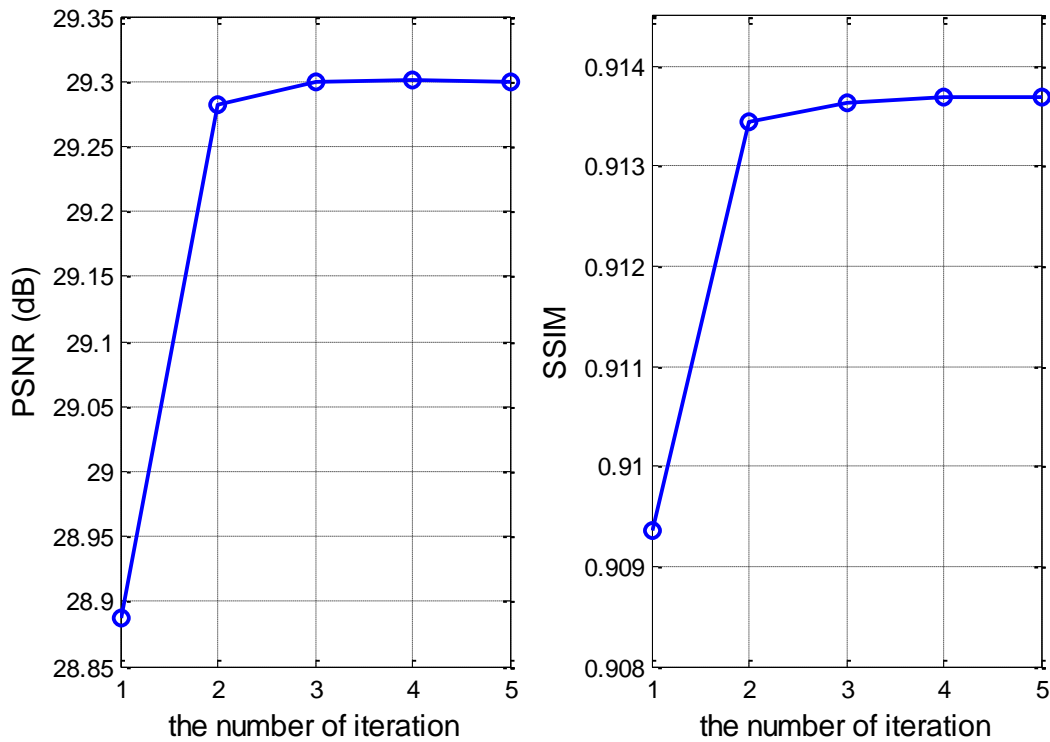


图 4-8 迭代次数对重构图像 PSNR 和 SSIM 客观结果的影响。

4.3.4.1 迭代次数 $maxIter$ 的影响

如图4-8所示，我们给出了所有40张测试人脸的平均PSNR和SSIM值随迭代次数的变化曲线。我们可以发现：（1）随着迭代次数的增加，提出的HR-LiSR方法相对于其它对比方法的增益变得越来越明显。这表明高分辨率流形结构指导的 K 近邻搜索方法对于揭示样本间的流形结构关系和回归模型的构建是至关重要的。也就是说，准确的高分辨率图像块的估计将进一步

增加支撑集的可信程度，进而提升回归模型的预测性能；（2）提出的HR-LiSR方法只需要经过少数几次迭代（例如，在实验中，我们选取3次）便可以快速地收敛，因而算法复杂度比较低。这也说明我们的方法可能用于实际的人脸图像合成处理或者人脸识别等应用中。

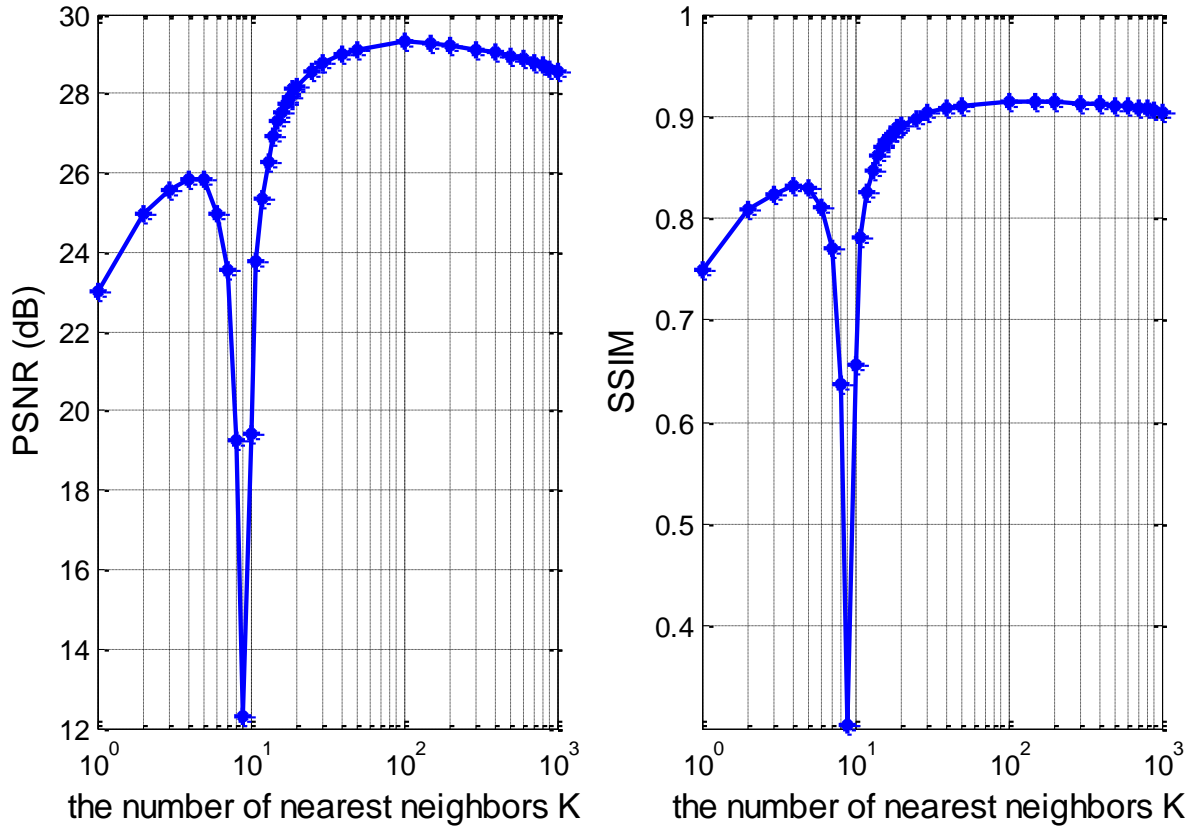


图 4-9 平均 PSNR 和 SSIM 客观指标随着近邻个数 K 的变化。

表 4-3 不同支撑集大小即近邻个数 K 下的 PSNR (dB) 性能。

支撑集大小 K	3	6	9	12	15	20	50	100	500	1000
PSNR	25.55	24.94	12.28	25.34	27.30	28.18	29.12	29.30	28.94	28.55
SSIM	0.8243	0.8111	0.3036	0.8257	0.8700	0.8903	0.9099	0.9137	0.9103	0.9046

4.3.4.2 近邻个数 K 的影响

图4-9给出了所有40张测试人脸图像在不同近邻个数 K 值下的客观PSNR和SSIM指标。从该图可以看出，PSNR和SSIM指标曲线呈现一个明显的“先上升-后下降-再上升”的特性。此外，我们也通过表4-3列出了几个具体 K 值下的PSNR和SSIM值。可以看出，该参数对算法性能有非常大的影响。在 $K=9$ 左右，算法性能的下降主要是由于输入图像块被选择的少数样本图像块过度拟合所引起的。因为在我们的实验中，输入低分辨率人脸图像块为 3×3 个像素，当 $K=9$ 时，

回归模型将对训练数据过拟合（这一现象在文献^[89]中有所报道），从而导致不合理的高分辨率重构结果。为了避免对输入待重构图像的过拟合问题，在实验中，我们将支撑集大小即近邻个数 K 设为100。

4.3.4.3 支撑集产生方式的影响

在本小节，我们测试了支撑集的产生方式对人脸超分辨率重构结果的影响。特别地，我们设计了四种不同的支撑集生成方法：估计高分辨率图像块指导的 K 近邻支撑生成方法（HR-LiSR）、低分辨率图像指导的 K 近邻支撑生成方法（LR-LiSR）、随机支撑集生成方法和估计高分辨率图像块指导的稀疏支撑生成方法。图4-10给出了上述四种支撑集产生方式下40张测试人脸图像的平均PSNR和SSIM值随支撑集大小即近邻个数 K 的变化情况。HR-LiSR方法、LR-LiSR方法、随机支撑集生成方法和稀疏支撑集生成方法的最优PSNR和SSIM结果分别是29.30 dB, 28.95 dB, 28.54 dB, 27.98 dB和0.9137, 0.9096, 0.9046, 0.8924。当 $K=1000$ 时，上述四种方法的PSNR和SSIM值分别是28.55 dB, 28.16 dB, 28.55 dB, 27.98 dB和0.9046, 0.8979, 0.9046, 0.8924。

通过设置一个合适的支撑集大小即近邻个数 K （ K 设成100），可以看出HR-LiSR方法获得了最好的性能。与LR-LiSR方法相比，平均PSNR和SSIM值分别提升了0.35 dB和0.0041。这主要是因为低分辨率图像在降质过程中会有信息丢失，高分辨率图像块流形的几何结构比低分辨率图像块的流形结构更加可信。HR-LiSR方法和LR-LiSR方法的性能好于另外两种支撑集的产生方式的性能，这主要是HR-LiSR方法和LR-LiSR方法成功地引入了局部约束项。在实验中，对于估计高分辨率图像块指导的稀疏支撑生成方法，我们仔细地调整了稀疏正则约束项参数来使其达到最好的性能，水平的红线是该方法的结果，其性能明显比提出的HR-LiSR方法和LR-LiSR方法差。以上结果表明：在人脸超分辨率重构这一问题上，局部性先验是优于稀疏性先验的。

当 $K=1000$ 时，HR-LiSR方法和随机支撑集生成方法退化成同样的支撑集产生方式，他们都选择了整个训练集作为支撑集，因而此时两种方法的性能是一致的。这也表明，通过选取合适的近邻样本集合作为支撑样本，提出的HR-LiSR方法将更加可靠合理，人脸超分辨率重构结果也更加可靠。由此我们可以得出两点结论：首次，对于支撑集的产生而言，局部性是一种很重要的先验信息，引入这种先验可以获得稳定且准确的重构结果；其次，引入其它不合理的先验将使人脸超分辨率重构性能的下降。显然地，这种由于不恰当的引入先验而对算法性能造成的负面影响会随着 K 值的增加而减少。这主要是因为一个较大的近邻个数 K 就意味着支撑集的影响对最终重构结果更少一些。当 $K=1000$ 时，HR-LiSR方法和LR-LiSR方法的唯一区别就是：前者有迭代过程，而后者没有迭代过程。从实验结果可以看出，HR-LiSR方法额外的迭代过程将为它带来PSNR上0.39 dB和SSIM上0.0067的性能提升，这也再一次证明了我们提出迭代策略的有用性。

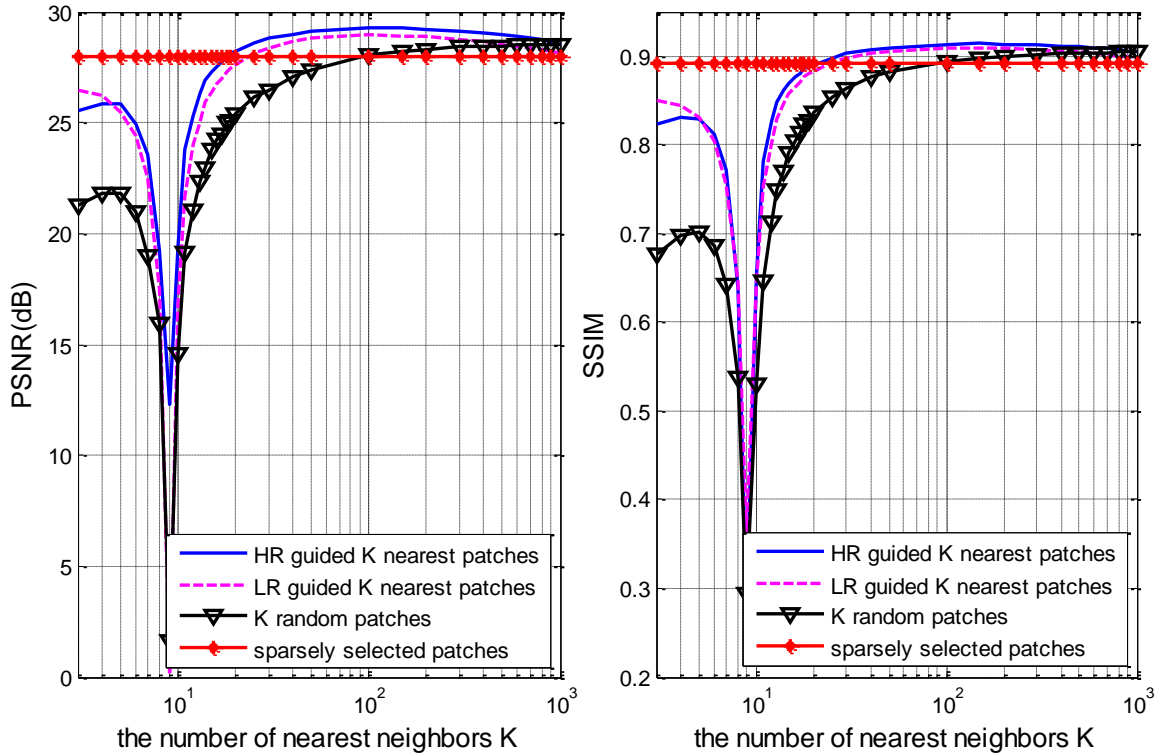


图 4-10 四种不同支撑集产生方式的 PSNR (左) 和 SSIM (右) 性能对比。

4.3.4.4 训练集大小的影响

以上实验结果表明, 为了约束图像超分辨率重构这一病态问题, Yang 等人提出方法中的稀疏先验^[8]和我们之前提出的 LcR 方法^[137, 138]以及本章提出的 LR-LiSR 方法和 HR-LiSR 方法的局部先验都是非常有效的。在上述的所有实验中, 训练集大小是 1000。很显然地, 训练集越大, 样本的表达能力越强, 因而可以获得更加准确的估计, 但大的训练集同时也会增加运行的时间。在本小节中, 我们测试了训练集大小对人脸超分辨率结果的影响。我们随机选择了三个训练子集 (训练子集大小分别为 100, 200 和 500), 而保持输入低分辨率人脸图像不变。表 4-4 给出了四种不同训练集大小下的 40 张低分辨率测试图像超分辨率重构结果的平均 PSNR 指标 (由于 KRR 方法^[113]和 GSCR 方法^[114]不受训练集大小的影响, 因此该实验不考虑这两种方法)。可以看出, 随着训练集大小的减小, 所有方法的平均性能指标都在下降。其中, 下降最快的方法是 Wang 等人的特征转换全局脸方法^[54]。这主要是因为全局脸方法是基于统计模型的, 在训练样本偏少的情况下, 这类方法无法正确揭示数据的真实分布。我们同时也注意到, 随着训练集的增大, 我们提出的 HR-LiSR 方法和 LR-LiSR 方法相较于另外三种方法的增益在逐渐提升。当训练集大小很小的时候, 提出的方法很难获得准确、合理的支撑集。因此, 学习得到的回归模型可能并不适用于输入的低分辨率人脸图像。

表 4-4 不同对比方法在不同训练集大小下的 PSNR (dB)指标。

训练集大小	Wang ^[54]	NE ^[83]	SR ^[8]	LcR ^[137, 138]	LLT ^[65]	LR-LiSR	LiSR
1000	26.62	27.91	28.27	28.84	27.06	28.96	29.30
500	25.27	27.71	28.15	28.72	27.62	28.65	28.95
200	23.38	27.44	27.88	28.39	27.50	28.20	28.43
100	22.05	27.23	27.65	28.05	27.41	27.85	28.00

4.3.4.5 全局约束的影响

由于基于图像块的方法对人脸图像进行分块地合成，没有考虑全局重建约束。为了进一步提升人脸超分辨率重构结果的性能，近年来图像超分辨率领域相关学者提出了许多后处理方法。这些后处理方法利用全局重建约束来优化基于图像块方法的重构结果，通过增加一个全局约束，从而保证了恢复的高分辨率人脸图像经过降质之后与重构后的低分辨率人脸图像一致。在实验中，我们观察到，提出的局部诱导回归模型对细节的重构是非常有效的，而全局约束后处理的引入对算法性能的提升非常有限。我们利用与文献^[65]相同的全局约束方法来优化本章提到的所有基于图像块的图像超分辨率重构方法。表4-5的结果给出了基于图像块方法和基于图像块方法联合全局约束模型方法。可以看出，SR方法^[8]、LcR方法^[137, 138]、GSCR方法^[114]和提出的LiSR方法的增益是非常有限的；相反，Bicubic插值方法、NE方法^[83]和LLT方法^[65]的增益是十分明显的。这主要是因为后三种方法本身无法重构出人脸的特征细节信息（请参见图4-5和图4-7），因而它们需要一个全局增强的过程来补偿人脸特征细节，从而优化超分辨率重构的结果。

在上述的全局约束中，我们假设图像的降质过程与输入低分辨率图像的降质过程是一致的。即要求重构的高分辨率人脸经过平滑（尺寸 4×4 的一个平均滤波器）和下采样（4倍双三次插值，因而低分辨率人脸图像的大小是 28×25 像素）后与输入低分辨率人脸图像保持一致。然而在实际情况下，我们无法提前知道降质模型的相关参数。为了测试降质过程对优化重构结果的影响，我们不改变窗口的类型，只改变平滑窗口的大小，测试降质模型不一致情况下，后处理对算法效果的影响。实验发现，在这种情况下，Bicubic插值方法，NE方法^[83]和LLT方法^[65]很难得到一个稳定的增益。图4-11给出了利用不同窗口大小作为全局约束参数，LLT方法^[65]的增益变化情况。当模糊核平滑窗口的大小不等于4的情况时，LLT方法无法得到较好的增益。由于在实际应用中，我们是无法提前知道真实降质过程的。因此，全局约束的好处是非常有限的。

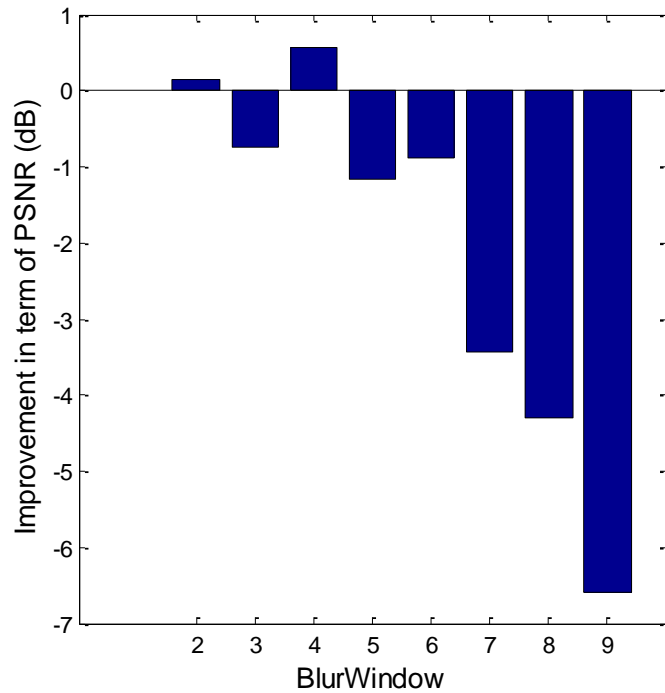


图 4-11 增加了全局约束（不同的窗口大小的模糊核）之后 LLT 方法^[65]相对于无全局约束的 LLT 方法^[65]的 PSNR (dB)增益。

表 4-5 有全局约束后处理和无后处理的图像块方法 PSNR (dB)对比。

方法名称	Bicubic	NE ^[83]	SR ^[8]	LcR ^[137,138]	KRR ^[113]	LLT ^[65]	LiSR
提出的方法	24.49	27.91	28.27	28.84	25.31	27.06	29.30
提出的方法加全局约束	26.17	28.28	28.29	28.86	26.91	27.64	29.32
PSNR 值增益	1.68	0.37	0.02	0.03	1.60	0.58	0.02

4.3.5 真实世界人脸图像上的重建结果

上述实验的输入低分辨率人脸图像都是通过简单的平滑和下采样原始高分辨率图像获得，无法反应真实情况下高分辨率图像与降质之后的低分辨率图像之间复杂的空间关系。因此，为了进一步验证提出方法在实际情况下的有效性。我们对在实际监控环境下获得的四张低分辨率人脸图像和七张来自CMU+MIT人脸库的人脸图像进行了测试。图4-12给出了原输入图像、提取出来的低分辨率人脸图像和提出方法超分辨率重构之后的高分辨率人脸图像。我们可以看出，尽管输入的低分辨率图像的成像环境（降质过程）各异，与样本库人脸图像差异很大且被强噪声干扰，但是提出的方法仍然可以产生合理的人脸超分辨率重构效果。



图 4-12 监控环境拍摄得到人脸图像的超分辨率重建结果。共(a)、(b)、(c)和(d)四组，每一组第一行为监控摄像头拍摄得到的人脸图像；第二行左边人脸图像为抠取出来的低分辨率人脸图像经过手工对齐和光照调整之后的人脸图像，右边为对应的人脸超分辨率重构结果图像。



图 4-13 CMU+MIT 人脸库图像的人脸超分辨率重建结果。

4.3.6 讨论

4.3.6.1 全局约束和局部约束对比分析

全局人脸超分辨率方法（如Wang等人提出的特征转换全局脸方法^[54]）可以很好地保证重构人脸图像的全局结构信息。然而，该方法的合成精度有限，且在重构的高分辨率人脸图像的嘴巴、眼睛和人脸轮廓等部位有很明显的人工效应。因此，这类方法往往还需要一个额外的残差补偿策略来增加重构高分辨率人脸图像的特征细节。由于基于局部图像块的方法将整个人脸图像进行分块合成，与基于全局脸的方法相比，增强了样本空间的表达能力。

此外，局部性这一约束对于图像块表示而言也是至关重要的。通过局部性约束的引入，我们之前提出的方法，例如LcR方法^[137,138]、本章提出的LR-LiSR方法和HR-LiSR方法的性能均好于其它对比方法（基于稀疏表示局部图像块方法或者基于联合表示局部图像块方法）。

4.3.6.2 高分辨率流形结构保持和低分辨率流形结构保持对比分析

由于高分辨率图像与降质之后的低分辨率图像之间的对应关系是“多对一”的映射关系，低分辨率样本图像块空间的近邻关系不再反应真实的图像块之间的近邻关系。因此，高低分辨率图像块流形空间具有一致局部流形结构的假设在实际情况下并不成立。传统基于流形保持的超分辨率方法均是通过把低分辨率流形空间的几何关系（样本间的重构关系）保持到高分辨率流形空间，来推测目标高分辨率人脸图像。显然，基于以上分析，由于高低分辨率流形空间局部几何结构的不一致性，这种直接邻域嵌入的方式在实际情况下是不合理的。合理的做法应该是将原始高分辨率流形空间的几何结构特性保持到重构后的高分辨率图像空间中。然而，在实际情况下我们又无法提前知道与输入低分辨率图像对应的目标高分辨率图像，进而无法知道原始的高分辨率流形空间的几何结构信息。为此，我们设计了一种迭代保持的策略，通过逐次优化估计得到的高分辨率图像，并以每次迭代过程中更新得到的高分辨率图像来估计高分辨率流形空间的几何结构特性，实现超分辨率重构过程中保持原始高分辨率流形空间几何结构的目的。

4.3.6.3 一揽子方案和迭代方法对比分析

传统人脸超分辨率方法（包括我们之前提出的LcR方法^[137, 138]）都试图通过一揽子方案来获得目标高分辨率人脸图像。其核心思路均是将低分辨率图像块流形空间的几何结构特性保持到重构后的高分辨率图像块流形空间中。因而，它们其实只是考虑了一个流形空间（即低分辨率图像块流形空间），而忽略了不受降质过程影响、流形结构更加可靠的原始高分辨率图像块流形空间。从我们以上的实验可以看到，迭代优化过程对高低分辨率图像块流形空间的保持和进一步增强高分辨率重构结果是非常重要的。换言之，估计得到的高分辨率图像块可为高低分辨率图像块流形空间的邻域保持提供基本的保证；反过来，高低分辨率流形空间一致性邻域保持学习将进一步优化重构高分辨率图像。

4.4 总结与未来工作展望

在本章中，我们提出一种基于局部诱导支撑回归的单张正面人脸超分辨率方法。给定一张待重构的低分辨率人脸图像，我们首先得到与之对应的支撑集和在此支撑集上的映射函数。然后，目标高分辨率图像可以通过支撑集上的映射函数预测得到。为了得到支撑集，我们设计了两种生成策略：低分辨率图像块诱导的支撑集生成方法和高分辨率图像块诱导的支撑集生成方法。对应的，我们提出两种人脸超分辨率方法：基于低分辨率图像块诱导的支撑集回归人脸超分辨率方法（LR-LiSR）和基于高分辨率图像块诱导的支撑集回归人脸超分辨率方法（HR-LiSR）。由于我们无法提前获得原始高分辨率图像，因此也就无法得到原始高分辨率流形结构信息。为此，我们设计了一种迭代保持策略，迭代地学习低分辨率图像空间到高分辨

率图像空间的映射和生成支撑集。提出的基于局部诱导支撑回归模型应用于单张正面人脸图像的超分辨率重构，并且获得了优于前沿方法的性能。

本章所有实验都是针对正面人脸图像。然而，在实际应用中，拍摄得到的低分辨率人脸图像往往包括各种姿势的变化。为了对非正面的人脸图像进行超分辨率重构，最直接的方式是学习与输入低分辨率人脸图像姿势相同的非正面低分辨率人脸训练样本与正面高分辨率人脸训练样本之间的对应关系，然后学习它们之间的对应关系，最后利用学习得到的关系预测目标高分辨率图像。但是，由于实际获得的低分辨率人脸图像可能是任意的，我们不可能根据输入低分辨率图像提前准备任意姿势的低分辨率人脸训练样本。在这种情况下，通常做法是把输入低分辨率人脸图像归到样本库中存在的某一种姿势，然后利用该姿势的低分辨率训练样本和对应的高分辨率训练样本来学习、重构建目标正面高分辨率人脸图像。此外，在实际监控环境中，我们通过可以获得多张低分辨率图像（一段序列），因此，我们不光可以从高低分辨率样本对空间学习重构的约束条件，还可以从视频序列帧间冗余信息中获得超分辨率重构所需要的额外特征细节信息。结合基于学习方法和基于多帧重建方法进行人脸超分辨率重构将是我们下一步的工作重点。

第五章 基于局部迭代嵌入和字典渐进更新的人脸超分辨率算法

5.1 引言

图像降质模型是模拟真实高分辨率场景如何经过一系列复杂的物理过程而退化到低分辨率、低质量的降质图像。如果将图像的降质过程模型看作一个降质系统，并假设输入原始高分辨率图像为 H ，经图像降质作用后输出的低分辨率、低质量的降质图像为 L ，降质过程只受随机加性噪声 n 的影响。那么，降质过程的模型示意如图5-1所示，用公式表示为^[144]：

$$L = DBMH + n, \quad (5-1)$$

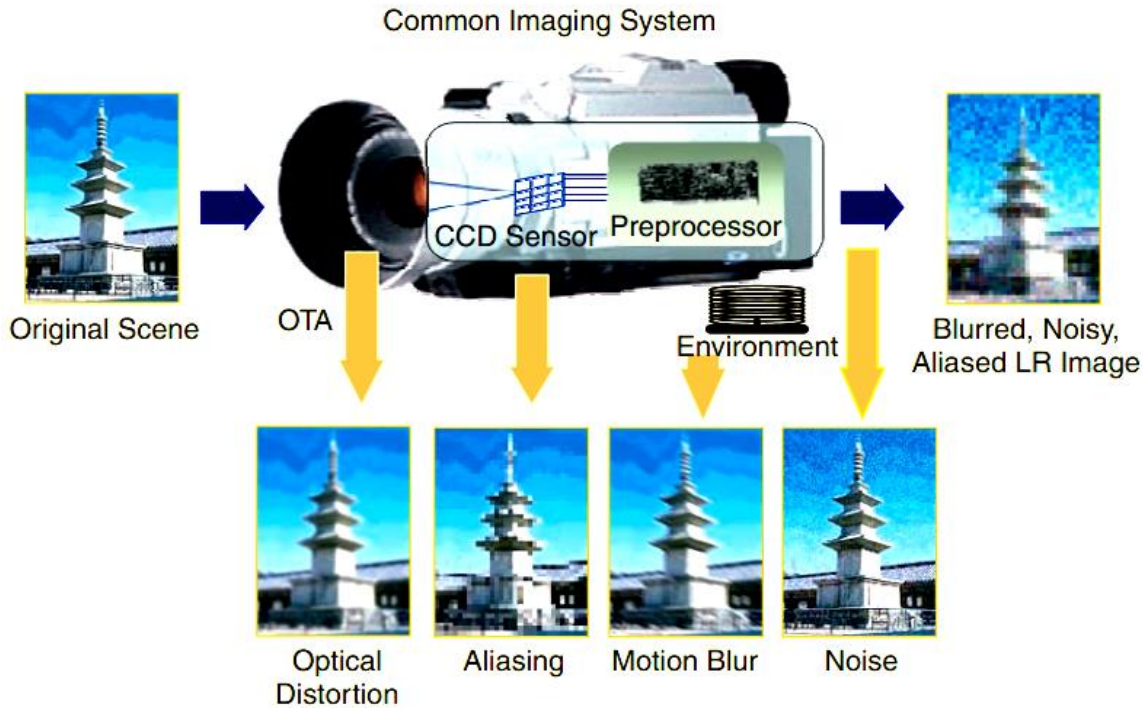


图 5-1 图像降质过程示意^[145]。

上式中 M 为 L 相对于 H 的运动矩阵， B 是模糊矩阵， D 是下采样抽值矩阵， n 为噪声项。图像超分辨率的目标是：给定一幅降质之后的低分辨率图像或图像序列，恢复高分辨率图像。图像超分辨率方法主要分为基于多帧重建的超分辨率方法和基于单帧学习的超分辨率方法。基于多帧重建的超分辨率方法主要包括三个步骤：校准、重构和复原。由于帧间的校准非常困难、帧间的互补信息难以保证，这类方法重建的结果往往不是非常理想，且超分辨率放大倍数有限^[27]。基于单帧学习的超分辨率方法可以看作是基于多帧重建超分辨率方法的一种特殊情况，

与基于多帧重建的超分辨率方法通过挖掘帧间的互补信息不同，这类基于单帧学习的方法的输入如果只有一幅图像，那么另外可用的额外信息只有图像本身的统计特性（平滑性、自相似性和稀疏性等等）和关于图像的一些先验知识。

近年来，根据给定一组高低分辨率训练样本对，学习它们之间显示映射关系或者隐式编码关系，并利用此关系来预测输入低分辨率图像对应高分辨率图像的图像超分辨率方法引起了图像处理领域专家的广泛关注^[38]。受流形学习思想的启发^[84, 141]，基于高低分辨率流形空间具有一致局部几何结构的假设，学者们提出了一系列的图像超分辨率方法^[1, 2, 8, 43, 54, 71, 83, 94, 120]。例如，Chang等人^[83]提出一种基于局部线性嵌入流形学习的图像超分辨率方法，他们首次将流形学习思想引入到图像超分辨率重构中。Ma等人^[1, 43]提出一种基于位置块的人脸超分辨率方法，对给定低分辨率人脸图像某个位置上的图像块，利用训练集人脸图像中同一位置上的所有图像块对其进行最小二乘线性合成。该方法避免了流形学习或者特征提取等步骤，提高了重构的效率，同时也提升了合成人脸图像的质量。2011年Jung等人^[2]提出一种基于凸优化的位置图像块人脸超分辨率方法，将稀疏约束加入到图像块求解表示中，利用稀疏正则化方法获得人脸超分辨率的最佳重建权重。2012年，Jiang等人^[137, 138]进一步改善了基于位置块的人脸超分辨率方法，在对图像块进行重建过程中，用流形的局部几何约束替换了文献^[2]中的稀疏约束，使重建结果同时具有稀疏性和局部性。

无论是使用最小二乘表示方法^[1, 43]、稀疏表示方法^[2]还是局部约束表示方法^[137, 138]，它们均试图将低分辨率的流形结构保持到重构后的高分辨率空间中。然而，由于高低分辨率图像之间是“多对一”的对应关系，高低分辨率图像块之间的流形不具有 consistency（特别是在降质过程非常复杂的情况下更是如此^[95]），最终，重构后的高分辨率图像流形与原始高分辨率流形空间不一致。但是，无论降质过程如何复杂，原始高分辨率图像块的流形结构是不受降质过程影响的，而现有方法均没有充分挖掘不受降质过程影响的原始高分辨率流形空间的几何结构信息，这也使得它们重建出来的结果缺乏可靠性。

为了解决上述问题，本章提出一种基于多层次局部约束迭代嵌入和字典渐近更新的“由粗到精”人脸超分辨率算法（图5-2为提出方法的基本框架），该算法同时考虑了高分辨率人脸图像空间局部几何结构保持和高低分辨率图像空间之间近邻嵌入学习两大过程。我们提出方法的“由粗到精”主要体现在以下两个方面：第一，重构高分辨率人脸图像不断迭代优化。对于每一层中间低分辨率人脸训练集和原始高分辨率人脸训练集，利用预估得到的高分辨率人脸图像与原始高分辨率训练样本之间的距离作为约束，约束下一次超分辨率中输入图像块表示的重构权重，不断优化重构结果；第二，低分辨率人脸训练集的更新学习。对于输入的原始低分辨率人脸训练集和原始高分辨率人脸训练集，学习得到多层中间低分辨率人脸训练集，逐渐减小低分辨率训练空间与原始高分辨率训练空间的维度差，从而能够更加准确地进行邻域嵌入学习与预测。

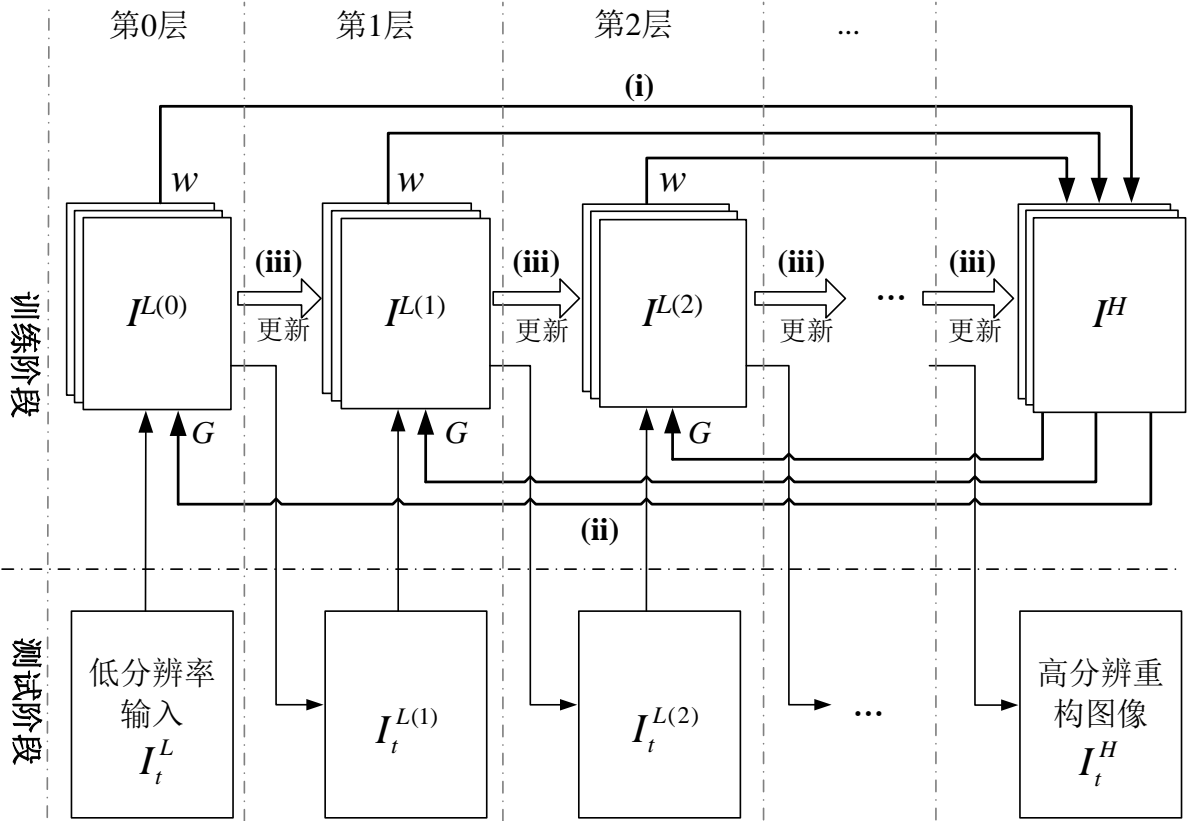


图 5-2 基于字典渐近更新的“由粗到精”人脸超分辨率算法的流程图。

与以往的人脸超分辨率方法相比，本章提出方法的贡献可以概括为以下几点：

(1) 利用高低分辨率图像流形局部几何结构一致性假设，许多传统的人脸超分辨率方法试图通过保持低分辨率图像块空间的局部几何结构，将此几何结构保持到重构后的高分辨率图像块空间，最终进行基于邻域嵌入的人脸超分辨率重构。但是，它们只利用了低分辨率图像块流形空间的几何结构，而忽略了不受图像退化过程影响的原始高分辨率图像块流形空间的几何结构。为了充分利用原始高分辨率图像块流形空间的几何结构信息，与传统方法不同，本章所提出的方法可以重复利用原始高分辨率图像块空间的几何结构，并用其来约束输入低分辨率图像块在低分辨率图像块空间的重构权重，不断优化该权重，最终得到最优的重构权重和超分辨率重构结果；

(2) 传统基于学习的人脸超分辨率方法均是通过学习训练样本中低分辨率图像（低分辨率训练集是固定的）与高分辨率图像之间显示的映射关系或者隐式的编码关系，并利用该学习得到的关系来预测合成高分辨率人脸图像。然而，由于高分辨率图像与低分辨率图像之间是“多对一”的对应关系，低分辨率图像流形空间与原始高分辨率图像流形空间在局部流形结构上存在严重的不一致性。为此，我们提出对低分辨率人脸图像训练集进行更新，生成多个中间低分辨率字典库，使得生成之后的低分辨率训练集与原始高分辨率训练集在流形结构上的一致性得到提升，人脸超分辨率重构可以在一个更加一致流形空间进行。

本章接下来的安排如下：第5.2节介绍了基于图像块表示人脸超分辨率的相关工作及其存在的问题。在这一小节，我们首先介绍了本章节所涉及的一些形式化符号表示，然后给出了基于图像块表示人脸超分辨率方法的基本框架，最后总结了这类方法存在的两个问题。第5.3节提出“由粗到精”人脸超分辨率方法。首先给出了基于局部约束迭代邻域嵌入的人脸超分辨率方法，然后在些基础上详细地描述了本章提出的基于多层局部约束迭代邻域嵌入的人脸超分辨率方法，最后，简要分析了提出人脸超分辨率方法的复杂度。第5.4节测试了提出算法和前沿算法的性能对比。具体的，我们首先介绍了本章所用到的中国人脸图像库，然后分析了提出方法所涉及到的几个关键参数对算法性能的影响，这些参数包括：局部约束参数、迭代次数和中间低分辨率字典层数。接着，我们给出了提出算法和前沿对比算法的主观对比结果、客观对比结果（PSNR和SSIM指标）和复杂度对比。为了验证提出算法对实际人脸图像进行超分辨率重构的有效性，我们也给出了一组在实际环境下拍摄得到的低分辨率人脸图像的重构结果。最后，讨论了实际环境下的人脸超分辨率还存在的问题以及它与人脸识别之间的关系，以及在下一阶段可能的改进方向。第5.5节对本章进行了总结。

5.2 相关工作

5.2.1 符号定义及问题提出

在本节中，我们给出了提出方法的具体细节。我们首先介绍本章用到的一些术语及对应的数学符号。给定一张输入的测试低分辨率人脸图像 I_t^L （下标“ t ”是为了区分测试图像和训练图像），人脸超分辨率算法的目的就是通过学习低分辨率人脸图像训练集 $I^L = \{I_1^L, I_2^L, \dots, I_N^L\}$ 与高分辨率人脸图像训练集 $I^H = \{I_1^H, I_2^H, \dots, I_N^H\}$ 之间的关系，预测输入低分辨率测试人脸图像对应的高分辨率人脸图像 I_t^H 。其中， N 是训练集中样本的个数。对于局部图像块人脸超分辨率方法，我们根据预先定义的图像块大小和相邻图像块之间的交叠像素，将输入的低分辨率人脸图像划分成 M 个图像块 $\{x_i(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}$ ， $x_i(p, q)$ 表示输入低分辨率人脸图像位置 (p, q) 上的一个小图像块， U 代表一张人脸图像上每一列的图像块数， V 代表一张人脸图像上每一行的图像块数，那么有 $M=UV$ 。类似地，我们将所有 N 张高分辨率的训练人脸图像和 N 张低分辨率的训练人脸图像也按照同样的划分方式划分成 M 个图像小块， $\{y_i(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}_{i=1}^N$ 和 $\{x_i(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}_{i=1}^N$ 。 $x_i(p, q)$ 表示第 i 张低分辨

率人脸训练样本在位置 (p, q) 上的一个小块, $y_i(p, q)$ 则表示第 i 张高分辨率人脸训练样本在位置 (p, q) 上的一个小块。

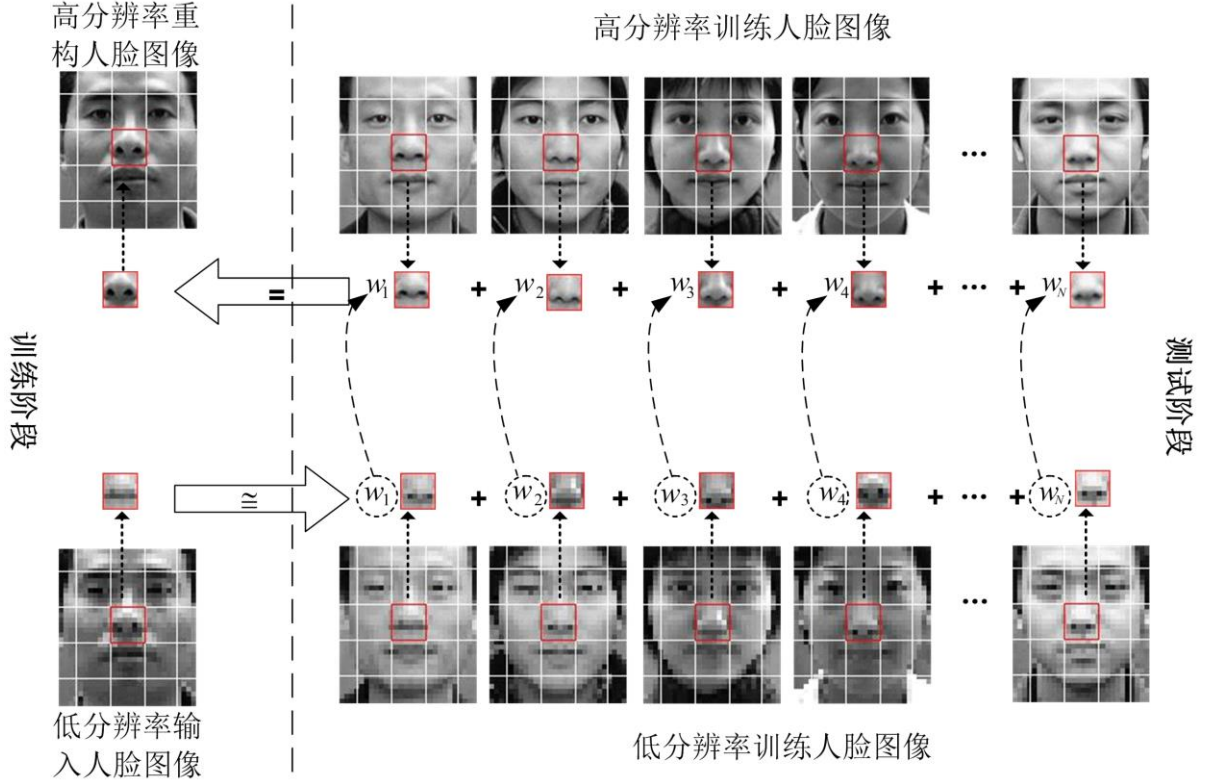


图5-3 基于位置块的人脸超分辨率算法框架。

5.2.2 图像块表示方法

对于输入的低分辨率人脸图像位置 (p, q) 上的一个低分辨率小块 $x_i(p, q)$ ，如图5-3所示的一个鼻子位置上的一个图像块，基于图像块的方法试图在低分辨率样本图像空间获得它的最优重构权重：

$$J(w(p, q)) = \|x_i(p, q) - \sum_{i=1}^N w_i(p, q)x_i(p, q)\|_2^2 + \tau \Omega(w(p, q)), \quad (5-2)$$

式(5-2)的第一项表示数据重构误差项，第二项表示重构权重的约束，它可以使重构权重具有某种性质。式(5-2)中的正则化参数 τ 用来平衡观察数据的重构误差和重构权重的先验知识。一旦获得了最优重构权重 $w^*(p, q) = \arg \min_{w(p, q)} J(w(p, q))$ ，目标高分辨率图像块 $y_i(p, q)$ 可

可以通过高分辨率人脸图像训练库中的所有相同位置的图像块合成得到：

$$y_t(p, q) = \sum_{i=1}^N w_i^*(p, q) y_i(p, q). \quad (5-3)$$

当我们得到所有超分辨率重构的高分辨率图像块 $\{y_t(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}$ ，最终的目标高分辨率人脸图像 I_t^H 可以根据图像块在人脸上的位置融合得到（直接对相邻图像块交叠部分的像素求平均值）。从上述人脸超分辨率重构过程，我们可以发现，这类基于局部块的方法的关键是如何获得最优的重构权重 w^* （如图5-3所示）。在接下来的章节，为了描述方便起见，在不引起混淆的情况下，我们在进行形式化陈述时，去掉了位置标识项 (p, q) 。

最有代表性的局部图像块超分辨率方法由Chang等人提出，他们假设高分辨率图像块流形空间和低分辨率图像块流形空间具有相似的局部几何结构。依据局部线性嵌入的思想，他们利用高分辨率图像块流形空间中的 K 个近邻样本点来线性估计的高分辨率图像块，线性重构权重则是通过最小化输入低分辨率图像块重构误差得到，该方法不需要大量的样品就可以获得比较好的效果。在文献^[68]中，Zhang和Cham进一步提出了DCT空间邻域嵌入方法来推测人脸图像的局部特征。Hu等人^[66]提出了一种基于局部结构学习的人脸超分辨率的方法，该方法假定两张相似的人脸图像具有类似的局部像素结构，每个像素可以由它相邻的像素点线性组合得到。Li等人^[59, 97]指出，高分辨率图像样本空间和低分辨率图像样本空间具有一个隐式公共的流形空间，在此假设的基础上，他们提出一种基于流形对齐的人脸超分辨率方法。

人脸图像分析与合成的相关研究成果表明：在人脸分析与合成的过程中，人脸是一类高度结构化的物体，人脸图像块所在的位置信息是至关重要的。受此思想的启发，Ma等人^[1, 43]提出一种基于位置块的人脸超分辨率方法。该方法利用训练库中同样位置上的所有样本图像块对输入低分辨率图像块进行联合表示。然而，当训练集的样本数比输入低分辨率图像块的维数还要大时，重构权重系数不唯一。为了克服这一问题，稀疏约束被引入到块重构的目标函数中，也即是式（5-2）中的先验约束变成 $\Omega(w) = \|w\|_1$ 。稀疏约束的引入不仅可以保证原欠定方程有唯一的解，而且减小了重构误差。然而，该方法也有一个缺陷：它过分强调重构权重的稀疏性，而忽略了对揭示数据空间非线性流形非常有用的局部性先验。因而，基于稀疏表示的方法所获得的图像块表示远非最优。为了解决上述问题，我们在文献^[137, 138]中进一步提出一种基于改进的基于位置块人脸超分辨率方法。通过引入流形局部约束（摒弃非最优的稀疏约束），该方法的目标函数是：

$$J(w) = \|x_t - \sum_{i=1}^N w_i x_i\|_2^2 + \tau \|dist \square w\|_2^2, \quad (5-4)$$

式（5-4）中，“ \square ”表示两个向量的点积运算， $dist$ 是一个 N 维的局部自适应子，它可以根据输入图像块 x_t 与低分辨率训练图像块 x_i 之间的距离，计算得到每个低分辨率训练图像块的自

由度（惩罚）。因此， $dist$ 可以简单定义为输入低分辨率图像块与低分辨率训练样本图像块之间的欧式距离：

$$\|x_t - x_i\|_2, 1 \leq i \leq N. \quad (5-5)$$

通过对低分辨率训练样本图像块设定不同的自由度，也即是说，离输入低分辨率图像块距离近的样本图像块更容易被选中用来作输入低分辨率图像块的表达，反之，那些离输入低分辨率图像块距离远的样本图像块则被惩罚（其对应的表示权重很小）。通过选取少数一些近邻图像块，上述表示方法可以同时达到稀疏性和局部性。

说明1：改进的图像块表示方法无论是联合表示、稀疏表示还是局部表示，他们均是基于高低分辨率图像块流形具有相似的局部几何结构这一前提假设。虽然局部保持假设看起来是正确的，但是该假设在实际情况下并不成立。这主要是因为高低分辨率图像之间存在“多对一”的映射关系（由于图像降质过程有信息损失，观察到的一张低分辨率图像可能会有多张高分辨率场景与之对应）。如图5-4所示，随着下采样倍数的增大或者加入泊松噪声，高低分辨率图像块流形之间的局部保持率急剧下降。例如，当 $K=20$ 时，局部保持率从下采样两倍的76%急剧下降到下采样8倍加泊松噪声的25%。

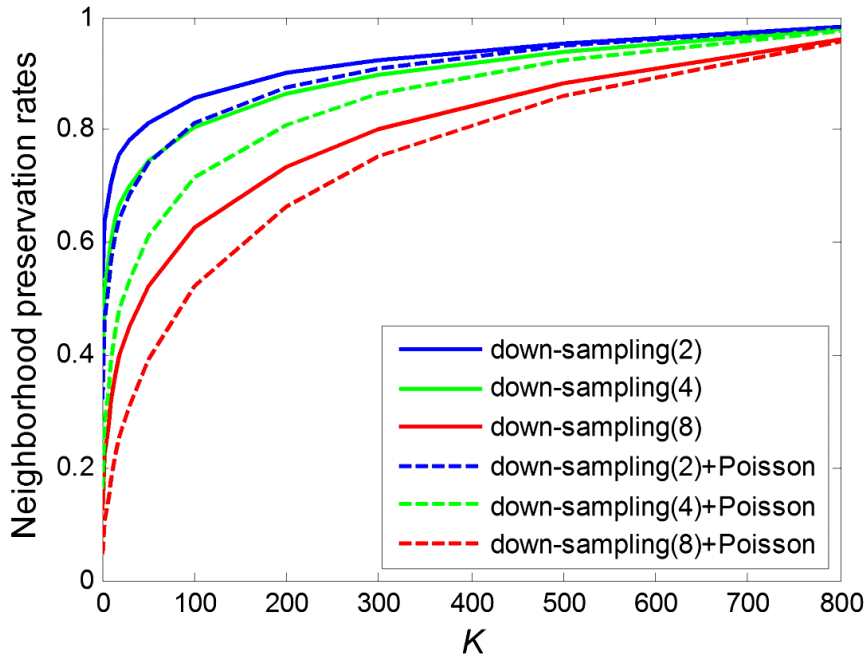


图5-4 在不同近邻数 K 下，退化后的低分辨率图像块流形与原始高分辨率图像块流形之间的局部保持率。每个图例括号里面的数值表示下采样的倍数，“down-sampling(2)+Poisson”表示下采样两倍且加入泊松噪声。

说明2：以上方法的重建权重获得和目标高分辨率图像块的预测是分别进行的，它们本质上只考虑了低分辨率图像块流形，完全没有利用和挖掘高分辨率图像块的几何结构信息。为了证明高分辨率图像块流形几何结构的有用性，我们设计了不同流形保持方法对人脸超分辨

率结果的影响。一种方法是低分辨率流形结构保持方法，另一种方法是原始高分辨率流形结构保持方法（目标高分辨率图像是预先知道的，因此可以看作理想情况）。从图5-5的对比结果可以，流形保持策略对基于流形假设的局部块人脸超分辨率方法是至关重要的，传统基于低分辨率图像块流形保持方法还有很大的提升空间。在图5-4中，当 K 等于9是，局部保持率急剧下降，这可以解释为最小二乘解对观察图像块的“过拟合”^[89]。

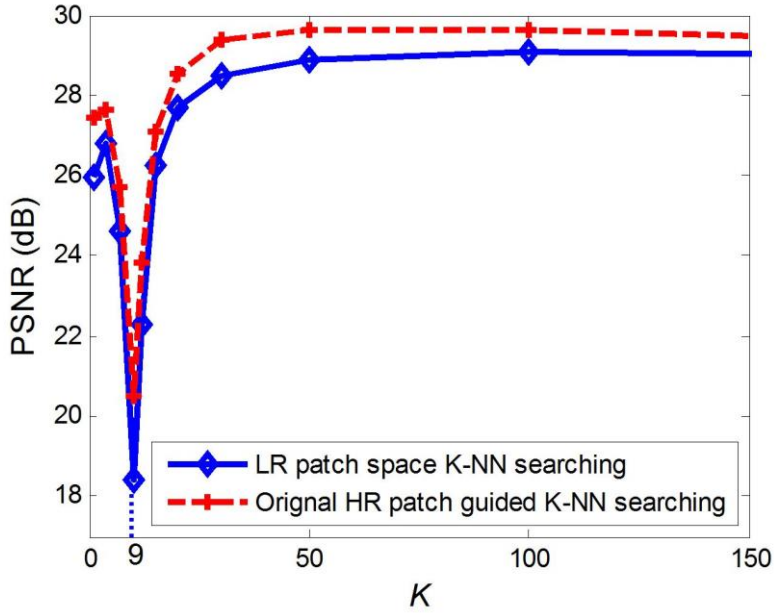


图5-5 在不同近邻数 K 下，两种不同流形保持方法的人脸超分辨率性能对比。蓝色实线：传统基于 K 邻域嵌入的流形保持学习方法；红色虚线：原始高分辨率图像块流形保持学习方法（理想情况）。

5.3 提出的“由粗到精”人脸超分辨率方法

鉴于上面的讨论，我们试图解决以下两个问题：（1）在表示输入低分辨率图像块时，如何保持高分辨率人脸图像块流形的几何结构；（2）如何减少高低分辨率图像块流形之间的一致性。

为了解决第一个问题，我们提出一种迭代的邻域嵌入策略，迭代更新输入低分辨率图像块的表示和估计得到的高分辨率图像块（5.3.1节）。具体来说，我们在高分辨率图像块空间计算一个局部约束子，估计得到的高分辨率图像块与训练集中每个高分辨率样本图像块的距离。输入低分辨率图像块的重建权重和新估计得到的高分辨率图像块不断地被迭代更新，在此“由粗到精”的更新方式下，我们可以期待一个很好的高分辨率图像块估计结果。

为了解决第二个问题，我们提出一种低分辨率训练集更新的方法来逐渐地减小高低分辨率流形空间的不一致性（5.3.2节）。在图像空间中，我们假设低分辨率图像和对应的高分辨率

图像之间存在一个虚拟的路径。人脸图像逐渐地从低分辨率图像变化到高分辨率图像的过程中将在图像空间形成一条平滑的转移路径。如图5-6所示，沿着该路径建立一些中间字典可以更好地揭示低分辨率图像潜在的图像降质过程。因此，我们可以期待根据中间字典逐渐地优化估计得到的高分辨率图像。

5.3.1 局部约束迭代邻域嵌入方法

局部约束表示方法通过在低分辨率图像空间定义一个局部约束子，然后利用一个一揽子的方案获得目标高分辨率图像。与该方法不同，我们提出的局部约束迭代嵌入方法首先在高分辨率空间计算局部约束子，然后对输入低分辨率图像块的重构权重和目标高分辨率图像块进行迭代求解。具体地，重构权重和高分辨率图像块所对应的双变量目标函数构造如下：

$$J(w, y_t) = \|x_t - \sum_{k \in C_K(y_t)} w_k x_k\|_2^2 + \tau \|dist|_K \square w\|_2^2, \quad (5-6)$$

其中， $C_K(y_t)$ 表示 y_t 在高分辨率训练图像集中的 K 个近邻的索引号所组成的集合，因此

$$C_K(y_t) = \text{support}(dist|_K), \quad (5-7)$$

其中， $dist|_K$ 表示 $dist$ 中 K 个最小的值所构成的集合， $dist \in \mathbb{R}^N$ 是估计高分辨率图像块与同样位置高分辨率训练集图像块的欧式距离：

$$dist = \|y_t - y_i\|_2, \quad 1 \leq i \leq N. \quad (5-8)$$

为了求解式 (5-6)，对于第 s 个循环，我们首先将 $y_{t(s)}$ 预设为其的Bicubic插值版本，然后通过最小化式 (5-6) 来迭代优化重构权重 $w_{(s)}$ 与估计高分辨率图像块 $y_{t(s)}$ 。给定 $y_{t(s)}$ ，求解 $w_{(s)}$ 可以看作求解一个带约束的最小二乘问题。而它的求解则可以由文献^[135]提出的方法得到一个解析解。

一旦得到最优低分辨率图像块表示 w^* ，我们就可以利用该表示系数构建对应的高分辨率图像块：

$$y_t = \sum_{k \in C_K(y_t)} w_k^* x_k. \quad (5-9)$$

根据图像块在人脸上的位置，并按照光栅扫描顺序，从左到右从上到下，对输入低分辨率人脸图像上的所有图像块分别进行处理。为了消除由于分块分别处理可能带来的块效应问题，我们相邻图像块对施加兼容性约束，即对于相邻的图像块的交叠部分，我们对其进行加权平均。提出的局部约束迭代邻域嵌入人脸超分辨率方法的整个流程如算法5-1所示 ($patch_size$ 和 $overlap$ 分别表示图像块大小和相邻图像块之间的交叠像素个数)。

算法5-1 基于局部约束迭代邻域嵌入的人脸超分辨率方法

1. **输入:** 高分辨率人脸图像训练集 $I^H = \{I_1^H, I_2^H, \dots, I_N^H\}$, 低分辨率人脸图像训练集 $I^L = \{I_1^L, I_2^L, \dots, I_N^L\}$, 输入低分辨率人脸图像 I_t^L , 图像块大小 $patch_size$, 近邻图像块之间的交叠像素 $overlap$, 近邻个数 K , 正则化参数 τ , 最大迭代次数 $maxIter$;
2. **输出:** 重构后的高分辨率人脸图像 I_t^H ;
3. 计算每行(列)的图像块数目:

$$U = \text{ceil}((imrow - overlap)/(patch_size - overlap));$$

$$V = \text{ceil}((imcol - overlap)/(patch_size - overlap));$$
4. 将高分辨率人脸图像训练集、低分辨率人脸图像训练集和输入低分辨率人脸图像按照同样的划分策略分别分成 M 个图像块, $\{x_i(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}_{i=1}^N$, $\{y_i(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}_{i=1}^N$ 和 $\{x_t(p, q) | 1 \leq p \leq U, 1 \leq q \leq V\}$;
5. 初始化: $s=0$, $I_{t(0)}^H = \text{Bicubic}(I_t^L)$;
6. **For** $p=1$ to U **do**
7. **For** $q=1$ to V **do**
8. **Repeat**
9. 计算 (p, q) 位置上上一步估计得到的高分辨率图像块 $y_{t(s)}(p, q)$ 和训练集中所有 M 个高分辨率样本图像块 $\{y_i(p, q)\}_{i=1}^N$ 之间的欧式距离:

$$dist = \|y_t - y_i\|_2, \quad 1 \leq i \leq N;$$
10. 求解优化目标函数(5-6)得到 $y_{t(s)}(p, q)$ 的最优表示权重 $w_{(s)}^*(p, q)$.
11. 通过下式估计高分辨率图像块 y_t :

$$y_{t(s)}(p, q) = \sum_{k \in C_K(y_{t(s)}(p, q))} w_{k(s)}^*(p, q) y_k(p, q);$$
12. $s=s+1$;
13. **Until** $s > maxIter$
14. **End for**
15. **End for**
16. 将所有加权重构出的高分辨率图像块按照其在人脸图像上的位置进行叠加, 然后除以每个像素位置被交叠的次数, 最终重构出高分辨率人脸图像 I_t^H .

5.3.2 多层局部约束迭代邻域嵌入方法

在局部约束迭代邻域嵌入方法中，通过引入高分辨率图像块流形的几何结构和邻域嵌入迭代策略，从低分辨率流形中获得最优的重构权重。如式（5-6）所示，最优重构权重是通过最小化输入低分辨率图像块 x_i 在低分辨率训练集上的重构误差所得到的。而在实际情况下，低分辨率训练集的几何结构与高分辨率训练集的几何结构是不一致的。因此，基于局部约束迭代邻域嵌入方法的人脸超分辨率方法的性能不可避免的受到降质低分辨率图像流形结构的影响。

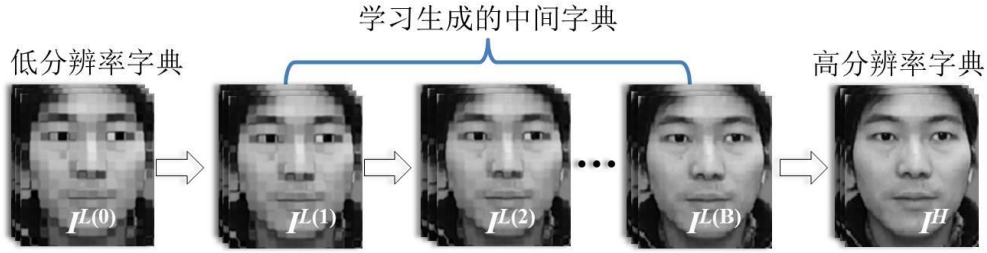


图5-6 给定原始低分辨率人脸图像训练集 $I^{L(0)}$ 和高分辨率人脸图像训练集 I^H ，提出的方法可以学习得到一系列中间低分辨率字典 $I^{L(0)}$ ， $I^{L(2)}$ ， \dots ， $I^{L(B)}$ 来近似图像的成像过程，从而可以使邻域嵌入学习可以在一个更加一致的高低分辨率流形空间进行。

为了解决这一问题，我们对基于局部约束迭代邻域嵌入的人脸超分辨率方法进行扩展，提出一种多层局部约束迭代邻域嵌入方法。在此方法中，降质之后的低分辨率训练样本会被不断的更新，得到一系列的中间低分辨率字典库。因此，我们提出多层局部约束迭代邻域嵌入方法的目标函数可以写成如下形式：

$$J(w^b, y_i^b, X^b) = \|x_i^b - \sum_{k \in C_K(y_i^b)} w_k^b x_k^b\|_2^2 + \tau \|dist^b|_K \square w^b\|_2^2, \quad b = 0, 1, \dots, B. \quad (5-10)$$

其中， b 表示中间低分辨率字典层的序号， B 表示中间低分辨率字典总的层数（如图5-6所示）。当 $b=0$ 时， x_i^b 表示输入低分辨率图像块 $x_i^0 = x_i$ ， $X^0 = [x_1, x_2, \dots, x_N]$ 是原始低分辨率训练图像块集合。随着低分辨率训练集的更新，人脸图像的超分辨率重构过程可以在一个更加一致的高低分辨率流形空间进行，因而可以获得相较于传统近邻保持方法更好的结果。

式（5-10）是一个多变量的优化问题，通常情况下很难获得最优解。庆幸的是，一旦我们将低分辨率训练集 X^b 固定， w^b 和 y_i^b 的优化就退化成了优化之前的式（5-6）。因此，求解式（5-10）所剩下的问题就是如何学习得到中间层的低分辨率字典。

为了获得中间低分辨率字典，对于每一层，利用5.3.1中提出的局部约束迭代邻域嵌入人脸超分辨率方法更新低分辨率字典里的所有低分辨率人脸图像，得到一层新的中间低分辨率字典。具体地，对于第 b 层低分辨率人脸图像训练集 $I^{L(b)}$ ，采用“去一法”对其中每张低分辨率人脸样本图像分别进行超分辨率重构，得到第 $b+1$ 层低分辨率人脸图像训练集 $I^{L(b+1)}$ ，将其作

为第 $b+1$ 层低分辨率中间字典； $b=0$ 时，第0层低分辨率人脸图像训练集采用预先构建的低分辨率人脸图像训练集 I^L ，具体包括以下步骤：

第一步，取第 b 层低分辨率人脸图像训练集 $I^{L(b)} = \{I_1^{L(b)}, I_2^{L(b)}, \dots, I_N^{L(b)}\}$ 中的第 i 张低分辨率人脸样本图像 $I_i^{L(b)}$ 作为输入低分辨率人脸图像，第 b 层低分辨率人脸图像训练集中剩余的所有低分辨率人脸样本图像作为新的低分辨率人脸图像训练集， $\hat{I}^{L(b)} = \{I_j^{L(b)} \mid j=1, 2, \dots, i-1, i+1, \dots, N\}$ ，高分辨率人脸图像训练集 $I^H = \{I_1^H, I_2^H, \dots, I_N^H\}$ 中相应剩余的所有高分辨率人脸样本图像作为新的高分辨率人脸图像训练集 $\hat{I}^H = \{I_j^H \mid j=1, 2, \dots, i-1, i+1, \dots, N\}$ ；

第二步，以第一步中所得新的第 b 层低分辨率人脸图像训练集 $\hat{I}^{L(b)}$ 、新的第 b 层高分辨率人脸图像训练集 \hat{I}^H 作为输入，对第一步所取的输入低分辨率人脸图像 $I_i^{L(b)}$ 进行超分辨率重构，估计得到输入低分辨率人脸图像 $I_i^{L(b)}$ 对应的高分辨率人脸图像 y_i^b ，将 y_i^b 作为第 $b+1$ 层低分辨率人脸图像训练集 $I^{L(b+1)}$ 中第 i 张低分辨率人脸样本图像 $I_i^{L(b+1)}$ 。

最后，判断是否 $b=B-1$ ，若否则令 $b=b+1$ ，以上述过程得到的第 $b+1$ 层低分辨率人脸图像训练集 $I^{L(b+1)}$ 作为输入，进行下一次迭代得到下一层的低分辨率中间字典；若是则得到多层低分辨率中间字典 $I^{L(0)}, I^{L(2)}, \dots, I^{L(B)}$ 。算法5-2为中间低分辨率字典学习的算法步骤。

算法 5-2 中间低分辨字典学习

1. **输入：** 高分辨率人脸图像训练集 $I^H = \{I_1^H, I_2^H, \dots, I_N^H\}$ ，低分辨率人脸图像训练集 $I^L = \{I_1^L, I_2^L, \dots, I_N^L\}$ ，图像块大小 $patch_size$ ，近邻图像块之间的交叠像素 $overlap$ ，近邻个数 K ，正则化参数 τ ，最大迭代次数 $maxIter$ ；
 2. **输出：** 中间低分辨率字典 $I^{L(1)}, I^{L(2)}, \dots, I^{L(B)}$ ；
 3. **初始化：** $b=0$ ， $I^{L(0)} = I^L$ ；
 4. 对于低分辨率训练集中的每一张低分辨率人脸图像 $I_i^{L(b)}$ ；
 5. 低分辨率人脸图像训练集中剩余所有低分辨率人脸样本图像作为新的低分辨率人脸图像训练集， $\hat{I}^{L(b)} = \{I_j^{L(b)} \mid j=1, 2, \dots, i-1, i+1, \dots, N\}$ ；
 6. 高分辨率人脸图像训练集中相应剩余所有高分辨率人脸样本图像作为新的高分
-

分辨率人脸图像训练集 $\hat{I}^H = \{I_j^H | j = 1, 2, \dots, i-1, i+1, \dots, N\}$;

7. 利用算法5-1, 并根据输入 $\hat{I}^{L(b)}$ 、 \hat{I}^H 和 $I_i^{L(b)}$ 超分辨率重构 $I_i^{L(b)}$ 所对应的高分辨率人脸图像 $I_i^{*H(b)}$;
8. 重复上述步骤4到步骤7, 直至低分辨率训练集中所有低分辨率人脸图像被超分辨率重构完毕;
9. 判断是否 $b=B-1$, 若否则令 $b=b+1$, 以上述过程得到的第 $b+1$ 层低分辨率人脸图像训练集 $I^{L(b+1)}$ 作为输入, 进行下一次迭代得到下一层的低分辨率中间字典; 若是则得到多层低分辨率中间字典 $I^{L(1)}, I^{L(2)}, \dots, I^{L(B)}$ 。

5.3.3 计算复杂度

在此小节中, 我们讨论了本章提出的多层局部约束迭代邻域嵌入方法的计算复杂度。由于中间字典可以通过离线学习得到, 而且是一劳永逸的, 本小节仅讨论在线超分辨率重建过程的计算复杂度。如算法5-1所述, 提出超分辨率方法最耗时部分主要集中在两个环节: (i) K 最近邻搜索; (ii) 重构权重计算。有四个主要的因素决定着提出算法计算复杂性: 最近邻数 K 、训练集大小 N 、一张图像里的图像块数据 M 和一个图像块的维度 p^2 。

据文献^[87], 对于 K 最近邻搜索, 计算复杂度是 $O(Kp^2N)$; 对于重构权重计算阶段, 计算复杂度是 $O(K^3p^2)$ 。因此, 每次计算式 (5-6) 时的计算复杂度是 $O(Kp^2N + K^3p^2)$ 。此外, 由于每一层的最大迭代次数是 \maxIter , 而总的层数为 B , 这些都将直接影响最终的计算复杂度。因此, 提出算法总的计算复杂度是 $O(\maxIter(Kp^2N + K^3p^2)B)$ 。在接下来的第5.4.3小节, 我们将对比提出方法和目前最好方法的计算复杂度 (CPU时间)。

5.4 实验结果

在本节中, 我们给出了实验的具体细节, 证明了提出方法对人脸超分辨率重构的有用性。本章实验主要旨在回答以下问题:

- 是否有必要迭代计算高分辨率流形几何结构和低分辨率图像块重构权重? 此迭代过程是收敛的么?
- 是否有必要更新迭代低分辨率字典来获得中间低分辨率字典? 此更新迭代过程是收敛的么?
- 与前沿人脸超分辨率方法相对, 提出方法的性能如何?

- 提出方法可以真正提高实际人脸图像的质量么？

5.4.1 人脸库描述

本章的所有实验均是在中国科学院计算技术研究所提供的 CAS-PEAL-R1 人脸库^[4]上进行的，该数据库包括 1040 个对象的 30871 张图像。在本章中，我们利用其中 1040 个对象中性表情和正常光照下的人脸图像，并且随机选择了其中 1000 张图像作为训练集，将剩下的 40 张图像作为测试图像。因此，实验中用到的所有测试图像均不在训练集中。上述所有 1040 张图像均通过两个人眼中心位置进行对齐，然后统一裁剪到 128×112 像素（如图 5-7 所示）。低分辨率图像通过一个窗口为 4×4 的平均滤波器平滑和 Bicubic 下采样算法 4 倍下采样高分辨率图像所得，因此低分辨率图像尺寸为 32×28 像素。

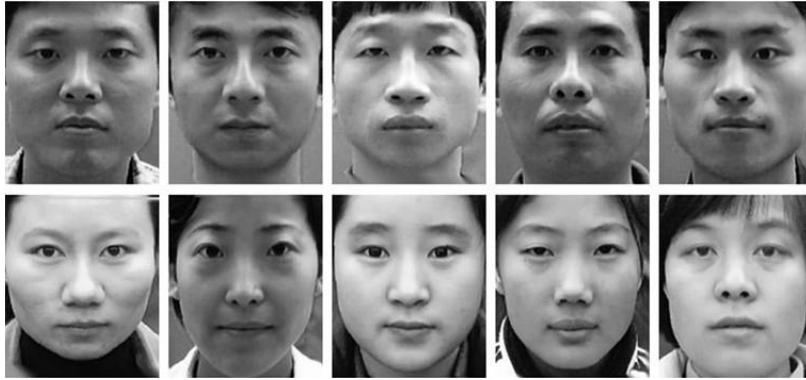


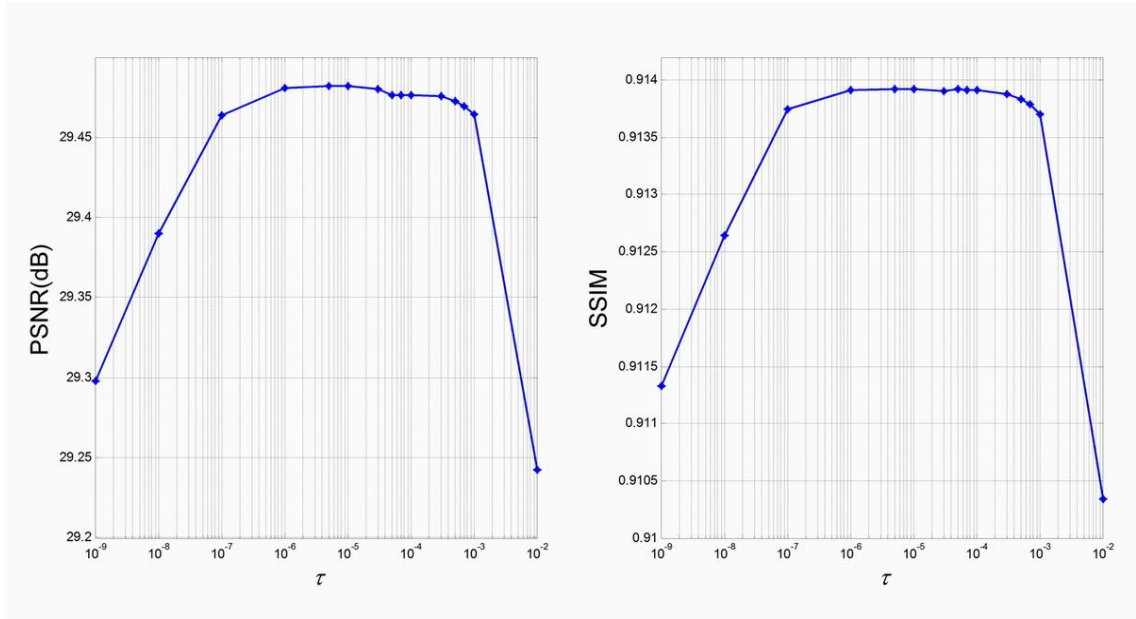
图 5-7 中国人脸图像库样本示例。

5.4.2 参数分析

在本小节中，我们探讨了局部约束参数 τ ，最大迭代次数 $maxIter$ ，中间字典的层数 B 三个参数对提出方法的影响。

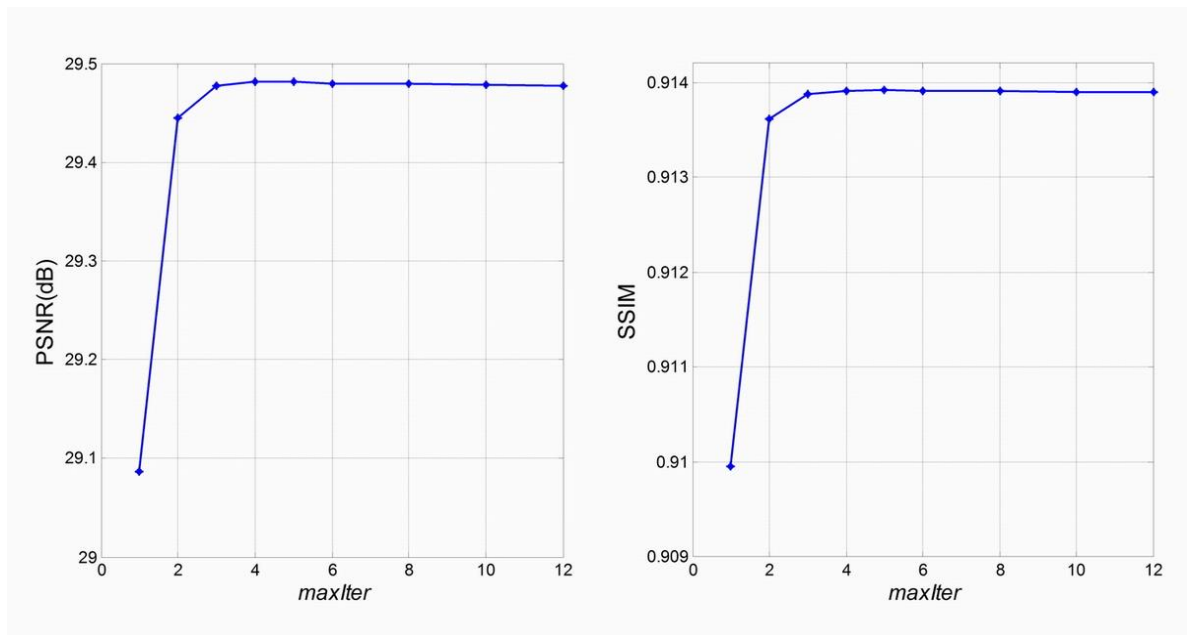
(1) 局部约束参数的影响。

为了评估局部约束参数 τ 对提出算法的影响，我们测试了提出方法在不同 τ 值下的性能。图 5-8 给出了不同 τ 值下的所有 40 张测试图像的平均 PSNR 和 SSIM 性能。从该图我们可以出，局部约束参数对提出算法的性能有很大的影响：随着 τ 值的增大，提出方法的增益变得越来越明显，这说明了局部性对图像块表示的重要性。当 τ 的取值在 $1e-7$ 和 $1e-3$ 之间时，我们提出的方法可以持续获得稳定的性能。为了平衡重构误差， τ 的取值不能太高。因此，在实验中，我们将提出算法的局部约束参数 τ 值设为 $1e-5$ 。


 图 5-8 提出方法在不同 τ 值下的 PSNR 和 SSIM 性能指标。

(2) 迭代次数的影响。

为了评估迭代次数 $maxIter$ 对提出算法的影响, 我们给出了提出方法在不同 $maxIter$ 值下的性能指标。如图 5-9 所示, 我们画出了不同 $maxIter$ 值下的所有 40 张测试图像的平均 PSNR 和 SSIM 性能。从该图我们可以得到如下几点结论: (1) 随着迭代次数 $maxIter$ 增加, 提出方法的增益变得越来越大, 这说明引入高分辨率流形结构, 进行高分辨率流形结构保持对基于邻域嵌入的图像超分辨率是非常重要的; (2) 提出方法只需要经过几次迭代便可以快速地收敛。例如, 迭代次数设成 4 即可。


 图 5-9 提出方法在不同 $maxIter$ 值下的 PSNR 和 SSIM 性能指标。

另外，我们也测试了 $C_K(y_i)$ 在每次迭代过程中的变化情况。特别地，我们定义了一个近邻相似率指标：

$$NSR = \frac{|C_K(y_i) \cap C_K(y^*)|}{|C_K(y^*)|}, \quad (5-11)$$

式 (5-11) 中的 y_i 表示根据输入低分辨率图像块估计得到的高分辨率图像块， y^* 表示输入低分辨率图像块对应的理想高分辨率图像块。 $|\cdot|$ 是一个求集合中元素个数的算子。NSR 反应了估计得到的高分辨率图像块选择近邻的准确程度，NSR 越大说明估计得到的高分辨率图像块越准确。表 5-1 给出了不同迭代次数下的 NSR 指标。可以看出，通过几次迭代，支撑集 $C_K(y_i)$ 就不会有太大的变化，逼近了理想的情况（只相差 4.93%）。

表5-1 不同迭代次数下的NSR指标。

迭代次数	1	2	3	4	5
NSR	83.33%	94.82%	95.05%	95.07%	95.07%

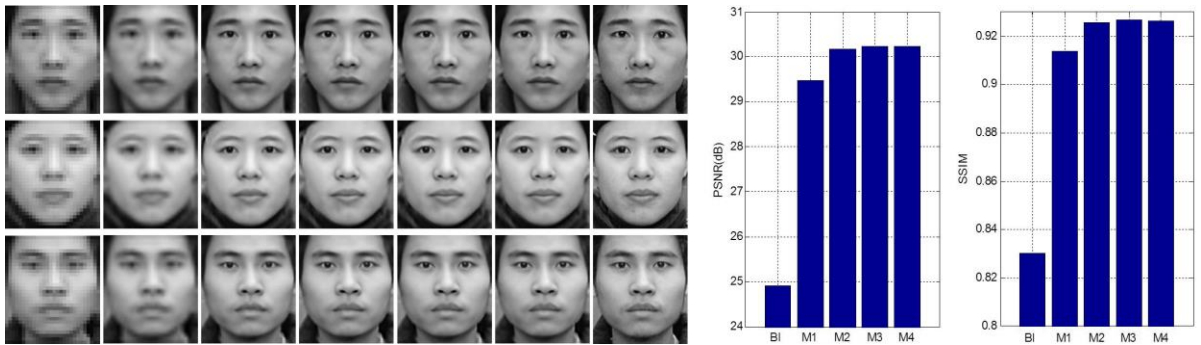


图5-10 不同层字典下的人脸超分辨率重构结果对比。左图为三组主观人脸超分辨率重构结果，从左到右依次为：输入低分辨率人脸图像，Bicubic插值（BI）结果，一层字典（M1）下的人脸超分辨率结果，两层字典（M2）下的人脸超分辨率结果，三层字典（M3）下的人脸超分辨率结果，四层字典（M4）下的人脸超分辨率结果，以及原始高分辨率人脸图像。右图为不同方法下的PSNR和SSIM客观结果，BI (PSNR=24.92 dB, SSIM=0.8301), M1 (PSNR=29.48 dB, SSIM=0.9138), M2 (PSNR=30.16 dB, SSIM=0.9256), M3 (PSNR=30.22 dB, SSIM=0.9265), M4 (PSNR=30.23 dB, SSIM=0.9265)。

(3) 中间低分辨率字典层数的影响

图 5-10 显示了不同层数下的主观结果（由于篇幅所限，我们只给 40 组测试人脸图像中的三组人脸图像超分辨结果）和客观结果。双三次插值（BI）被视作对比基准算法。图 5-10 的图题中的“M1”表示基于单层低分辨率训练集和高分辨率训练集学习的人脸超分辨率方法，“M2”表示两层低分辨率训练集和高分辨率训练集的方法等等。从主观结果可以看出，

BI 方法插值出来的结果非常平滑且模糊，而我们提出的方法可以恢复大部分的面部细节。同时，随着层数的增加，脸部轮廓和纹理逐渐变得更加清晰。在图 5-10 的图题中，我们在每种方法后面的括号中列出了该算法所有 40 个测试人脸图像的平均 PSNR 和 SSIM 值。从图 5-10 的右半部分可知，提出方法相对于 BI 方法的增益是显著的（在 PSNR 上有 4.56 dB 的提升，在 SSIM 上也有 0.0837 的提升）。随着层数的增加，提升方法相对于 BI 插值方法的增益变得越来越显著。在 PSNR 和 SSIM 两个指标上，基于多层字典学习的方法相对于基于单层字典学习的方法的增益分别是 0.74 dB 和 0.0127。当层数大于 3 时，所提出的方法将达到性能稳定，这也说明其可以快速收敛。



图 5-11 不同算法在 CAS-PEAL-R1 人脸库上超分辨率重构结果对比。从左到右依次为：输入低分辨率人脸图像、Wang 等人提出的特征转换全局脸方法^[54]的重构结果、Chang 等人提出的邻域嵌入方法^[83]的重构结果、Yang 等人提出的稀疏表示方法^[8]的重构结果、Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]的重构结果和我们提出方法的重构结果。

5.4.3 与前沿算法的对比

在本节中，我们将与 Wang 等人提出的特征转换全局脸方法^[54]以及三种基于局部图像块的方法进行对比。这三种基于局部图像块的方法分别是 Chang 等人提出的邻域嵌入方法^[83]，Yang 等人提出的稀疏表示方法^[8]和 Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]。为了得到最佳的性能，我们为每种对比方法进行了细致的参数调节，以达到它们各自最好的性能。特别地，对于 Wang 等人提出的特征转换全局脸方法^[54]，我们选取了 99% 的主成分（约 250 个特征脸）。Chang 等人提出的邻域嵌入方法^[83]中的最近邻个数 K 设成 50。对于 Yang 等人提出的稀疏表示方法^[8]，所有参数的设置均参照作者所给的源代码。对于 Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]，最近邻个数 K 设成 200。对于我们所提出的方法，最近邻个数 K 设成 150。对于所有的基于局部图像块方法（Chang 等人提出的邻域嵌入方法^[83]、Yang 等人提出的稀疏表示方法^[8]、Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]和我们提出的方法），我们建议高分辨率图像块大小设成 12×12 像素，近邻图像块之间的交叠为四个像素。实验中我们发现，相邻图像块之间越多的交叠将可以获得越好的性能。为了公平起见，对于所有基于局部图像块的方法，我们对它们设置了同样的图像块大小和相邻图像块交叠。

(1) 主观结果对比

图 5-11 给出了不同方法超分辨率重构的结果。第一列为输入低分辨率人脸图像，最后一列为对应的原始真实高分辨率人脸图像，第二列到第六列为五种不同人脸超分辨率方法重构的高分辨率人脸图像。从这些主观视觉超分辨率结果可以看出，提出的方法获得了最具竞争力的结果，其重构结果具有更多的人脸特征细节，尤其是在鼻子、嘴巴和人脸轮廓等部分，而且与原始高分辨率人脸图像更加相似。Wang 等人提出的特征转换全局脸方法^[54]有比较明显的“鬼影”效应，特别是在人脸的边缘和嘴巴等部位。Chang 等人提出的邻域嵌入方法^[83]可能引入出一些原始高分辨率原本不存在的特征细节，这主要是由于 K 最邻域嵌入的求解过拟合或者拟合不当所引起的。Yang 等人提出的稀疏表示方法^[8]和 Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]均可获得较好的性能，但是它们重构的结果要么是太平滑，要么是在人眼和人脸轮廓有明显的人工效应。

(2) 客观结果对比

为了进一步评估不同人脸超分辨率算法的性能，我们也给出了五种方法的客观指标——PSNR 和 SSIM 值（如图 5-12 所示）。从该图可以看出，在这五种对比方法中，本章提出的方法获得了最好的性能。图 5-12 的图题给出了不同方法在所有 40 张测试图像上的平均 PSNR 和 SSIM 值。由此结果可以得出如下几点结论：（1）基于局部图像块的方法始终优于 Wang 等人提出的特征转换全局脸方法^[54]，这主要归因于基于局部图像块方法的分块策略，通过分块大大增强原始训练库的表达能力；（2）通过引入人脸图像的位置先验信息，基于位置块的人脸超分辨率方法（Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]和我们提出的方法）

好于 Chang 等人提出的邻域嵌入方法^[83]，因为他们不考虑人脸位置先验信息；（3）在所有五种方法中，我们提出方法相对于第二好的方法（Yang 等人提出的稀疏表示方法^[8]）是非常显著的（在 PSNR 和 SSIM 上分别提升了 1.66 dB 和 0.0303）。我们将这一优越性归因于提出方法的迭代邻域嵌入策略和中间字典学习策略。

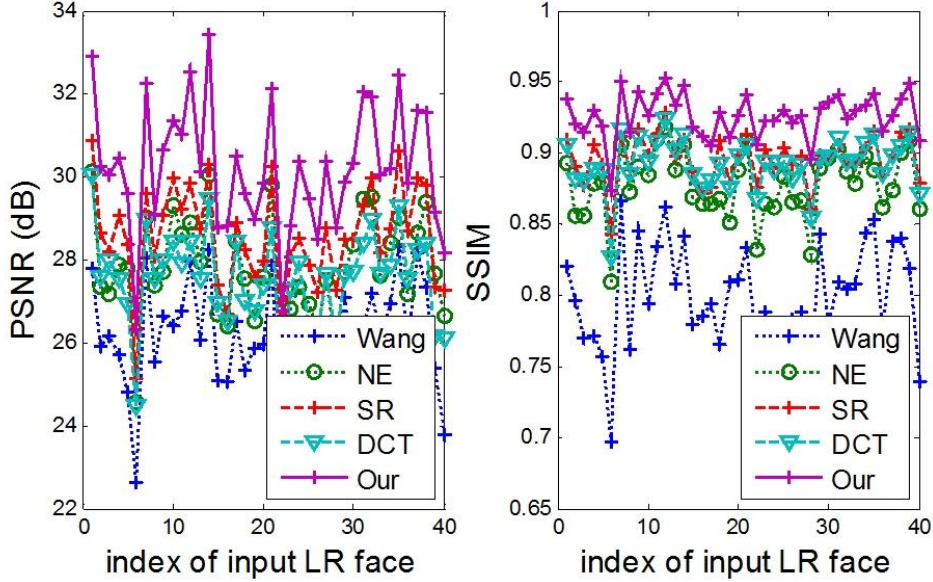


图 5-12 不同算法在 40 张测试人脸图像重构结果的平均 PSNR 和 SSIM 值对比。Wang 等人提出的特征转换全局脸方法^[54] (PSNR=26.04 dB, SSIM=0.7989)、Chang 等人提出的邻域嵌入方法^[83] (PSNR=27.76 dB, S-SIM=0.8841)、Yang 等人提出的稀疏表示方法^[8] (PSNR=28.56 dB, SSIM=0.8962)、Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68] (PSNR=27.59 dB, SSIM=0.8929)和我们提出方法 (PSNR=30.22 dB, SSIM=0.9265)。

(3) 算法复杂度对比。

本小节，我们对比了提出方法和几种前沿人脸超分辨率方法的复杂度（CPU 时间）。图 5-13 给出了所有 40 张测试图像在 Windows 平台 Matlab (R2012a) 软件、Intel Xeon 8 核 CPU (2.13GHz 和 16G 内存) PC 机下的平均 CPU 运行时间。与 Chang 等人提出的邻域嵌入方法^[83]、Yang 等人提出的稀疏表示方法^[8]、Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]相对，由于提出方法是一种“由粗到精”的迭代策略，它需要消耗更多的时间来逐渐优化重构结果。Wang 等人提出的特征转换全局脸方法^[54]是最快的，它只需要几步简单的矩阵相乘和相加。基于局部图像块表示的方法展示出很好的超分辨率重构性能，然而以增加额外的计算复杂度为代价。由于在基于局部图像块表示的人脸超分辨率过程中，每个图像块的表示与超分辨率重构是相互独立的，我们可以通过采取并行处理的方式进行加速。

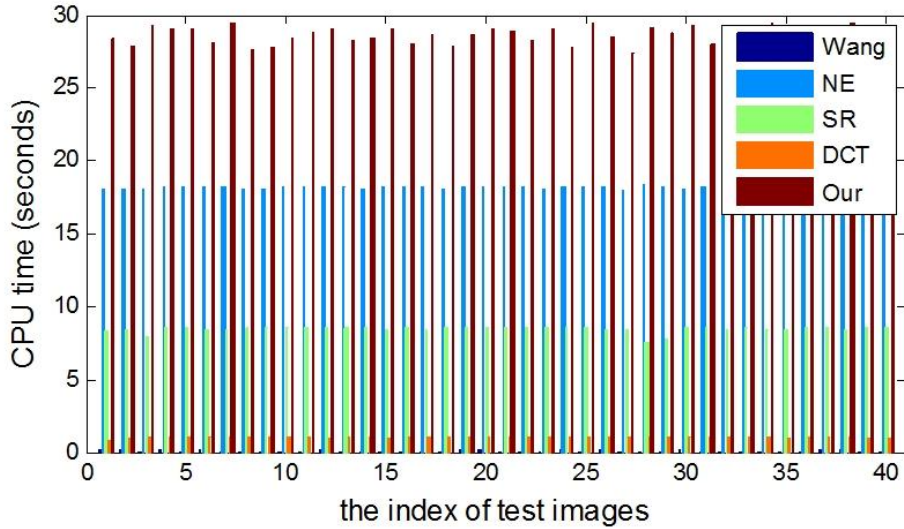


图 5-13 不同算法在 40 张测试人脸图像的平均 CPU 时间对比。Wang 等人提出的特征转换全局脸方法^[54]、Chang 等人提出的邻域嵌入方法^[83]、Yang 等人提出的稀疏表示方法^[8]、Zhang 等人提出的基于 DCT 域的邻域嵌入方法^[68]和我们提出方法的平均 CPU 时间分别是 0.1249 秒, 18.18 秒, 8.480 秒, 1.051 秒 and 28.62 秒。

5.4.4 实际人脸超分辨率重构结果

以上实验的输入低分辨率人脸图像均是通过对原始高分辨率人脸图像进行平滑和下采样操作来获得, 这种模拟的降质过程无法反应高分辨率图像和实际应用中得到的低分辨率图像之间的真实空间关系。在实际情况下, 要完全的模拟真实的降质过程是非常困难的(降质过程如何影响图像的结构和统计特性是不清楚的)。

因此, 为了进一步测试提出方法的有效性, 我们对实际环境下获得的低分辨率图像进行超分辨率重构。图 5-14 为实际拍摄的一幅照片, 其中每个人脸的尺寸约为 30 个像素。我们首先用自动人脸检测方法^[146]来检测上述照片中的人脸, 然后利用检测得到的两个人眼中间位置进行人脸对齐操作。由于人类对光照改变比较敏感, 我们只是利用彩色图像的亮度成分进行超分辨率重构, 而对另外的两个彩色通道直接进行 Bicubic 上采样。图 5-14 的第二到第五列显示了部分人脸图像的超分辨率结果。对于每组结果, 左图为检测并对齐的低分辨率人脸图像, 右图为超分辨率重构的高分辨率人脸图像。从第二行到第四行(对应照片中标记为绿色框的人脸)的重构结果可以看出, 我们提出的方法可以重构出非常合理的结果, 尽管输入低分辨率人脸图像与训练库中的人脸图像(图 5-7 所示)差异很大。然而, 我们同时也发现, 重构人脸图像的人眼边缘存在一些人工效应。这种人工效应可以解释为测试图像和训练图像的不一致性。例如, 测试人脸图像中的人脸都是眯眼的, 而训练库中的人脸图像均是睁开眼睛的(图 5-7 所示)。图 5-14 中的最后一列(对应照片中标记为红色框的人脸)为超分辨率重构失败的情况, 可以看出, 这组重构结果无法提升输入低分辨率人脸图像的质量。以

上的实验结果也说明了基于学习方法存在一定的局限性，即这类方法要求测试数据和训练数据应该具有一定的相似性，否则重构的结果将不佳^[44]。

5.4.5 讨论

从以上的实际实验结果可以看出，人脸超分辨率方法在这种条件下重构出来的人脸图像仍然有一些噪声和明显的人工效应，这些结果相对来说不如在中国人脸图像库上的仿真实验结果（测试图像和训练集都是来自于中国人脸库）。这也说明了基于学习的一个缺陷，它要求训练集和输入的测试图像有足够的统计相似性（例如相似的拍摄条件）。特别地，在实际应用中（对监控人脸图像进行超分辨率重构），拍摄得到的人脸图像可能不是完全的正面人脸（有一定的角度或者是任意姿势的）、光照强度有限、存在一定的遮挡等。图 5-15 给出了输入为侧面偏转 30 度情况下的低分辨率人脸图像的超分辨率重构结果和输入为正面低分辨率人脸图像情况下的超分辨率重构结果。可以看出，输入是侧面人脸时的重构结果明显差于输入是正面人脸时的情况。这主要是因为侧面低分辨率人脸图像与原始高分辨率人脸图像之间的局部流形结构的不一致性更加明显，基于流形一致保持的人脸超分辨率方法的学习性能严重下降所导致的。

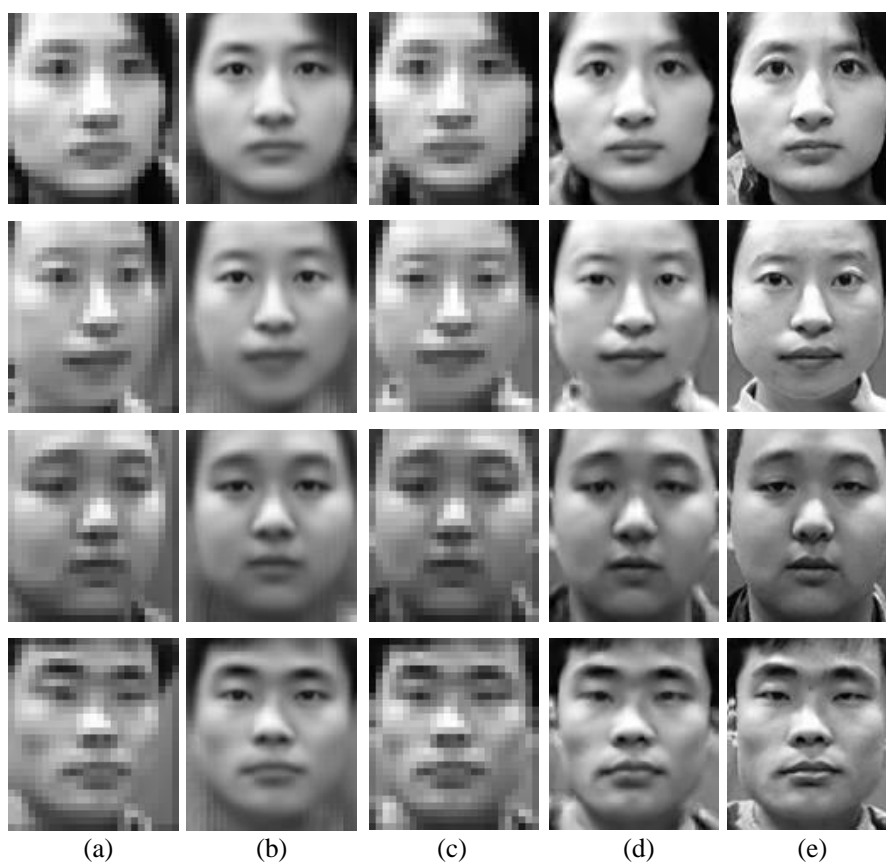


图 5-15 输入为侧面人脸和正面人脸下的超分辨率结果。(a) 和 (c) 分别为输入侧面低分辨率人脸和正面低分辨率人脸，(b) 和 (d) 分别是对应输入下的超分辨率重构结果，(e) 为原始高分辨率人脸图像。

上述实际是在仿真实验条件下的结果，低分辨率图像均是侧面偏转 30 度和下采所得，而实际的成像过程更加复杂。因此，感兴趣监控人脸的增强是一个极其困难和复杂的问题，对于这类实际应用问题要想得到一个满意的结果，需要多种技术的组合，例如，近来提出的稠密图像匹配^[122]，人脸成分分解^[72]等。除此之外，在实际监控环境中，我们通常可以得到一组人脸图像序列。因此，除了利用高低分辨率人脸图像样本对所提供的先验信息，时域多帧图像之间的互补信息也同样可以用来作为超分辨率重构的信息来源。通过联合基于多帧重建的方法（挖掘时域多帧图像所提供的信息）和基于学习的方法（挖掘训练集所提供的信息）来恢复一幅高质量的高分辨率人脸图像将是我们进一步的工作，时域多帧信息的引入也为解决人脸超分辨率重构过程中的遮挡问题提供了可能。



图 5-14 实际拍摄照片超分辨率结果示意。第一行为拍摄得到的照片，第二行到第五行为超分辨率重构结果。对于每组结果，左图为检测并对齐的低分辨率人脸图像，右图为超分辨率重构的高分辨率人脸图像。

近年来,同时人脸超分辨率重构和人脸识别的研究结果指出,当前的人脸超分辨率方法可以获得视觉上满意的结果,但是由于他们不是为人脸识别所设计的,因而,人脸超分辨率重构的结果可能不适用于接下来的识别任务^[11, 97, 108, 119, 147-164]。尽管这样,我也不能得出人脸超分辨率重构对人脸识别没有任何意义的结论。因为人脸识别算法的性能严重依赖于人脸描述子(人脸图像的表达)。正如 Liang 等人^[56]的论文所指出的,基于全局特征描述(Eigenface Fisherface 或者 KPCA)的方法很难表示出对人脸识别有用的图像高频成分。因此,如果人脸识别算法利用全局的人脸特征来表示一幅人脸图像时,确实是没有必要在人脸识别之前进行人脸超分辨率重构,尤其是在输入人脸图像的尺寸大小已经大于 30×30 像素的时候(更大尺寸下的基于全局特征表示的人脸识别方法的性能没有明显的提升)。从整体来看,人脸都是非常相似的,只有在微观细节上存在一定的差异,而这种差异可以利用一些局部描述子进行表达,例如 Local Binary Pattern (LBP), Scale-Invariant Feature Transform (SIFT) and Histogram of Oriented Gradients (HOG)等。因此,一种可以有效地重构出高频特征细节的人脸超分辨率方法将有助于提高基于局部描述子的人脸识别算法的性能。因此,设计一种可以重构更多具有判别特征的高效人脸超分辨率将是我们将要考虑的第二个问题。

5.5 结论

在此章节中,我们提出一种全新的人脸超分辨率框架——多层次局部约束迭代邻域嵌入方法。在此方法中,我们首先利用 Bicubic 插值获得初始高分辨率人脸图像,然后计算估计得到的高分辨率图像块和训练集高分辨率图像块之间的距离,并在高分辨率空间查找 K 个最近邻的高分辨率图像块。接着,我们利用计算得到的距离来惩罚距离大的高分辨率样本图像块,而尽量选择那些距离比较近的图像块。最后,根据同样的重构权重和高分辨率训练样本库中对应的 K 个近邻高分辨率图像块线性估计目标高分辨率图像块。经过几步迭代,可以重构出最终高分辨率图像块。将所有加权重构出的高分辨率图像块按照其在人脸图像上的位置进行叠加,然后除以每个像素位置被交叠的次数,最终重构出高分辨率人脸图像。在上述单层局部约束迭代嵌入算法的基础上,我们对其进行了扩展,通过中间低分辨率人脸训练字典学习,进而提出了一种基于多层局部约束迭代嵌入的人脸超分辨率算法。基于这种多层策略,提出的方法可以很好的逼近图像潜在的降质过程,并且进一步优化超分辨率重构结果。实验结果表明,与前沿算法相对,提出的方法在主客观性能上均达到了最优。

第六章 总结与展望

视频监控系统作为目前最为有效的安防措施之一，被非常广泛地使用着，视频监控信息正以其无法比拟的优势在刑事侦查中发挥着越来越重要的作用。人脸图像作为最直观的线索，在刑侦破案过程中起到非常重要的作用，是刑侦人员最关注的目标之一。视频监控中，由于距离、压缩、下采样、噪声等影响，拍摄得到的人脸图像缺乏必要的辨识特征。当目标嫌疑人距监控摄像头较远时，人脸目标占据的有效像素很少，提供的信息有限；由于监控设备自身的限制，监控系统对人脸图像的下采样也会产生相似的影响；视频图像的压缩过程会给拍摄得到的人脸图像加上模糊效果，影响对人脸图像本身的特征辨识；人脸图像上的噪声会影响对人脸特征细节的判断，由于噪声影响的粒度和成像器件的大小密切相关，因此随着图像目标分辨率降低，噪声对于特征辨识的影响急剧增加。人脸超分辨率技术可以克服已有成像系统的不足，突破成像环境的各种条件限制，借助高低分辨率人脸图像样本训练对提供的先验知识，利用以较低的成本重建高质量、高分辨率、可辨识的人脸图像，进而提高公安机关破案率，为社会安全视频监控图像信息深度应用提供从基础理论。

6.1 本文工作总结

本文通过借鉴人类视觉感知和认知的机理完成人脸图像库的扩充表达、人脸图像的噪声鲁棒表示、高低分辨率图像流形空间一致流形关系的学习，最后到图像观测模型的逼近过程。主要研究成果为：

(1) 提出基于最近特征线的人脸超分辨率重构方法

在人脸图像库表达能力扩充层面上，通过引入最近特征线思想，连接任意两个样本图像（块）形成特征线，将人脸图像库中的两个样本图像（块）扩充为特征线上的无数多个样本，从而实现从基于样本点的图像表示到基于特征线的图像表示。通过保持相对传统基于局部线性嵌入流形学习方法更小局部内的线性关系来揭示高低分辨率人脸流形空间的非线性关系，更加体现了流形学习的本质。现有人脸超分辨率算法都是基于一个较小规模的人脸图像数据库，这些人脸图像库只能构成高维人脸流形空间的一个稀疏空间，即使最相邻的样本点也难以称其为局部，欠采样人脸样本空间的局部线性假设并不成立。因此，局部线性嵌入流形学习思想从本质上并不适合于欠采样人脸样本数据的图像超分辨率重构。

(2) 提出基于局部约束表示的噪声鲁棒人脸超分辨率重构方法

在图像鲁棒表示层面上，通过研究符合人类视觉先验知识的鲁棒图像表示方法，利用局部性来约束图像块的表示系数，在图像块样本空间中自适应地选取与输入图像块距离近的样本图像块，惩罚距离远的样本图像块，通过赋予相似图像块更大的权重系数，使得图像块的

表示对噪声更加鲁棒，实现人脸图像从稀疏表示到局部约束表示的转变。现有人脸超分辨率的图像块表示建立在信号稀疏表示理论上，采用尽可能少的图像块对观测图像块进行表示，未考虑样本图像块空间的局部相似性这一视觉先验知识，制约了图像块表示的精确性和稳定性的进一步提升。

(3) 提出基于高分辨率空间流形结构保持的人脸超分辨率重构方法

在高低分辨率图像空间邻域嵌入学习层面上，通过研究基于高分辨率图像流形空间约束的邻域嵌入学习方法，利用不会受下采样和噪声影响的高分辨率流形结构作为先验知识，约束重构后的高分辨率图像流形空间，使重构后的高分辨率流形结构特征与原始高分辨率流形结构一致，实现从保持低分辨率图像流形结构到高分辨率图像流形约束的一致流形学习。现有人脸超分辨率方法本质上都是将低分辨率图像的局部几何结构保持到重建后的高分辨率图像流形中，由于低分辨率图像获取会有信息损失，高低分辨率图像之间存在“多对一”的不确定性关系，导致重建高分辨率流形空间与原始真实高分辨率空间的局部几何结构不一致。尤其降质过程比较复杂（放大倍数过大且有模糊或者噪声的影响）导致高低分辨率图像流形不一致的条件下，现有人脸超分辨率无法有效恢复出高分辨率人脸图像。

(4) 提出基于局部迭代嵌入和字典渐进更新的人脸超分辨率重构方法

在成像过程的多层次协同建模层面上，借鉴人类认识事物的一般规律——由粗到精、由全局到局部，首次将低分辨率人脸字典更新的思想引入到人脸超分辨率重构中，渐近地学习得到多层中间低分辨率人脸字典，逐渐减小低分辨率训练空间与原始高分辨率训练空间的维度差。在渐近更新后的低分辨率人脸空间学习高低分辨率图像之间的关系，从而能够更加准确地进行邻域嵌入学习与预测，实现从单一层次流形保持学习到多层次流形保持学习。现有人脸超分辨率方法均是直接一次性地学习低分辨率图像空间到高分辨率空间的显式映射关系或隐式编码关系，然而高低分辨率图像之间存在“多对一”的不确定性，特别是在下采样倍数较大或者低分辨率空间受噪声因素影响时，高低分辨率图像空间的维度差将非常大（表现为高分辨率流形不一致），一次性的映射学习方法无法有效地对高低分辨率图像空间的关系进行准确建模。

6.2 未来工作展望

人脸超分辨率重建是当前图像处理领域非常活跃的研究课题之一，它涉及信号处理、机器学习、优化问题求解以及认识科学等诸多学科的交叉研究，具有重要的理论研究和实际应用价值。本文对实际监控环境下的人脸图像超分辨方法进行了一些尝试，提出了几种新的人脸超分辨率重建方法，为未来的深入研究打下了一定的基础。

然而，由于实际监控环境下，低分辨率人脸图像获取过程极其复杂，成像过程受距离、压缩、下采样、噪声等多方面因素的影响，监控人脸图像的超分辨率重构远未达到成熟，其中有不少值得继续研究的方面，仍然存在许多问题有待进一步研究与探索：

- (1) **高低分辨率图像流形空间一致特征的提取。**现有方法要么直接利用原始像素特征，要么利用梯度信息作为高低分辨率图像的表示，然而低分辨率图像获取过程会有信息丢失，高低分辨率图像流形空间的几何结构不具有 consistency。因此，提取更加一致的高低分辨率图像特征，或者将高低分辨率图像投影到一个更加一致的空间是这类基于流形假设人脸图像超分辨率方法的关键。
- (2) **充分挖掘人脸这一特定对象的先验信息。**人脸图像不同于自然图像，它具有很强的结构先验。在对人脸图像进行超分辨率重构时，现有方法根据自然图像超分辨率重构一样，对整个人脸图像划分为统一大小的图像块，并对每个图像块进行分别处理，完全忽略了人脸的五官位置信息。因此，结合人脸图像的五官位置信息，对人脸图像进行合理分块，并利用人脸的对称性，对人脸图像进行超分辨率重构是下将来需要进一步深入研究的问题。
- (3) **基于人脸先验的人脸参数图构建。**现在方法对不同位置的人脸图像块进行建模表示时，采用了统一的模型参数，然后人脸图像不同位置的纹理信息不同，例如额头、脸部位置比较平均，而眉毛、眼睛和嘴巴等位置的纹理较丰富，应该采取不同的模型参数。根据人脸图像的位置信息，提取训练得到一个关于特定模型的参数图，再利用该参数图指导测试人脸图像在进行图像块表达时的参数选取，从而提高模型的适用性，是一条非常有效的途径。
- (4) **视频人脸超分辨率重构。**基于多帧重建的方法中约束信息主要来源于多帧低分辨率图像之间存在的互补信息（时空相关性约束），基于学习的方法中约束信息则主要来源于大量高分辨率图像学习样本（内容相关性约束）。对于监控视频超分辨率而言，这两种约束信息各有优缺点：时空相关性约束信息来源于实际视频中的相邻图像序列，因此在对于刑侦业务非常重要的保真度方面更好，但是由于相邻图像帧数有限，帧间互补信息也无法保证，高频信息来源有限，在分辨率放大倍数较大时在细节的清晰度方面要更差；而内容相关性约束信息来源于大量高分辨率图像学习样本，因此在细节的清晰度方面要更好，但是由于输入图像与样本图像是实际视频中的相邻图像，差异较大，因此在保真度方面更差。传统超分辨率方法仅以时空相关性（基于重建的方法）或内容相关性（基于学习的方法）来约束高分辨率图像的重建过程，约束信息来源单一，难以兼顾清晰度和保真度，重建效果不理想。将时空相关性和内容相关性结合起来对超分辨率重构过程进行多约束，同时提高重建结果的清晰度和保真度，为更全面地提高监控视频人脸的质量提供了一个方向。

附录 部分推导与数学证明

附录 A 图像噪声强度与正则化参数的关系

下面将说明局部约束正则参数 τ 和输入低分辨率人脸图像的噪声水平 σ 之间的关系。

假设 $w_D(i, j) = Dw(i, j)$, 目标函数 (3-8) 可重新写作:

$$w_D^*(i, j) = \arg \min_{w_D(i, j)} \left\{ \|X(i, j) - Y_D(i, j)w_D(i, j)\|_2^2 + \tau \|w_D(i, j)\|_2^2 \right\}, \quad (\text{A-1})$$

其中, $Y_D(i, j) = Y(i, j)D^{-1}$, 最优表示权重可写作 $w(i, j) = D^{-1}w_D(i, j)$ 。为了方便起见, 在不引起混淆的情况下, 我们舍掉位置信息 (i, j) 。那么, 式 (A-1) 可写作:

$$w_D^* = \arg \min_{w_D} \left\{ \|X - Y_D w_D\|_2^2 + \tau \|w_D\|_2^2 \right\}. \quad (\text{A-2})$$

根据Bayesian理论, 目标函数 (A-2) 可写作如下形式:

$$\begin{aligned} w_D^* &= \arg \max_{w_D} \{ \log P(w_D | X) \} \\ &= \arg \min_{w_D} \{ -\log P(w_D) - \log P(X | w_D) \}. \end{aligned} \quad (\text{A-3})$$

其中, $P(X | w_D)$ 是条件概率, $P(w_D)$ 是表示权重向量 w_D 的先验分布. 由于 X 是已知观察, 那么 $P(X)$ 是一个可以忽略的常数. 为了计算最大后验估计, 我们假设观察到的低分辨率人脸图像是受加性高斯噪声干扰的, 那么我们就有:

$$P(X | w_D) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} \|X - Y_D w_D\|_2^2\right). \quad (\text{A-4})$$

先验分布 $P(w_D)$ 可以通过独立同分布的高斯分布来刻画:

$$P(w_D) = \frac{1}{\nu\sqrt{2\pi}} \exp\left(-\frac{1}{2\nu^2} \|w_D\|_2^2\right), \quad (\text{A-5})$$

其中, ν 是 w_D 的标准差. 将 $P(X | w_D)$ 和 $P(w_D)$ 代入式 (A-3), 我们可以得到:

$$\tau = \frac{\sigma^2}{\nu^2}. \quad (\text{A-6})$$

假设 w_D 的标准差 ν 是固定的, 输入图像的噪声越大, 即 σ 越大, 那么所需要的 τ 就必须越大。

参考文献

1. Ma, X., J. Zhang, and C. Qi, *Hallucinating face by position-patch*. Pattern Recognition, 2010. **43**(6): p. 2224-2236.
2. Jung, C., et al., *Position-Patch Based Face Hallucination Using Convex Optimization*. IEEE Signal Proc. Let., 2011. **18**(6): p. 367-370.
3. Huang, H., et al., *Super-resolution of human face image using canonical correlation analysis*. Pattern Recognition, 2010. **43**(7): p. 2532-2543.
4. Gao, W., et al., *The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations*. IEEE Trans. Syst. Man Cybern. Part A-Syst. Hum., 2008. **38**(1): p. 149-161.
5. Thomaz, C.E. and G.A. Giraldi, *A new ranking method for principal components analysis and its application to face image analysis*. Image and Vision Computing, 2010. **28**(6): p. 902 - 913.
6. Wang, Z., et al., *Image quality assessment: from error visibility to structural similarity*. IEEE Trans. Image Process., 2004. **13**(4): p. 600-612.
7. Martinez, A.M. and R. Benavente, *The AR face database*. Computer Vision Center, Barcelona, Spain, Tech. Rep. 24, 1998.
8. Jianchao, Y., et al., *Image Super-Resolution Via Sparse Representation*. Image Processing, IEEE Transactions on, 2010. **19**(11): p. 2861-2873.
9. 安防视频监控系统贡献知多少, in 中国安防网. 2010.
10. 黎智辉, 面向视频侦查的视频内容可辨识度实时增强技术研究. 2013, 公安部物证鉴定中心: 上海.
11. Wang, Z., et al., *Low-resolution face recognition: a review*. The Visual Computer, 2013. **30**(4): p. 1-28.
12. 卢涛, 低质量监控视频人脸超分辨率技术研究. 2013, 武汉大学.
13. J., H., *Diffraction and resolving power*. Journal of the Optical Society of America, 1964. **54**(7): p. 931-936.
14. Goodman, J., *Introduction to fourier optics*. . 1968, New York: Mc Graw-Hill.
15. Huang, T.S. and R. Y. Tsay, *Multiple frame image restoration and registration*. 1984, JAI: Greenwich.
16. Zhao, W., et al., *Face recognition: A literature survey*. ACM Comput. Surv., 2003. **35**(4): p. 60.
17. Wang, N., et al., *A Comprehensive Survey to Face Hallucination*. International Journal of Computer Vision, 2013. **106**(1): p. 1-22.
18. 卓力, 王素玉, and 李晓光, *图像视频的超分辨率复原*. 2011, 北京: 人民邮电出版社.
19. Baker, S. and T. Kanade. *Hallucinating faces*. in FG. 2000.
20. Yue, H., et al., *Landmark Image Super-Resolution by Retrieving Web Images*. 2013.
21. Kopf, J. and D. Lischinski. *Depixelizing pixel art*. in ACM Transactions on Graphics (TOG). 2011. ACM.
22. Džaja, B., M. Bonković, and L. Malešević, *Solving a two-colour problem by applying probabilistic approach to a full-colour multi-frame image super-resolution*. Signal Processing: Image Communication, 2013. **28**(5): p. 509-521.
23. Li, X., et al., *A multi-frame image super-resolution method*. Signal Processing, 2010. **90**(2): p. 405-414.
24. Cheng, M.-H., H.-Y. Chen, and J.-J. Leou, *Video super-resolution reconstruction using a mobile search strategy and adaptive patch size*. Signal Processing, 2011. **91**(5): p. 1284-1297.

25. Farsiu, S., et al., *Advances and challenges in super-resolution*. International Journal of Imaging Systems and Technology, 2004. **14**(2): p. 47-57.
26. Milanfar, P., *Super-resolution imaging*. 2010.
27. Lin, Z. and H.-Y. Shum, *Fundamental limits of reconstruction-based superresolution algorithms under local translation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2004. **26**(1): p. 83-97.
28. Lin, Z. and H.-Y. Shum, *Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation*. IEEE Trans. Pattern Anal. Mach. Intell., 2004. **26**(1): p. 83-97.
29. Freeman, W., E. Pasztor, and O. Carmichael, *Learning Low-Level Vision*. International Journal of Computer Vision, 2000. **40**(1): p. 25-47.
30. Baker, S. and T. Kanade, *Limits on super-resolution and how to break them*. IEEE Trans. Pattern Anal. Mach. Intell., 2002. **24**(9): p. 1167-1183.
31. Elad, M., *Sparse and redundant representations: from theory to applications in signal and image processing*. 2010: Springer.
32. Li, J., et al., *Single-image super-resolution reconstruction based on global non-zero gradient penalty and non-local Laplacian sparse coding*. Digital Signal Processing, 2014. **26**(0): p. 101-112.
33. Haichao, Z., et al., *Generative Bayesian Image Super Resolution With Natural Image Prior*. Image Processing, IEEE Transactions on, 2012. **21**(9): p. 4054-4067.
34. Peleg, T. and M. Elad, *A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution*. IEEE Transactions on Image Processing, 2014. **23**(6): p. 2569-2582.
35. Jian, S., et al., *Gradient Profile Prior and Its Applications in Image Super-Resolution and Enhancement*. Image Processing, IEEE Transactions on, 2011. **20**(6): p. 1529-1542.
36. Heng, S., et al., *Spatially Adaptive Block-Based Super-Resolution*. Image Processing, IEEE Transactions on, 2012. **21**(3): p. 1031-1045.
37. 张凯兵, *基于广义稀疏表示的图像超分辨重建方法研究* 2012, 西安电子科技大学.
38. Freeman, W.T., T.R. Jones, and E.C. Pasztor, *Example-based super-resolution*. Computer Graphics and Applications, IEEE, 2002. **22**(2): p. 56-65.
39. van Ouwerkerk, J.D., *Image super-resolution survey*. Image and Vision Computing, 2006. **24**(10): p. 1039-1052.
40. Liu, C., H.-Y. Shum, and W.T. Freeman, *Face hallucination: Theory and practice*. International Journal of Computer Vision, 2007. **75**(1): p. 115-134.
41. Zeyde, R., M. Elad, and M. Protter, *On single image scale-up using sparse-representations*, in *Curves and Surfaces*. 2012, Springer. p. 711-730.
42. Ni, J., et al., *Example-Driven Manifold Priors for Image Deconvolution*. TIP, 2011. **20**(11): p. 3086-3096.
43. Ma, X., J. Zhang, and C. Qi. *Position-based face hallucination method*. in ICME. 2009.
44. Liu, C., H.-Y. Shum, and C.-S. Zhang. *A two-step approach to hallucinating faces: global parametric model and local nonparametric model*. in CVPR. 2001.
45. Abdi, H. and L.J. Williams, *Principal component analysis*. Wiley Interdisciplinary Reviews: Computational Statistics, 2010. **2**(4): p. 433-459.
46. Shlens, J., *A tutorial on principal component analysis*. arXiv preprint arXiv:1404.1100, 2014.

47. Cross, G.R. and A.K. Jain, *Markov random field texture models*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1983(1): p. 25-39.
48. Jia, K. and S. Gong, *Generalized Face Super-Resolution*. IEEE Trans. Image Process., 2008. **17**(6): p. 873-886.
49. Wu, W., Z. Liu, and X. He, *Learning-based super resolution using kernel partial least squares*. Image and Vision Computing, 2011. **29**(6): p. 394-406.
50. An, L. and B. Bhanu, *Face image super-resolution using 2D CCA*. Signal Processing, 2014. **103**(0): p. 184-194.
51. Zhuang, Y., J. Zhang, and F. Wu, *Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation*. Pattern Recognition, 2007. **40**(11): p. 3178-3194.
52. Park, J.-S. and S.-W. Lee, *An Example-Based Face Hallucination Method for Single-Frame, Low-Resolution Facial Images*. IEEE Trans. Image Process., 2008. **17**(10): p. 1806-1816.
53. Chakrabarti, A., A.N. Rajagopalan, and R. Chellappa, *Super-Resolution of Face Images Using Kernel PCA-Based Prior*. IEEE Trans. on Multimedia, 2007. **9**(4): p. 888-892.
54. Wang, X. and X. Tang, *Hallucinating face by eigentransformation*. IEEE Trans. Syst. Man Cybern. Part C-Appl. Rev., 2005. **35**(3): p. 425-434.
55. Capel, D. and A. Zisserman. *Super-resolution from multiple views using learnt image models*. in CVPR. 2001.
56. Liang, Y., X. Xie, and J.-H. Lai, *Face hallucination based on morphological component analysis*. Signal Processing, 2013. **93**(2): p. 445-458.
57. Hu, Y., et al., *A novel kernel-based framework for facial-image hallucination*. Image and Vision Computing, 2011. **29**(4): p. 219-229.
58. Jiang, J., et al., *Two-step superresolution approach for surveillance face image through radial basis function-partial least squares regression and locality-induced sparse representation*. Journal of Electronic Imaging, 2013. **22**(4): p. 041120.
59. Li, B., et al., *Aligning Coupled Manifolds for Face Hallucination*. IEEE Signal Proc. Lett., 2009. **16**(11): p. 957-960.
60. Jingang, S. and Q. Chun. *Face hallucination based on PCA dictionary pairs*. in Image Processing (ICIP), 2013 20th IEEE International Conference on. 2013.
61. Li, Y., et al., *Face hallucination based on sparse local-pixel structure*. Pattern Recognition, 2014. **47**(3): p. 1261-1270.
62. Hongliang, L., X. Linfeng, and L. Guanghui, *Face Hallucination via Similarity Constraints*. Signal Processing Letters, IEEE, 2013. **20**(1): p. 19-22.
63. Zhongyuan, W., et al., *Face Hallucination Via Weighted Adaptive Sparse Regularization*. IEEE Transactions on Circuits and Systems for Video Technology, 2014. **24**(5): p. 802-813.
64. Zhongyuan, W., et al. *Face hallucination via weighted sparse representation*. in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. 2013.
65. Huang, H. and N. Wu, *Fast Facial Image Super-Resolution via Local Linear Transformations for Resource-Limited Applications*. IEEE Transactions on Circuits and Systems for Video Technology, 2011. **20**(10): p. 1363-1377.
66. Yu, H., et al., *From Local Pixel Structure to Global Image Super-Resolution: A New Face Hallucination*

- Framework*. Image Processing, IEEE Transactions on, 2011. **20**(2): p. 433-445.
67. Tao, L., et al. *From local representation to global face hallucination: A novel super-resolution method by nonnegative feature transformation*. in *Visual Communications and Image Processing (VCIP), 2013*. 2013.
68. Wei, Z. and C. Wai-Kuen, *Hallucinating Face in the DCT Domain*. Image Processing, IEEE Transactions on, 2011. **20**(10): p. 2769-2779.
69. Gao, G. and J. Yang, *A novel sparse representation based framework for face image super-resolution*. Neurocomputing, 2014. **134**(0): p. 92-99.
70. Tao, L., et al. *Robust super-resolution for face images via principle component sparse representation and least squares regression*. in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. 2013.
71. Ma, X., et al., *Sparse representation and position prior based face hallucination upon classified over-complete dictionaries*. Signal Processing, 2012. **92**(9): p. 2066-2074.
72. Yang, C.-Y., S. Liu, and M.-H. Yang, *Structured Face Hallucination*. in *CVPR*. 2013.
73. Qiao, L., S. Chen, and X. Tan, *Sparsity preserving projections with applications to face recognition*. Pattern Recognition, 2010. **43**(1): p. 331-341.
74. Lee, D.D. and H.S. Seung, *Learning the parts of objects by non-negative matrix factorization*. Nature, 1999. **401**(6755): p. 788-791.
75. Ding, C., T. Li, and M.I. Jordan, *Convex and semi-nonnegative matrix factorizations*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010. **32**(1): p. 45-55.
76. Hardoon, D., S. Szedmak, and J. Shawe-Taylor, *Canonical correlation analysis: An overview with application to learning methods*. Neural computation, 2004. **16**(12): p. 2639-2664.
77. Krishnan, A., et al., *Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review*. Neuroimage, 2011. **56**(2): p. 455-475.
78. 黄华, et al., *基于粒子滤波的人脸图像超分辨率重建方法* 软件学报, 2006(12): p. 2529-2536.
79. 吴炜, et al., *一种新颖的人脸图像超分辨率技术* 光学精密工程 2008(5): p. 815-821.
80. 张雪松, 江静, and 彭思龙, *特征子空间规整化的人脸图像超分辨率重建* 计算机辅助设计与图形学学报 2010(3): p. 487-493.
81. 兰诚栋, et al., *低质量监控图像鲁棒性人脸超分辨率算法* 计算机辅助设计与图形学学报 2011(9): p. 1474-1480.
82. Sun, J., et al. *Image hallucination with primal sketch priors*. in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*. 2003. IEEE.
83. Chang, H., D.-Y. Yeung, and Y. Xiong. *Super-resolution through neighbor embedding*. in *CVPR*. 2004.
84. Roweis, S.T. and L.K. Saul, *Nonlinear Dimensionality Reduction by Locally Linear Embedding*. Science, 2000. **290**(5500): p. 2323-2326.
85. Yeung, D.-Y. and W. Fan. *Image hallucination using neighbor embedding over visual primitive manifolds*. in *CVPR*. 2007.
86. Kumar, B.G.V. and R. Aravind. *Face hallucination using OLPP and Kernel Ridge Regression*. in *ICIP*. 2008.
87. Gao, X., et al., *Image Super-Resolution With Sparse Neighbor Embedding*. IEEE Trans. Image Process., 2012. **21**(7): p. 3194-3205.

88. Jiang, J., et al. *A super-resolution method for low-quality face image through RBF-PLS regression and neighbor embedding*. in *Proc. IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. 2012.
89. Bevilacqua, M., et al. *Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding*. in *BMVC*. 2012.
90. Liang, C., et al. *A joint learning based face hallucination approach for low quality face image*. in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. 2013.
91. Yuanhong, H. and Q. Chun, *Face Hallucination Based on Modified Neighbor Embedding and Global Smoothness Constraint*. *Signal Processing Letters, IEEE*, 2014. **21**(10): p. 1187-1191.
92. Li, X., et al., *Single image super-resolution via subspace projection and neighbor embedding*. *Neurocomputing*, 2014. **139**(0): p. 310-320.
93. Chen, X. and C. Qi, *Nonlinear neighbor embedding for single image super-resolution via kernel mapping*. *Signal Processing*, 2014. **94**(0): p. 6-22.
94. Yang, J., et al. *Image super-resolution as sparse representation of raw image patches*. in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. 2008. IEEE.
95. Gao, X., et al., *Joint Learning for Single Image Super-resolution via Coupled Constraint*. *IEEE Trans. Image Process.*, 2012. **21**(2): p. 469-480.
96. Zhang, K., et al., *Single Image Super-Resolution With Multiscale Similarity Learning*. *IEEE Trans. Neural Netw. Learn. Syst.*, 2013. **24**(10): p. 1648-1659.
97. Bo, L., et al., *Low-Resolution Face Recognition via Coupled Locality Preserving Mappings*. *Signal Processing Letters, IEEE*, 2010. **17**(1): p. 20-23.
98. Zeng, X. and H. Huang, *Locally nonlinear transformation for facial image superresolution*. *Optical Engineering*, 2013. **52**(2): p. 027001-027001.
99. Huang, C., et al., *Generalized joint kernel regression and adaptive dictionary learning for single-image super-resolution*. *Signal Processing*, 2014. **103**(0): p. 142-154.
100. Dong, W., et al., *Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization*. *TIP*, 2011.
101. Jianchao, Y., et al. *Bilevel sparse coding for coupled feature spaces*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012.
102. Jie, R., L. Jiaying, and G. Zongming, *Context-Aware Sparse Decomposition for Image Denoising and Super-Resolution*. *Image Processing, IEEE Transactions on*, 2013. **22**(4): p. 1456-1469.
103. Shuyuan, Y., et al., *Dual-Geometric Neighbor Embedding for Image Super Resolution With Sparse Tensor*. *IEEE Transactions on Image Processing*, 2014. **23**(7): p. 2793-2803.
104. Weisheng, D., et al., *Nonlocally Centralized Sparse Representation for Image Restoration*. *Image Processing, IEEE Transactions on*, 2013. **22**(4): p. 1620-1630.
105. Weisheng, D., et al., *Sparse Representation Based Image Interpolation With Nonlocal Autoregressive Modeling*. *Image Processing, IEEE Transactions on*, 2013. **22**(4): p. 1382-1394.
106. Xiaoqiang, L., et al. *Geometry constrained sparse coding for single image super-resolution*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012.

107. Shuyuan, Y., et al., *Single-Image Super-Resolution Reconstruction via Learned Geometric Dictionaries and Clustered Sparse Coding*. Image Processing, IEEE Transactions on, 2012. **21**(9): p. 4016-4028.
108. Shenlong, W., et al. *Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012.
109. Kui, J., W. Xiaogang, and T. Xiaoou, *Image Transformation Based on Learning Dictionaries across Image Spaces*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013. **35**(2): p. 367-380.
110. Kaibing, Z., et al. *Multi-scale dictionary for single image super-resolution*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012.
111. Jianchao, Y., et al., *Coupled Dictionary Training for Image Super-Resolution*. Image Processing, IEEE Transactions on, 2012. **21**(8): p. 3467-3478.
112. He, H. and W. Siu. *Single Image Super-Resolution using Gaussian Process Regression*. in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*. 2011.
113. Kim, K.I. and Y. Kwon, *Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior*. IEEE Trans. Pattern Anal. Mach. Intell., 2010. **32**(6): p. 1127-1133.
114. Tang, Y., et al., *Greedy regression in sparse coding space for single-image super-resolution*. Journal of Visual Communication and Image Representation, 2013. **24**(2): p. 148-159.
115. Zhou, F., B. Wang, and Q. Liao, *Super-resolution for facial image using multilateral affinity function*. Neurocomputing, 2014. **133**(0): p. 194-208.
116. Timofte, R., V. De, and L.V. Gool. *Anchored Neighborhood Regression for Fast Example-Based Super-Resolution*. in *Computer Vision (ICCV), 2013 IEEE International Conference on*. 2013. IEEE.
117. Ho, T.-C. and B. Zeng, *Image super-resolution by curve fitting in the threshold decomposition domain*. Journal of Visual Communication and Image Representation, 2012. **23**(1): p. 208-221.
118. Nasir, H., V. Stanković, and S. Marshall, *Singular value decomposition based fusion for super-resolution image reconstruction*. Signal Processing: Image Communication, 2012. **27**(2): p. 180-191.
119. Xiao, Z. and H. Hua, *Super-Resolution Method for Multiview Face Recognition From a Single Image Per Person Using Nonlinear Mappings on Coherent Features*. Signal Processing Letters, IEEE, 2012. **19**(4): p. 195-198.
120. Ma, X., et al., *A simple approach to multiview face hallucination*. Signal Processing Letters, IEEE, 2010. **17**(6): p. 579-582.
121. Chen, Y.-W., S. Sasatani, and X.-H. Han, *Alignment-Free and High-Frequency Compensation in Face Hallucination*. The Scientific World Journal, 2014. **2014**.
122. Tappen, M. and C. Liu, *A Bayesian Approach to Alignment-Based Image Hallucination*. ECCV, 2012: p. 236-249.
123. Liang, Y., et al., *Face hallucination with imprecise-alignment using iterative sparse representation*. Pattern Recognition, 2014. **47**(10): p. 3327-3342.
124. Vrigkas, M., C. Nikou, and L.P. Kondi, *Robust maximum a posteriori image super-resolution*. Journal of Electronic Imaging, 2014. **23**(4): p. 043016-043016.
125. Innerhofer, P. and T. Pock, *A Convex Approach for Image Hallucination*. arXiv preprint arXiv:1304.7153, 2013.

126. 黄丽, et al., 基于多尺度和多方向特征的人脸超分辨率算法 计算机辅助设计与图形学学报, 2004. **16**(7): p. 953-961.
127. 张雪松, 江静, and 彭思龙, 人脸图像超分辨率的自适应流形学习方法 计算机辅助设计与图形学学报, 2008. **20**(7): p. 856-863.
128. 马祥 and 齐春, 全局重建和位置块残差补偿的人脸图像超分辨率算法 西安交通大学学报, 2010(4): p. 9-12.
129. 吴炜, et al., 核偏最小二乘算法的图像超分辨率算法 电子科技大学学报, 2011. **40**(1): p. 105-110.
130. 兰诚栋, 陈亮, and 卢涛, 利用位置权重稀疏表示的人脸超分辨率算法 武汉大学学报: 信息科学版, 2013(1): p. 27-30.
131. 兰诚栋, 陈亮, and 卢涛, 采用后验信息构建稀疏原子库的超分辨率人脸重建 北京工业大学学报, 2013. **39**(007): p. 1072-1077.
132. Li, S.Z. *Face recognition based on nearest linear combinations*. in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*. 1998. IEEE.
133. Chen, Y.-N., et al., *Face recognition using nearest feature space embedding*. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2011. **33**(6): p. 1073-1086.
134. Rowley, H.A., S. Baluja, and T. Kanade, *Neural network-based face detection*. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1998. **20**(1): p. 23-38.
135. Wang, J., et al. *Locality-constrained linear coding for image classification*. in *CVPR*. 2010.
136. Yu, K., T. Zhang, and Y. Gong. *Nonlinear Learning using Local Coordinate Coding*. in *NIPS*. 2009.
137. Jiang, J., et al. *Position-Patch Based Face Hallucination via Locality-Constrained Representation*. in *ICME*. 2012.
138. Junjun, J., et al., *Noise Robust Face Hallucination via Locality-Constrained Representation*. *Multimedia, IEEE Transactions on*, 2014. **16**(5): p. 1268-1281.
139. Tappen, M.F., B.C. Russell, and W.T. Freeman. *Exploiting the Sparse Derivative Prior for Super-Resolution and Image Demosaicing*. in *IEEE Workshop on Statistical and Computational Theories of Vision*. 2003.
140. Ni, K.S. and T.Q. Nguyen, *Image Superresolution Using Support Vector Regression*. *IEEE Trans. Image Process.*, 2007. **16**(6): p. 1596-1610.
141. Tenenbaum, J., V.D. Silva, and J. Langford, *A global geometric framework for nonlinear dimensionality reduction*. *Science*, 2000. **290**(5500): p. 2319-2323.
142. Von Luxburg, U., *A tutorial on spectral clustering*. *Statistics and computing*, 2007. **17**(4): p. 395-416.
143. Zheng, M., et al., *Graph regularized sparse coding for image representation*. *IEEE Trans. Image Process.*, 2011. **20**(5): p. 1327-1336.
144. Elad, M. and A. Feuer, *Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images*. *Image Processing, IEEE Transactions on*, 1997. **6**(12): p. 1646-1658.
145. Park, S.C., M.K. Park, and M.G. Kang, *Super-resolution image reconstruction: a technical overview*. *IEEE Signal Process. Magazine*, 2003. **20**(3): p. 21-36.
146. Everingham, M., J. Sivic, and A. Zisserman. *Hello! My name is... Buffy -- Automatic Naming of Characters in TV Video*. in *BMVC*. 2006.

147. Biswas, S., G. Aggarwal, and P.J. Flynn. *Pose-robust recognition of low-resolution face images*. in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. 2011.
148. Hennings-yeomans, P.H., S. Baker, and B.V.K.V. Kumar. *Simultaneous super-resolution and feature extraction for recognition of low-resolution faces*. in *CVPR*. 2008.
149. Chuan-Xian, R., D. Dao-Qing, and Y. Hong, *Coupled Kernel Embedding for Low-Resolution Face Image Recognition*. Image Processing, IEEE Transactions on, 2012. **21**(8): p. 3770-3783.
150. Zou, W.W.W. and P.C. Yuen, *Very Low Resolution Face Recognition Problem*. Image Processing, IEEE Transactions on, 2012. **21**(1): p. 327-340.
151. Berretti, S., P. Pala, and A. Bimbo, *Face Recognition by Super-Resolved 3D Models From Consumer Depth Cameras*. Information Forensics and Security, IEEE Transactions on, 2014. **9**(9): p. 1436-1449.
152. Hua, H. and H. Huiting, *Super-Resolution Method for Face Recognition Using Nonlinear Mappings on Coherent Features*. Neural Networks, IEEE Transactions on, 2011. **22**(1): p. 121-130.
153. Biswas, S., et al., *Pose-Robust Recognition of Low-Resolution Face Images*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013. **35**(12): p. 3037-3049.
154. Biswas, S., K.W. Bowyer, and P.J. Flynn, *Multidimensional Scaling for Matching Low-Resolution Face Images*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012. **34**(10): p. 2019-2030.
155. Pong, K.-H. and K.-M. Lam, *Multi-resolution feature fusion for face recognition*. Pattern Recognition, 2014. **47**(2): p. 556-567.
156. Wheeler, F.W., R.L. Weiss, and P.H. Tu. *Face recognition at a distance system for surveillance applications*. in *Proc. IEEE Conf. on Biometrics: Theory Applications and Systems (BTAS)*. 2010.
157. Maeng, H., et al. *NFRAD: Near-Infrared Face Recognition at a Distance*. in *Proc. IEEE Joint Conf. on Biometrics (IJCB)*. 2011.
158. Zuodong, Y., et al. *Local consistency preserved coupled mappings for low-resolution face recognition*. in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*. 2013.
159. Fookes, C., et al., *Evaluation of image resolution and super-resolution on face recognition performance*. Journal of Visual Communication and Image Representation, 2012. **23**(1): p. 75-93.
160. Junjun, J., et al. *Graph discriminant analysis on multi-manifold (GDAMM): A novel super-resolution method for face recognition*. in *Image Processing (ICIP), 2012 19th IEEE International Conference on*. 2012.
161. Xu, Y., W. Zuo, and Z. Fan, *Supervised sparse representation method with a heuristic strategy and face recognition experiments*. Neurocomputing, 2012. **79**(0): p. 125-131.
162. Yoshida, T., et al. *Robust Face Super-Resolution Using Free-Form Deformations for Low-Quality Surveillance Video*. in *Multimedia and Expo (ICME), 2012 IEEE International Conference on*. 2012.
163. Zhang, D., J. He, and M. Du, *Morphable model space based face super-resolution reconstruction and recognition*. Image and Vision Computing, 2012. **30**(2): p. 100-108.
164. Kong, Y., S. Zhang, and P. Cheng, *Super-resolution reconstruction face recognition based on multi-level FFD registration*. Optik - International Journal for Light and Electron Optics, 2013. **124**(24): p. 6926-6931.

已获得研究成果情况

• 发表期刊论文

- [1] **J. Jiang**, R. Hu, Z. Wang, and Z. Han. “Face Super-Resolution via Multilayer Locality-constrained Iterative Neighbor Embedding and Intermediate Dictionaries Learning,” IEEE Transactions on Image Processing, vol.23, no.10, pp. 4220-4231, Oct 2014. (CCF A)
- [2] **J. Jiang**, C. Chen, X. Ma, Z. Wang, Z. Wang, and R. Hu. “SRLSP: A Face Image Super-Resolution Algorithm Using Smooth Regression with Local Structure Prior,” IEEE Transactions on Multimedia, DOI 10.1109/TMM.2016.2601020, 2016. (CCF B)
- [3] **J. Jiang**, X. Ma, C. Chen, L. Tao, Z. Wang, and J. Ma. “Single Image Super-Resolution via Locally Regularized Anchored Neighborhood Regression and Nonlocal Means,” IEEE Transactions on Multimedia, DOI 10.1109/TMM.2016.2599145, 2016. (CCF B)
- [4] **J. Jiang**, R. Hu, Z. Wang, and Z. Han. “Noise Robust Face Hallucination via Locality-constrained Representation,” IEEE Transactions on Multimedia, vol. 16, no. 5, pp. 1268-1281, Aug 2014. (CCF B)
- [5] **J. Jiang**, J. Ma, C. Chen, X. Jiang, and Z. Wang. “Noise Robust Face Image Super-resolution through Smooth Sparse Representation,” IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2016.2594184, 2016. (CCF B)
- [6] **J. Jiang**, C. Chen, K. Huang, Z. Cai, and R. Hu. “Noise Robust Position-Patch based Face Super-Resolution via Tikhonov Regularized Neighbor Representation,” Information Sciences, vol. 367-368, pp. 354-372, 2016. (CCF B)
- [7] **Jiang**, R. Hu, Z. Wang, Z. Han, and Jiayi Ma. “Facial Image Hallucination through Coupled-Layer Neighbor Embedding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 9, pp. 1674-1684, 2016. (CCF B)
- [8] **J. Jiang**, X. Ma, R. Hu, and Z. Cai. “Sparse Support Regression for Image Super-Resolution,” IEEE Photonics Journal, vol. 7, no.5, pp. 1-11, Oct. 2015. (中科院二区)
- [9] **J. Jiang**, R. Hu, Z. Wang, and Z. Cai. “CDMMA: Coupled Discriminant Multi-Manifold Analysis for Matching Low-Resolution Face Images,” Signal Processing, vol. 124, pp. 162-172, 2016. (CCF C)
- [10] **J. Jiang**, R. Hu, C. Liang, Z. Han, and C. Zhang. “Face Image Super-Resolution through Locality-induced Support Regression,” Signal Processing, vol. 103, pp.168–183, Oct 2014. (CCF C)
- [11] Z. Wang, R. Hu, C. Liang, Y. Yu, **J. Jiang**, M. Ye, J. Chen, and Q. Leng. “Zero-shot Person Re-identification via Cross-view Consistency,” IEEE Transactions on Multimedia, vol. 18, no. 2, pp. 260-272, 2016. (CCF B)
- [12] J. Ma, H. Zhou, J. Zhao, Y. Gao, **J. Jiang**, and J. Tian. “Robust feature matching for remote sensing image registration via locally linear transforming,” IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 12, pp. 6469 - 6481, 2015. (*ESI Hot Paper and ESI Highly Cited Paper*) (CCF B)
- [13] Z. Wang, R. Hu, S. Wang, and **J. Jiang**. “Face Hallucination via Weighted Adaptive Sparse Regularization,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 5, pp. 802-813, May 2014. (CCF B)

• 发表会议论文

- [1] **J. Jiang**, C. Chen, X. Song, and Z. Cai. "Hyperspectral Image Classification using Set-To-Set Distance," in Proceedings of the 41th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai, China, pp. 3346-3350. (*CCF B*)
- [2] **J. Jiang**, Z. Wang, C. Chen, and T. Lu. "L1-L1 norms for face super-resolution with mixed Gaussian-impulse noise," in Proceedings of the 41th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai, China, pp. 2089-2093. (*CCF B*)
- [3] **J. Jiang**, J. Fu, T. Lu, R. Hu and Z. Wang. "Locally Regularized Anchored Neighborhood Regression for Fast Super-Resolution," in Proceedings the 15th International Conference on Multimedia and Expo (ICME 2015), Turin, Italy, Jun 2015. (*CCF B*)
- [4] **J. Jiang**, R. Hu, Z. Han, L. Chen, and J. Chen. "Coupled Discriminant Multi-Manifold Analysis (CDMMA) with Application to Low-Resolution Face Recognition," in Proceedings of the 21th International Conference on Multimedia Modelling (MMM 2015), Sydney, Australia, pp. 37-48, Jan 2015. (*Best Student Paper Runner-up Award*) (*CCF C*)
- [5] **J. Jiang**, R. Hu, Z. Han, Z. Wang, T. Lu, and J. Chen. "Locality-constraint iterative neighbor embedding for face hallucination," in Proceedings of the 14th International Conference on Multimedia and Expo (ICME 2013), San Jose, California, USA, Jul 2013. (*CCF B*)
- [6] **J. Jiang**, R. Hu, L. Chen, Z. Han, T. Lu, and J. Chen. "Coupled-layer Neighbor Embedding for Surveillance Face Hallucination," in Proceedings of the 20th International Conference on Image processing (ICIP 2013), Melbourne, Australia, pp. 2802-2806, Sep 2013. (*CCF C*)
- [7] **J. Jiang**, R. Hu, Z. Wang, and Z. Han. "Manifold regularized sparse support regression for single image super-resolution," in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), Vancouver, Canada, pp. 1429-1433, May 2013. (*CCF B*)
- [8] **J. Jiang**, R. Hu, Z. Wang, Z. Xiong, and Z. Han. "Support-Driven Sparse Coding for Face Hallucination," in Proceedings of the 47th International Conference on Circuits and Systems (ISCAS 2013), Beijing, China, pp. 2980-2983, May 2013. (*CCF C*)
- [9] **J. Jiang**, R. Hu, Z. Han, K. Huang, and T. Lu. "Efficient Single Image Super-Resolution via Graph Embedding," in Proceedings of the 13th International Conference on Multimedia and Expo (ICME 2012), Melbourne, Australia, pp. 610-615, Jul 2012. (*CCF B*)
- [10] **J. Jiang**, R. Hu, Z. Han, T. Lu, and K. Huang. "Position-Patch Based Face Hallucination via Locality-constrained Representation," in Proceedings of the 13th International Conference on Multimedia and Expo (ICME 2012), Melbourne, Australia, pp. 212-217, Jul 2012. (*CCF B*)
- [11] **J. Jiang**, R. Hu, Z. Han, K. Huang, and T. Lu. "Graph Discriminant Analysis on Multi-Manifold (GDAMM): A Novel Super-Resolution Method for Face Recognition," in Proceedings of the 19th International Conference on Image processing (ICIP 2012), Orlando, Florida, U.S.A., pp. 1465-1468, Sep 2012.
- [12] **J. Jiang**, R. Hu, Z. Han, T. Lu, and K. Huang. "A Super-Resolution Method for Low-Quality Face Image through RBF-PLS Regression and Neighbor Embedding," in Proceedings of the 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan, pp.1253-1256, Mar 2012. (*CCF B*)

- [13] **J. Jiang**, R. Hu, Z. Han, T. Lu, and K. Huang. “Surveillance Face Hallucination via Variable Selection and Manifold Learning,” in Proceedings of the 46th International Conference on Circuits and Systems (ISCAS 2012), Seoul, Korea, pp. 2681-2684, May 2012. (*CCF C*)
- [14] Z. Han, **J. Jiang***, R. Hu, T. Lu, and K. Huang. “Face Image Super-Resolution via Nearest Feature Line,” in Proceedings of the 20th ACM international conference on Multimedia (ACM MM 2012), Nara, Japan, pp. 769-772, Nov 2012. (* indicates equal contribution). (Accepted as short paper). (*CCF A*)

● 获得国家发明专利

- [1] 导师第一发明人、本人第二发明人. “一种基于局部约束表示的人脸超分辨率重建方法”, 申请号: 201110421452.3, 申请日: 2011-12-16. (已授权)
- [2] 导师第一发明人、本人第二发明人. 一种基于最近特征线流形学习的人脸超分辨率方法”, 申请号: 201110421817.2, 申请日: 2011-12-16. (已授权)
- [3] 导师第一发明人、本人第二发明人. “一种 ATM 机广角摄像机镜头畸变校正方法”, 申请号: 201010608963.1, 申请日: 2010-12-28. (已授权)
- [4] 导师第一发明人、本人第二发明人. “基于多流形图判别分析超分辨率的人脸识别方法”, 申请号: 201210164069.9, 申请日: 2012-05-24. (已授权)
- [5] 导师第一发明人、本人第二发明人. “一种基于局部约束迭代邻域嵌入的人脸幻构方法”, 申请号: 201310147620.3, 申请日: 2013-04-24. (已授权)
- [6] 导师第一发明人、本人第二发明人. “一种基于流形正则稀疏支撑回归的单帧图像超分辨率重建方法”, 申请号: 201310147510.7, 申请日: 2013-04-24. (已授权)
- [7] 导师第一发明人、本人第二发明人. “基于支撑集驱动稀疏编码的人脸超分辨率重建方法”, 申请号: 201210523889.2, 申请日: 2012-12-07. (已授权)
- [8] 导师第一发明人、本人第二发明人. “一种基于字典渐近更新的人脸超分辨率算法”, 申请号: 201410323594.X 申请日: 2014-07-09. (已授权)
- [9] 导师第一发明人、本人第二发明人. “一种基于反向图保持的单张图像超分辨率方法”, 申请号: 201210145054.8, 申请日: 2012-05-10. (受理中)

• 获奖情况

- 2015.01, The 21th International Conference on Multimedia Modelling “Best Student Paper Runner-up Award” ;
- 2014.12, 国家媒体软件工程技术研究中心杰出博士生 (1/28)
- 2014.12, 国家媒体软件工程技术研究中心学术贡献奖 (1/28)
- 2014.12, 首届“武昌区政府奖学金” (全校 7 名)
- 2014.09, IBM 中国优秀学生奖学金 (全国 77 名)
- 2013.10, 教育部, 研究生国家奖学金
- 2013.09, 雷军奖学金 (全校 30 名本科生、10 名硕士生和 10 名博士生获得该奖)
- 2013.07, 国家媒体软件工程技术研究中心杰出博士生 (1/28)
- 2013.07, 国家媒体软件工程技术研究中心学术贡献奖 (1/28)
- 2013.06, 武汉大学“十大学术之星提名奖” (全校 10 人获得“十大学术之星”, 5 人获得“十大学术之星提名奖”)
- 2013.06, 武汉大学研究生学术创新奖三等奖
- 2012.12, 教育部, 研究生国家奖学金
- 2012.12, 武汉大学优秀研究生标兵 (1%)
- 2012.12, 武汉大学优秀研究生甲等奖学金 (4%)
- 2012.11, 武汉大学计算机学院“第二届 E 鸣博士生论坛” 第一名 (1/74)
- 2012.06, 武汉大学“研究生出国参加国际学术会议资助” (全校 19 名)
- 2012.01, 武汉大学工程技术研究中心杰出博士生 (2/33)
- 2012.01, 武汉大学工程技术研究中心学术贡献奖 (1/33)

• 学术交流与学术服务

攻读博士学位以来，受邀前往日本京都（ICASSP 2012）、澳大利亚墨尔本（ICME 2012）、美国奥兰多（ICIP 2012）、中国北京（ISCAS 2012）、中国合肥（“多媒体领域未来研讨会”）等地作大会分会口头汇报和海报展示（4次口头汇报和4次报展）：

- 2013.05, 中国北京国际电路与系统研讨会（ISCAS2013）展报一次
- 2013.01, 中国北京智慧地球大会（UV2013）口头汇报一次
- 2012.10, 美国奥兰多国际图像处理大会（ICIP2012）展报一次
- 2012.07, 澳大利亚墨尔本多媒体与扩展大会（ICME2012）口头汇报和展报各一次
- 2012.05, 中国合肥多媒体领域未来研讨会上做“低质量人脸超分辨率研究”特邀大会报告
- 2012.03, 日本京都国际语音、音频和信号处理大会（ICASSP2012）展报一次

担任以下期刊的审稿人：

- IEEE Trans. on Image Processing (CCF 列表 A 类国际期刊)
- Computer Vision and Image Understanding (CCF 列表 B 类国际期刊)
- Signal Processing (CCF 列表 C 类国际期刊)
- Neurocomputing (CCF 列表 C 类国际期刊)
- Signal Processing: Image Communication (CCF 列表 C 类国际期刊)
- IET Computer Vision (CCF 列表 C 类国际期刊)
- IEEE Signal Processing Letter (CCF 列表 C 类国际期刊)
- IET Biometrics

致 谢

时光荏苒，岁月如梭，转眼之间我的博士研究生生涯即将接近尾声。我在武汉大学五年多的学习和生活也即将落下帷幕，感谢武汉大学对我的培养。回首匆匆逝去的日子，心里难以平静，这段岁月将成为我一生中最宝贵的财富。

在本论文即将完成之际，首先衷心感谢我的导师胡瑞敏教授，感谢胡老师5年来对我的悉心指导和亲切关怀。胡老师以其敏锐的学术眼光指引我进入图像超分辨率处理这一研究领域，并不断指导和鼓励我对这一课题进行研究和创新。本论文是在胡老师的悉心指导下完成，无论是在选题调研阶段、研究提炼阶段，还是在编辑撰写阶段，胡老师都倾注了大量的心血和精力。期间，胡老师严谨的治学态度、卓越的思辨能力、求实的工作作风、以及对理想的孜孜不倦的追求都给我留下了非常深刻的印象。跟随胡老师求学这五年中，无论在学术上，还是在生活中，胡老师都给予了我很大的教诲和帮助，导师的言传身教让我成长，为我的人生留下宝贵的财富，让我终生受益。

感谢国家多媒体软件工程技术研究中心的陈军教授、陈世鸿教授、王中元副教授、杨玉红副教授、怀念副研究员、张晋东高工、涂卫平副教授、高戈副教授、朱荣副教授、艾浩军副教授、韩镇讲师、梁超讲师、王晓晨讲师、肖晶讲师、傅佑铭讲师、陈强国讲师和谭小琼讲师在学习和工作中给予我的指导和帮助，你们渊博的专业知识、严谨的治学态度和谦逊的为人品质给我留下了深深的印象，你们的勉励和关心让我难以忘怀。另外还要感谢工程中心办公室的夏仁凤老师、寇宁老师、朱兵老师、刘莉老师、李瑛老师和袁红老师，谢谢你们在生活中给予我的关照。

感谢课题组的师兄弟、师姐妹以及实验室的所有同学，因为有了你们在科研工作上给予我的大力帮助和支持，才使我能够更加顺利地完成博士论文的写作，和你们一起让我原本单调枯燥的科研生活变得充实、有趣。

最厚重的感谢要献给我的父母和家人，感谢你们在我求学道路上付出的心血和汗水。感谢你们多年来对我生活无微不至的关心和照顾，从小到大，从幼稚到成熟，每一次升学的压力，复习的寂寞，成功的喜悦，都是你们陪我走过，这种爱无以为报。你们的关心和守护给了我努力学习、快乐生活的信心和勇气，成为我人生的坚强后盾和精神动力。感谢我的爱人付吉灿对我的理解和关心，在学习上鼓励我，在生活上照顾我，陪我一起走过困难和挫折。

二十余载的求学生涯，回顾走过的每一步，在此真诚的对每一个帮助过我的人说声谢谢，谢谢你们！

江俊君

2014年冬于珞珈山