

DOI: 10.11992/tis.201804019

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180611.1118.004.html>

计算视觉核心问题: 自然图像先验建模研究综述

孙必慎¹, 石武祯², 姜峰²

(1. 中国电子科技集团公司 第 27 研究所, 河南 郑州 450005; 2. 哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

摘 要: 视觉先验是计算机视觉的核心问题之一, 是认知心理层面、系统神经层面与计算视觉层面研究的交合点, 涉及各个层面研究的理解与综合。视觉先验功能模拟方面, 以自然图像信息为对象, 挖掘自然图像一般性规律并将其数学形式化为可计算的图像模型, 为众多图像处理与计算机视觉智能应用提供算法和支撑。本文对自然图像先验建模研究各流派工作进行了全面的剖析, 并展示了自然图像先验建模工作在视觉信息增强和编码等方向的前瞻性应用。

关键词: 计算机视觉; 图像先验; 稀疏表示; 局部平滑; 非局部自相似; 压缩感知; 深度学习; 卷积神经网络

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2019)01-0071-11

中文引用格式: 孙必慎, 石武祯, 姜峰. 计算视觉核心问题: 自然图像先验建模研究综述 [J]. 智能系统学报, 2019, 14(1): 71-81.

英文引用格式: SUN Bishen, SHI Wuzhen, JIANG Feng. Core problem in computer vision: survey of natural image prior models[J]. CAAI transactions on intelligent systems, 2019, 14(1): 71-81.

Core problem in computer vision: survey of natural image prior models

SUN Bishen¹, SHI Wuzhen², JIANG Feng²

(1. No.27 Institute, China Electronic Technology Corporation (CETC), Zhengzhou 450005, China; 2. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

Abstract: One of the core problems in computer vision is that the visual prior is the point of intersection of the cognitive psychological level, systematic neural level, and computer vision level, and requires an understanding and synthesis of the three. Simulations of the visual prior function are performed to explore and formalize the general rules for natural images that support various applications in image processing and computer science. In this paper, we comprehensively analyze the work of various schools of natural image priori modeling and discuss the prospective application of natural image prior modeling in visual information enhancement and coding.

Keywords: computer vision; image prior; sparse representation; local smoothness; non-local self-similarity; compressed sensing; deep learning; convolutional neural network

视觉先验是认知心理层面、系统神经层面与计算视觉层面研究的交合点, 涉及各个层面研究的理解与综合。根据各领域切入点的不同, 对视觉先验研究主要可以分为两条路线, 即认知心理层面与系统神经层面的机理测试以及计算层面的功能模拟。视觉先验机理测试方面通过非侵入式

的测量仪器, 在自然图像刺激条件下检测人脑神经细胞的响应特性, 反推大脑对自然图像一般性规律进行归纳与记忆的区域结构及其功能原理。随着人类对自身视觉系统的研究逐步深入, 从初级视皮层到高级视觉区域, 都取得了许多重要的研究成果^[1-4]。尽管研究者一致认为初级视皮层的视觉处理过程受环境统计特性的影响, 但怎样在两者之间建立准确的数学关系一直是一个难题。功能性磁共振成像 (functional magnetic reson-

收稿日期: 2018-04-15. 网络出版日期: 2018-06-11.

基金项目: 国家自然科学基金项目 (61572155, 61672188, 61272386); 国家重点基础研究发展计划项目 (2015CB351804).

通信作者: 姜峰. E-mail: fjiang@hit.edu.cn.

ance imaging, fMRI) 只能有限地反映大脑皮层局部区域的系统动力, 必须从理论上进行抽象, 才能将“原理”嵌入到我们的工具中去^[5]。同时, 从可计算的层面上去解释, 是一种合理、必须的选择。

1 传统低层计算视觉层面的自然图像先验模型

视觉先验功能模拟方面, 以自然图像为对象, 挖掘自然图像一般性规律并将其数学形式化为可计算的图像模型, 主要包括自然图像光滑性先验模型、自然图像统计规律先验模型、自然图像视觉编码稀疏性先验模型、自然图像非局部自相似模型等, 并为众多图像处理与计算机视觉智能应用^[6-9]提供算法和支撑。

自然图像光滑性先验建模的方式是约束图像的一阶或高阶导数及其非线性变换, 以形成图像处理的正则化能量模型, 从而构成基于变分方法的偏微分方程。当约束取为梯度模的平方时, 导出经典的热扩散算法; 当约束取为梯度模时, 导出经典的全变分 (total variational)^[10] 方法; 当约束取为梯度的非线性函数时, 导出各向异性扩散算法; 当约束取为高阶导数的模时, 导出高阶各项异性扩散算法^[11]。这些方法着重刻画了图像的局部光滑程度, 忽略了全局的相似块之间的本质关系, 比如自相似性关注于图像像素点的一个多阶邻域的重构与恢复, 抹平了图像的细节, 不能有效地给出光滑先验和数据精度项之间的关系。

自然图像统计规律先验建模的方式是考察自然图像滤波响应率并对其分布进行统计建模。基于 Markov 随机场理论, 特别是通过 Gibbs 分布来建模自然图像滤波响应统计规律的方法, Mumford 等^[12] 提出的 FRAME 模型考虑图像 MRF 建模中的邻域系统和势函数的选取问题, 将传统的以导数滤波器为基础的平滑性先验进行了推广, 将 MRF 模型推向高维的形态。Freeman 等^[13] 提出的 MRF 模型, 被应用于构建图像推理的先验模型以及基于样例的纹理合成中。Welling 等^[14] 提出使用专家乘积模型 (product of experts, PoE) 去解决高维图像数据的先验概率建模问题。在 PoE 的基础上, Roth 等^[15] 设计了均一性势函数形式的专家场模型 (fields of experts, FoE), 弥补了 PoE 模型的缺陷, 可以学习任意尺寸的自然图像的先验, 后续工作 (如文献^[16]) 通过贝叶斯最小均方误差方法取代最大后验概率方法计算恢复图像, 从一定层面克服了到目前为止 MRF 的一些缺点。自然图像统计规律先验建模研究的是自然图

像所含最少成分的统计规律, 工作中心可总结为 2 点: 如何找到自然图像中的最少成分, 以及如何描述一个庞大的自然图像集合在此最小成分方向上的边缘概率分布。目前的方法面临着共有的问题: 统计涨落问题无法规避, 不同图像间以及图像不同区域之间的统计涨落严重地束缚了自然图像统计先验模型的功能; 需要投入大量代价在如何描述其分布响应情况, 描述手段复杂, 影响模型的学习及后续的优化过程, 进一步影响自然图像统计先验模型的实用性能。

自然图像视觉编码稀疏性先验建模源于“有效编码假说”^[17]。现有的稀疏编码方法在图像表示、物体识别等领域的应用受到了较高计算复杂度的约束。不少工作已经开始着手提高稀疏编码的效率和鲁棒性^[18-26]。尽管稀疏概念的一种来源出自于滤波器对视觉信息的滤波响应, 但最终的关注点却主要在图像块的重构与恢复, 并没有很多地考虑整幅图像的响应统计, 无法提供表示一幅完整图像的统计模型, 忽略了相似块之间的本质关系, 比如自相似性, 导致得到的稀疏编码系数不够准确; 同时, 自适应字典学习过程中需要求解一个具有非常高计算复杂度的大规模优化问题; 此外, 在编码测量中, 为了确保信号的线性投影能够保持信号的原始结构, 对投影矩阵的设计及稳定性有严格要求 (如约束等距性)。

自然图像另一个重要的特性是非局部自相似性。Protter 等^[27] 提出非局部均值 (nonlocal means, NLM) 的高效先验模型并将其用于图像超分辨, 假设去模糊得到的图像跟观测到的模糊图像同样具有非局部自相似, 提出了用来去模糊的基于 NLM 的正则项。受 NLM 启发, 基于全图的 NLM 模型^[28] 和基于调序的 NLM 模型^[29] 被提出。受到图拉普拉斯 (graph laplacian) 理论的启发, Gilboa 等^[30] 根据非局部算子定义了变分框架。为了能够取得更好的性能, 自然图像的稀疏性和非局部自相似性通常联合起来。在文献^[31] 中, 最终设计的目标函数由刻画局部稀疏性和非局部自相似性两个正则项组成, 进而获得更高的复原图像质量。在文献^[32] 中, 同时稀疏编码 (simultaneous sparse coding, SSC) 被用来约束图像中相似的图像块在同一个字典下应该具有相似的稀疏分解, 从而使得稀疏编码系数变得更加鲁棒和准确。同理, 利用相同的思路, 将非局部图分别跟 MS 模型、MRF 模型、AR 模型、KR 模型结合, 就会得到非局部 MS 模型^[33]、非局部 MRF 模型^[34]、非局部 AR 模型^[35] 以及非局部 KR 模型。通过分析总结不难发现,

目前以上非局部模型利用图像非局部自相似性都是在原来局部模型的基础之上加入了非局部权重。但通过加权的方式就不可避免地导致结果中出现扰乱或不准确现象。所以寻找一个能够充分利用图像非局部自相似性的模型,而不是仅仅采用加权的方式,在图像处理领域仍然是一个非常重要的挑战。

以上传统的低层计算视觉层面的先验功能建模方法,已广泛用于各种低层计算机视觉问题当中,都取得了一定显著的效果。近三四年,尽管在以上的方法流派上陆续有新方法出现,但几乎都是在显著增加计算复杂程度的代价下,带来极为有限的性能的提升。显然,对于传统的低层计算视觉层面的先验功能建模方法来讲,各种流派和方法都面临着巨大的挑战。对于这些思想的回顾,不难有以下认识。

1) 传统的低层计算视觉层面的先验功能建模方法对于自然图像一般性规律的认识、解释较为片面和孤立,表现为不同的数学形式和理论解释,对于方法与方法之间的内在本质关联,不同先验之间的耦合关系,尽管已有的一些工作研究了不同先验之间的协作方法,但找出一种可以囊括所有孤立规律的数学建模形式几乎是不可能的。以上这些原因,实际上对于这些先验方法和思想的综合运用造成了较大的障碍。

2) 对于不同类别和不同范畴的自然图像,不同的先验模型作用效果也是不同的。这一点在原来的自然图像建模工作中,并没有被客观地重视和研究。从以上对各流派的分析,如果将自然图像空间就信息熵进行划分,各种先验方法和流派在不同的等熵子空间明显会体现出不同效能。不难得出,在面对真实应用或对于更大规模的视觉信息,如视频编码,这个问题会更加严重。

2 基于深度学习重建与增强方法

低层计算视觉层面,第一个开拓性的基于深度学习的图像重建方法是 Dong 等^[36-37]提出的单图像超分辨率网络 SRCNN。这个网络通过模拟基于稀疏表示的单图像超分辨率方法的低分辨率与高分辨率图像间的映射关系,实现低分辨率图像到高分辨率图像的端到端映射。得益于深度网络强大的学习能力,基于深度学习的重建和增强方法在具有较低的计算复杂度的同时,重建和增强效果显著提高。紧随其后,一系列基于深度学习的图像重建和增强方法相继提出。为了降低网络的复杂度并且进一步改进网络的性能, Dong 等^[38]提出快速

版本的 SRCNN。改进版本的网络通过在输出端使用转置卷积层实现图像的上采样,使得低分辨率图像可以直接作为网络的输入,从而降低了计算复杂度。为了克服较深的网络训练过程中收敛困难的问题, Kim 等^[39]提出学习图像的残差来加速网络的收敛,并且通过增加网络深度来增强网络的感受野和非线性,从而改进网络重建的性能。为了训练出更加有效的网络,一些技术如残差学习^[40]、批正则化^[41]、梯度裁剪^[39]等相继被提出,这些方法改善了深度网络的训练过程,使得网络容易收敛。

另外,为了在深度网络中融入传统图像先验信息来改进重建效果,已有一些方法采用正则化求解优化问题的思路,将不同的自然图像先验信息,如局部平滑、非局部自相似和稀疏表示等,用于设计网络的目标函数。Wang 等^[42]使用深度网络来模拟基于稀疏表示方法的求解过程,实现了深度网络和稀疏表示方法的有效结合。Gu 等^[43]也将稀疏编码融入深度网络,提出了一个卷积稀疏编码网络,解决了重叠块间像素的一致性问题。Liang 等^[44]认为直接训练 SRCNN 模型非常耗时,因此,在网络末端引入了 Sobel 算子提取的边缘信息。实验数据表明,引入的边缘先验加速了网络的收敛过程,而且一定程度上改善了重建效果。与此同时,随着对抗式网络(generative adversarial network, GAN)^[45]在监督学习中的兴起, Ledig 等^[46]和 Johnson 等^[47]将 GAN 引入到超分辨率重建问题中,通过共同训练生成网络和对抗网络,大大提升了超分辨率重建性能,重建后的图像在视觉效果上得到显著提升。

这些工作充分说明并验证深度学习对于低层计算机视觉层面先验功能模拟有着极大的潜力,但是这方面的研究总体还处于起步阶段,很多问题尚待解决。比如:深度网络结构中哪些神经元对最终的图像复原效果有效的问题尚没有人关注过;现在的工作发现底层的视觉任务中非残差学习很难训练得到较深的网络,但是已有的基于残差学习的工作也只是简单地将映射目标由原来的目标图像变成估计残差而已,在什么条件下可以做到非残差学习也可以训练一个很深的网络,以及更加有效的残差学习方式都是有待解决的问题;由于图像先验对图像复原效果有很大的帮助,如何有效地融合深度学习技术和自然图像先验来改善图像复原质量也是值得探讨的。

3 基于自然图像先验模型的编码框架

20 世纪 80 年代, ISO、IEC、ITU 三大国际标

准组织开始制定数字视频编码标准,形成了基于预测(包括帧内、帧间、视间)、变换(变换+量化)和熵编码等三大类关键技术的混合编码框架,制定了以MPEG-2为代表的第一代视频编码标准。2003年,MPEG-4 AVC/H.264第二代国际标准出台,实现了压缩效率翻番,中国同期也开发制定了视频编码国际标准AVS。进而为了进一步提高视频编码效率,IEO/IEC和ITU-T推出了HEVC/H.265视频编码标准。随着硬件技术的飞速发展,超高清视频对视频编码技术提出了新需求。代表性的超高清编码方法有四叉树编码单元划分、预测和变换(已被HEVC和AVS2编码标准采用)、基于超分辨率的编码。近年来图像视频编码的研究热点还包括无线网络视频编码以及多视点视频编码等。

在基于自然图像先验模型的编码方面,我们关注近期视觉先验知识同编码框架融合、合理的视觉先验知识介入的途径与方式、以及图像和视频在采集和显示过程中的增强等问题的研究进展。在Stankovic等^[48]提出的视频感知编码方案中,视频帧按关键帧和非关键帧分类,分别以传统和压缩感知方法采样,在保证重构质量的前提下最高节省50%的采样量。为了克服只依赖稀疏性而没有考虑时间域运动的问题,Park等^[49]提出在解码端使用运动估计和补偿的技术。Prades等^[50]提出针对CVS的分布式字典方案:关键帧采用传统技术编解码,非关键帧分块压缩采样及量化。在解码端,由已解码的关键帧获得字典,在字典的辅助下恢复非关键帧。在Chen等^[51]提出的自适应字典的方案中,关键帧提取基于帧的压缩感知观测值,压缩感知帧提取基于块的压缩感知观测值;在解码端,从已经重建的相邻帧及生成的边信息中学习得到字典。帧块的重构都能被表示成最小范数问题,利用可分离逼近算法(SpaRSA)^[52]来稀疏重构。这些方法在构建码本方面都是使用降质图像块作为寻找高质量图像块的索引重建图像,这是一个病态无确定解问题,极易造成误匹配,最终效果很大程度取决于对图像先验知识的掌握与介入程度^[53-55]。

对于图像视频编码,深度学习技术已展现出了强大的生命力。一方面,深度学习技术可以把编解码器进行联合优化,使得编解码器性能达到最优。另一方面,利用深度学习技术可以提供多样化的编解码方法,方便针对不同的任务对图像视频实现智能编解码。根据与现有编码框架相互

兼容的程度,可以将基于深度学习的图像视频编码分为两大类:与现有编码框架兼容的图像视频编码和以深度学习为核心的图像视频编码。

与现有编码框架兼容的图像视频编码方面,Jiang等^[56]提出了兼容传统图像压缩标准的深度学习压缩框架,在编码器前端和解码器后端分别加入卷积神经网络对编解码器联合优化,并给出了前后两个神经网络联合训练的算法,大大提高了编解码器的压缩性能。Zhao等^[57]进一步提出了通过学习一个虚拟编解码器神经网络来近似从原始图像的有效描述图像到后处理压缩图像的投影,这使得训练网络的时候梯度可以从后处理神经网络有效地反向传播到特征描述神经网络。HEVC (high efficiency video coding)^[58]与深度学习相结合的编码框架也开始受到研究者的关注,该类编码框架将深度学习技术引入到HEVC中,是原有编码框架的扩充。目前,深度学习技术在HEVC中的应用涉及帧内预测^[59]、帧间预测^[60]、CU预测模式决策^[61]、变换^[62]、熵编码^[63]、后处理^[64]等技术环节。帧内预测与帧间预测是视频编码框架中最核心的模块之一,预测又是深度学习技术擅长的。对于帧内预测,Li等^[59]使用了全连接神经网络,利用更多的上下文信息来预测当前块的像素值,然而这种方法增加了一种新的预测模式,与原有的编码框架兼容性较差,并且这种全连接神经网络在一定程度上忽略了图像的局部平滑特性,缺少局部感受野的考虑,导致预测的效果不稳定。对于帧间预测,HEVC一旦选择了帧间预测的一种模式,就会完全依据参考帧来预测当前像素块,这种预测模式只是参考了时域信息,从而完全忽略了空域信息的参考价值。Yan等^[60]利用卷积神经网络做分像素插值,在一定程度上提升了帧间预测的准确性,这种通过改进像素插值来提升预测质量的方法忽略了空域上下文,只参考了时域的信息。

以深度学习为核心的图像视频编码主要包括自编码器和循环神经网络(recurrent neural network, RNN)这两种常用的结构。Toderici等^[65]提出了一种基于循环神经网络的图像压缩框架,这是一种渐进式的编码方法;在此基础上,又提出一种全分辨率图像压缩的循环神经网络框架^[66],该框架包含基于RNN的编码器和解码器及一个基于神经网络的熵编码器,并利用了感知误差,在GRU和ResNet的基础上提出了新的深度结构,取得了很好的重建效果。随后,Johnston等^[67]

在之前工作的基础上对网络结构进行了改进。最近, Theis 等^[68]和 Balle 等^[69]分别提出了一种基于神经网络的编解码框架。Theis 等^[68]为了有效地估计编码系数的分布和比特率, 引入了一个高斯多尺度混合模型 (gaussian scale mixture), 实现了有效的码率控制。在 Balle 等^[69]的网络中, 由于受到生物神经元的启发, 采用了联合非线性来实现局部增益的控制。如何有效地量化也是此类问题中一个具有挑战性的问题, Agustsson 等^[70]在连续松弛的量化和熵的基础上, 提出了一种 Soft-to-Hard 量化方法。由于图像局部内容的不相似性, Li 等^[71]通过引入重要性图 (importance map) 提出了一种基于内容自适应的量化方法。受到对抗神经网络的启发, Rippel 等^[72]引入了对抗训练模块, 有效地提升了重建效果。随着 Inpainting 相关技术的发展, 基于 Inpainting 的压缩框架开始受到研究者的关注, Baig 等^[73]利用深度学习技术结合 Inpainting 理念提出一个新的图像压缩框架, 该框架也是基于一种渐进式重建的思想。

4 自然图像先验建模应用示例

本节采用两个例子来展示自然图像先验建模在视觉信息增强和编码领域的前瞻性应用。

4.1 基于多尺度深度网络的单图像超分辨

单图像超分辨是一个经典的却仍然很热门的研究方向。很多单图像超分辨方法已经被提出, 例如非深度学习方法 A+^[74]、SRF^[75], 以及基于深

度学习的方法 SRCNN^[76]、SCN^[77]、FSRCNN^[78]。前人的工作表明充分利用图像的多尺度信息可以有效改善图像超分辨的效果。但是 SRCNN^[76]、SCN^[77]、FSRCNN^[78]都没有考虑到多尺度信息。Shi 等^[79]提出构建基于膨胀卷积的 inception 模块来学习多尺度信息。

膨胀卷积可以定义为

$$(F *_l k)(p) = \sum_{s+t=p} F(s)k(t)$$

膨胀卷积是先对卷积核进行不同尺度的膨胀之后再与输入图像进行卷积操作。具体来说, d 度膨胀卷积就是在进行卷积操作之前先在滤波器核的元素之间插入 $d-1$ 个 0 得到新的滤波器核再与输入图像进行卷积操作。因为插入了多个 0, 所以不同膨胀度的卷积就像对图像进行降采样得到小尺度的图像再进行卷积一样。因此, 不同膨胀度的卷积学习到不同尺度的信息, 然后再将这些不同尺度的信息进行融合就可以达到对多尺度信息的有效利用。

仿照 GoogLeNet 提出的 inception 模块, 可以构建基于膨胀卷积的多尺度学习 inception 模块。图 1 展示了基于膨胀卷积的 inception 模块与 GoogLeNet 提出的 inception 模块的对比。GoogLeNet 的 inception 模块解决的是滤波器尺寸的选择问题, 而基于膨胀卷积的 inception 模块更希望利用不同尺度的输入信息。因此基于膨胀卷积的 inception 模块具有利用多尺度信息的性质。

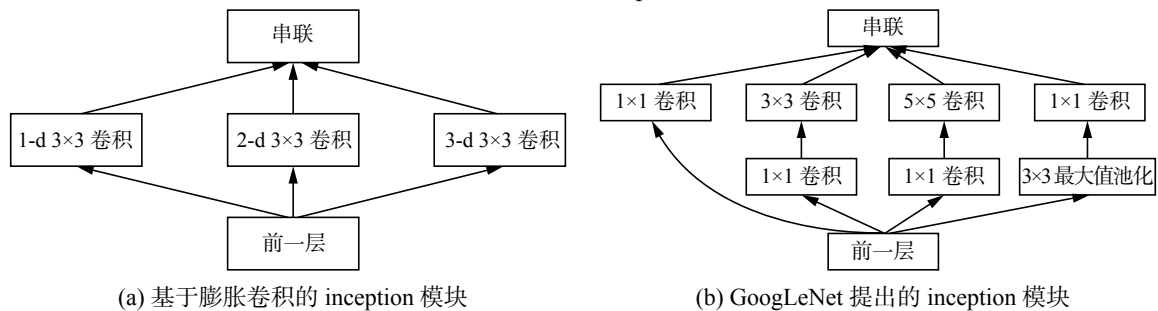


图 1 基于膨胀卷积的起始模块与 GoogLeNet 提出的起始模块的对比

Fig. 1 Comparison between the inception module based on dilated convolution and the inception in GoogLeNet

利用基于膨胀卷积的 inception 模块来学习多尺度信息, 可以构建有效的深度网络 (命名为 MSSRNet) 进行单图像超分辨方法, 如图 2 所示。概括起来, MSSRNet 由若干个基于膨胀卷积的 inception 模块级联构成端到端的网络来学习图像残差。这个网络的特点是: 利用基于膨胀卷积的 inception 模块来学习多尺度信息; 通过级联多个模块来增加网络深度来达到增加网络感受野大小和非线性; 通过学习残差而不是直接学习高分辨图

像来加快网络收敛速度以及改善网络性能。MSSRNet 与多个流行的单图像超分辨方法进行定量比较的结果如表 1 所示。表 1 给出了不同方法在 3 个数据库 (Set5、Set14 和 BSD200) 中 3 种放大倍数 (分别放大 2 倍、3 倍和 4 倍) 平均 PSNR 结果。如表 1 所示, MSSRNet 在 3 个数据库中的 3 种放大倍数都取得最佳的重建效果。关于该方法的更多具体的细节, 请读者参考文献 [79]。

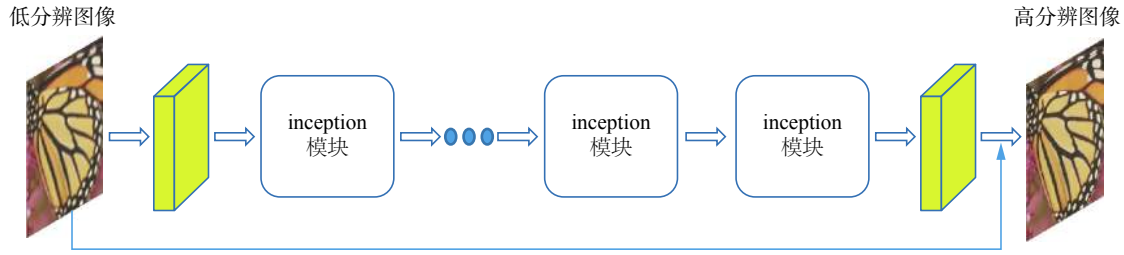


图 2 基于膨胀卷积起始模块的单图像超分辨网络结构

Fig. 2 Single-image super-resolution network structure based on the inception module based on dilated

表 1 不同方法在 3 个标准数据集上 3 种放大倍数的定量比较

Table 1 Quantitative comparison between various methods based on three benchmarks over three magnifications

数据集	放大倍数	双三次插值	A+ ^[74]	SRF ^[75]	SRCNN ^[76]	SCN ^[77]	FSRCNN ^[78]	MSSRNet ^[79]
Set5	2	33.66/0.929 9	36.55/0.954 4	36.87/0.955 6	36.34/0.952 1	36.76/0.954 5	37.00/0.955 8	37.33/0.958 1
	3	30.39/0.929 9	32.59/0.908 8	32.71/0.909 8	32.39/0.903 3	33.04/0.913 6	33.16/0.914 0	33.38/0.917 8
	4	28.42/0.810 4	30.28/0.860 3	30.35/0.860 0	30.09/0.850 3	30.82/0.872 8	30.71/0.865 7	31.10/0.877 7
Set14	2	30.23/0.868 7	32.28/0.905 6	32.51/0.907 4	32.18/0.903 9	32.48/0.906 7	32.63/0.908 8	32.89/0.911 7
	3	27.54/0.773 6	29.13/0.818 8	29.23/0.820 6	29.00/0.814 5	29.37/0.822 6	29.43/0.824 2	29.57/0.828 2
	4	26.00/0.701 9	27.32/0.747 1	27.41/0.749 7	27.20/0.741 3	27.62/0.757 1	27.59/0.753 5	27.83/0.763 1
BSD200	2	29.70/0.862 5	31.44/0.903 1	31.65/0.905 3	31.38/0.928 7	31.63/0.904 8	31.80/0.907 4	32.08/0.911 8
	3	27.26/0.763 8	28.36/0.807 8	28.45/0.809 5	28.28/0.803 8	28.54/0.811 9	28.60/0.813 7	28.78/0.818 8
	4	25.97/0.694 9	26.83/0.735 9	26.89/0.736 8	26.73/0.729 1	27.02/0.743 4	26.98/0.739 8	27.17/0.748 9
平均值		28.80/0.815 1	30.53/0.849 1	30.67/0.850 5	30.39/0.847 4	30.81/0.854 2	30.88/0.853 7	31.13/0.859 6

4.2 基于深度网络的快速准确的压缩采样重建

尽管压缩感知技术有很好的应用前景, 最近几年却发展非常缓慢, 其中一个主要原因是之前的压缩重建算法往往需要一个非常复杂的解优化过程, 重建速度非常慢而影响其实用性。如图 3 所示, 传统的方法 DWT^[80]、TV^[81]、MH^[82]、CoS^[83] 和 GSR^[84] 运行速度非常慢, 因此研究快速的压缩采样和重建算法显得非常必要。经过这么多年的发展, 采样矩阵的设计以及快速压缩重建算法依然是压缩感知技术的两个核心问题。Shi 等^[85] 提出使用深度网络 (CSNet) 来解决这两个问题, 也就是用深度网络来学习一个高效的采样算子, 以及设计一个快速却重建效果好的网络实现图像快速端到端重建。

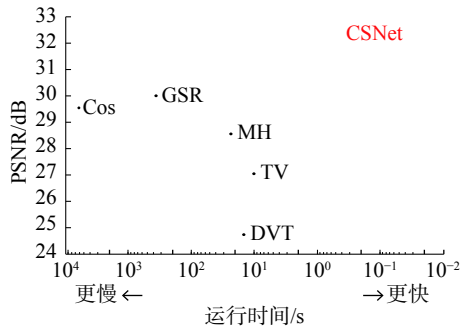


图 3 基于深度学习的压缩感知网络框架与传统的基于块的压缩感知方法的对比

Fig. 3 Comparison between compressed sensing framework based on deep learning and the ones based on traditional blocks

图 4 给出了基于深度网络的压缩感知方法与传统的 BCS-SPL (block based compressed sampling and smoothed projected landweber) 方法的对比。如图所示, BCS-SPL 算法包括 3 个过程: 压缩采样、初始重构和非线性信号重构。

在 BCS 中, 一个图像会被裁切为 $B \times B$ 大小的图像块, 然后用一个合适大小的测量矩阵进行采样。如果采样率为 M/N , 每个块需要 $n_B = \lfloor M/N \cdot B^2 \rfloor$ 个采样值。因此采样矩阵 Φ_B 是一个 $n_B \times B^2$ 的矩阵。假设 \mathbf{x}_j 是第 j 块的向量, 则相应的测量值可以表示为 $\mathbf{y}_j = \Phi_B \mathbf{x}_j$ 。这就实现了图像的压缩采样过程。另一个需要解决的问题是, 在给定采样值的情况下, 怎么将其准确恢复出原来的图像。原来的 BCS-SPL 方法使用最小均方误差线性估计来估计初始解。这个初始重构过程可以表示为

$$\tilde{\mathbf{x}}_j = \tilde{\Phi}_B \mathbf{y}_j$$

$$\tilde{\Phi}_B = R_{xx} \Phi_B^T (\Phi_B R_{xx} \Phi_B^T)^{-1}$$

式中 R_{xx} 是输入图像信号的自相关函数。明显地, $\tilde{\Phi}_B$ 是一个 $B^2 \times n_B$ 的矩阵。在得到初始解后, BCS-SPL 方法往往会进行多阶段的平滑投影操作进行非线性重构来进一步改善图像的质量。

CSNet 模拟传统的基于分块的压缩感知采样和重建过程。图 4 的上部是一个传统的 BCS-SPL 方法的采样和重建框架, 下部是 CSNet 的网络框架。如图 4 所示, 传统的 BCS-SPL 方法包括压缩采样、初始重构和非线性信号重构 3 个过程。

CSNet 设计一个采样子网络、一个初始重构子网络和一个深度重构子网络来模拟传统 BCS-SPL 的 3 个过程。采样子网络用一个卷积层模拟传统的采样矩阵 Φ 。假设块划分大小为 $B \times B$, 采样率为 M/N , 那么 Φ 是一个 $n_B = \lfloor M/N \cdot B^2 \rfloor$ 行 B^2 列的矩阵。那么采样子网络将会是一个有 $n_B = \lfloor M/N \cdot B^2 \rfloor$ 个 $B \times B$ 大小滤波器的卷积层。初始重构子网络

包括一个卷积层和一个合并重构层。初始重构子网络的卷积层有 B^2 个空间维度为 1×1 大小滤波器, 而合并重构层是一个拼接和串联操作。深度重构子网包括多个卷积操作。这个方法很好地模拟了传统的压缩重构过程, 并且将复杂的解优化过程隐含于简单的深度重构子网中, 可以得到更好的重构质量和更快的运行速度。

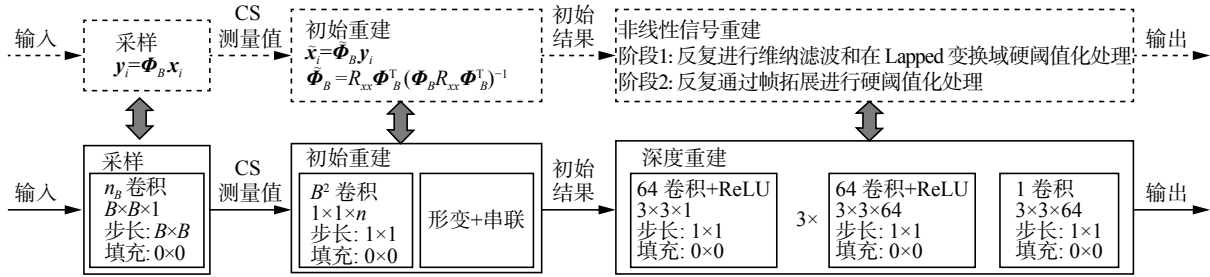


图 4 CSNet 与流行算法的运行速度和重构 PSNR 的对比

Fig. 4 Comparison of the running speed and PSNR between CSNet and the state-of-the-art methods

表 2 给出基于深度学习的压缩感知方法 (CSNet) 与其他流行的压缩感知方法在 Set14 数据集中 5 种采样率下的重建结果的平均 PSNR 和 SSIM 结果的对比。如表 2 所示, CSNet 在各种采样率下与现有的方法相比都取得了显著的重建效

果提升。图 3 给出了 CSNet 与流行算法在 Set5 数据集上采样率为 0.1 的重建图像的 PSNR 与运行速度的对比。结果显示, CSNet 不仅重建效果好运行速度也显著提升。更多关于 CSNet 的细节, 请读者参考文献 [75]。

表 2 不同算法在 Set14 数据集中 5 种压缩采样率下的重建结果的平均 PSNR 和 SSIM 对比

Table 2 Averages of PSNR and SSIM over 5 sampling rates between various methods on Set14

采样率	DWT ^[80]	TV ^[81]	MH ^[82]	CoS ^[83]	GSR ^[84]	CSNet ^[85]
0.1	24.16/0.679 8	25.24/0.688 7	26.38/0.728 2	27.20/0.743 3	27.50/0.770 5	29.13/0.816 9
0.2	28.13/0.788 2	28.07/0.784 4	29.47/0.823 7	30.07/0.827 8	31.22/0.864 2	32.15/0.894 1
0.3	30.38/0.838 9	30.12/0.842 4	31.37/0.869 4	32.03/0.873 2	33.74/0.907 1	34.34/0.929 7
0.4	31.99/0.875 3	32.03/0.883 7	33.03/0.900 9	34.00/0.908 4	35.78/0.933 6	36.16/0.950 2
0.5	33.54/0.904 4	33.84/0.914 8	34.52/0.923 9	35.84/0.931 4	37.66/0.952 2	37.89/0.963 1
平均值	29.64/0.817 3	29.86/0.822 8	30.95/0.849 2	31.83/0.856 8	33.18/0.885 5	33.93/0.910 8

5 结束语

综上所述, 从理论模型到应用方法, 纵观各层面的根本任务、彼此联结及发展现状, 有如下认识: 在低层计算机视觉层面, 深度结构可以从丰富的感知信息中归纳或解释复杂的结构和建立数据中内在的表征, 特别是为视觉信息的表达提供了高效的手段与工具; 与传统自然图像先验功能模拟相比, 以数据为根本驱动, 以精确描述和解释自然图像子空间为目标的自然图像先验深度模型, 可先天克服传统先验建模方法片面、孤立的缺点, 这使得我们跳出以往较为单一先验建模方法, 在更高的层面思考问题。

借助计算层面的自然图像先验模型是编码技术的必然选择。先验知识在自然图像认知过程中

起着至关重要的作用。基于数字信号处理的传统编码技术并没有直接从人类的视觉感知过程出发, 而压缩图像最终要服务于人的感知; 尽管认知心理层面的视觉先验机理测试给出一些重要的发现与启示, 但是仅仅依靠其去解释大脑对自然图像一般性规律的归纳与认知是不现实的。因此从计算层面的视觉先验功能模拟角度去解释, 才能够将视觉信息编码从图像空间过渡到自然图像子空间, 这是一种合理、必须的选择。

参考文献:

- [1] PARK S, KONKLE T, OLIVA A. Parametric coding of the size and clutter of natural scenes in the human brain[J]. *Cerebral cortex*, 2015, 25(7): 1792–1805.
- [2] FÖLDIÁK P. Neural control: closed-loop human brain

- reading[J]. *Current biology*, 2011, 21(2): 80–81.
- [3] KAY K N, GALLANT J L. I can see what you see[J]. *Nature neuroscience*, 2009, 12(3): 245.
- [4] KIM M K, KIM M, OH E, et al. A review on the computational methods for emotional state estimation from the human EEG[J]. *Computational and mathematical methods in medicine*, 2013, 22(12): 3127–3142.
- [5] JIAN Muwei, QI Qiang, DONG Junyu, et al. Saliency detection using quaternionic distance based weber local descriptor and level priors[J]. *Multimedia tools and applications*, 2018, 77(11): 14343–14360.
- [6] JIAN Muwei, LAM K M, DONG Junyu, et al. Visual-patch-attention-aware saliency detection[J]. *IEEE transactions on cybernetics*, 2015, 45(8): 1575–1586.
- [7] JIAN Muwei, LAM K M, DONG Junyu. Facial-feature detection and localization based on a hierarchical scheme[J]. *Information sciences*, 2014, 262(3): 1–14.
- [8] JIAN Muwei, QI Qiang, DONG Junyu, et al. Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection[J]. *Journal of visual communication and image representation*, 2018, 53: 31–41.
- [9] TREVES A, TASHIRO A, WITTER M P, et al. What is the mammalian dentate gyrus good for?[J]. *Neuroscience*, 2008, 154(4): 1155–1172.
- [10] CHANTAS G, GALATSANOS N P, MOLINA R, et al. Variational bayesian image restoration with a product of spatially weighted total variation image priors[J]. *IEEE transactions on image processing*, 2010, 19(2): 351–362.
- [11] CHAN T F, OSHER S, SHEN J. The digital TV filter and nonlinear denoising[J]. *IEEE transactions on image processing*, 2001, 10(2): 231–241.
- [12] ZHU Songchun, WU Yingnian, MUMFORD D. Filters, random fields and maximum entropy (FRAME): towards a unified theory for texture modeling[J]. *International journal of computer vision*, 1998, 27(2): 107–126.
- [13] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[J]. *International journal of computer vision*, 2000, 40(1): 25–47.
- [14] WELLING M, OSINDERO S, HINTON G E. Learning sparse topographic representations with products of student-t distributions[C]//Advances in Neural Information Processing Systems. Vancouver, Canada, 2003: 1383–1390.
- [15] ROTH S, BLACK M J. Fields of experts[J]. *International journal of computer vision*, 2009, 82: 205.
- [16] SUN Deqing, ROTH S, BLACK M J. A quantitative analysis of current practices in optical flow estimation and the principles behind them[J]. *International journal of computer vision*, 2014, 106(2): 115–137.
- [17] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. *The journal of physiology*, 1962, 160(1): 106–154.
- [18] XU Mai, LI Shengxi, LU Jianhua, et al. Compressibility constrained sparse representation with learnt dictionary for low bit-rate image compression[J]. *IEEE transactions on circuits and systems for video technology*, 2014, 24(10): 1743–1757.
- [19] SUN Yipeng, TAO Xiaoming, LI Yang, et al. Dictionary learning for image coding based on multisample sparse representation[J]. *IEEE transactions on circuits and systems for video technology*, 2014, 24(11): 2004–2010.
- [20] SUN Fuming, TANG Jinhui, LI Haojie, et al. Multi-label image categorization with sparse factor representation[J]. *IEEE transactions on image processing*, 2014, 23(3): 1028–1037.
- [21] RABBANI H, GAZOR S. Local probability distribution of natural signals in sparse domains[J]. *International journal of adaptive control and signal processing*, 2014, 28(1): 52–62.
- [22] ELHAMIFAR E, VIDAL R. Sparse subspace clustering: Algorithm, theory, and applications[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(11): 2765–2781.
- [23] 肖迪, 马青青, 王兰, 等. 基于稀疏表示的云协助安全数字水印技术 [J]. *信息网络安全*, 2017, 37(1): 1–7.
- XIAO Di, MA Qingqing, WANG Lan, et al. Cloud-assisted secure digital watermarking based on sparse representation[J]. *Netinfo security*, 2017, 37(1): 1–7.
- [24] 熊金波, 马蓉, 张媛媛, 等. 面向社交网络的图片信息隐藏方法与实现 [J]. *信息网络安全*, 2017, 37(3): 6–13.
- XIONG Jinbo, MA Rong, ZHANG Yuanyuan, et al. Image information hiding method and implementation for social network[J]. *Netinfo security*, 2017, 37(3): 6–13.
- [25] 徐杰, 贺敏, 包秀国. 基于压缩感知的视频台标识别研究 [J]. *信息网络安全*, 2015(5): 77–81.
- XU Jie, HE Min, BAO Xiuguo. Research on TV logo detection based on compressive sensing[J]. *Netinfo security*, 2015(5): 77–81.
- [26] 任栋, 宋伟, 于京, 等. 特殊视频内容检测算法研究综述 [J]. *信息网络安全*, 2016(9): 184–191.
- REN Dong, SONG Wei, YU Jing, et al. A survey on special video content detection algorithms[J]. *Netinfo security*, 2016(9): 184–191.
- [27] PROTTER M, ELAD M, TAKEDA H, et al. Generalizing the nonlocal-means to super-resolution reconstruction[J]. *IEEE transactions on image processing*, 2009,

- 18(1): 36–51.
- [28] TALEBI H, MILANFAR P. Global image denoising[J]. *IEEE transactions on image processing*, 2014, 23(2): 755–768.
- [29] RAM I, ELAD M, COHEN I. Image processing using smooth ordering of its patches[J]. *IEEE transactions on image processing*, 2013, 22(7): 2764–2774.
- [30] GILBOA G, OSHER S. Nonlocal operators with applications to image processing[J]. *Multiscale modeling and simulation*, 2008, 7(3): 1005–1028.
- [31] ZHANG Lei, DONG Weisheng, ZHANG D, et al. Two-stage image denoising by principal component analysis with local pixel grouping[J]. *Pattern recognition*, 2010, 43(4): 1531–1549.
- [32] MAIRAL J, BACH F, PONCE J, et al. Non-local sparse models for image restoration[C]//Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan, 2009: 2272–2279.
- [33] JUNG M, BRESSON X, CHAN T F, et al. Nonlocal Mumford-Shah regularizers for color image restoration[J]. *IEEE transactions on image processing*, 2011, 20(6): 1583–1598.
- [34] SUN Jian, TAPPEN M F. Learning non-local range Markov random field for image restoration[C]//CVPR 2011. Colorado Springs, CO, USA, 2011: 2745–2752.
- [35] GAO Xinwei, ZHANG Jian, JIANG Feng, et al. Image interpolation via local autoregressive and nonlocal 3-D sparse regularization[C]//Visual Communications and Image Processing. San Diego, USA, 2012.
- [36] DONG Chao, LOY C C, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution[C]//Proceedings of the 13th European Conference On Computer Vision. Cham, Germany, 2014: 184–199.
- [37] DONG Chao, LOY C C, HE Kaiming, et al. Image super-resolution using deep convolutional networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2016, 38(2): 295–307.
- [38] DONG Chao, LOY C C, TANG Xiaoou. Accelerating the super-resolution convolutional neural network[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 391–407.
- [39] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 1646–1654.
- [40] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, United States, 2016: 770–778.
- [41] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification[C]//Proceedings of IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1026–1034.
- [42] WANG Zhaowen, LIU Ding, YANG Jianchao, et al. Deep networks for image super-resolution with sparse prior[C]//Proceedings of IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 370–378.
- [43] GU S, ZUO W, XIE Q, et al. Convolutional sparse coding for image super-resolution[C]// IEEE International Conference on Computer Vision. IEEE Computer Society, 2015: 1823–1831.
- [44] LIANG Yudong, WANG Jinjun, ZHOU Sanping, et al. Incorporating image priors with deep convolutional neural networks for image super-resolution[J]. *Neurocomputing*, 2016, 194: 340–347.
- [45] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA, USA, 2014: 2672–2680.
- [46] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, 2017: 105–114.
- [47] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 694–711.
- [48] STANKOVIĆ V, STANKOVIĆ L, CHENG S. Compressive video sampling[C]//Processing of 16th European Signal Processing Conference. Lausanne, Switzerland, 2008: 1–5.
- [49] PARK J Y, WAKIN M B. A multiscale framework for compressive sensing of video[C]//Processing of Picture Coding Symposium. Chicago, USA, 2009: 1–4.
- [50] PRADES-NEBOT J, MA Yi, HUANG T. Distributed video coding using compressive sampling[C]//Processing of Picture Coding Symposium. Chicago, USA, 2009: 1–4.
- [51] CHEN H W, KANG Liwei, LU C S. Dictionary learning-based distributed compressive video sensing[C]// Processing of Picture Coding Symposium. Nagoya, Japan, 2010: 210–213.

- [52] WRIGHT S J, NOWAK R D, FIGUEIREDO M A T. Sparse reconstruction by separable approximation[J]. *IEEE transactions on signal processing*, 2009, 57(7): 2479–2493.
- [53] PONG K H, LAM K M. Multi-resolution feature fusion for face recognition[J]. *Pattern recognition*, 2014, 47(2): 556–567.
- [54] PYATYKH S, HESSER J. Salt and pepper noise removal in binary images using image block prior probabilities[J]. *Journal of visual communication and image representation*, 2014, 25(5): 748–754.
- [55] YIN Zhouping, ZHANG Hongmei. Research on saliency prior based image processing algorithm[J]. *Journal of multimedia*, 2014, 9(2): 294–301.
- [56] JIANG Feng, TAO Wen, LIU Shaohui, et al. An end-to-end compression framework based on convolutional neural networks[J]. *IEEE transactions on circuits and systems for video technology*, 2018, 28(10): 3007–3018.
- [57] ZHAO Lijun, BAI Huihui, WANG Aahong, et al. Learning a virtual codec based on deep convolutional neural network to compress image[EB/OL]. arXiv preprint arXiv: 1712.059 69, 2017. <https://arxiv.org/abs/1712.05969>
- [58] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. *IEEE transactions on circuits and systems for video technology*, 2012, 22(12): 1649–1668.
- [59] LI Jiahao, LI Bin, XU Jizheng, et al. Intra prediction using fully connected network for video coding[C]//IEEE International Conference on Image Processing. Beijing, China, 2017: 1–5.
- [60] YAN Ning, LIU Dong, LI Houqiang, et al. A convolutional neural network approach for half-pel interpolation in video coding[C]//Proceedings of 2017 International Symposium on Circuits and Systems. Baltimore, MD, USA, 2017: 1–4.
- [61] LIU Zhenyu, YU Xianyu, CHEN Shaolin, et al. CNN oriented fast HEVC intra CU mode decision[C]//IEEE International Symposium on Circuits and Systems. Montreal, QC, Canada, 2016: 2270–2273.
- [62] LIU Dong, MA Haichuan, XIONG Zhiwei, et al. CNN-Based DCT-like transform for image compression[C]//Proceedings of the 24th International Conference on Multimedia Modeling. Bangkok, Thailand, 2018: 61–72.
- [63] SONG Rui, LIU Dong, LI Houqiang, et al. Neural network-based arithmetic coding of intra prediction modes in HEVC[C]//IEEE Visual Communications and Image Processing Conference. Petersburg, FL, USA, 2017: 1–4.
- [64] DAI Yuanying, LIU Dong, WU Feng. A convolutional neural network approach for post-processing in HEVC intra coding[C]//Proceedings of the 23rd International Conference on Multimedia Modeling. Reykjavik, Iceland, 2017: 28–39.
- [65] TODERICI G, O' MALLEY S M, HWANG S J, et al. Variable rate image compression with recurrent neural networks[EB/OL]. arXiv preprint arXiv: 1511.060 85, 2016. <https://arxiv.org/abs/1511.06085>
- [66] TODERICI G, VINCENT D, JOHNSTON N, et al. Full resolution image compression with recurrent neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii, USA, 2017: 5435–5443.
- [67] JOHNSTON N, VINCENT D, MINNEN D, et al. Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks[EB/OL]. arXiv preprint arXiv: 1703.101 14, 2017. <https://arxiv.org/abs/1703.10114>
- [68] THEIS L, SHI Wenzhe, CUNNINGHAM A, et al. Lossy image compression with compressive autoencoders [EB/OL]. arXiv preprint arXiv: 1703.003 95, 2017. <https://arxiv.org/abs/1703.00395>
- [69] BALLÉ J, LAPARRA V, SIMONCELLI E P. End-to-end optimization of nonlinear transform codes for perceptual quality[C]//Proceedings of 2016 Picture Coding Symposium. Nuremberg, Germany, 2016: 1–5.
- [70] AGUSTSSON E, MENTZER F, TSCHANNEN M, et al. Soft-to-hard vector quantization for end-to-end learning compressible representations[C]//Advances in Neural Information Processing Systems. Long Beach, CA, USA, 2017: 1141–1151.
- [71] LI Mu, ZUO Wangmeng, GU Shuhang, et al. Learning convolutional networks for content-weighted image compression[EB/OL]. arXiv preprint arXiv: 1703.105 53, 2017. <https://arxiv.org/abs/1703.10553>
- [72] RIPPEL O, BOURDEV L. Real-time adaptive image compression[EB/OL]. arXiv preprint arXiv: 1705.058 23, 2017. <https://arxiv.org/abs/1705.05823>
- [73] BAIG M H, KOLTUN V, TORRESANI L. Learning to Inpaint for Image Compression[C]//Advances in Neural Information Processing Systems. Long Beach, CA, USA, 2017: 1246–1255.
- [74] TIMOFTE R, DE SMET V, VAN GOOL L. A+: adjusted anchored neighborhood regression for fast super-resolution[C]//Asian Conference on Computer Vision. Springer, Cham, 2014: 111–126.
- [75] SCHULTER S, LEISTNER C, BISCHOF H. Fast and accurate image upscaling with super-resolution forests

- [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CV-PR2015. Boston, Massachusetts, USA, 2015: 3791–3799.
- [76] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(2): 295–307.
- [77] WANG Z, LIU D, YANG J, et al. Deeply improved sparse coding for image super-resolution. arXiv preprint [EB/OL]. arXiv preprint arXiv: 1507.08905, 2015. <https://arxiv.org/abs/1507.08905v1>
- [78] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]//European Conference on Computer Vision. Springer, Cham, 2016: 391–407.
- [79] SHI W, JIANG F, ZHANG F, et al. Single image super-resolution with dilated convolution based multi-scale information learning inception module[C]// IEEE International Conference on in Image Processing (ICIP2017). Beijing, China, 2017:1–5.
- [80] MUN S, FOWLER J E. Block compressed sensing of images using directional transforms[C]// 2009 16th IEEE International Conference on Image Processing IEEE, 2009: 3021–3024.
- [81] LI, C, YIN W, ZHANG Y. TVAL3: TV minimization by augmented lagrangian and alternating direction algorithm [EB/OL]. <https://www.caam.rice.edu/~optimization/L1/TVAL3/>
- [82] CHEN C, TRAMEL E W, FOWLER J E. Compressed-sensing recovery of images and video using multihypothesis predictions[C]// 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers. IEEE, 2011: 1193–1198.
- [83] ZHANG J, ZHAO D, ZHAO C, et al. Compressed sensing recovery via collaborative sparsity[C]//Processing of 2012 Data Compression Conference. Snowbird, Utah, USA, 2012: 287–296.
- [84] ZHANG J, ZHAO D, GAO W. Group-based sparse representation for image restoration[J]. IEEE transactions on image processing, 2014, 23(8): 3336–3351.
- [85] SHI W, JIANG F, ZHANG S, et al. Deep networks for compressed image sensing[C]// Processing of 2017 IEEE International Conference on Multimedia and Expo. Hongkong, China, 2017: 877–882.

作者简介:



孙必慎, 男, 1976 年生, 高级工程师, 主要研究方向为光电火控、侦察、预警及测量技术研究。主持多个国家重点型号装备研制, 多次获得国家、部级技术进步奖。发表学术论文 50 余篇。



石武祯, 男, 1989 年生, 博士研究生, 主要研究方向为图像处理、计算机视觉以及图像编解码。发表学术论文 10 余篇。



姜峰, 男, 1978 年生, 教授, 博士生导师, 主要研究方向为图像处理、计算机视觉以及人工智能。主持项目 20 余项, 多次获得省部级技术进步奖。发表学术论文 100 余篇。