

基于残差学习的多阶段图像压缩感知神经网络

杨春玲 裴翰奇

(华南理工大学 电子与信息学院, 广东 广州 510640)

摘 要: 在传统的图像压缩感知研究中存在两个主要问题: 在采样端, 传统的线性采样方式存在一定程度的局限性; 在重构端, 传统的优化重构算法有着高耗时的缺陷。新发展的图像压缩感知神经网络框架虽然很好地解决了重构速度问题, 但重构效果一般。为了解决上述问题, 文中提出一种基于残差学习的多阶段图像压缩感知网络——MSRes-ICS, 其中具体包含 3 个子网络: 采样子网络, 初始重构子网络与图像增强子网络。在观测端, 提出的非线性采样子网络, 利用残差学习打破了传统采样方法的局限性, 在采样值中保留了更加丰富的图像信息; 在重构端, 通过引入插值卷积, 初始重构子网络从采样值中进行特征整合以获取质量优良的初始重构图像, 最终利用残差学习与插值卷积, 提出多阶段图像增强子网络以进一步细化重建图像, 提升最终效果。大量仿真实验结果表明, 与现有最优的图像压缩感知重构算法相比, MSResICS 拥有更加优良的图像压缩感知重构精度。

关键词: 图像压缩感知; 残差学习; 深度网络; 采样机制

中图分类号: TN919.81

文章编号: 1000-565X(2020)05-0082-10

奈奎斯特-香农采样定理中指出: 为了无失真的恢复原始信号, 采样信号的频率不能低于原始信号的两倍带宽。2006 年, Donoho 等^[1-3]提出的压缩感知理论 (CS) 证明了当原始信号具备稀疏特性时, 采样频率可以远低于奈奎斯特定理的要求。在压缩感知中, 采样和压缩的过程同步完成, 有效解决了传统采样中先采样后压缩的策略所造成的资源浪费。凭借以上优势, 压缩感知理论在学术界和工业界引起了广泛的关注和研究。

在图像压缩感知中, 随机观测矩阵由于符合等距性限制条件 (RIP)^[2], 且具有较好的重构效果, 被广泛应用在图像的采样中。同时, 分块压缩感知理论 (BCS)^[4-5] 因其在存储与计算方面的优势而被广泛用于图像压缩感知中。在这个框架下, 图像被分成多个不重叠的图像块单独采样与重构, 以减小

整体的计算量。但是, 不容忽视的一点是, 分块压缩感知框架引起了不良的块效应: 在恢复图像中, 由于不同分块之间相对独立的观测与重建, 在各个块之间出现了大量不连续的边界。块效应降低了重构图像的视觉质量, 特别是在低采样率下, 这个问题尤为严重。

图像压缩感知中的另外一个核心问题是图像的重构算法, 这部分在很大程度上决定了最终恢复结果的优劣。基于图像的结构特性, 文献 [5] 中提出的平滑投影 Landweber 算法 (SPL) 将自然图像的恢复结果提升到了一个较高的水平; 通过利用不同图像块之间的相似信息, 基于组稀疏表示的重构算法 (GSR)^[6] 获得了目前最好的图像重构性能, 并且对图像中的噪声有着一定的抑制效果。但是, 由于需要大量的迭代, 这些恢复算法都有着高

收稿日期: 2019-05-30

基金项目: 广东省自然科学基金资助项目 (2017A030311028, 2016A030313455)

Foundation items: Supported by the Natural Science Foundation of Guangdong Province (2017A030311028, 2016A030313455)

作者简介: 杨春玲 (1970-), 女, 教授, 主要从事图像/视频压缩编码、图像质量评价研究。E-mail: eeclyang@scut.edu.cn

耗时的缺点,使得图像压缩感知难以实际应用。

最近,基于卷积神经网络的图像压缩感知框架的出现,为图像压缩感知的研究和应用带来了新的思路。

2012年,文献[7]中提出的 AlexNet 展现出了卷积神经网络(CNN)的性能潜力,并引起了一次新的研究热潮。过去的几年中,深度学习(DL)领域的研究得到了进一步的发展,在多种高阶语义下的计算机视觉任务中取得了大量的研究成果,如对象检测、多目标跟踪和动作识别等。同时,基于卷积神经网络的低阶语义计算机视觉任务的研究也引起了关注,例如在图像超分辨率重建(SR)^[8-12]与图像去雾算法(Image Dehazing)^[13]等病态逆问题的研究中,深度学习均取得了不错的结果,展现出其性能的优越。

近年出现的一些基于深度学习的图像压缩感知(ICS)算法的研究,表现出了深度学习强大的学习能力和极快的运行速度。Kulkarni等^[14]首次提出基于卷积神经网络的图像压缩感知重构算法,取得了不错的重构效果。文献[15]中提出在重构神经网络中引入 Resnet 的结构,从而在图像压缩感知的恢复过程中学习更加丰富的图像特征。Zhang等^[16]设计了一种新颖的网络结构,能够较好的模仿传统图像压缩感知中的迭代收缩重构算法。受到基于稀疏表示的图像压缩感知重构算法启发,文献[17]中 ConvCSNet 尝试将稀疏性嵌入神经网络的设计之中。文献[18]中提出一种基于深度学习的多尺度残差网络结构来学习图像中的多尺度信息,增强了重构算法对噪声的鲁棒性。文献[19]中提出一种端到端的神经网络结构,成功复现了传统图像压缩感知中的经典重构算法——平滑投影 Landweber 算法。基于文献[19]的工作,Cui等^[20]提出了压缩感知观测矩阵的神经网络学习方法,弥补了传统观测矩阵设计过程中的不足。总的来说,与传统图像压缩感知相比,基于神经网络的压缩感知方法显著地降低了重构算法的时间成本,但仍未解决重构图像中的块效应问题,并且重构端网络结构缺乏针对性,限制了最终的重构性能。

为了解决上述问题,文中提出了一种针对图像压缩感知的全新神经网络结构——MSResICS:在采样端,放弃了传统的分块线性映射采样框架,与之对应的是一个基于残差学习的全卷积图像采样子网络,相比之下,该子网络结构不仅避免了分块压缩感知带来的块效应,而且成功地在采样部分引入

非线性的特征提取方法,保留了更加丰富的图像信息;在恢复端,首先引入一个基于插值卷积的初始重构子网络,从采样值中整合特征并恢复初始重构图像,其次为了进一步提升重构图像的效果,在前两个子网络的基础上,加入多阶段图像增强子网络,以生成多幅残差图像,对恢复结果做进一步的细化;最后通过实验仿真来分析 MSResICS 的具体性能。

1 图像压缩感知基本原理

在图像压缩感知中,有两个主要问题:1) 图像的高效采样;2) 图像的高质量重构。在经典的分块压缩感知中,观测值由观测矩阵的线性映射得到:

$$y_b = \Phi_b x_b \quad (1)$$

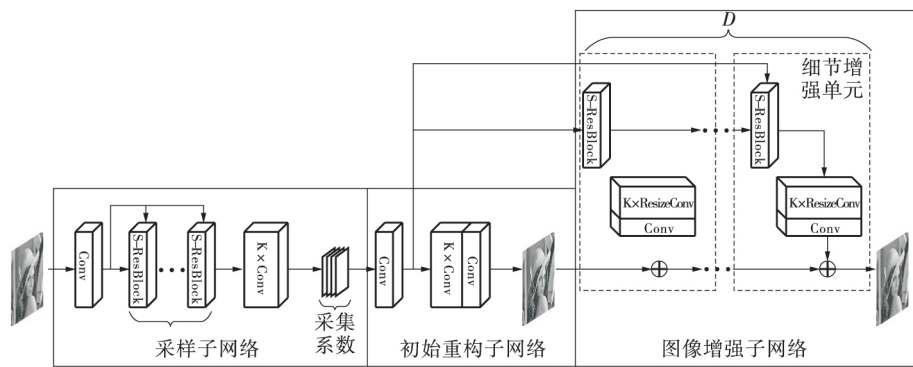
式中, x_b 为 $N \times 1$ 维的原始图像块列矢量, y_b 为 $M \times 1$ 维的分块采样信号, Φ_b 为 $M \times N$ 维的随机观测矩阵。在压缩感知中,由于 $M \ll N$,从采样信号 y_b 中重构原始信号 x_b 是一项病态逆问题。但压缩感知原理指出,如果目标信号 x_b 在变换域 Ψ 是稀疏的,可以通过求解以下优化问题来重构 x_b :

$$\tilde{x}_b = \underset{x}{\operatorname{argmin}} \|\Phi_b x_b - y_b\|_2^2 + \lambda \|\Psi x_b\|_1 \quad (2)$$

式中, λ 为正则化参数。基于图像信号的结构特性,许多文献提出了不同的优化算法来求解式(2),并取得了不错的效果,例如平滑投影 Landweber 算法^[5]与基于组稀疏表示的重构算法^[6]。但传统的图像压缩感知算法计算量巨大,难以用于实际应用。在图像压缩感知神经网络的研究工作中,则是基于深度学习方法对采样系数与原始图像间的映射关系进行学习。这种方法主要分为两类:1) 以 ReconNet 为代表,应用于图像压缩感知重构的卷积神经网络;2) 以 CSNet 为代表,同时应用于图像压缩感知观测域重构的卷积神经网络。以上提到的两种类型的神经网络结构解决了传统算法中高耗时的缺陷,并取得了不错的恢复效果,但同时也存在着一定的缺陷:1) 分块压缩感知框架引起的块效应降低了最终图像的恢复效果;2) 缺少针对性的恢复端网络结构限制了网络的重构性能。

2 实现图像压缩感知的多阶段残差学习神经网络

在本节中,将详细讨论本文提出的图像压缩感知神经网络(MSResICS)。MSResICS 由3部分组成:采样子网络,初始重构子网络与图像增强子网络,具体网络细节参见图1。



Conv 代表卷积层, S-ResBlock 代表简化残差学习块, ResizeConv 代表插值卷积层

图1 MSResICS 的网络结构

Fig. 1 Structure of proposed MSResICS

2.1 基于残差学习的采样子网络

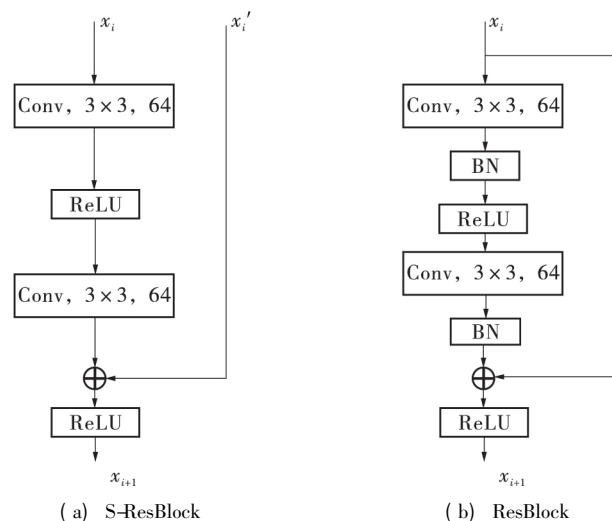
为了解决分块压缩感知框架下带来的块效应问题, MSResICS 放弃了传统的分块线性映射采样方法, 提出一种直接对整幅图像采样的全卷积采样子网络。受文献 [10] 与 [11] 启发, 文中利用了一种简化残差学习块 (S-ResBlock), 以使观测值中包含更加丰富的图像信息, 且网络参数数据量小, 占用内存更少, 如图 1 所示。具体结构参考图 2, S-ResBlock 和原始 ResBlock^[21] 相比主要存在两点不同: 1) 删除了批量标准化层 (BN), 可以更大程度上保留丰富的图像信息, 并减少了网络整体计算量; 2) 恒等映射分支的输入不再为固定输入, 选择更加灵活, 保证网络输出的观测值包含不同语义维度下的图像信息。在后文中, 将展示关于 S-ResBlock 与 ResBlock 的对比实验, 以进一步验证该结构带来的提升。

如图 1 所示, 在全卷积采样子网络中, 第 1 个卷积层用来从输入图像中提取出对应的高维特征图, 得出的特征紧接着被输入到 R 个级联的 S-ResBlock。值得注意的是, 所有 S-ResBlock 的恒等映射分支均采用第 1 个卷积层输出的高维特征图作为输入。最后一个 S-ResBlock 的输出被输入到连续 K 个步长为 S 的卷积层中进行下采样特征提取, 得到最终的观测值。因此, 对原始大小为 $M \times N$ 的图像, 采样子网络的输出大小将是 $\frac{M}{S^K} \times \frac{N}{S^K} \times C$, 其中 C 为最后一个卷积层中的卷积核个数。

相比传统的分块压缩感知和线性映射采样方式, 该采样子网络成功避免了块效应问题, 并且将非线性结合到了采样过程之中, 更有效地对特征信息进行提取, 获得信息更加丰富的观测值。

2.2 基于插值卷积的初始重构子网络

在传统图像压缩感知的信号恢复过程中, 通常



Conv 代表卷积层, ReLU 代表线性整流激活函数, BN 代表批量标准化层

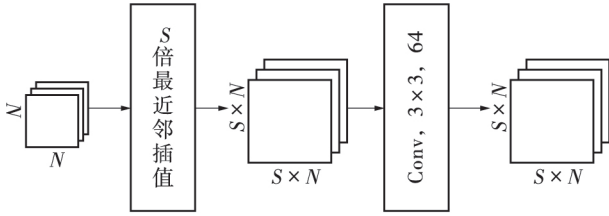
图2 两种不同的残差学习结构

Fig. 2 Two different residual learning block

先利用观测值和简单的重构算法获得图像的初始重构, 然后利用初始重构图像和观测值, 通过较复杂的重构算法进一步提升重构图像的质量。在这种模式下, 通过重构算法的迭代逐渐恢复图像中的细节, 可以取得优秀的重构效果。文中提出的卷积神经网络重构算法即参考这种策略。

基于插值卷积 (Resize Convolution)^[22], 文中提出了一个初始重构子网络, 用以生成高质量的初始重构图像。如文献 [22] 所述, 插值卷积是对特征图进行上采样的一种直观而简单的方法。在结构上, 插值卷积分为两层: 首先是一层具有 S 倍上采样效果的最近邻插值层, 再后接一个常用的 3×3 卷积层以进行特征提取, 如图 3 所示。与另外一种神经网络中常用的上采样方法——反卷积 (Deconvolution) 相比, 插值卷积成功避免了反卷积操作中容易带来的棋盘效应, 提升了恢复图像的效果。初始重构子网络主要由 3 部分结构构成: 第 1

部分是单层卷积层, 用来从采样系数中提取高维特征图; 第2部分是 K 个插值卷积, 对得到的特征图进行上采样操作, 使特征图恢复到原始图像尺寸; 最后一部分是一层只有单个卷积核的卷积层, 对插值卷积得出的特征图进行整合, 输出最后的图像初始重构结果。



Conv 代表卷积层

图3 插值卷积结构

Fig. 3 Structure of resize convolution

在得出图像的初始重构之后, 接下来的工作便以此为基础, 逐渐对图像进行精细化的细节恢复。

2.3 多阶段图像增强子网络

为了进一步细化图像结果, 文中设计了一个多阶段图像增强子网络, 利用初始重构图像和观测值逐渐恢复更多的图像纹理细节。

如图1所示, 该多阶段图像增强子网络的基本结构为细节增强单元, 其输入是从采样系数中提取的高维特征图。在每个细节增强单元中, 先由S-ResBlock对输入进行处理, 提取深层特征, 再利用插值卷积来重构得出残差图像, 与上一级的重构图像进行叠加, 得出新的重构图像。通过使用 D 个细节增强单元, 神经网络逐步从采样系数中提取出多幅包含图像细节信息的残差图像, 在初始重构图像的基础上不断提升重构质量。最后, 在 D 阶段图像增强后, 得出MSResICS的最终输出图像。

与之前的神经网络重构算法相比, 文中提出的网络结构利用残差学习作为网络结构基础, 对图像进行更加精细的多阶段恢复, 恢复的图像细节更加丰富, 进一步提升了重构图像质量。

2.4 损失函数

均方误差(MSE)是有效的图像质量客观评判标准。文中在MSResICS网络的训练过程中, 同样采用均方误差函数作为网络的优化目标函数:

$$L(\theta) = \frac{1}{2N} \sum_{i=1}^N \|x_i - f(x_i, \theta)\|^2 \quad (3)$$

式中, x_i 代表训练集图像, N 为训练集样本的个数, θ 为网络参数, $f(x_i, \theta)$ 为网络的整体输出。

3 网络实现细节和实验结果

3.1 网络实现和训练细节

考虑到重构性能和内存占用之间的权衡, 在实验中, 结构参数设置为: $R=3$, $K=2$, $S=2$, $D=4$ 。但在计算资源允许的情况下, 网络越深, 重构效果将会越好。在细节上, 还有几点需要注意: 1) MSResICS中的每一层卷积层中卷积核的大小均设置为 3×3 , 并且都带有线性整流激活函数(ReLU); 2) 假如没有明确的描述, 每层卷积层都含有64个卷积核; 3) 关于网络的初始化方法, 文中采用了文献[23]给出的针对ReLU激活函数的参数设置策略, 以获得最优的效果。

考虑到图像的数量和类别, 文中采用MSCOCO数据集^[24]作为MSResICS的训练集和验证集: 通过从MSCOCO数据集中随机抽取大约10000张图像, 对其进行灰度化处理, 并将原图像的中心部分裁剪为 256×256 大小的子图像。此外, 文中同时采用了数据增广对训练数据进行进一步的扩增。在训练过程中, Batchsize的大小设为32, 以在训练速度和重构结果上均获得最佳的性能表现。神经网络的优化算法选用的是在许多其他工作中均有较好表现的Adam Optimizer^[25], 其中的参数设为: $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1 \times 10^{-8}$ 。为了更好地收敛, 网络总共训练100个epoch, 一个epoch中循环3000次。在训练中, 分别以3种不同的学习率对网络逐渐优化: 初始30个epoch学习率为0.0001, 接下来40个epoch为0.00001, 最后30个epoch为0.000001。在实验中, 网络具体实现的软件框架为: python3.6 + tensorflow1.12。关于时间消耗, 在GPU-GTX1080ti上, 一次完整的网络模型训练大概需要3d的时间。

3.2 仿真实验结果及性能对比

在本节中, 分别设计了3组对比实验, 以从多种角度分析说明MSResICS在不同条件下的重构性能表现。

3.2.1 不同算法间的重构性能对比

为了具体分析MSResICS的重构性能表现, 文中将MSResICS的图像重构结果与5种具有代表性的图像压缩感知方法进行对比, 其中包含两种传统的优化算法与3种基于神经网络的方法, 分别是: 多假设重构算法(Multi-Hypothesis, MH)^[26]、基于组稀疏表示的重构算法^[6]、ReconNet^[14]、ISTA-

Net⁺^[15]和CSNet^[19]。多假设重构算法与基于组稀疏表示的算法均是效果比较优秀的传统算法,特别是基于组稀疏表示的重构算法,是目前性能最好的图像压缩感知重构算法。而ReconNet、ISTA-Net⁺和CSNet都是具有代表性的压缩感知神经网络结构,后两者的恢复效果在神经网络方法中达到了最佳。

为了评估不同算法的重构性能,文中使用由8张标准灰度图像组成的测试组作为评测参照,使用的图像分别来自两个网络来源: <https://homepages.cae.wisc.edu/~ece533/images/>, <https://www.io.csic.es/PagsPers/JPortilla/image-processing/bls-gsm/63-test-images>, 具体图像参照图4。文中给出了0.5、0.25、0.125不同采样率下的实验结果。8幅图像的峰值信噪比(PSNR)/结构相似性(SSIM)对照结果被统一汇总在表1中。

从表1中可以看出,MSResICS在所有测试情况下的表现均优于其他算法。与优化算法中性能较优秀的基于组稀疏表示的重构算法进行对比,MSResICS在0.125~0.500的采样率下PSNR分别

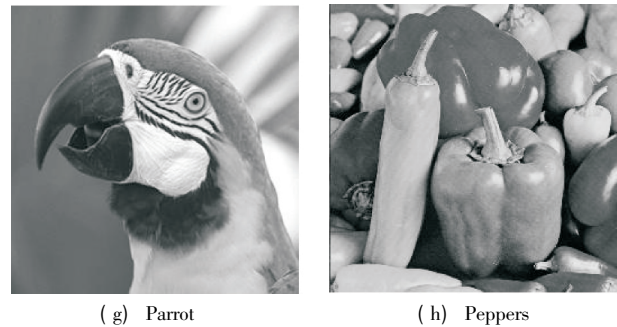


图4 8张测试图像

Fig. 4 Eight test images

可以提升平均大约3.21、1.81、0.96 dB,而相对于多假设重构算法提升更为明显,分别为5.15、4.48、4.64 dB。从实验结果可以看出,MSResICS的效果不仅超出传统的优化方法,而且大大提高了在低采样率下的图像重构质量。在与基于卷积神经网络的方法对比时,在0.125、0.250、0.500采样率下,MSResICS在PSNR上比ReconNet平均提升了6.25、6.44、7.34 dB,比ISTA-Net⁺提升大约3.29、2.22、1.32 dB,比CSNet提升2.09、1.96、1.74 dB,展现出其性能的优越性。SSIM作为一种更为先进的衡量图像相似度的指标,在一定程度上进一步体现了算法对于恢复图像的质量好坏。文中提出的MSResICS在SSIM上相较于对比算法也都有着不少的提升,并同样在低采样率下的效果更为出色。

在不同的图像上,MSResICS的性能表现也有所不同:例如在细节信息较多的图像Goldhill、House上,MSResICS相较于其他算法的PSNR平均提升分别为3.34、3.51 dB,而在细节信息较少的图像Peppers上平均提升为2.92 dB,在SSIM指标上也保持着类似的规律,证明在细节部分,MSResICS的恢复效果表现更好。在图5和图6中,文中进一步展示了一些更加详细的视觉结果图。在图5 House的0.5采样率恢复图像中,MSResICS在房子烟囱部分的砖块细节恢复上纹理更加明显。在图6 Monarch的0.125采样率恢复图像中对比的5种算法均有较为明显的块效应,而MSResICS的恢复结果整体平滑;而且在细节部分,比如蝴蝶的主体、背景的花卉,MSResICS的恢复图像均更加清晰。从整体上来看,在传统的分块压缩感知框架下,无论是基于优化的算法还是基于神经网络的算法,恢复图像中图像块的边界容易出现不自然的分界线,尤其在低采样率下这个问题尤为明显。文中提出的MSResICS则成功地避免了这个问题,在重构图像整体的视觉效果上都均匀平滑。除此之外,在细节



表1 MSResICS 与几种不同重构算法的 PSNR/SSIM 恢复结果

Table 1 Recovery PSNR/SSIM results for MSResICS and other different algorithms

测试图像	采样率	PSNR (dB) /SSIM					
		MH ^[26]	GSR ^[6]	ReconNet ^[14]	ISTA-Net ⁺ [15]	CSNet ^[19]	MSResICS
Boat	0.125	27.32/0.7907	30.21/0.8907	26.15/0.7429	28.85/0.8414	30.12/0.8852	32.39/0.9207
	0.250	31.26/0.8906	35.39/0.9535	29.41/0.8421	33.59/0.9299	34.25/0.9422	36.52/0.9630
	0.500	35.85/0.9478	41.13/0.9822	33.75/0.9241	39.86/0.9776	39.24/0.9755	42.16/0.9858
Cameraman	0.125	23.01/0.7361	24.07/0.8041	22.56/0.6878	25.01/0.7934	26.13/0.8353	28.35/0.8807
	0.250	27.16/0.8411	28.09/0.8756	24.84/0.7654	29.02/0.8826	29.16/0.8876	31.16/0.9240
	0.500	31.69/0.9195	33.09/0.9446	28.30/0.8581	34.06/0.9487	32.49/0.9489	35.03/0.9641
Goldhill	0.125	26.84/0.7362	27.70/0.7801	26.49/0.6997	27.97/0.7679	29.81/0.8683	31.22/0.8691
	0.250	30.08/0.8441	31.76/0.8887	29.25/0.8086	31.64/0.8824	33.62/0.9174	34.20/0.9337
	0.500	33.68/0.9245	36.84/0.9589	32.71/0.9024	37.15/0.9601	38.52/0.9752	39.43/0.9789
Lena	0.125	27.28/0.8287	28.73/0.8789	25.99/0.7538	28.68/0.8554	29.55/0.8826	31.87/0.9215
	0.250	30.68/0.8976	32.54/0.9354	28.84/0.8403	32.66/0.9271	33.32/0.9423	35.22/0.9594
	0.500	35.46/0.9521	38.93/0.9782	32.53/0.9164	38.28/0.9748	38.37/0.9801	39.71/0.9838
House	0.125	30.97/0.8425	34.90/0.8971	28.20/0.7746	32.16/0.8577	32.84/0.8786	34.99/0.8936
	0.250	34.94/0.9025	38.34/0.9438	31.22/0.8341	36.27/0.9082	36.42/0.9124	38.23/0.9399
	0.500	38.54/0.9517	42.32/0.9756	34.82/0.8939	40.31/0.9569	41.12/0.9726	43.61/0.9822
Monarch	0.125	24.61/0.8101	26.82/0.8981	23.69/0.7573	27.43/0.8746	29.06/0.9156	31.40/0.9451
	0.250	28.25/0.8853	32.78/0.9573	27.14/0.8502	33.10/0.9525	33.25/0.9602	34.81/0.9715
	0.500	32.93/0.9402	39.19/0.9828	31.28/0.9217	39.09/0.9836	37.77/0.9834	39.30/0.9879
Parrot	0.125	25.17/0.8256	26.86/0.8615	24.70/0.7786	27.71/0.8741	28.51/0.8968	31.20/0.9242
	0.250	30.08/0.9123	32.32/0.9365	28.02/0.8557	31.30/0.9283	32.42/0.9401	34.26/0.9556
	0.500	35.70/0.9548	38.57/0.9723	31.54/0.9142	37.14/0.9689	37.87/0.9783	38.54/0.9804
Peppers	0.125	26.26/0.7569	28.09/0.8218	24.87/0.7055	28.47/0.8161	29.91/0.8589	31.21/0.8857
	0.250	29.62/0.8388	31.83/0.8867	27.68/0.7911	32.59/0.8892	31.41/0.9051	33.50/0.9224
	0.500	33.23/0.9056	36.36/0.9424	30.58/0.8600	36.13/0.9432	34.92/0.9374	36.43/0.9566
平均结果	0.125	26.43/0.7909	28.42/0.8540	25.33/0.7375	28.29/0.8351	29.49/0.8776	31.58/0.9051
	0.250	30.26/0.8765	32.88/0.9222	28.30/0.8234	32.52/0.9125	32.98/0.9259	34.74/0.9462
	0.500	34.64/0.9370	38.30/0.9671	31.94/0.8989	37.76/0.9642	37.54/0.9689	39.28/0.9775

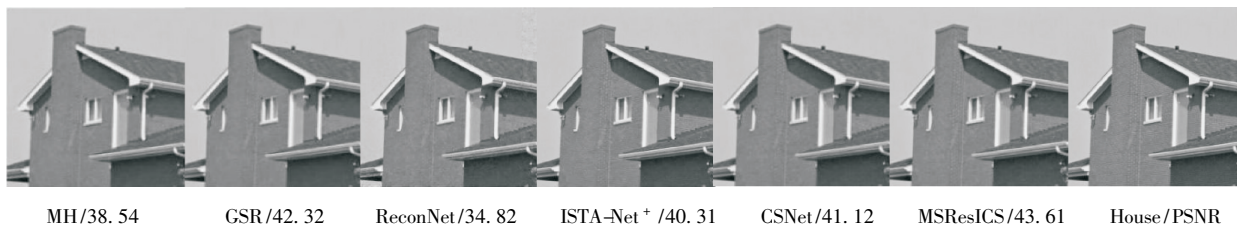


图5 0.5 采样率下不同算法对图像 House 的视觉恢复图像

Fig. 5 Visual recovery images of different methods on image House under sampling ratio 0.5



图6 0.125 采样率下不同算法对图像 Monarch 的视觉恢复图像

Fig. 6 Visual recovery images of different methods on image Monarch under sampling ratio 0.125

的恢复质量对比上, MSResICS 也比其他算法保留了更多的高频信息和边缘信息, 纹理更加清晰。同时, 文中还将 MSResICS 与文献 [18] 中提出的 MSNet 的部分结果进行了一定的对比, 参照表 2, 可以看出 MSResICS 在重构图像的 PSNR 上要远远优于 MSNet。

表 2 0.25 采样率下 MSResICS 与 MSNet 恢复结果 PSNR 对比

Table 2 Comparison of recovery PSNR results for MSResICS and MSNet with sampling ratio 0.25

重构算法	PSNR/dB			
	Boat	Lena	Monarch	Peppers
MSNet ^[18]	30.74	30.21	28.90	29.51
MSResICS	36.52	35.22	34.81	33.50

3.2.2 在小数据集上的重构性能对比

为了公平的对比, 进一步考察 MSResICS 在小数据集上的表现, 文中基于文献 [19] 中使用到的 BSDS500 数据集^[27]对 MSResICS 进行了重新训练, 其中训练集总共由 400 张图片构成, 分别来自 BSDS500 中的训练集 (200 张图片) 与测试集 (200 张图片), 验证集则由来自 BSDS500 验证集的 100 张图片构成。除数据集不同外, 其余训练设置则与上述实验保持一致。表 3 中展示了在不同采样率下, 基于 BSDS500 数据集训练的 MSResICS 对上述八幅图像的平均恢复结果, 并以 CSNet 作为基准进行对比。在低采样率的情况下, MSResICS (BSDS500) 在恢复图像的 PSNR 与 SSIM 上相较于 CSNet 仍然有明显提升。但与基于 MSCOCO 训练的版本对比, MSResICS (BSDS500) 的恢复图像质量有所下降。产生以上结果的主要原因为: MSResICS 在训练过程中, 输入的是整幅图像, 需要的图片量更为庞大, 而 CSNet 对于图像分块进行训练, 在较小的图片数据集上即可取得不错的效果; 同时, 由于 MSResICS 的输入是整幅图像的原因, 它克服了 CSNet 在低采样率下恢复图像中的块效应问题, 提升了恢复图像的质量。

表 3 基于 BSDS500 训练的 MSResICS 恢复结果平均 PSNR/SSIM 比较

Table 3 Comparison of recovery average PSNR/SSIM results on MSResICS trained on BSDS500

训练方式	PSNR (dB) /SSIM		
	0.125 ¹⁾	0.250 ¹⁾	0.500 ¹⁾
CSNet	29.49/0.8776	32.98/0.9259	37.54/0.9689
MSResICS (BSDS500)	31.45/0.9024	33.34/0.9296	37.27/0.9654
MSResICS (MSCOCO)	31.58/0.9051	34.74/0.9462	39.28/0.9775

1) 为采样率。

3.2.3 不同网络结构的重构性能对比

在保证网络其余结构一致的情况下, 表 4 中展示了 S-ResBlock 与原版 ResBlock 在不同采样率下的对比实验结果。在不同图像上, S-ResBlock 的效果均优于 ResBlock, PSNR 结果提升 0.6 ~ 1.2 dB, SSIM 大约提升 0.02 ~ 0.06, 证实了前者针对压缩感知任务的有效性。

表 4 S-ResBlock 与 ResBlock 结构恢复结果 PSNR/SSIM 比较

Table 4 Comparison of recovery PSNR/SSIM results between S-ResBlock and ResBlock structures

测试图像	采样率	PSNR (dB) /SSIM	
		ResBlock	S-Resblock
Cameraman	0.125	27.41/0.8207	28.35/0.8807
	0.250	30.02/0.8925	31.16/0.9240
	0.500	34.37/0.9492	35.03/0.9641
Lena	0.125	30.71/0.9156	31.87/0.9215
	0.250	34.59/0.9501	35.22/0.9594
	0.500	38.82/0.9775	39.71/0.9838
House	0.125	33.81/0.8823	34.99/0.8936
	0.250	37.52/0.9215	38.23/0.9399
	0.500	42.46/0.9754	43.61/0.9822

为了验证采样子网络和图像增强子网络对重构性能的影响, 本小节在保持训练条件不变的情况下重新训练了两种网络结构, 分别为: ①随机高斯观测 + 初始重构子网络 + 图像增强子网络, ②采样子网络 + 初始重构子网络, 并与完整的 MSResICS 进行重构效果上的对比。表 5 中给出了对应的实验对比。

表 5 不同子网络结构设定下神经网络之间的恢复结果 PSNR/SSIM 比较

Table 5 Comparison of recovery PSNR/SSIM results between networks with different subnetwork settings

测试图像	采样率	PSNR (dB) /SSIM		
		①随机高斯观测 + 初始重构子网络 + 图像增强子网络	②采样子网络 + 初始重构子网络	完整结构
Peppers	0.125	27.93/0.8122	30.48/0.8769	31.21/0.8857
	0.250	30.18/0.8426	32.85/0.9154	33.50/0.9224
	0.500	33.37/0.9212	35.62/0.9540	36.43/0.9566
Monarch	0.125	26.15/0.8215	30.62/0.9335	31.40/0.9451
	0.250	29.11/0.8955	33.95/0.9665	34.81/0.9715
	0.500	33.41/0.9549	37.77/0.9852	39.30/0.9879
Parrot	0.125	27.94/0.8822	29.91/0.9162	31.20/0.9242
	0.250	30.29/0.9171	33.36/0.9501	34.26/0.9556
	0.500	34.66/0.9676	37.62/0.9787	38.54/0.9804

比结果。在不同采样率下,完整的 MSResICS 都要优于两组对比试验: 恢复图像 PSNR/SSIM 相较于①平均提升 4.2 dB/0.057,而对比②平均提升 0.94 dB/0.006。部分视觉效果图对比如图 7 所示,图中 3 列从左至右分别为①、②和完整结构 3 种情况下的重构图像。可以看出,与两个对照组相比,完整结构的网络在细节和轮廓上的恢复效果均较好,

例如 Goldhill 的远景部分, Lena 的帽子部分。以上的结果表明,采样子网络对于网络的性能有着关键的影响,相比于传统高斯采样矩阵,其大大强化了图像压缩感知采样过程中的效率,提升了网络的重构效果;而图像增强子网络在初始重构图像的基础上进一步增强了恢复图像中的细节信息,提高了最终恢复图像的质量。



图7 不同子网络结构设定下神经网络之间的视觉恢复结果对比

Fig. 7 Comparison of visual recovery results between networks with different sub-network settings

4 结论

在文中,提出了一种新颖的网络结构——MSResICS 来处理图像压缩感知中的两个重要问题:采样和恢复。首先,文中放弃了在一些情况下会限制恢复性能的传统分块压缩感知框架,提出了基于残差学习的全卷积采样方法,并运用插值卷积,设计了对应的初始重构子网络。为了细致深入的图像恢复,基于前两个子网络的工作,文中进一步引入了多阶段图像增强子网络以稳定地提升重构图像质量。通过 end-to-end 的训练方式,提出的 MSResICS 获得了令人满意的重构结果。与几种最先进的重构算法相比,MSResICS 在恢复图像的 PSNR 上平均提高了 2~3db,在视觉效果上也有着不错的提升。在未来的工作,作者将尝试将 MSResICS 结合到视频压缩感知框架下,并寻求更多的突破。

参考文献:

- [1] DONOHO D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [2] CANDLES E J, WAKIN M B. An introduction to compressive sampling [J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.
- [3] 石光明,刘丹华,高大化,等. 压缩感知理论及其研究进展 [J]. 电子学报, 2009, 37(5): 1070-1081.
SHI Guang-ming, LIU Dan-hua, GAO Da-hua, et al. Advances in Theory and Application of Compressed Sensing [J]. Journal of Electronics & Information Technology, 2009, 37(5): 1070-1081
- [4] GAN L. Block compressed sensing of natural Images [C] // 2007 15th International Conference on Digital Signal Processing. Cardiff: IEEE, 2007: 403-406.
- [5] MUN S, FOWLER J E. Block compressed sensing of

- images using directional transforms [C] // 2010 Data Compression Conference. Snowbird, UT: IEEE, 2010: 547-547.
- [6] ZHANG J, ZHAO D, GAO W. Group-based sparse representation for image restoration [J]. IEEE Transactions on Image Processing, 2014, 23(8): 3336-3351.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] 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.
- [9] SAJJADI M S M, SCHÖLKOPF B, HIRSCH M. EnhanceNet: single image super-resolution through automated texture synthesis [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 4501-4510.
- [10] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C] // Proceedings of the IEEE conference on computer vision and pattern recognition workshops. Honolulu, Hawaii: IEEE, 2017: 136-144.
- [11] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network [C] // Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, Hawaii: IEEE, 2017: 3147-3155.
- [12] HUANG Y, WANG W, WANG L. Video super-resolution via bidirectional recurrent convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 1015-1028.
- [13] LI B, PENG X, WANG Z, et al. AOD-net: all-in-one dehazing Network [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017.
- [14] KULKARNI K, LOHIT S, TURAGA P, et al. Recon-net: non-iterative reconstruction of images from compressively sensed measurements [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE, 2016: 449-458.
- [15] YAO H, DAI F, ZHANG D, et al. DR2-net: deep residual reconstruction network for image compressive sensing [EB/OL]. [2019-6-1]. <https://arxiv.org/abs/1702.05743>.
- [16] ZHANG J, GHANEM B. ISTA-Net: interpretable optimization-inspired deep network for image compressive sensing [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018: 1828-1837.
- [17] LU X, DONG W, WANG P, et al. ConvCSNet: a convolutional compressive sensing framework based on deep learning [EB/OL]. [2019-6-1]. <https://arxiv.org/abs/1801.10342>.
- [18] 练秋生, 富利鹏, 陈书贞, 等. 基于多尺度残差网络的压缩感知重构算法 [J]. 自动化学报, 2019, 45(11): 2082-2091.
- LIAN Qiu-sheng, FU Li-peng, CHEN Shu-zhen, et al. A compressed sensing algorithm based on multi-scale residual reconstruction network [J]. ACTA AUTOMATICA SINICA, 2019, 45(11): 2082-2091.
- [19] SHI W, JIANG F, ZHANG S, et al. Deep networks for compressed image sensing [C] // 2017 IEEE International Conference on Multimedia and Expo (ICME). Hong Kong: IEEE, 2017: 877-882.
- [20] CUI W, JIANG F, GAO X, et al. Deep neural network based sparse measurement matrix for image compressed sensing [C] // 2018 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece: IEEE, 2018: 3883-3887.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE, 2016: 770-778.
- [22] ODENA A, DUMOULIN V, OLAH C. Deconvolution and checkerboard artifacts [EB/OL]. [2019-6-1]. <https://distill.pub/2016/deconv-checkerboard/>.
- [23] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C] // 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1026-1034.
- [24] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [C] // Computer Vision-ECCV 2014 lecture notes in computer science. Zurich: Springer, 2014: 740-755.
- [25] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. [2019-6-1]. <https://arxiv.org/abs/1412.6980>.
- [26] CHEN C, TRAMEL E W, FOWLER J E. Compressed-sensing recovery of images and video using multi-hypothesis predictions [C] // 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR). Pacific Grove, CA: IEEE, 2011: 1193-1198.
- [27] PABLO A, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.

Multistage Image Compressive Sensing Neural Network Based on Residual Learning

YANG Chunling PEI Hanqi

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, Guangdong, China)

Abstract: There are two main problems in traditional Image Compressive Sensing (ICS): in sampling aspect, traditional linear sampling methods have some limitations; in reconstruction aspect, optimization-based reconstruction methods are highly time-consuming. Newly proposed ICS Neural Network can successfully deal with the speed problem in reconstruction, but lacks the accuracy of traditional algorithms. To solve this problem, a novel multi-stage ICS network based on residual learning (MSResICS) was proposed. It consists of three sub-networks, namely, sampling sub-network, initial reconstruction sub-network, and image enhancement sub-network. In sampling stage, with the help of residual learning, a nonlinear sampling sub-network was designed, which breaks the limitation of conventional sampling method and retains richer image information in samples. In reconstruction stage, the initial reconstruction sub-network extracts features from samples and obtain an initial reconstructed image of high quality by introducing interpolation convolution. With residual learning and interpolation convolution, a multi-stage image enhancing sub-network was proposed to further refine the reconstruction image and improve the quality of final result. Extensive simulations show that MSResICS has a better reconstruction accuracy than the existing optimal ICS reconstruction methods.

Key words: image compressive sensing; residual learning; deep network; sampling mechanism

(上接第 81 页)

Parameter Estimation of 2D-GTD Model Based on the Improved 2D-ESPRIT Algorithm

ZHANG Xiaokuan¹ ZHENG Shuyu² ZHAO Weichen³ ZONG Bin Feng¹ LI Jiang²

(1. Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, Shaanxi, China;

2. Graduate School, Air Force Engineering University, Xi'an 710051, Shaanxi, China;

3. Jilin University of Finance and Economics, Changchun 130022, Jilin, China)

Abstract: The classical 2D-ESPRIT algorithm is sensitive to signal-to-noise-ratio (SNR) when estimating the parameters of two dimensional (2D) geometric theory of diffraction (GTD) model. To solve the problem, an improved 2D-ESPRIT algorithm, which improves the anti-noise performance and parameter estimation performance effectively, was proposed. Firstly, the improved algorithm obtains a new matrix \mathbf{Y} containing the conjugate information of the original matrix \mathbf{E}_e by constructing an exchange matrix. Secondly, a new covariance matrix \mathbf{R} can be obtained by auto-correlation, cross correlation, superposition and average processing. Finally, the parameters of 2D-GTD model can be estimated by calculating the new covariance matrix \mathbf{R} . Simulation results show that the improved 2D-ESPRIT algorithm has a better anti-noise performance, and it owns higher estimation precision and more stable performance than the classical 2D-ESPRIT algorithm accurately in the case of low signal-to-noise ratio. The impact of other factors, such as matrix pencil parameters and paring variable β , on parameter estimation precision was also studied, which can be provide references for future simulations.

Key words: parameter estimation; scattering center; 2D-GTD model; improved 2D-ESPRIT algorithm; conjugate matrix