光电技术及应用

DOI：10.16818/j.issn1001－5868.2021.03.021

基于深度学习的单像素相机超分辨率成像优化

魏子然[1,2,3]，杨　威[1]，张建林[1,3]，徐智勇[1,3]，刘　永[2]，王盛杰[4*]

（1. 中国科学院光电技术研究所 光束控制重点实验室，成都 610209；2. 电子科技大学 光电科学与工程学院，成都 610054；
3. 中国科学院大学，北京 100049；4. 中国民用航空飞行学院，四川 广汉 618307）

摘　要： 基于压缩感知和单像素成像的基本原理，设计了一种用于图像超分辨率重建的新型深度卷积神经网络架构。这种单像素超分辨率成像算法成功地将深度学习图像超分辨率重建技术与压缩感知单像素成像技术相结合，从而发展出一种全新的深度学习单像素成像优化方法。与传统的常规压缩感知图像重构算法相比，该算法有效提升了图像超分辨率重建精度和单像素成像质量。通过图像重建的仿真实验和单像素相机的成像实验验证，结果表明这种基于深度学习的新型单像素相机成像方式具有良好的性能表现。

关键词： 单像素成像；压缩感知；深度学习；超分辨率；图像重构

中图分类号： TP391　**文章编号：** 1001－5868(2021)03－0412－06

Super-Resolution Imaging Optimization of Single Pixel Camera Based on Deep Learning

WEI Ziran[1,2,3], YANG Wei[1], ZHANG Jianlin[1,3], XU Zhiyong[1,3], LIU Yong[2], WANG Shengjie[4]

(1. Institute of Optics and Electronics of the Chinese Academy of Sciences, Chengdu 610209, CHN;
2. School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, CHN; 3. University of Chinese Academy of Sciences, Beijing 100049, CHN;
4. Civil Aviation Flight University of China, Guanghan 618307, CHN)

Abstract: Based on the basic principles of compressed sensing (CS) and single pixel camera imaging, a new deep convolutional neural network architecture for image super-resolution reconstruction is redesigned. The new single pixel super-resolution imaging algorithm successfully combines deep learning image super-resolution reconstruction technology with compressed sensing single pixel imaging technology to develop an entirely new deep learning based optimized imaging method of single pixel camera. Compared with the conventional CS-based algorithms, the accuracy of image super-resolution reconstruction and the imaging quality of single pixel camera are improved effectively. The good performance of the new proposed deep learning imaging method of single pixel camera has been verified by simulation experiments of image reconstruction and imaging experiments of single pixel camera.

Key words: single pixel imaging; compressed sensing; deep learning; super-resolution; image reconstruction

0　引言

Donoho 和 Candes 于 2006 年提出了压缩感知(Compressed Sensing，CS)理论[1-2]，其基本原理过程是：通过测量矩阵以比奈奎斯特(Nyquist)采样频率更低的采样速率对原始信号进行线性测量，从而直接获得原始信号的压缩表示，然后根据所得的测量值通过相应的 CS 重建算法重建出原始信号[1]。压缩感知的核心思想是在信号处理过程中将信号采样和信号压缩通过对信号的测量过程合并进行。但

收稿日期：2020－11－10.

* 通信作者：王盛杰

是，压缩感知重建信号要求一定的理论前提：原始信号是稀疏的或者能够进行稀疏表示。压缩感知信号处理可以有效地缩小信号尺寸，实现对信号数据的压缩，无需进行奈奎斯特采样的中间阶段，这在很大程度上节约了信号传输和存储的成本，并降低了计算复杂度[2-3]。压缩感知理论主要包括信号的稀疏表示、测量矩阵的构造和信号重构算法三个部分[4-6]。压缩感知对信号的低采样率和高压缩比使它在信号处理中具有很大的优势，这使得 CS 在医学成像、无线通信、模式识别、雷达探测、光学/遥感成像和图像超分辨重建等领域受到高度关注[7]。

压缩感知理论为单像素相机的图像超分辨率重建和成像提供了理论基础。美国莱斯大学正是基于压缩传感技术，在 2008 年成功开发了世界上第一台单像素相机[8]。在此基础上，许多高校和科研机构都对单像素成像进行了一系列的改进和优化[9-11]。本实验室搭建了一套单像素成像系统，其中采用数字微镜阵列(Digital Mirror Device，DMD)来实现对成像光信号的编码采样。对于单像素成像而言，关键技术还是成像过程中的信号(图像)重构算法。目前，压缩感知信号重建常用的传统算法有：正交匹配追踪算法(OMP)[12]、正则化正交匹配追踪算法(ROMP)[13]、L1 正则化最小二乘法(L1_LS)[14]、L1 范数基追踪算法(L1_BP)[15]、迭代加权最小二乘法(IRLS)[16]。本文将深度学习图像超分辨率重建技术与压缩感知单像素成像技术相结合，基于压缩感知理论和单像素成像原理[8,17-18]，提出了一种通过深度学习方式在单像素成像系统中进行图像超分辨率重建的新型成像算法。需要指出的是，本文中所述的超分辨率并非传统意义上突破光学衍射极限的超分辨成像，而是指以更低的图像采样分辨率实现更高分辨率的图像重建过程。

图像超分辨率(SR)，尤其是单图像超分辨率(SISR)，在图像恢复重建的研究中引起了越来越多的关注，而深度学习技术可以显著提升图像超分辨率重建的性能[19-21]，但是深度卷积神经网络在结构优化方面存在一些不足与局限性。首先，针对图像重建的神经网络模型性能对网络结构的微小变化非常敏感，同一模型在不同的初始化和训练方式条件下具有不同的性能水平。因此，具有良好设计和结构优化的神经网络架构对于网络的训练和学习显得至关重要。本文提出了一种基于残差块(Residual Block，ResBlock)的深度卷积神经网络架构，并将这种深度学习方法应用在我们的单像素成像系统中。与单像素相机使用的传统压缩感知图像重建算法相比，深度学习方式可提高图像超分辨率重建精度和单像素成像质量。

1　深度卷积神经网络结构优化

卷积神经网络(CNN)可以应用于许多计算机视觉任务，许多基于 CNN 的方法被相继提出[22-24]。在 Dong 等提出了基于深度学习的图像超分辨率构造方法之后[25]，残差网络(Residual Network，ResNet)[26]被引入深度学习中，以训练更深并且具有更好性能的网络结构[19-20]，从而优化图像超分辨率重建结果。然而，深度学习方式很少被应用在单像素超分辨率成像中。因此，针对单像素成像，本文提出了一种基于改进残差块(ResBlock)的新型深度卷积神经网络架构，通过这种深度学习方式提升单像素成像中图像超分辨重建的性能。残差网络作为网络的主体结构，用来对图像进行特征提取(Feature Extractor)，然后利用 Pixel Shuffle[22] 来对特征进行上采样(Upsample Module)，具体结构如图 1 所示。图 2 表示图 1 中的残差块，它由三个不同的卷积层组成，其中尺度因子 κ 为学习参数之一，conv 表示卷积操作运算。

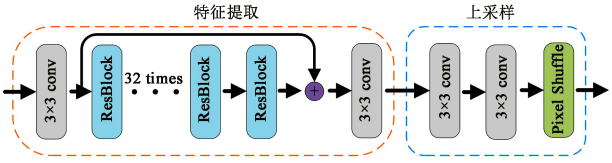


图 1　深度卷积神经网络主体架构

如图 2 所示，与文献[26]中的传统残差网络相比，优化后的残差块去掉了 BN 层(Batch Normalization)。BN 层的主要功能是对特征进行归一化，减少网络模型对初始化的依赖，但同时 BN 层在一定程度上会让网络更加庞大和臃肿。然而，图像超分辨率重建的目的是精确重建图像的细节和纹理，对图像以及图像特征之间的相关性和一致性要求并不高，而且 BN 层还会消耗大量的计算内存。因此，去掉 BN 层不仅可以节省 GPU 的内存，而且还可以提高重建图像细节的能力。因此，与传统的 ResNet 网络结构相比，在计算资源有限的情况下，在图 2 所示的新 ResNet 基础上我们可以建立一个更庞大、更复杂的网络，提升网络模型的性能。

在深度学习中，更复杂的网络结构意味着更多

的网络参数，而增加网络参数数量在一定程度上可以更好地改善网络的性能。因此，通常可以通过堆叠更多的网络层数或增加滤波器数量来增加网络参数的数量。假设 CNN 结构的网络深度为 B（网络层数）、宽度为 F（特征数），此时网络参数的数量为 $O(BF^2)$，消耗的计算内存为 $O(BF)$。因此，增加网络的宽度 F 而不增加网络的深度 B 可以在有限的计算资源下更好地提升网络模型的容量。但是，如果将特征参数的数量增加到超过一定范围，则会使网络的训练过程变得不稳定[24]。由于在图 2 所示的网络中使用了大量的滤波器，并且网络中有很多特征图（Feature Maps），因此我们选择加上一个残差尺度因子 κ 来稳定训练过程，κ 位于最后一个卷积层之后，如图 2 所示。κ 的初始值为 0.1，它作为网络中众多学习参数之一，可以在网络训练过程中自适应地进行学习和调整。基于以上分析可知，图 1 所示的网络具有更高的计算效率，且消耗计算内存更少。

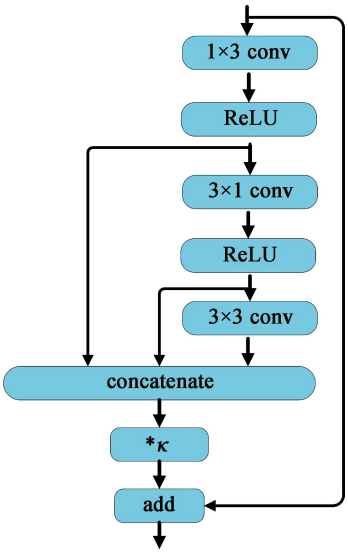


图 2 图 1 网络中残差块结构

2 网络学习过程

在本文的单像素成像系统中，重建图像尺寸为 256×256，采用图像行列双侧投影方式进行光信号采样，信号采样率为 1/9，一幅成像场景图像的采样次数为 7 396（86×86）。因此，通过这样以更低的采样分辨率实现更高分辨率的图像重建，从这个意义上说实现了单像素相机的超分辨率成像。由于 DMD 的微镜单元只能实现二进制的硬件状态，为了使单像素相机能够更加有效地使用 DMD 对光信号进行编码采样，以及便于通过 DMD 对测量矩阵的硬件实现和降低计算复杂度，本文使用稀疏的 0-1 二进制对角分块矩阵[27-28]作为测量矩阵。这种 0-1 二进制测量矩阵被固定，其中的“0”元素和“1”元素具有恒定的分布状态，通过模拟这种二进制分块对角测量矩阵使用双侧投影方式采样和重建图像的过程，来对图 1 中的网络结构和参数进行训练学习，从而实现对图像的超分辨率重建。把这种深度学习算法应用到本文的单像素成像系统中，从而通过深度学习技术来优化和提升单像素超分辨成像的性能。

本文通过深度学习方式对上述网络的训练在四块 GTX 1080Ti GPU 上进行，使用 PyTorch[29]框架来构建和训练网络，训练数据集为 DIV2K[30]，该数据集包含约 900 张高分辨率图片，适合作为图像超分辨率重建的数据集。在 DIV2K 数据集上对网络训练了 720 次，并使用随机裁剪（Random Crop）、随机镜像（Random Mirror）和随机翻转（Random Flip）方式对数据集进行扩容和增强。图像裁剪大小为 256×256，因此网络的标签图像尺寸为 256×256，对应的输入数据尺寸为 86×86（7 396），相应相机的采样率为 1/9。为了进行灰度平衡，将图像中的所有像素灰度值都减去 128，然后再送入到网络。在训练网络的过程中，在使用峰值信噪比（PSNR）和结构相似度（SSIM）评估图像重建质量条件下，L1 损失函数比 L2 损失函数性能更优[31]。因此，使用 L1 范数作为损失函数，并使用随机梯度下降方法（SGD）来更新网络参数，其中动量值（Momentum）为 0.9，权值衰减（Weight Decay）参数为 1×10^{-4}，初始学习率（Learning Rate）为 1×10^{-4}，为了使网络训练更快地收敛拟合，网络每训练 120 次对学习率进行减半。

3 实验分析与讨论

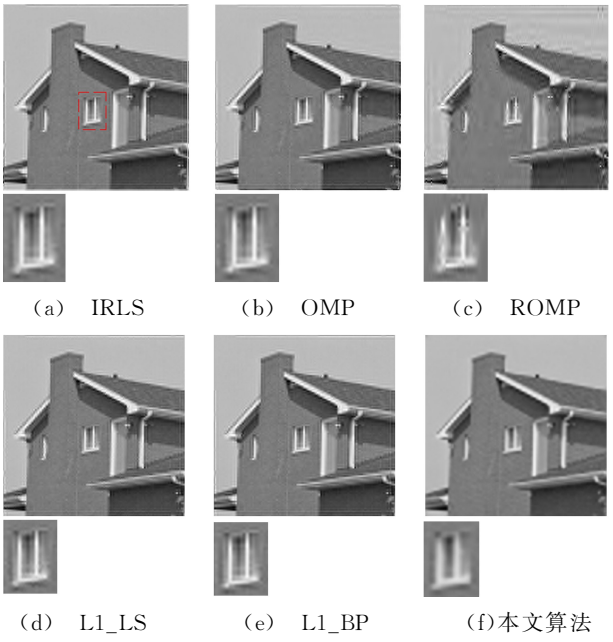
为了验证新提出的深度学习算法对图像超分辨率重建的性能提升效果，以及更好地模拟单像素相机的成像过程，在将深度学习算法应用在单像素成像系统之前，首先进行图像重建的仿真实验。在实验中，使用峰值信噪比（PSNR）和结构相似度（SSIM）作为重建图像质量的评价指标。由于我们的单像素成像场景为远处的建筑物，因此选择 Building 和 House 两张代表房屋建筑的图像进行图像重建。对于 1/9 的成像采样率，也就是一幅图

像在采样测量次数为 7 396 的情况下进行重建，分别使用传统的 IRLS，OMP，ROMP，L1_LS，L1_BP 算法和本文的深度学习算法对 Building 和 House 图像进行超分辨率重建，重建图像的 PNSR 和 SSIM 实验结果如表 1 所示。

表 1　不同算法以 1/9 的采样率重建 Building 和 House 图像的 PSNR 和 SSIM

图像		IRLS	OMP	ROMP	L1_LS	L1_BP	本文算法
Building	PSNR/dB	28.757 2	28.167 6	27.175 0	28.466 4	28.245 8	29.639 3
	SSIM	0.865 68	0.846 13	0.794 72	0.854 86	0.850 65	0.871 39
House	PSNR/dB	28.822 1	28.378 9	26.956 8	28.623 7	28.425 2	29.745 7
	SSIM	0.821 94	0.801 46	0.730 94	0.815 74	0.807 81	0.845 10

根据表 1 所示实验结果的统计数据可知，与传统算法相比，利用本文提出的深度学习算法重建图像的精度得到了有效提升。将表 1 中 House 图像的重建结果进行了展示，并对重建图像中红色虚线框标记的局部区域进行了放大显示，如图 3 所示。



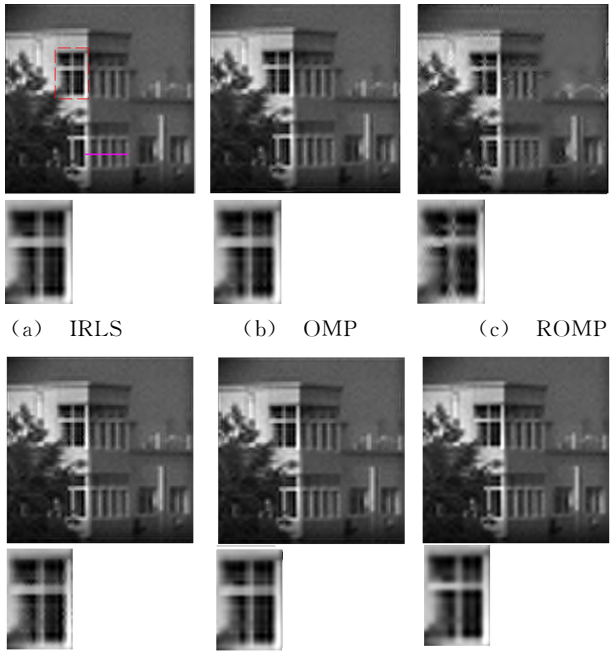
(a) IRLS　(b) OMP　(c) ROMP

(d) L1_LS　(e) L1_BP　(f)本文算法

图 3　不同算法重建 House 图像结果

如图 3 所示，并结合表 1 中的实验数据，与其余五种传统算法相比，本文提出的算法重建 Building 和 House 图像的 PSNR 增加了约 1～2 dB，SSIM 增加了约 0.02～0.10。对于 House 图像的重建，使用本文方法重建的图 3(f)，马赛克现象得到有效控制，抑制了重建图像的伪影和噪声，图像变得更加平滑，重建图像中的轮廓边缘相对于其他算法失真和扭曲更小，图像结构更加均衡稳定，具有更高的峰值信噪比和结构相似度。

在进行图像超分辨率重建的仿真实验后，使用单像素相机进行实际超分辨率成像的图像重建实验。基于单像素相机成像时采集储存的成像场景测量数据，同样使用上述不同算法重建了成像场景图像。由于无法提供成像场景的原始图像作为参考基准来评估成像质量，因此不便于对成像重建图像进行信噪比的定量分析。为了进行直观的视觉对比，比较了单像素相机使用不同算法重建图像的结果，并给出成像场景中红色虚线框标记的窗户区域的放大图，如图 4 所示。在图 4(a)中，在成像场景中品红色实线的标记位置代表具有一定频率和亮暗变化的灰度波动区域。对不同算法重建图像中所标记位置的灰度分布通过曲线进行了展示，以评估图像质量，如图 5 所示。此外，为了更清晰地显示灰度分布的细节，将图 5 中红色虚线框标记的曲线局部区域进行了放大展示。



(a) IRLS　(b) OMP　(c) ROMP

(d) L1_LS　(e) L1_BP　(f) 本文算法

图 4　通过不同算法对成像场景进行单像素成像的图像重建结果

图 4 所展示的实际成像效果跟图 3 中 House 图像重建结果相似，与其他五种传统算法相比，采用

本文提出的深度学习算法重建的图像具有更好的视觉效果，图像轮廓更加清晰，结构更加均衡，失真更少，噪声也更低，没有明显的马赛克效应和视觉伪迹，更加接近真实的成像场景。另外，通过图 5 来分析图 4(a)中品红实线标记位置的灰度分布情况，和五种传统算法相比，本文方法重建图像没有明显的尖锐波动噪声点，灰度分布更加均衡和平滑，相应的灰度曲线基本上不偏不倚地处于所有灰度曲线的居中位置，相当于起到了过滤噪声的平滑滤波效果，从而使得噪声变得相对较低。通过图 5 中曲线局部区域的细节展示可知，对于亮暗变化区域的灰度分布，本文方法成像重建图像的灰度变化频率更加稳定均衡，没有明显尖锐和畸形的波动。因此，通过以上实验结果可以证明，基于新型神经网络架构的深度学习算法在单像素成像中对图像超分辨率重建具有更高的重建精度和更好的成像质量。

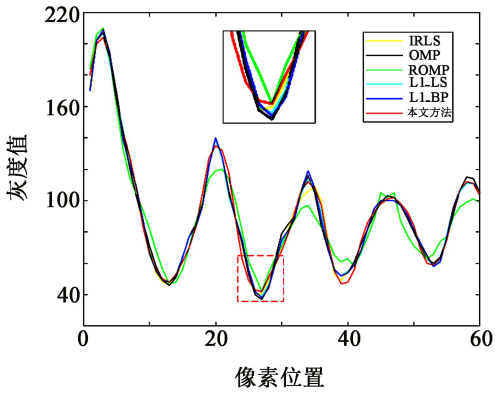


图 5 品红色实线标记成像场景相应位置的灰度分布情况

4 总结

本文提出了一种基于深度卷积神经网络并针对单像素超分辨率成像的优化方法。新的单像素成像算法将深度学习图像超分辨率重建技术与压缩感知单像素成像技术相结合，与传统的压缩感知图像重建算法相比，有效提升了单像素相机的图像重建精度和成像质量。当然，深度学习技术在单像素成像上的进一步应用目前还存在一些问题需要解决，但本文提出的深度学习方式为单像素成像的优化开辟了一种新的思路。

参考文献：

[1] Donoho D L. Compressed sensing [J]. IEEE Trans. on Information Theory, 2006, 52(4): 1289-1306.

[2] Candes E. Compressive sampling[C]// Proc. of the International Congress of Mathematicians, 2006, 3: 1433-1452.

[3] 封晨波，覃亚丽，陈 辉，等. 基于改进非局部均值的分数阶全变分算法[J]. 计算机工程，2019，45(4)：241-247.
Feng Chenbo, Qin Yali, Chen Hui, et al. Fractional total variation algorithm based on improved non-local means [J]. Computer Engin., 2019, 45(4): 241-247.

[4] 赵鸿图，李 成. 基于马尔科夫链的随机测量矩阵研究[J]. 计算机工程，2020，46(4)：241-246.
Zhao Hongtu, Li Cheng. Research on random measurement matrix based on Markov chain[J]. Computer Engin., 2020, 46(4): 241-246.

[5] 王 楠，李 智，程欣宇，等. 基于均值比例与压缩感知的视频双水印算法[J]. 计算机工程，2019，45(3)：250-255，261.
Wang Nan, Li Zhi, Cheng Xinyu, et al. Video dual watermarking algorithm based on mean ratio and compressive sensing[J]. Computer Engin., 2019, 45(3): 250-255, 261.

[6] Baraniuk R G. Compressive sensing [J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.

[7] 焦李成，杨淑媛，刘 芳，等. 压缩感知回顾与展望[J]. 电子学报，2011，39(7)：1651-1662.
Jiao Licheng, Yang Shuyuan, Liu Fang, et al. Development of prospect of compressive sensing[J]. Acta Electronica Sinica, 2011, 39(7): 1651-1662.

[8] Duarte M F, Davenport M A, Takbar D, et al. Single-pixel imaging via compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 83-91.

[9] Sun M J, Edgar M P, Phillips D B, et al. Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning [J]. Opt. Express, 2016, 24(10): 10476-10485.

[10] Zhao Y, Chen Q, Sui X, et al. Super resolution imaging based on a dynamic single pixel camera[J]. IEEE Photon. J., 2017, 9(2): 1-11.

[11] Mao T, Chen Q, He W, et al. Improving signal-to-noise performance of compressive imaging based on spatial correlation[J]. Opt. Rev., 2016, 23(4): 571-578.

[12] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Trans. Inf. Theory, 2007, 53(12): 4655-4666.

[13] Needell D, Vershynin R. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit [J]. IEEE J. Sel. Top. Signal Process, 2010, 4(2): 310-316.

[14] Kim S J, Koh K, Lustig M, et al. A interior-point method for large-scale L1-Regularized least square[J]. IEEE J. Sel. Top. Signal Process, 2007, 1(4): 606-617.

[15] Blumensath T, Davies M E. Iterative hard thresholding for compressed sensing[J]. Appl. & Computational Harmonic Analysis, 2008, 27(3): 265-274.

[16] Daubechies I, DeVore R, Fornasier M. Iteratively reweighted least squares minimization for sparse recovery[J]. Commun. Pure Appl. Math., 2010, 63: 1-38.

[17] Wakin M B, Laska J N, Duarte M F, et al. Compressive imaging for video representation and coding[C]// Proc. of the Picture Coding Symposium, 2006.

[18] Romberg J. Imaging via compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 14-20.

[19] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]// Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition, 2016: 1646-1654.

[20] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution [C]// Proc. IEEE Confer. on Computer Vision and Pattern Recognition, 2016.

[21] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proc. IEEE Confer. on Computer Vision and Pattern Recognition, 2017.

[22] Shi W, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2016: 1874-1883.

[23] Ledig C, Wang Z, Shi W, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]// Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2017: 105-114.

[24] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition, 2018.

[25] Chao D, Chen C L, He K, et al. Learning a deep convolutional network for image super-resolution[C]//Proc. European Conf. on Computer Vision, 2014: 184-199.

[26] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2016: 770-778.

[27] Szegedy C, Ioffe S, Vanhoucke V, et al. Inceptionv4, inception-resnet and the impact of residual connections on learning[C]// IEEE Conf. on Computer Vision and Pattern Recognition, 2016.

[28] Gan L. Block compressed sensing of natural images[C]// Proc. of the Inter. Conf. on Digital Signal Processing, 2007: 403-406.

[29] Fowler J E. Block-based compressed sensing of images and video[J]. Foundations and Trends in Signal Processing, 2012, 4(4): 297-416.

[30] Paszke A, Gross S, Chintala S, et al. Automatic differentiation in PyTorch [C]// NIPS 2017 Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques, 2017.

[31] Zhao H, Gallo O, Frosio I, et al. Loss functions for neural networks for image processing[J]. Computer Science, 2015: 1511.08861.

作者简介：

魏子然(1994—)，男，博士研究生，研究方向为信号压缩与恢复、压缩感知信号处理与图像重建技术，以及单像素成像技术。

(上接第 370 页)

[29] Bejan A. Designed porous media: maximal heat transfer density at decreasing length scales[J]. Int. J. Heat Mass Transfer, 2004, 47(14-16): 3073-3083.

[30] Azoumah Y, Neveu P, Mazet N. Constructal design combined with entropy generation minimization for solid-gas reactors[J]. Int. J. Therm. Sci., 2006, 45(7): 716-728.

[31] Lorente S, Bejan A. Heterogeneous porous media as multiscale structures for maximum flow access[J]. J. Appl. Phys., 2006, 100(11): 114909.

[32] Zhou S B, Chen L G, Sun F R. Constructal entropy generation minimization for heat and mass transfer in a solid-gas reactor based on triangular element[J]. J. Phys. D: Appl. Phys., 2007, 40(11): 3545-3550.

作者简介：

李文龙(1992—)，男，硕士研究生，研究方向为电子器件热管理与热设计；

谢志辉(1977—)，男，博士，教授，博士生导师，主要研究方向为现代热力学、先进热管理理论与技术。

(上接第 379 页)

[20] Look D C, Molnar R J. Degenerate layer at GaN/sapphire interface: influence on Hall-effect measurements[J]. Appl. Phys. Lett., 1997, 70(25): 3377-3379.

[21] Van de Walle C G, Neugebauer J. First-principles calculations for defects and impurities: applications to Ⅲ-nitrides[J]. J. of Appl. Phys., 2004, 95(8): 3851-3879.

[22] Li G, Chua S J, Wang W. The Hall mobility and its relationship to the persistent photoconductivity of undoped GaN[J]. Solid State Communications, 1999, 111(11): 659-663.

作者简介：

李　佳(1995—)，男，湖北钟祥人，硕士研究生，主要从事 GaN 宽禁带半导体薄膜材料的制备及性能研究；

季小红(1972—)，女，江苏南通人，博士，教授，博士生导师，主要研究方向为光电材料的制备、性能和应用研究。