

端对端深度学习无损去图像散射研究

漆建军

(广东外语外贸大学 信息科学技术学院, 广东 广州 510006)

摘要: 针对光线通过介质会产生散射现象, 提出了一种基于神经网络的端对端图像去散射方案, 对于散射退化的图片进行退散射的处理, 此方案不需要复杂的光学设备, 应用场景广泛, 在仿真和实际实验中都取得了预期的结果。提出了一套利用拍摄屏幕来建立散射退化图片和无散射图片的数据集的方法, 且这一方法在许多其他的图像处理图像恢复工作中有应用价值。

关键词: 去散射; 深度学习; 端对端; 图像恢复

中图分类号: TP751

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.09.010

引用格式: 漆建军. 端对端深度学习无损去图像散射研究[J]. 信息技术与网络安全, 2020, 39(9): 49-55.

End-to-end deep learning based non-invasive image descattering

Qi Jianjun

(School of Information Science and Technology, Guangdong University Foreign Studies, Guangzhou 510006, China)

Abstract: The image quality will decay when light goes through the scattering media or reflect on the rough surface. In this paper, we propose an end-to-end deep learning descattering method. This method does not need complex optical device and perform very well in complex scene. This method has a state-of-art performance in both the simulation dataset and real experiment dataset. In this paper, we also creatively put forward a pipeline to collect the hazing and ground truth dataset by capturing the screen, which can also be used in many other field.

Key words: descattering; deep learning; end to end; image reconstruction

0 引言

在采集自然界图片的时候经常会遇见散射造成的图像损失, 对于散射有很多情况, 比如: 当光线经过细小不规则分布的颗粒组成的介质的时候, 光子与细小颗粒发生碰撞, 造成行进的线路发生改变, 产生散射, 最后到达传感器平面的时候不同线路延时的光子叠加, 会造成图像模糊, 造成对比度与清晰度下降的现象; 光照射在粗糙的物体表面, 发生了漫反射, 反射的图像会有类似散斑的现象, 这样的衰减图片对于自动驾驶、计算机视觉、医学成像都有非常不利的影[1]。为了恢复图像, POPOFF S M 等利用测定传输矩阵解卷[2]和生成共轭波前补偿, 但是测量传输矩阵需要费时的测量, 硬件相对复杂, 使用环境相对单一, 而且只能在照明光具有一定相干性的前提下进行[3]。为此, 特提出一种使用深度学习的方法, 通过学习大量散射图

与无散射图对比, 让网络学习到散射图和无散射图之间的映射关系, 避免了费时的传输矩阵测量。并且这样的方法不需要对于样本进行染色或者标记, 不会对样本产生破坏[4]。为此, 将传统的双层卷积结构进行改进, 采用了 dense 结构的 Unet 网络, 不管是在去散射的指标效果还是视觉效果上都有比较大的提升。因为深度学习需要大量的训练数据, 训练数据的好坏对于学习结果起到至关重要的作用, 还提出了一套通过拍摄屏幕的方法来搜集数据集的方法, 主要创新有如下几点。

(1) 提出了一套相机拍摄屏幕创建深度学习训练数据的流程, 包含屏幕相机校准和图像匹配的方法, 此方法也可以广泛用于其他图像处理和图像恢复的任务。

(2) 提出了一种基于深度学习的端对端去散射网络, 相比于传统测定传输矩阵的方法更加省时省

力,且对于成像的样本没有损害,不需要染色或者其他标记。

(3)使用环境相对广泛,采集装置相对简单,仅仅使用普通手机就可以进行采集,且对于照明光源没有太多要求。

(4)改进了像素到像素图像恢复双层卷积的 Unet 结构,采用了 dense 的连接结构,使网络具有更好的恢复图像性能。

1 文献综述

1988 年 FENG S 等人提出了光学的记忆效应 (memory effect)^[5],说明当通过随机的均匀的散射介质的时候,系统具有一定角度的平移不变性。LIU M 等人进一步提出了双层相位掩膜的模拟模型^[6],揭示了记忆效应的范围角度与介质厚度成反比,与波长成正比。EDREI E 等人提出在记忆效应的范围内,光学系统的输出可以写成目标与 PSF 函数卷积的形式^[7],那么从散斑图恢复目标图只用解卷就可以了。POPOFF S M 等人通过散射介质的测量传输矩阵^[2],然后求逆来进行目标恢复。Katz 等人将透过散光片相机采集到的数据进行自相关,这样可以使散射介质的扩散函数高峰更高,拖尾更小,使得求解逆问题更加容易。这样实现了不需要测量波前信息的无损散射介质成像。Vellekoop 等人通过改变每一个 SLM 像素的相位值,使得最后经过波前调制的散射光可以重新汇聚在一点上。ANTIPA N 等人在相机传感器前面加了一个散光镜 (diffuser)^[8]来制作无透镜相机,然后通过 ADMM 的优化算法来进行解卷。MONAKHOVA K 等人在其硬件上提出半经典半深度学习的 Le-ADMM 网络^[9],自动学习先验和数据项之间的权重,从而达到最优。

雾是由空气中许多自用移动的颗粒形成的,因为光子在传播过程中和这些悬浮的颗粒发生了相互作用,导致一部分光被散射,所以从本质上说雾也是散射介质,而且是时变的散射介质。McCartney 提出了雾的大气散射模型,NARASIMHAN S G 和 NAYAR S K 进一步发展了它^[10],基于散射模型的去散射方法都是估计散射模型中的介质传输函数。He Kaiming 等人在 2010 年提出利用了自然界图片普遍存在的图像先验信息暗通道进行去雾^[11]。Cai 等人提出了 DehazeNet 依据 dark prior 的 Max 优化问题,选择了 Max 的激活函数。NAH S 等人在网络中加入了多尺度的信息^[12],多尺度的信息被证明在去

散射的任务中具有作用。

2 散射图像模型

光经过散射介质进入传感器的过程可以看做重建场景和点扩散函数的卷积,如式(1)所示。

$$i(x, y) = \int c(\lambda) \cdot [s(x, y, d, \lambda, i) \times o(x, y)] d\lambda + n(x, y) \quad (1)$$

其中 $o(x, y)$ 是需要重建的目标物体; s 是系统的点扩散函数,当散射介质十分薄的时候,点扩散函数可以近似看作空域不变的^[5]; c 为传感器对与不同光谱的吸收系数; n 为系统噪声; i 为相机传感器采集到的图像信息。已知测量量 $i(x, y)$,目的是求解获取 $o(x, y)$ 的真实值。这是一个典型的病态逆问题。如果直接求解卷,因为这是一个欠定的问题,所以会有无限多可能的解,传统的优化的解决方案是通过图像的先验来限制可行解的范围,然后再通过优化的方法进行求解^[13]。然而这样的缺点是需要手工定制图像的先验,很难找到一个任何场景下都准确的图像先验。另外手动调节损失函数中不同先验之间的权重也是一项十分费力的工作,由此提出了基于深度学习的去散射成像方案,深度学习可以自动学习图像的先验信息^[14],避免繁杂的手动调节参数^[15]。可以看作深度学习的神经网络学习了成像系统的点扩散函数,做了一个解卷的工作。

3 网络结构与参数

Unet 结构是在 pixel2pixel 的图像恢复中是一种十分流行的网络结果^[16]。因为其形状像“U”形所以称为 Unet,这过程分为特征提取和特征重建。图 1 是其网络结构图。前半部分(左边)为特征提取部分,把图像的特征信息缩减到高维。

基本的卷积结构如图 2 所示,由 BatchNorm2d 层、ReLU 层、 3×3 的二维卷积层(padding)和 Dropout 层组成。一个特征尺度到另外一个特征尺度之间选用 2×2 的 Max Pooling 池化层来过滤主要的信息。

维度的变换是采用一个 3×3 的二维卷积接一个 dense 结构连接的卷积。其中 dense 结构的卷积如图 3 所示,一共有五个 3×3 的卷积,第二个卷积以后,每个卷积的输出都与前面每一的卷积做合并操作(concatenation),然后作为下一个卷积的输入。对于特征提取部分,每一个卷积的输出第一都是 32 维。所以 dense 结构的增长率(growth rate)是 16,一个 dense 结构一共有 4 次卷积,所以增加的维度是 64。

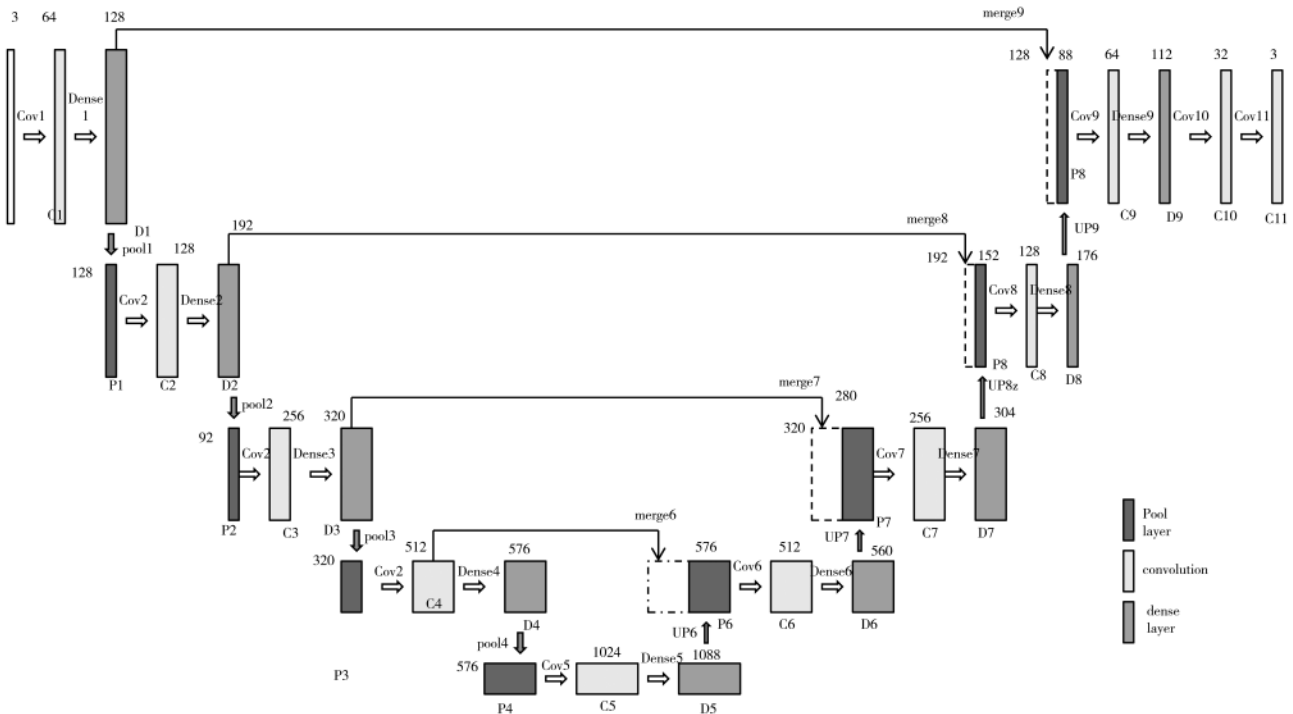


图 1 Unet 结构图

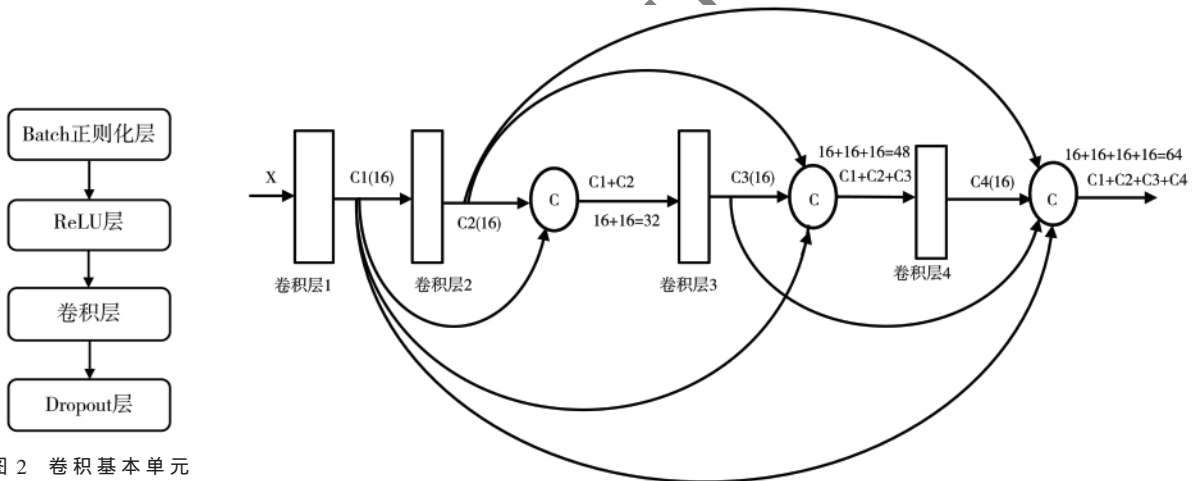


图 2 卷积基本单元

输出层数=输入层数+卷积层数×增长率
卷积层数=4
增长率=16

图 3 dense 的连接结构

对于特征重建部分, dense 结构的增长率为 16, 一共有 3 次卷积, 所以增加的维度是 $3 \times 16 = 48$ 。这样做的优点是和每一个卷积之间都有跳连, 把每一层的信息都利用上了, 同时还能防止因为深度过深而产生的学习梯度消失的问题。

对于 Unet 的重建, 如图 1 右边部分, 可以采用双线性插值的上采样方法, 而不是转置卷积的方法。这样做可以避免转置卷积带来的网格状的噪声, 取得了较好的效果。Unet 的网格噪声和 DescatteringNet

的对比图如图 4 所示。

从图 1 可以看出, Unet 在特征提取和特征重建之间采用合并方式的跳连, 让抽取的每一级信息都可以提供给重建用, 防止特征的丢失。最后输出加了 sigmoid 层, 让输出的图像范围在 0 和 1 之间。

4 仿真结果以及预训练

图 5 是实际实验中光透过散射介质测定的中心视场的点扩散二维函数的标定图与点扩散函数横截线标定图。实验中采用的散射介质是采用磨砂

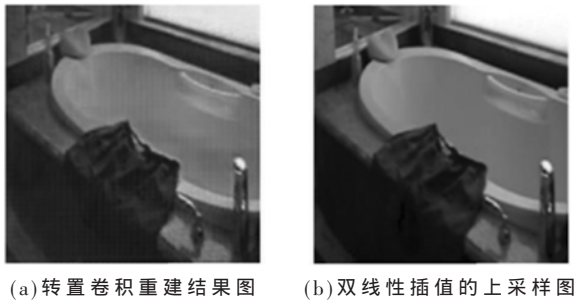
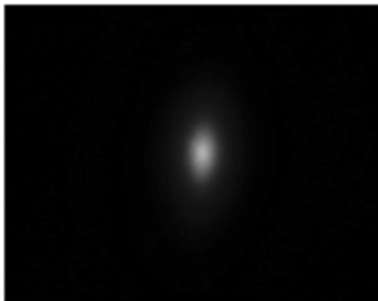
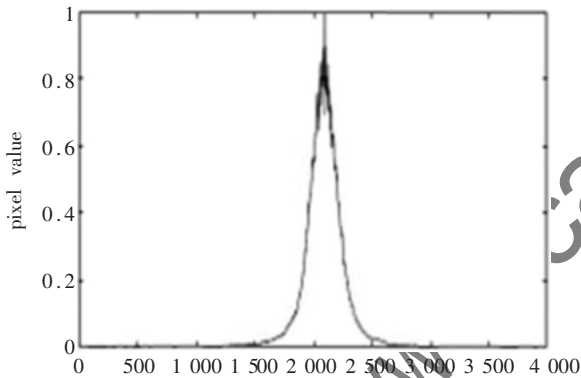


图 4 Unet 的网格噪声和 DescatteringNet 的对比图



(a) 中心视场点扩散二维函数的标定图



(b) 点扩散函数横截线标定图

图 5 两种函数的标定图

纸均匀随机打磨过的薄塑料片。

图 5(a)所示是中心视场的 PSF 标定图,拍摄的是一个直径为 2 mm 的点光源,可以看见点光源经过散射介质以后在传感器上扩散到了很远的距离。图 5(b)这个是中心视场 PSF 的横截线的标定图,可以看出 PSF 核的拖尾很长,能量比较集中的区域就有接近 1 000 个像素。根据实测的 PSF 生成了模拟的散射介质成像的合成数据。仿真采用的 PSF 是由高斯函数和一个长拖尾的锯齿波函数构成的,选用的扩散核在 900 个像素左右,如图 6 所示。

使用仿真数据来进行算法的测试以及使用仿真出来的数据作为实测实验的预训练数据。将使用

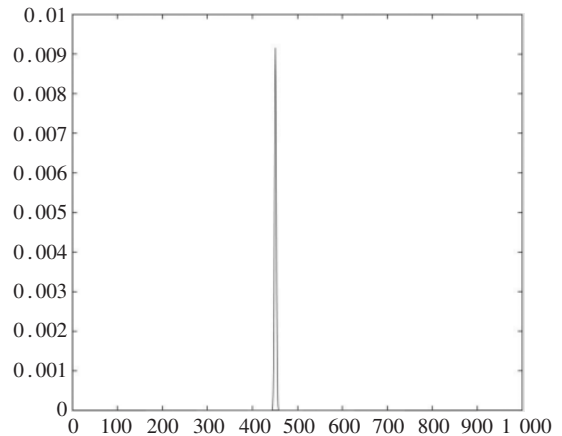


图 6 仿真的点扩散函数

DescatteringNet 和使用 Unet 来去散射的结果对比,得到的结果如图 7 所示。



图 7 使用 DescatteringNet 和使用 Unet 来去散射的结果对比图

从图 7 可以看见 Unet 的恢复结果还是有些灰,颜色不太对,DescatteringNet 的结果明显好于 Unet。可以发现传统的双层卷积的 Unet 的结果并不好,特别是 SSIM 结构信息恢复很差,甚至不如没有恢复前的。出现这种现象是因为 Unet 采用的是转置卷积,而且有很严重的网格噪声,导致结构信息被破坏。现将两种方法的结果指标进行对比,如表 1 所示。

对于更加严重的散射退化的图像,进行了对比

表 1 轻度散射仿真结果 PSNR 值与 SSIM 值比较

	PSNR	SSIM
散射图	22.568 9	0.815 7
Unet	23.681 3	0.737 4
DescatteringNet	29.784 0	0.838 0

实验,图 8 为 Unet 和 DescatteringNet 严重散射仿真结果对比图,从图中可以看出传统 Unet 网络存在很严重的色彩溢出问题,并且有光晕噪声;DescatteringNet 颜色的连续性更强,也更加接近原图。



图 8 Unet 和 DescatteringNet 严重散射仿真结果对比图

同时,将 UNet 和 DescatteringNet 严重散射 PSNR 值与 SSIM 值进行比较,得到的结果如表 2 所示。

表 2 Unet 和 DescatteringNet 严重散射 PSNR 值与 SSIM 值比较

	PSNR	SSIM
散射图	19.788 8	0.718 3
Unet	20.560 6	0.600 4
DescatteringNet	21.232 8	0.729 8

从表 2 可以明显看出,DescatteringNet 比 Unet 提升了接近 0.7 dB,SSIM 提升了接近 10%。

5 实际训练数据采集和恢复结果

训练数据的采集采用拍摄屏幕的方法,实验数

据收集装置如图 9 所示,采用手机在黑暗的环境中拍摄电脑屏幕。为了尽量模拟实际应用,采用华为荣耀 20i 的后置摄像头进行拍摄,手机拍摄的数据作为需要恢复的散射图,电脑屏幕上显示的图片作为真实图,手机后置摄像头分辨率为 5 072×6 896,拍摄的屏幕是分辨率是 1 920×1 080,在黑暗环境中进行拍摄。为了避免因为手机空间分辨率和显示器空间分辨率接近而产生的摩尔纹,将相机稍微远离屏幕一点点。全程采用手动模式,包括白平衡也是手动。先拍摄黑白的棋盘格图片,用来进行标定,然后拍摄全白和全黑的图片,用来计算图片的相对像素值。在实际的实验中,采用的散射介质是一块十分薄的表面粗糙的透明玻璃,使用透明胶紧贴相机镜头固定,拍摄的屏幕是 GS65 笔记本电脑高对比度屏幕。



图 9 获得实验数据的采集装置图

5.1 相机屏幕校准

相机屏幕校准直接影响训练数据的质量好坏,可以分为以下几个步骤进行相机屏幕校准。

首先需要对于屏幕色彩进行校准,采用的是 il Pro 的屏幕色彩校准仪进行校准。校准散射介质的镜头的畸变,采用的方式是将棋盘图显示在屏幕上,然后采用贴有散射介质的相机镜头对屏幕上的棋盘格各个角度拍摄多张照片,使用多项式对畸变进行拟合与矫正。

然后固定相机位置拍摄棋盘格图片,并拍摄收集到的训练数据集,数据集采用的是 Adobe4K 数据集,全程采用手动模式,相机的参数固定不变。除了拍摄以上图片以外,还需拍摄一张全黑和全白的图片,对于拍摄到的棋盘格图片,还需做角点检测,然后计算单应矩阵,对图片做变形,配准屏幕上的图片和拍摄图片数据。通过把拍摄的全黑的图片作为黑色 0 值,算出相对颜色的值,再通过全白色的图片做一个去暗角的操作。

5.2 实验结果

实验采集了 814 组实验数据,因为 GPU 显存限制,在实际操作的时候要对于图片进行裁剪,随机裁剪 512×512 的大小块送入网络进行训练,考虑到薄玻璃具有“光的记忆效应”,同一张图片上的点扩散函数应该具有空间不变性,所以这样分块学出来的点扩散函数与整体的图片应该是一样的,因此在推理阶段,需将整张图片送入内存,利用 CPU 进行推断。采用薄透明玻璃作为散射介质采集到的图片去散射推理结果如图 10 所示,通过观察发现推理阶段的 PSNR 和 SSIM 与训练阶段比较相似,说明关于点扩散函数的空间不变性假设得到了验证。

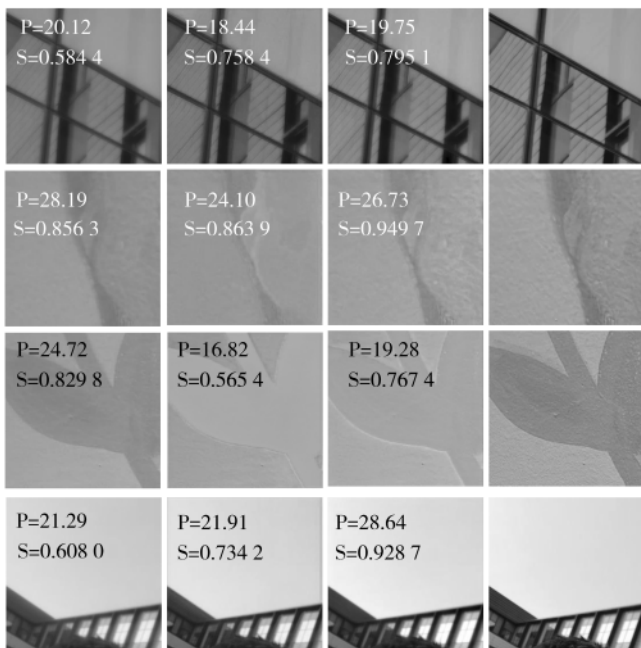


图 10 DescatteringNet 和 Unet 实测结果对比图

同时,将在这个过程中得到的 DescatteringNet 和 Unet 实测实验结果进行对比,如表 3 所示。

表 3 DescatteringNet 和 Unet 实测实验结果对比

	PSNR	SSIM
Unet	21.24	0.770 0
DescatteringNet	23.92	0.852 1

可以看出两种网络结构都可以起到恢复图像的作用,但是可以明显发现 Unet 有严重的光晕现象,而 DescatteringNet 的结果要好许多,在结果上 DescatteringNet 在实际散射图像的恢复上比 Unet 的结果

在 PSNR 上要好 1 dB 左右,SSIM 要好 0.1 左右。实验表明采用仿真数据的预训练的结果比直接从头开始训练的结果要好,因为预训练可以让优化避开许多的局部最优点,从而尽量落入全局最优点。为显示预训练与无预训练的结果差别,将预训练与无预训练结果的 PSNR 值与 SSIM 值进行对比,如表 4 所示。

表 4 预训练与无预训练结果 PSNR 值与 SSIM 值的对比

	PSNR	SSIM
无预训练	20.69	0.853 9
预训练	22.27	0.887 8

6 结论

通过提出一种端到端的深度学习去散射的方法,得到了一种可以避免传统方法需要手动设定图像先验的麻烦和避免复杂的设计系统,而只需要一套手机摄像头大小的传感器即可,且在仿真数据集和实测数据集上都做了详细的实验,得到的实验结果一致,验证了 DescatteringNet 拥有比之前 Unet 更好的表现。除此以外,这种处理方法对于样本没有损害,测试运算量小,速度快。从本质上来说,可以理解为神经网络学习了散射图和参考图之间的点扩散函数,然后进行了一套解卷的操作。除此以外,还提出了一套拍摄屏幕搜集创建数据集的工作流程,且这套流程在大部分图像处理和图像恢复的任务中都可以使用。

值得说明的是,DescatteringNet 虽然比 Unet 在视觉效果和指标上都更好一些,但是其与原图还是有一定的差距,特别是在颜色恢复和细节处理上还有一定的差距。这也是目前这个领域中需要进一步探索的问题。

参考文献

- [1] DAI Q, CHOPP H, POUYET E, et al. Adaptive image sampling using deep learning and its application on X-Ray fluorescence image reconstruction[J]. arXiv: 1812.10836, 2018.
- [2] POPOFF S M, LEROSEY G, CARMINATI R, et al. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media[J]. Physical Review Letters, 2010, 104(10): 100601.
- [3] KATZ O, HEIDMANN P, FINK M, et al. Non-invasive

- real-time imaging through scattering layers and around corners via speckle correlations[J].Nature Photonics, 2014, 8(10): 784-790.
- [4] BALTER S, HOPEWELL J W, MILLER D L, et al. Fluoroscopically guided interventional procedures: a review of radiation effects on patients' skin and hair[J]. Radiology, 2010, 254(2): 326.
- [5] FENG S, KANE C, LEE P A, et al. Correlations and fluctuations of coherent wave transmission through disordered media[J]. Physical Review Letters, 1988, 61(7): 834-837.
- [6] Liu Honglin, Liu Zhentao, Chen Meijun, et al. Physical picture of the optical memory effect[J]. Photonics Research, 2019, 7(11): 1323.
- [7] EDREI E, SCARCELLI G. Memory-effect based deconvolution microscopy for super-resolution imaging through scattering media[J]. Rep, 2016, 6(1): 33558.
- [8] ANTIPA N, KUO G, HECKEL R, et al. DiffuserCam: lensless single-exposure 3D imaging[J]. Optica, 2017, 5(1): 1-9.
- [9] MONAKHOVA K, YURTSEVER J, KUO G, et al. Learned reconstructions for practical mask-based lensless imaging[J]. Optics Express, 2019, 27(20): 28075.
- [10] NAYAR S K, NARASIMHAN S G. Vision in bad weather[C]. Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 2002.
- [11] HE K, SUN J, TANG X, et al. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(12): 2341-2353.
- [12] NAH S, KIM T H, LEE K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.
- [13] VENKATAKRISHNAN S V, BOUMAN C A, WOHLBERG B. Plug-and-play priors for model based reconstruction[C]. Global Conference on Signal & Information Processing. IEEE, 2014.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [15] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[J]. arXiv: 1703.10593, 2017.
- [16] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.

(收稿日期: 2020-06-04)

作者简介:

漆建军(1966-),男,博士,副教授,主要研究方向:光信息处理。

电子技术应用“射频微波主题宣传季”帷幕正式拉开

为了全面系统地介绍微波射频领域的最新研究成果,展示射频微波领域的最新技术与产品,促进射频微波领域产学研用各界的交流,《电子技术应用》杂志、《信息技术与网络安全》杂志、ChinaAET网站拟于2020年7月~9月共同举办“射频微波主题宣传季”活动。

作为国内射频微波领域重要的媒体平台,《电子技术应用》杂志及网站聚集了一大批优秀的射频与微波技术工程师及研发人员,影响力贯穿整个射频微波产业链。本次活动,旨在搭建射频微波电路、模块与通信系统的企业及科研院所的研究人员探讨新技术的交流平台。

活动期间,围绕“5G通信”、“低小慢飞行器设计与防御”、“高精度室内定位”等射频与微波领域的热点方向,主办方将推出形式多样、内容丰富的系列活动,全面介绍射频微波领域的学术研究成果、最新技术及产业动向、新产品及应用。活动内容包括期刊“射频与微波”特约技术专栏、ChinaAET.com网站“射频微波”专题、“RFMW论坛”系列线上直播等。

活动主站: <http://ec.chinaaet.com/show/rf2020/>,期待射频微波领域的从业人员大力关注与参与!

活动联系人: 王伟(电话 010-62311179; Email: wangw@chinaaet.com)。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所