

深度递归卷积神经网络的偏振应用

李林^{1,2}, 段济杰^{1,2}, 马辉^{1,3}, 常金涛³, 周嘉玲³, 韩瑞芳³

1. 清华大学深圳研究生院深圳市无损和微创医疗技术重点实验室, 广东 深圳 518055; 2. 清华大学生物医学工程系, 北京 100084;
3. 清华大学物理系, 北京 100084

【摘要】衡量焦平面分割(DoFP)偏振相机插值算法好坏时,有两个备受关注的误差量:线性偏振度(DoLP)和偏振角度(AoP)。将深度递归卷积网络(DRCN)插值算法引入到偏振相机的插值上来。DRCN插值算法是全卷积的一个网络。整个网络的卷积核大小为 3×3 ,在高分辨率图像重建部分采用16层的卷积递归,减少了网络的训练参数数目。在文中给出了DRCN单幅图像的插值误差、DoLP以及AoP的插值误差,误差采用常用的均方误差根进行计算。并将DRCN算法与双三次插值进行比较,从4幅中的单幅图像插值效果来看,该算法插值误差比双三次插值的误差小。同时在计算出来的DoLP、AoP的插值误差上,该插值算法也要好于传统的双三次插值算法。文章中给出了定性的比较,利用两个不同分辨率的DoFP偏振相机对同一场景进行拍摄,可以看出DRCN插值后计算出来的DoLP比双三次计算出来的DoLP更加接近于真实的DoLP。

【关键词】深度递归卷积神经网络;偏振;插值;线性偏振度;偏振角度

【中图分类号】Q-334

【文献标志码】A

【文章编号】1005-202X(2017)03-0240-06

Deeply recursive convolutional neural network for polarization

LI Lin^{1,2}, DUAN Chaijie^{1,2}, MA Hui^{1,3}, CHANG Jintao³, ZHOU Jialing³, HAN Ruifang³

1. Shenzhen Key Laboratory for Nondestructive and Minimal Invasive Medical Technologies, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China; 2. Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China; 3. Department of Physics, Tsinghua University, Beijing 100084, China

Abstract: To evaluate the interpolation algorithm for division-of-focal-plane (DoFP) polarization camera, two errors are highly concerned, degree of linear polarization (DoLP) and angle of polarization (AoP). Herein, we introduce deeply recursive convolutional network (DRCN) interpolation algorithm into the interpolation of the polarization camera. DRCN interpolation algorithm is a network of full convolution. The convolution kernel is of size 3×3 , and recursive convolution with 16 layers is used for the reconstruction of the high resolution image, which reduces the number of training parameters of the network. In this paper, we give the interpolation error of single image, interpolation error of DoLP and AoP, calculated by using root mean squared error. The DRCN algorithm is compared with the bicubic interpolation algorithm. In terms of the interpolation of a single image, the result showed that the interpolation error of the DRCN interpolation algorithm is less than that of bicubic interpolation algorithm. And the calculated interpolation errors of DoLP and AoP with DRCN interpolation algorithm are also less than those with bicubic interpolation algorithm. We also perform a qualitative comparison between the two algorithms, using two DoFP polarization cameras with different resolutions to shoot the same scene. Calculated DoLP with DRCN interpolation algorithm was more similar to the real DoLP than the calculated DoLP with bicubic algorithm.

Keywords: deeply recursive convolutional neural network; polarization; interpolation; degree of linear polarization; angle of polarization

【收稿日期】2016-12-10

【基金项目】国家自然科学基金(81230035, 61527826)

【作者简介】李林, 硕士研究生, 研究方向: 偏振光图像插值, E-mail: 1978382790@qq.com

【通信作者】段济杰, 助理研究员, 硕士生导师, 研究方向: 医学图像处理, E-mail: duan.chaijie@sz.tsinghua.edu.cn

前言

焦平面分割(Division of Focal Plane, DoFP)偏振相机因其结构独特,需要插值算法的支持^[1]。对于DoFP偏振相机的插值工作,已经有不少研究。最开始是线性插值,线性插值^[2]是利用已知像素点来线性计算未知的像素值,可以视为数据拟合。在众多的插值算法中,双三次系列^[3]的插值方法性能较好。但是因为插值的时候也是利用周边已知的像素点进行数据拟合,得到未知的像素点值,在边界处会存在较大的拟合误差。其实对于单幅图像来说,这种拟合误差是可以接受的,然而当双三次插值逐次应用到4幅偏振角度各异的图像,然后计算线性偏振度(Degree of Linear Polarization, DoLP)和偏振角度(Angle of Polarization, AoP)时,就可能会在边界处存在着较大误差。

纵观提出的各个算法,我们可以看到这些DoFP算法无一不是应用了单幅图像的插值算法。双线性和双三次插值算法的实现均是参考了文献^[4]。单幅图像的插值对DoFP偏振相机的插值应用具有指导性意义。如果我们将上边提及的4幅图像看成是独立的图像,那么单幅图像的所有插值算法均可以使用。对于偏振相机插值应用,我们衡量算法好坏的标准就是DoLP和AoP的插值误差。

卷积神经网络(Convolutional Neural Network, CNN)用于图像处理领域已经取得比较好的成绩。CNN可用于人脸识别^[5]、胸部组织分类^[6]、图像超分辨率重建^[7]。由于CNN在图像超分辨率重建上突出的性能表现,CNN在重建高分辨率图像的应用上的结构被改进。CNN卷积操作简便,使得CNN网络架构清晰明了。不少网络结构改成了CNN的模型,据称5层的CNN等效于稀疏编码器。CNN的网络参数对网络性能的影响也有不少研究^[8]。神经网络的参数会随着深度的增加而增加,CNN也不例外。为了减小网络参数,网络采用重复出现的结构^[9]。其实对于神经网络来说,我们希望网络越深越好,因为此时我们进行模式识别时所能看到的范围越广,那么做出来的决策可信度就会越高。当CNN用于图像高分辨率图像重建时,我们也希望接受野越大越好,也就是说网络深度越深越好。为了控制随网络深度增加而增加的参数量,高分辨率图像细节信息的重建部分采用一种递归型的推导层,也就是深度递归卷积网络(Deeply Recursive Convolutional Network, DRCN)^[10]。

目前关于DoFP偏振相机的研究,着重于提高算

法在边界上的处理性能。而DRCN方法是边界增强的算法,该算法在边界处的处理性能优越。本文在前人的基础上,结合偏振领域对插值图像精准度的要求,将DRCN调整为可以应用到DoFP偏振相机的插值算法。从DoLP和AoP的误差来看,该算法性能很好。该方法在单幅偏振图像上的性能比双三次插值性能要好的多,但是性能还有待提高。本文的意义在于将神经网络算法引入DoFP偏振相机的插值上,同时将该算法与双三次插值算法进行定性和定量的比较,这为我们进一步改进神经网络算法并应用于DoFP偏振相机的插值问题上提供了方向。

1 方法概述

1.1 卷积层

卷积层的主要任务是为了提取特征量。我们知道,一般的卷积核可以用来提取图像的纹理信息。如图1所示,空心箭头表示卷积操作,箭头上给出了卷积核,最左边是图像,空白处像素值为0。提取水平方向上的纹理信息,卷积核1,如图1a;提取竖直方向上的纹理信息,卷积核2,如图1b;提取45°方向上的纹理信息,卷积核3,如图1c;提取135°方向上的纹理信息,卷积核4,如图1d。第一次卷积运算,结果得到第一次卷积结果;再对卷积结果用同一卷积核进行卷积运算,得出的结果就是一个数。从各个卷积与图案的乘积结果来看,只要用修正线性单元(Rectified Linear Unit, RELU)将最终结果中为负值的部分(图1a、b)去掉,很容易得到,只有对角线方向的卷积核的卷积结果满足要求(图1c、d),也就是说判断出存在对角线两个方向上的纹理信息,这与图案中只存在对角线方向的纹理信息是一致的。

由于不同的卷积核和图像卷积可以提取出不同的纹理信息,而不同纹理信息的组合可以得到图像内部的边界信息。比如说,卷积核3和卷积核4的卷积结果组合,再加上它们在原图像中的相对位置,我们就很容易判断,原图像中的图案是‘X’。

1.2 RELU

RELU是网络实现非线性的重要环节。在前边的描述中,可以看到RELU可以辅助我们筛选出正确的结果。

1.3 DRCN

既然卷积核可以用来提取图像的边界,而我们传统的插值算法在边界处的插值误差是最大的。那么我们可以设计网络来校正传统插值方法在各种像素点分布下的插值误差,虽然,传统插值方法会在整

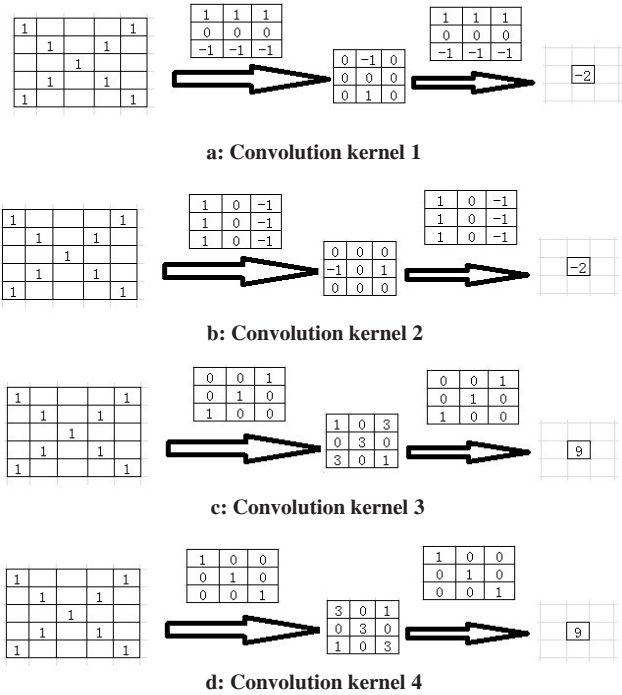


图 1 卷积核提取图像特征过程

Fig.1 Process of convolution kernel to extract image feature

幅图像范围内都有或多或少的插值误差,但是相比之下,边界处的插值误差会大一些。因此,最终的神经网络模型产生了这样的效果,给定一幅输入图像,网络模型的输出校正量将类似于提取到了图像的边界;因为模型在边界附近的校正量大,而在其它位置的校正量小。这是统计意义上的校正值输出。校正值输出加上传统插值的预估值,得到最终的高分辨率图像。与传统插值相同点之处是都利用了附近已知像素点的信息,不同之处在于,传统插值只是对数据的拟合进行插值,而 DRCN 则是统计意义上的校正插值。

1.4 DRCN 方法用于 DoFP 偏振图像

训练 DRCN,如文献[10]所描述,使用了文献[11]的 91 幅图。测试用了 4 个数据集: Set5 和 Set14^[12-13]、B100^[14]、Urban100^[15]。16 次递归,共 20 个卷积层(达到 41*41 的接收域),Momentum parameter 动力参数 0.9,权重衰减 0.000 1,3*3 滤波器每一层有 256 个。训练时使用的损失函数:

$$l_1(\theta) = \sum_{d=1}^D \sum_{i=1}^N \frac{1}{2DN} \left\| y^{(i)} - \hat{y}_d^{(i)} \right\|^2 \tag{1}$$

$$l_2(\theta) = \sum_{i=1}^N \frac{1}{2N} \left\| y^{(i)} - \sum_{d=1}^D \omega_d \cdot \hat{y}_d^{(i)} \right\|^2 \tag{2}$$

$$L(\theta) = \alpha l_1(\theta) + (1 - \alpha) l_2(\theta) + \beta \left\| \theta \right\|^2 \tag{3}$$

其中,公式(1)为中间各层的递归输出的偏差量,公

式(2)为最终的重建结果的偏差量,从最终的损失函数公式(3)来看,该损失函数不仅考虑了最终结果与真实图像的误差,还考虑了每一次迭代的中间结果与真实中间结果之间的误差,最后一项是衰减因子乘上参数集。

根据 DRCN 模型的框架,设置各层的初始化参数;然后根据各层功能的不一样按次序分别调用 MatConvNet 中相应的功能模块;各层得到的结果与真实结果之间偏差为依据,将各层的偏差取平均,以减小迅速梯度增长或迅速梯度下降的情况,根据偏差量依次得到网络中各层的参数校正量,直到偏差条件满足,结束训练。

使用 DRCN 模型:将训练得到的模型提取出来。使用模型时的网络结构与训练的一样,只是将参数校正部分去掉。由于训练的时候,输入图像是将真实的图像用双三次插值(Bicubic)核卷积升采样之后,再用双三次插值核卷积降采样之后得到的。也就是说,输入图像与输出图像同大小。那么,使用的时候,输入图像也是将真实的图像用双三次插值核卷积升采样之后,再用双三次插值核卷积降采样之后得到的。DoFP 相机得到的是 4 张不同偏振角度(即 0°、45°、135°、90°)下的偏振图,它们缺失信息的位置是不一致的。因此需要通过数据拟合的方式(比如说双三次拟合、双线性拟合)事先计算好各个偏振角度图片中的未知值,这就保留了全部的已知像素值。通过数据拟合的方式得出的 4 张“真实图”,严格来说不能当作“真实值”。

对数据拟合的方式得到的真实值进行分析:在平滑区域,数据拟合的结果将会与真实结果差别很小;而在边界区域,拟合结果与真实结果差别相对来说较大。

如果将数据拟合的结果(用前边提及到的数据拟合得到的)当作真实值,会不会导致 DRCN 网络最终的重建结果接近于数据拟合的结果,也就是说, DRCN 网络最终的重建结果可能比数据拟合的结果要差些? 答案是否定的。因为数据拟合方式保留了全部的已知像素点,也就是说,不是完全不知道真实的偏振图像信息的;其次,拟合数据的结果虽然与真实图像在某些像素点上有差别,但是它们在图片特征上的差别是很小的,而 DRCN 模型记录的是,根据提取到的图片特征给出相应的校正量。因此不管输入是否为真实图片,只要输入与真实差别不大,可以视为真实图片。因此,最终的结果不会无限近似于数据拟合结果。

$$\text{DoLP} = \sqrt{\frac{(I_0 - I_{90})^2 + (I_{45} - I_{135})^2}{(I_0 + I_{90})^2}} \tag{4}$$

$$\text{AoP} = 0.5 \times \arctan\left(\frac{I_{45} - I_{135}}{I_0 - I_{90}}\right) \tag{5}$$

其中, I_0 、 I_{45} 、 I_{135} 、 I_{90} 分别表示 0° 、 45° 、 135° 、 90° 偏振图像的强度。

根据图 2 所示,我们可以将方法归结如下: Step 1: DoFP 相机拆分出 4 张偏振图片 (0° 、 45° 、 135° 、 90°), 每张偏振图片大小与 DoFP 图片大小一样, 它们保留了已知像素值, 同时在未知像素值所在位置上填充 0; Step 2: 将 0° 、 45° 、 135° 、 90° 偏振图片用双三次数据拟合的方法, 得到“真实图”; Step 3: 将调用 Matlab 中 `imresize` 函数, 用 Bi-cubic 方法, 对 4 张“真实图”进行放大 2 倍, 然后缩小 2 倍; Step 4: 调用 DRCN 模型, 得到 4 张重建偏振图像; Step 5: 根据公式计算 DoLP 公式(4)和 AoP 公式(5)。

2 实验数据

用不同分辨率的 DoFP 相机在同一个角度拍摄同样的场景。用高分辨率相机得到的 DoFP 图片 (1 200*1 560) 的平均值线性插值结果作为真实的高分辨率图, 并计算出真实的 DoLP。而低分辨率相机得到的 DoFP 图片 (488*648) 是进行算法之间比较的对象, 分别用双三次和 DRCN 算法插值后, 并计算出 DoLP, 肉眼观察插值结果之间的优劣。

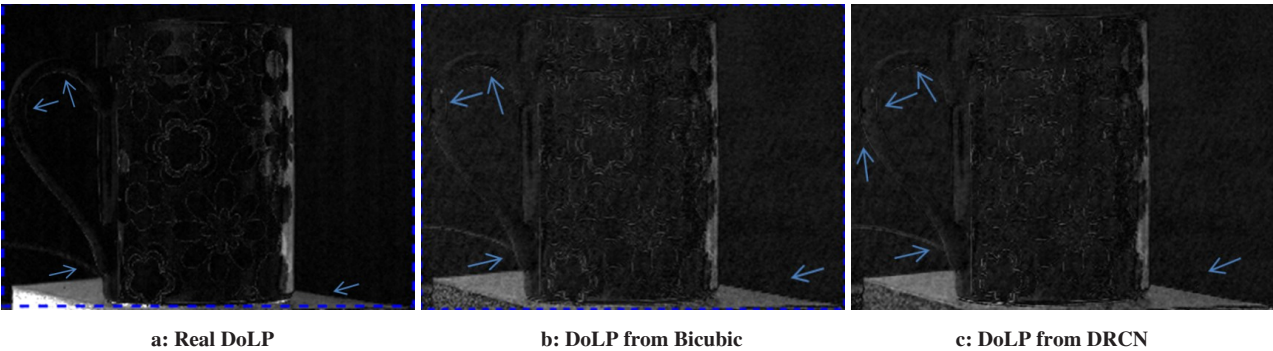
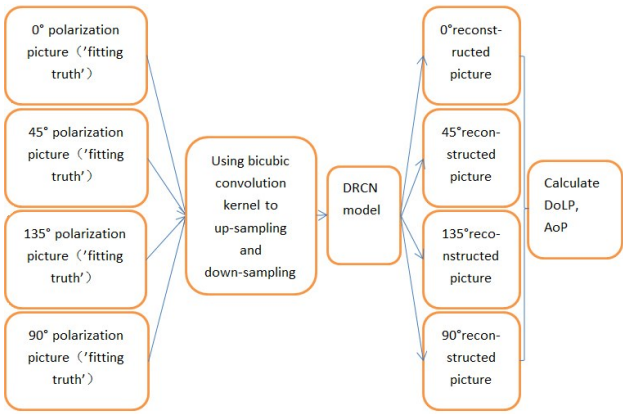


图 3 DoFP 相机拍摄的茶杯图片对应的 DoLP

Fig.3 DoLP for a cup image from DoFP camera

从参考文献[10]我们可以清楚地看出,DRCN 对于单幅图像插值来说在边界处的表现是相当好的, 在前面已经论述了该方法从理论上来说在边界处不存在跨边界插值情况, 因为它只是统计意义上的插值。在实际的 DoFP 偏振图像中, 我们得出 DoLP 图像比双三次插值要好。但是我们仍然需要一些定量的数据来说明。



DoFP: Division of focal plane; DRCN: Deeply recursive convolution network; DoLP: Degree of linear polarization; AoP: Angle of polarization

图 2 DoFP 偏振相机的插值流程图

Fig.2 Interpolation flow chart of DoFP polarization camera using DRCN interpolation algorithm

因为两款相机分辨率不一样, 视场不一样, 因此只取高分辨率和低分辨率图的公共部分进行比较。在这里, 我们只是将各种算法计算出的低分辨率的 DoLP 与高分辨率的 DoLP 进行定性比较。

图 3b 和图 3c 进行比较。如图上箭头所示, 图 3c 比图 3b 的锯齿状边界要少很多。将图像放大后可以清楚地看到, 图 3c 茶杯上的花纹更加相似于真实图 3a 的花纹。因此, DRCN 要好于双三次插值。

3 验证

偏振相机(灰度相机前加一个可旋转的偏振片), 采集 4 幅图像分别是 0° 、 45° 、 135° 、 90° 偏振图像, 它们作为真实的偏振图像。图像大小为 1 392*1 040。对真实图像进行降采样, 然后分别用双三次和 DRCN 插值回去。以 DRCN 算法为例, 利用训练好的 DRCN 模型, 根据上边的计算步骤, 进行实验。具体实验验

证方法参见文献[16]。

球晶样本的双三次插值和DRCN插值与真实偏振图像的各项误差如表1所示。

图4分别给出了真实的(图4a)、双三次插值(图

4b)、DRCN插值(图4c)的球晶样本的DoLP、AoP。球晶样本上的纹理信息非常丰富,所以这些图片从直观上看不出太多的差别。

表1 球晶样本的双三次插值和DRCN插值关于各项误差的定量比较

Tab.1 Quantitative root mean squared error (RMSE) comparison between bicubic interpolation algorithm and DRCN interpolation algorithm on the spherulite sample

Interpolation algorithm	0°	45°	135°	90°	err_D	err_A	av_single_im
Bicubic	2.573 142	4.804 586	2.256 014	2.565 373	0.032 722	0.148 347	3.049 778 366
DRCN	2.548 443	4.214 659	2.204 244	2.499 003	0.030 583	0.146 438	2.866 587 275

0°, 45°, 135° and 90° represent RMSE between the real image and image calculated by the two interpolation methods, respectively; err_D and err_A represent RMSE of DoLP and AoP between the real and interpolated results by the two methods; averaging RMSE of single image (0°, 45°, 135° and 90°) is abbreviated to av_single_im.

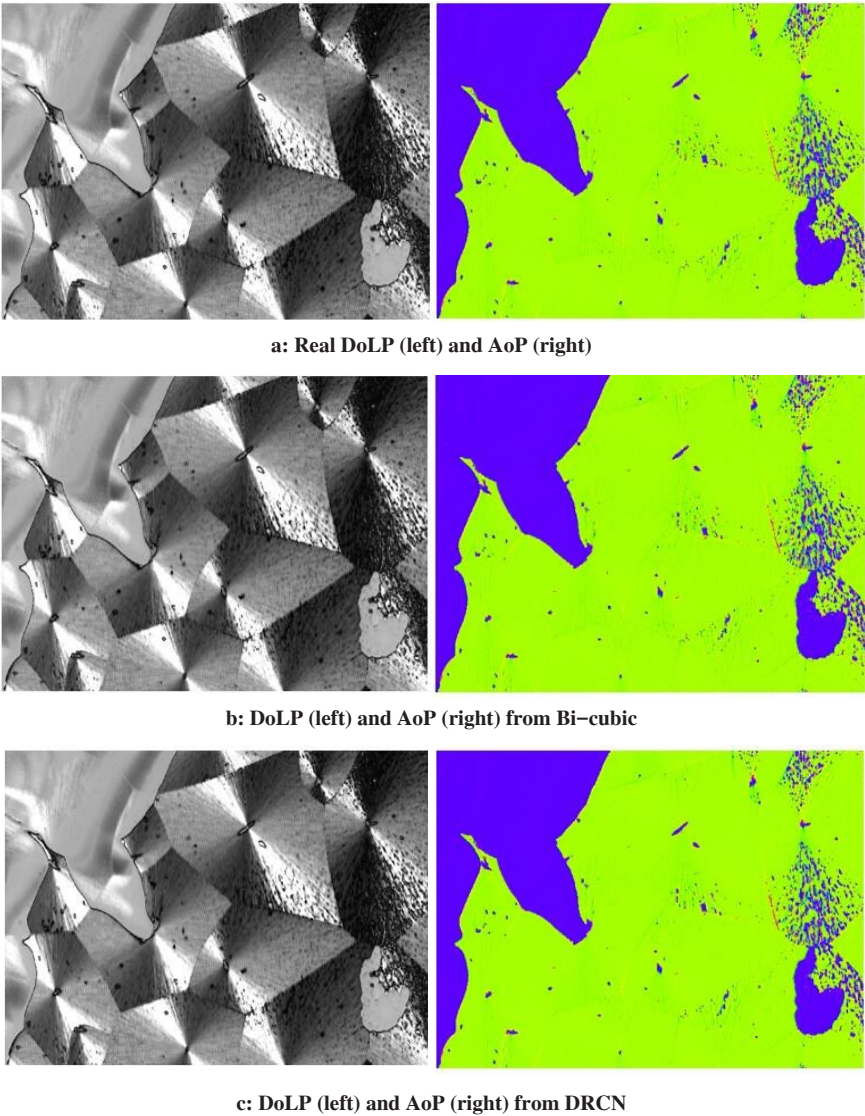


图4 球晶样本真实的双三次插值、DRCN插值计算出来的DoLP(左)和AoP(右)

Fig.4 DoLP and AoP of real and calculated by bicubic and DRCN for spherulite samples

如表1所示,对于单幅图像的插值效果来看,每一幅图DRCN都好于双三次插值方法。在表1中所给出的单幅图像平均误差可以看到,DRCN的RMSE误差大约要比双三次插值的小0.18。计算出来的DoLP和AoP误差DRCN比双三次插值要小些,单幅图像的误差比双三次更小些。可以看出DRCN与双三次插值方法误差接近,因为这些误差值本身就很小。虽然难以在图像上看出太大的差别,但是从定量结果可以看出,DRCN方法的确优于双三次,即使在细节信息如此丰富的图像上。

4 结 论

在本文中,我们将DRCN算法应用到了DoFP偏振相机的插值问题上。从实验结果来看,整体上来说,尤其是单幅图像,DRCN的性能要好于双三次插值。因为DRCN使用统计信息进行插值,所以插值的时候不用额外考虑边界附近像素点的插值策略。因此,插值出来的图像在边界处不存在跨边界插值的情况。当然也并不是每一个样本都会得到每一项的误差都要好于双三次插值。原因有两个:其一,因为偏振的4张图(0° 、 45° 、 135° 、 90°)明暗变化是很大,可能有的图整幅都偏暗,那么会受计算精度的影响;其二,同时该方法是统计意义上的插值,训练样本不可能面面俱到。DRCN除了依赖于训练样本,还有初始参数的设置等,从目前的结果来看,还是令人比较满意的。从理论上来说,该方法可以胜任的图像范围非常广,只要足够多的训练样本。从实用上来说,目前的训练数据可以在大多数场景中表现很出色。本文首次将神经网络应用到DoFP偏振相机的插值上,这对于后续的研究具有指导性意义。

【参考文献】

- [1] GRUEV V, KULKARNI M. Division-of-focal-plane spectral-polarization imaging sensor [J]. *Proc SPIE Int Soc Opt Eng*, 2012, 8364(3): 83640K-11.
- [2] RATLIFF B M, LACASSE C F, TYO J S. Interpolation strategies

- for reducing IFOV artifacts in microgrid polarimeter imagery [J]. *Opt Express*, 2009, 17(11): 9112-9125.
- [3] GAO S K, GRUEV V. Bilinear and bicubic interpolation methods for division of focal plane polarimeters [J]. *Opt Express*, 2011, 19(27): 26161-26173.
- [4] RUSSELL W S. Polynomial interpolation schemes for internal derivative distributions on structured grids [J]. *Appl Numer Math*, 1995, 17(2): 129-171.
- [5] LAWRENCE S, GILES C L, TSOI A C, et al. Face recognition: a convolutional neural-network approach [J]. *IEEE Trans Neural Netw*, 1997, 8(1): 98-113.
- [6] SAHINER B, CHAN H P, PETRICK N, et al. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images [J]. *IEEE Trans Med Imaging*, 1996, 15(5): 598-610.
- [7] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2016, 38(2): 295-307.
- [8] EIGEN D, ROLFE J, FERGUS R, et al. Understanding deep architectures using a recursive convolutional network [J]. *Comput Sci*, 2013.
- [9] LIANG M, HU X L. Recurrent convolutional neural network (RCNN) for object recognition [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2015: 3367-3375.
- [10] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2016: 1637-1645.
- [11] YANG J, WRIGHT J, HUANG T S, et al. Image super resolution via sparse representation [J]. *IEEE Trans Image Process*, 2010, 19(11): 2861-2873.
- [12] TIMOFTE R, SMET V D, GOOL L V. A+: adjusted anchored neighborhood regression for fast super-resolution: *Computer Vision--ACCV 2014* [C]. Berlin: Springer International Publishing, 2014.
- [13] TIMOFTE R, SMET V D, GOOL L V. Anchored neighborhood regression for fast example-based super-resolution [C]//*IEEE International Conference on Computer Vision*. Sydney: ICCV, 2013.
- [14] MARTIN D R, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological Statistics [C]//*IEEE International Conference on Computer Vision*. Barcelona: ICCV, 2001.
- [15] HUANG J B, SINGH A, AHUJA N. Single image superresolution using transformed self-exemplars [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2015: 5197-5206.
- [16] GAO S K, GRUEV V. Gradient-based interpolation method for division-of-focal-plane polarimeters [J]. *Opt Express*, 2013, 21(1): 1137-1151.

(编辑:薛泽玲)