

〈图像处理与仿真〉

基于残差编解码网络的红外图像自适应校正算法

牟新刚, 陆俊杰, 周 晓

(武汉理工大学 机电工程学院, 湖北 武汉 430070)

摘要: 针对基于场景的非均匀性校正算法存在非均匀性残余和鬼影等问题, 本文提出了一种基于残差编解码网络的红外图像自适应校正算法。该算法针对自适应校正问题的特点, 基于 UNet 结构, 通过多尺度采样学习残差映射生成非均匀性残差图像, 加入批标准化和 PReLU 激活函数提高校正效果, 最后使用全局跳跃连接得到最终的校正结果。通过对模拟红外图像序列和真实红外图像序列校正的实验结果表明, 相对于目前已有的非均匀性校正算法, 该方法在 PSNR (Peak Signal to Noise Ratio) 和粗糙度的客观数据上都有所提升, 主观视觉效果也更加清晰, 细节保留程度高。

关键词: 红外图像; 非均匀性校正; 多尺度采样; 残差学习

中图分类号: TP391.4

文献标识码: A

文章编号: 1001-8891(2020)09-0833-07

Adaptive Correction Algorithm of Infrared Image Based on Encoding and Decoding Residual Network

MOU Xingang, LU Junjie, ZHOU Xiao

(College of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan 430070, China)

Abstract: Traditional scene-based non-uniformity correction algorithms generally suffer from non-uniformity residuals and ghosts. In view of this, we propose an infrared image adaptive algorithm based on the encoding and decoding residual network. The algorithm focuses on the characteristics of the adaptive correction problem. Following the UNet structure, the residual image is generated through multiscale sampling and learning residual mapping. Batch normalization and PReLU are used to improve the correction effect. Finally, the global skip connection is used to obtain the final correction result. The experimental results of correcting the simulated non-uniform infrared image sequence and the real infrared image sequence showed that this method improved the objective data of the peak signal to noise ratio (PSNR) and roughness, compared with existing non-uniformity correction algorithms. Moreover, the subjective visual effect was clearer, and the degree of detail retention was high.

Key words: infrared image, non-uniformity, multi-scale sampling, residual learning

0 引言

随着社会的进步和发展, 各行各业对红外成像系统的需求变得越来越迫切, 但作为红外成像系统的核心器件红外焦平面阵列 (Infrared Focal Plane Array, IRFPA) 存在非均匀性响应^[1], 这种非均匀性响应会降低红外系统的成像质量, 所以在红外成像系统应用过程中必须对非均匀性进行校正。针对红外图像的非均匀性问题, 目前主要有基于定标和基于场景的两大

类校正算法。基于定标的非均匀性校正算法利用不同温度下获得的均匀辐射黑体图像进行标定得到探测单元响应的增益和偏置, 继而通过拟合计算待标定点的校正期望值。基于定标的算法计算量小, 硬件容易实现, 但是该类算法在某些需要连续工作的场景就不太适用。为了应对这个问题, 人们开始研究基于场景的非均匀性校正算法。基于场景的非均匀性校正算法是指红外图像的非均匀性校正参数通过场景获得而非通过黑体的定标获得, 如神经网络算法^[2]、时域高

收稿日期: 2020-02-19; 修订日期: 2020-07-26.

作者简介: 牟新刚 (1982-), 男, 博士, 副教授, 主要研究方向光电成像与信息处理、红外图像处理, E-mail: mouxingang@163.com。

基金项目: 国家基金项目 (61701357), 中央高校基本科研业务费专项资金资助 (183204007)。

通滤波算法^[3]和恒定统计算法^[4],这类算法虽然能够满足连续工作的要求,但是在场景和 IRFPA 运动过程中产生鬼影问题。针对鬼影问题, Qian 等人^[5]提出了基于空域低通-时域高通的校正算法, Zuo 等人^[6]利用双边滤波来分离红外图像, Rong 等人^[7]提出了基于引导滤波和自适应学习速率的红外焦平面阵列非均匀性校正算法, Mou 等人^[8]对红外探测器非均匀性响应分布特征进行统计与分析,提出基于图像块先验的单帧红外自适应校正算法。

近年来,随着深度学习的发展,卷积神经网络(Convolution Neural Network, CNN)在图像去噪领域表现优异,越来越多的 CNN 模型被用来解决图像去噪问题。针对基于场景的非均匀性校正算法存在非均匀性噪声残余和鬼影等问题,人们开始研究利用卷积神经网络进行非均匀性校正, He 和 Cao 等人^[9]提出了基于卷积神经网络和柱状固定模式噪声仿真模块的训练方案, Mou 等人^[10]利用真实红外探测器响应以及对应的两点校正结果制作训练集,提出了基于残差网络的非均匀性校正算法并取得了一定的效果。

目前基于深度学习的非均匀性校正算法研究较少,并且也存在计算效率不高,校正精度不高等问题。针对这些问题,本文基于 UNet^[11]结构,提出了残差编解码网络(Encoding-decoding Residual Network, EDRN),用来实现红外图像的自适应校正。该网络使用左右对称的编解码网络,通过多尺度采样提高网络的特征提取能力,加入残差映射生成残差图像减少非线性映射范围,减轻训练难度。通过实验结果证明,该网络模型克服了传统方法对红外图像非均匀性校正的鲁棒性差、会出现鬼影等局限性,对单帧红外图像的非均匀性校正的效果出色,不会产生鬼影现象,在很大程度上去除噪声的同时能够保留更多的边缘细节信息。

1 IRFPA 非均匀性响应模型

通常情况下, IRFPA 中每一个探测器单元的响应可以使用一个近似的线性模型表示:

$$y_{i,j}(n) = A_{i,j}(n) x_{i,j}(n) + B_{i,j}(n) \quad (1)$$

式中: $y_{i,j}(n)$ 和 $x_{i,j}(n)$ 分别表示探测器第 (i,j) 个像元的实际输出值和输入辐射值, $A_{i,j}(n)$ 和 $B_{i,j}(n)$ 分别表示为第 (i,j) 个像元的响应增益和偏置参数。

非均匀性校正就是从已知的实际输出值 $y_{i,j}(n)$ 中估计得到真实的输入辐射值 $x_{i,j}(n)$,从而去除非均匀性响应的影

$$x_{i,j}(n) = g_{i,j}(n) y_{i,j}(n) + o_{i,j}(n) \quad (2)$$

式中: $g_{i,j}(n)$ 和 $o_{i,j}(n)$ 分别为第 (i,j) 个像元的响应增益和

偏置的校正系数:

$$g_{i,j}(n) = \frac{1}{A_{i,j}(n)}, \quad o_{i,j}(n) = -\frac{B_{i,j}(n)}{A_{i,j}(n)} \quad (3)$$

由于红外探测器单元响应参数会随着时间漂移,使得基于定标的非均匀性校正算法不能彻底解决非均匀性的问题,基于场景的非均匀性校正算法也不能完全避免鬼影的产生,继而不能准确估计像元的响应增益和偏置的校正系数。本文提出的 EDRN 算法利用卷积神经网络估计 $g(n)$ 和 $o(n)$,从而去除非均匀性得到真实红外图像 $x(n)$ 。

2 残差编解码网络

UNet 结构最大的优点就是它能很好地将图像浅层特征与深层特征相结合,在图像处理领域获得了广泛的应用,这也是本文基于 UNet 结构进行网络设计的重要原因。针对红外图像对比度低,边缘模糊等特点,提出了一种基于 UNet 结构的残差编解码网络,具体内容包括:①使用左右对称的编解码网络,利用学习残差映射得到多尺度下非均匀性残差图像,最后利用全局跳跃连接将含非均匀性红外图像减去非均匀性残差图像得到最后的校正结果;②为了取得更好的校正效果,本文在卷积层之后加入批标准化^[12];③使用 PReLU 激活函数^[13]解决部分神经元坏死的问题。接下来本文将从网络框架,图像边界处理和损失函数具体介绍残差编解码网络。

2.1 网络框架

本文提出的 EDRN 通过下采样和上采样对图像的特征维度进行收缩和扩张,首先将原始的图像特征通过下采样逐步缩小,在更小的图像特征下进行卷积编码,然后将缩小的图像特征通过上采样逐步成为原始图像尺度,进行反卷积解码。由于编解码操作和上下采样都是成对出现,本框架可以进行任意尺度的扩展,本文在综合考虑网络复杂度和校正效果之后,选用了4个尺度的网络。图1给出了 EDRN 的结构框图,一共包括5个部分:4个编码单元,3个解码单元,3个上采样,3个下采样,一个全卷积层。整个网络包括15个卷积层,3个最大池化层和3个反卷积层,各层的参数信息如表1所示。

1) 上下采样

下采样使用一个 2×2 的最大池化层,最大池化层的步长设置为2,通过下采样可以减少输入数据的大小,网络本身也可以捕捉更有用的元素。而上采样使用反卷积层进行反卷积操作,将输入图像尺度扩大为原来的两倍。

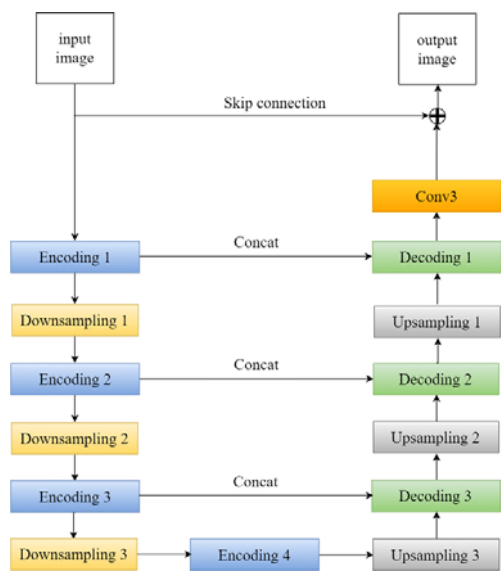


图1 EDRN 结构框图

Fig.1 EDRN framework

表1 EDRN 各层参数设置

Table 1 EDRN parameter settings

Layer	Filters	Input	Output
Conv	3×3×64	64×64×1	64×64×64
Conv	3×3×64	64×64×64	64×64×64
MaxPool	-	64×64×64	32×32×64
Conv	3×3×64	32×32×64	32×32×64
Conv	3×3×64	32×32×64	32×32×64
MaxPool	/	32×32×64	16×16×64
Conv	3×3×64	16×16×64	16×16×64
Conv	3×3×64	16×16×64	16×16×64
MaxPool	/	16×16×64	8×8×64
Conv	3×3×64	8×8×64	8×8×64
Conv	3×3×64	8×8×64	8×8×64
DeConv	/	8×8×64	16×16×64
Conv	3×3×128	16×16×196	16×16×128
Conv	3×3×128	16×16×128	16×16×128
DeConv	/	16×16×128	32×32×128
Conv	3×3×128	32×32×196	32×32×128
Conv	3×3×128	32×32×128	32×32×128
DeConv	/	32×32×128	64×64×128
Conv	3×3×128	64×64×196	64×64×128
Conv	3×3×128	64×64×128	64×64×128
Conv	3×3×1	64×64×128	64×64×1

2) 编解码单元

本文在残差学习^[14]的基础上，提出了两层结构的编码单元。编码单元结构如图2(a)所示，编码单元由两个卷积层和一个跳跃连接组成，通过学习非均匀性

残差特征，缩小输入输出的映射范围，能够学习更好的图像特征。解码单元与编码单元类似，唯一不同的是多出一个连接操作，连接本层输入与对应的编码单元输出特征，由于池化操作会丢失图像的部分信息，如果直接进行上采样解码操作的话，图像输出也会丢失信息导致模糊，连接操作就是用来解决这个问题。解码单元结构如图2(b)所示。

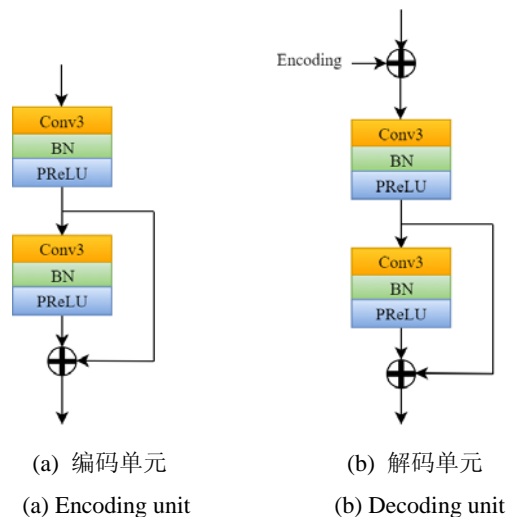


图2 编解码单元

Fig.2 Encoding and decoding unit

3) 全卷积层

全卷积层是一个3×3的卷积层，输出非均匀性残差图像。最后使用一个全局跳跃连接，将含非均匀性的红外图像减去非均匀性残差图像得到校正后的红外图像。

除了最后一层全卷积层之外，其他所有的卷积层使用批标准化^[12]和PReLU激活函数^[13]，批标准化能够加快网络收敛，提高训练速度，PReLU激活函数能够解决部分神经元坏死的问题。后面通过对ReLU、LReLU和PReLU三种激活函数的实验证明，PReLU激活函数更加适合EDRN模型，能够得到更好的结果。

2.2 图像边界处理

在进行卷积以及反卷积操作，采用对图像边缘补零的操作保证卷积以及反卷积前后的红外图像大小不变，这种方法能够抑制红外图像在边缘的退化，在实验中发现效果也更好。

2.3 损失函数

本文使用均方根误差作为损失函数，将带有非均匀性噪声的红外图像 $y(n)$ 输入网络模型中进行非均匀性校正得到校正后的红外图像 $x(n)$ ，均方误差可以表示如下：

$$L_{\text{MSE}} = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H [f(y_{i,j}) - x_{i,j}]^2 \quad (4)$$

式中： W 和 H 分别表示红外图像的宽度和高度； y 表示待校正的图像； f 代表整个网络的校正过程； x 代表不含非均匀性噪声的真实红外图像。

3 实验与结果分析

3.1 数据集的制备

本文选用 LTIR^[15] 红外图像标准数据集作为本文的数据集，该数据集包含丰富的标准红外图像，图像种类相对较多，图像质量较高，对于卷积神经网络模型来说，选取清晰丰富的图像数据集有利于模型的训练和提高模型的泛化性。考虑到计算机硬件性能条件，本文从 LTIR 数据集中随机选取 400 张 640×480 的红外图像作为训练集，选取 depthwise_crossing 目录下的红外图像作为本次的测试集，depthwise_crossing 目录是 851 张 640×480 的红外图像序列。

为了更好地训练，将 640×480 的红外图像裁剪为 64×64 的红外图像，按步长为 40 个像素的方式进行裁剪，并对裁剪之后的图像进行随机旋转 0°、90°、180°、270°，得到 254000 个 64×64 的子图像作为标准图。之前已经讨论过红外图像非均匀性响应模型，本章将向标准图随机加入均值为 1，标准差在 [0.05, 0.15] 区间的增益噪声，记为 σ_g ，以及均值为 0，标准差在 [5, 15] 区间的偏置噪声，记为 σ_o ，以此来模拟不同级别强度非均匀性响应。

3.2 训练参数设置与模型训练

本次 EDRN 算法的相关实验都是在深度学习框架 Pytorch 下进行的，训练的硬件平台为 Intel(R) Core(TM) i5-3570 CPU+NVIDIA 1080TI。

网络模型的超参数如表 2 所示，网络训练批次设置为 64，初始设置学习率为 0.001，学习率每迭代 50 次减半，优化器选用 Adam 优化器，使用正交矩阵初始化^[16]的方式进行网络卷积层权重初始化，一共训练 200 个回合。

3.3 实验结果分析

为了验证本文提出的 EDRN 算法非均匀性的校正效果，本文将分别在测试集红外图像序列和真实红外图像序列上进行实验分析，我们将与传统方法、深度学习方法进行比较，传统方法包括双边滤波算法^[6] (Bilateral Filters, BF)，导向滤波算法^[7] (Guided Filter, GF) 和传统神经网络算法^[2] (Neural Network, NN)，深度学习方法为 DLS 算法^[9] (Deep-learning-based Strip NUC Method, DLS)。值得注意的是，DLS 算法都是

用来去除条纹非均匀性，所以对它们加入本文模拟的非均匀性响应重新训练。

表 2 网络的超参数

Table 2 Hyper parameters of the network	
Hyperparameters	Value
Height	64
Width	64
Channel	1
Batch size	64
Initial learning rate	0.001
Optimizer	Adam
Epoch	200
Initialization	Orthogonal Regularization

本节主要使用 PSNR 作为定量评价算法的方法，使用粗糙度^[17]作为辅助评价方法。一般而言，PSNR 越大表明非均匀性校正效果越好，图像粗糙度越小表明非均匀性校正效果越好。

1) 网络整体训练情况

本节将分析网络整体训练情况，并且分析不同激活函数对于模型训练结果的影响，以此验证 PReLU 激活函数对于模型的有效性。

图 3 显示了 3 种不同激活函数对应的网络模型在训练过程中 L2 损失和 PSNR 变化趋势。从图 3(a)可以看出，使用 PReLU 激活函数在训练过程中 L2 损失变化相对更加平缓，而使用 LReLU 激活函数和 ReLU 激活函数则相对比较杂乱，由此可以得出使用 PReLU 激活函数的网络模型整体训练相对稳定，收敛性也更加好。另外，从图 3(b)来看，使用 PReLU 激活函数可以获得更高的 PSNR 数值指标，最后收敛于 39dB 左右。

表 3 显示了不同激活函数对应训练集的平均 PSNR，使用 PReLU 激活函数比使用 LReLU 激活函数提高了 0.11dB，比使用 ReLU 激活函数提高了 0.38 dB。

2) 模拟红外图像序列

本文使用测试集中的红外图像作为真值红外图像，图像的分辨率为 640×480，像素都是 8 位数据。实验中加入高强度非均匀性来模拟非均匀性响应，参数为 $\sigma_g=0.15$ ， $\sigma_o=15$ 。

本次对比实验将分别从客观数据评价和主观视觉评价两个方面来对比分析，客观数据包括峰值信噪比 PSNR 和粗糙度 ρ 。

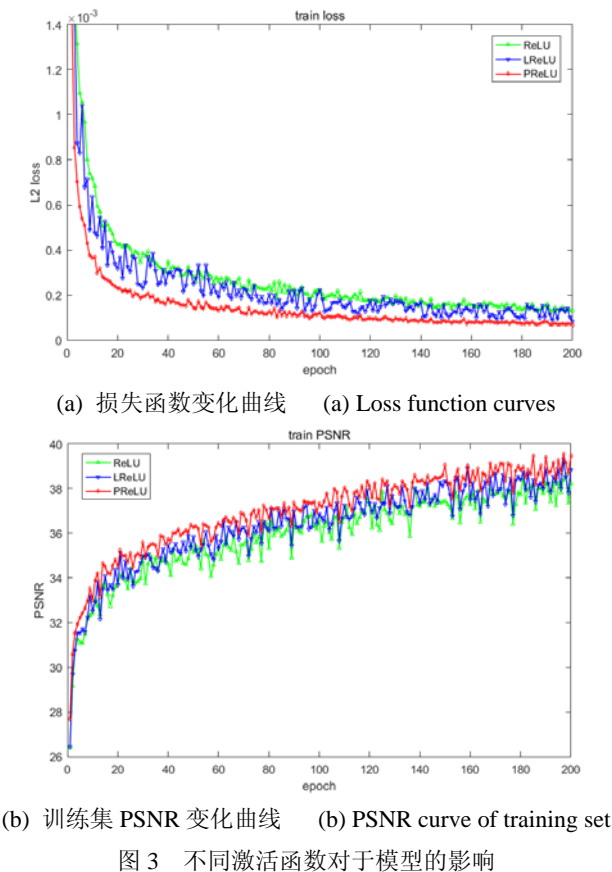


图 3 不同激活函数对于模型的影响

Fig.3 Effect of different activation functions on the model

表 3 不同激活函数对应训练集的平均 PSNR

Table 3 Mean PSNR (dB) of the training set corresponding to different activation functions

Activation functions	dB		
	PReLU	LReLU	ReLU
PSNR	36.30	36.19	35.92

图 4(c)~(g)显示了各算法对模拟非均匀性红外图像序列第 850 帧校正的结果,图 4(c)~(e)存在不同程度的竖条纹非均匀性,并且从局部放大图来看存在不同程度的噪点,边缘细节丢失比较严重,图 4(f)~(g)基本消除了竖条纹,但图 5(f)存在少量噪点没有去除干净,图 5(g)图像基本看不出噪点,图像质量最高。

图 5(a)~(b)为各算法对模拟非均匀性红外图像序列校正的 PSNR 和 ρ 曲线。从图 5(a)可以看出,传统 BF 算法、GF 算法和 NN 算法在 300 帧左右达到收敛状态,然后开始校正,而基于深度学习的方法 DLS 算法和本文 EDRN 算法直接进行校正状态,无需收敛,另外 EDRN 算法相比其他几种算法有更高的 PSNR 数值。类似于 PSNR 曲线变化规律,在图 5(b)中我们也能看出 EDRN 算法粗糙度变化更加稳定,数值也是最低。

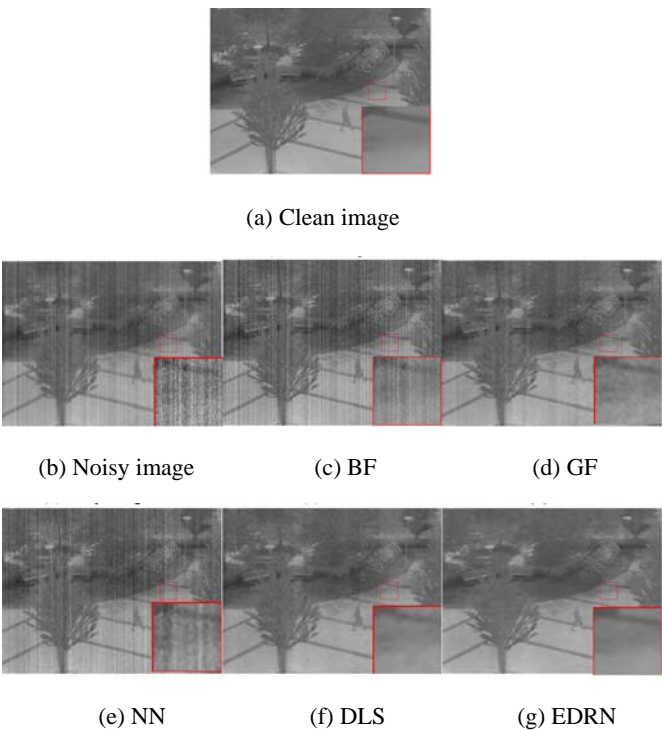


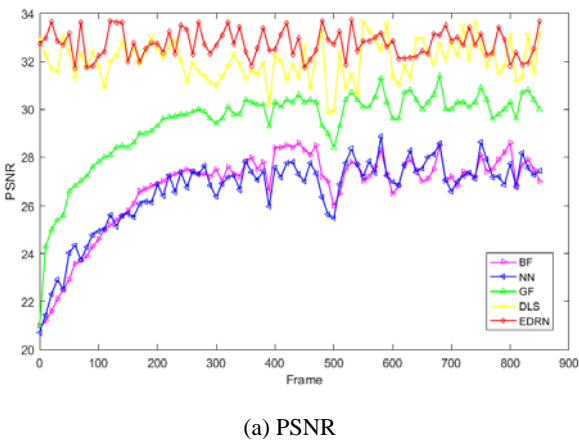
图 4 各算法对模拟非均匀性红外图像的校正结果

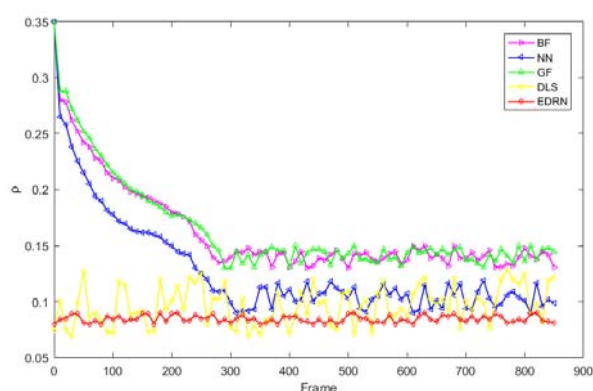
Fig.4 Effects of each algorithm on simulated non-uniform infrared image processing

从以上分析综合来看,本文 EDRN 算法相对比本文提出的其他算法在模拟非均匀性红外图像序列上有更好的校正效果。

3) 真实红外图像序列

为了验证 EDRN 算法对于真实红外图像序列的有效性,我们将使用两个真实红外图像序列进行非均匀性校正实验,包括 1500 帧 640×512 的真实红外图像序列和 350 帧 512×480 的真实红外图像序列。值得注意的是,真实红外图像序列采样精度为 14 bit,为了方便进行对比实验,本节将 14 bit 的红外图像序列映射为 8 bit 红外图像序列。





(b) ρ

图 5 各算法对模拟非均匀性红外图像序列校正的 PSNR 和 ρ 曲线

Fig.5 PSNR and ρ curves of each algorithm for simulating non-uniform infrared image sequence correction

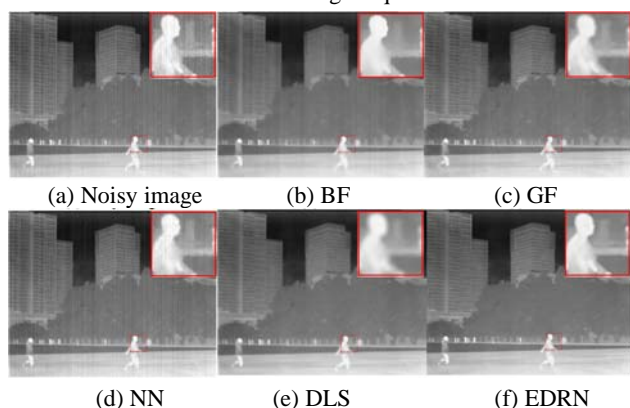


图 6 各算法对 640×512 的红外图像序列的校正结果

Fig.6 Correction effect of each algorithm on 640×512 infrared image sequence

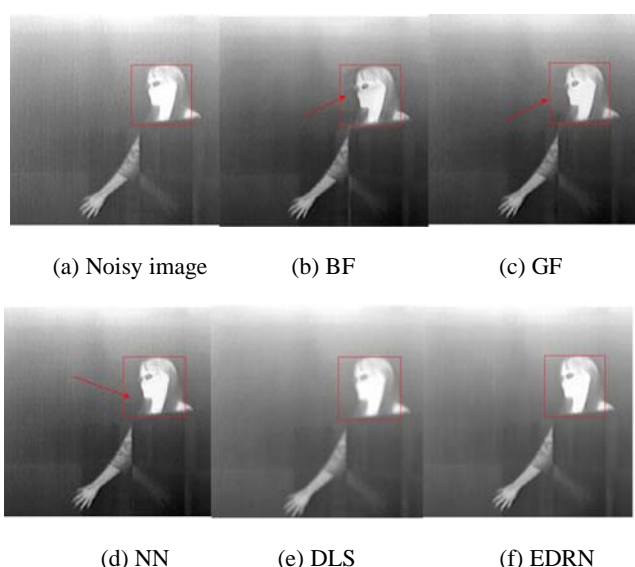


图 7 各算法对 512×480 的红外图像序列的校正结果

Fig.7 Correction effect of each algorithm on 512×480 infrared image sequence

图 6(b)~(f)为各算法对 640×512 的红外图像序列第 1200 帧的校正结果。从整体来看,图 6(b)和图 6(d)看到轻微的竖条纹非均匀性,图 6(c)(e)(f)都取得了一定的校正效果。从局部来看,图 6(b)在行人手臂部位丢失细节,图 6(c)在行人头部模糊,轮廓丢失,图 6(d)存在锯齿状颗粒,图 6(e)图像细节相对模糊,图 6(f)局部细节清晰,轮廓分明。

基于场景的非均匀性校正算法比较依赖场景的运动,容易受场景运动的影响而产生鬼影。图 7 显示了各算法对 512×480 的红外图像序列第 350 帧的校正结果。图 7(b)~(c)在人像的头部有明显的鬼影,图 7(d)在手臂以及头部也有轻微的鬼影,这类鬼影大部分都是由于局部滤波器对校正参数的不准确估计造成的,而图 7(e)~(f)基本没有鬼影,说明基于深度学习的非均匀性校正算法很好地克服了这一现象。

综合来看,本文 EDRN 算法校正过程中没有鬼影的产生,在图像细节方面也有很好地保留,即验证了本文 EDRN 算法在真实红外图像上也有良好校正效果。

4 结论

本文在深度学习模型 UNet 框架的启发下,提出了用于红外图像非均匀性校正的 EDRN 算法。EDRN 算法采用真值红外图像序列作为训练集,加入模拟非均匀性响应,目标是得到清晰的红外图像。根据对模拟非均匀性红外图像序列和真实红外图像序列校正的实验结果证明,相比于其他几种非均匀性校正的算法,EDRN 算法在 PSNR 和粗糙度 ρ 上有不同程度的提升,并在主观视觉方面也取得了更好的效果。

参考文献:

- [1] 陈钱. 红外图像处理技术现状及发展趋势[J]. 红外技术, 2013, 35(6): 311-318.
CHEN Qian. The Status and Development Trend of Infrared Image Processing Technology[J]. Infrared Technology, 2013, 35(6): 311-318.
- [2] Scribner D A, Sarkady K A, Krueer M R, et al. Adaptive nonuniformity correction for IR focal-plane arrays using neural networks[C]//International Society for Optics and Photonics, 1991: 100-109.
- [3] Scribner D A, Sarkady K A, Caldfield J T, et al. Nonuniformity correction for staring IR focal plane arrays using scene-based techniques[J]. Proceedings of SPIE - The International Society for Optical Engineering, 1990: 1308.

- [4] Redlich R, Figueroa M, Torres S N, et al. Embedded nonuniformity correction in infrared focal plane arrays using the Constant Range algorithm[J]. *Infrared Physics & Technology*, 2015, **69**: 164-173.
- [5] QIAN W, CHEN Q, GU G. Space low-pass and temporal high-pass nonuniformity correction algorithm[J]. *Optical Review*, 2010, **17**(1): 24-29.
- [6] ZUO C, CHEN Q, GU G, et al. New temporal high-pass filter nonuniformity correction based on bilateral filter[J]. *Optical Review*, 2011, **18**(2): 197-202.
- [7] Shenghui R, Huixin Z, Hanlin Q, et al. Guided filter and adaptive learning rate based non-uniformity correction algorithm for infrared focal plane array[J]. *Infrared Physics & Technology*, 2016, **76**: 691-697.
- [8] 牟新刚, 赵建新, 欧科君. 基于图像块先验的单帧红外自适应校正算法[J]. *激光与红外*, 2017, **047**(012): 1548-1552.
MOU X, ZHAO J, OU K. Single-frame infrared adaptive correction algorithm based on image patch priori[J]. *Laser & Infrared*, 2017, **047**(012): 1548-1552.
- [9] HE Zewei, CAO Yanpeng, DONG Yafei, et al. Single-image-based nonuniformity correction of uncooled long-wave infrared detectors:a deep-learning approach[J]. *Applied Optics*, 2018, **57**: 155-164.
- [10] MOU X, LU J, ZHOU X, et al. Single frame infrared image adaptive correction algorithm based on residual network[C]// *The 11th International Symposium on Photonics and Optoelectronics(SOPO)*. CRC Press / Balkema, 2018: 17-23.
- [11] 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.
- [12] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//*International Conference on International Conference on Machine Learning*. JMLR.org, 2015: 448-456.
- [13] HE K, ZHANG X, REN S, et al. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification[C]//*International Conference on Computer Vision*, 2015: 1026-1034.
- [14] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.
- [15] Berg A, Jörgen Ahlberg, Felsberg M. A thermal Object Tracking benchmark[C]// *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2015, **1**: 1-6.
- [16] Wang D, Cui P, Ou M, et al. Deep Multimodal Hashing with Orthogonal Regularization[C]// *IJCAI*. AAAI Press, 2015: 2291-2297.
- [17] Hayat, Majeed M, Torres, Sergio N, Armstrong, Ernest. Statistical Algorithm for Nonuniformity Correction in Focal-Plane Arrays[J]. *Applied Optics*, 1999, **38**(5): 772-780.