Infrared Physics and Technology 94 (2018) 142–150

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Regular article

Robust destriping method based on data-driven learning

Xiaodong Kuang, Xiubao Sui*, Yuan Liu, Chengwei Liu, Qian Chen, Guohua Gu

Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Keywords:
Image destriping
Data-driven learning
Convolutional neural networks

ABSTRACT

Destriping is the crucial first step of many multidetector imaging pipelines since the stripes greatly decrease the quality of obtained data and limit subsequent applications. It is also a severely ill-posed issue that estimates true gray value per pixel from a single stripe measurement. Existing approaches leverage hand-crafted filters or priors but show visually unsatisfactory results where some residual stripes still remain and the quantitative values of image data are lost. To address these problems, we propose a new data-driven method. We train a convolutional neural network on a large set of ground truth data instead of using hand-tuned filters. A UNet-like network is used to learn the regularity of complex stripe noise characteristics. To generate high-quality images, we combine a per-pixel loss and a perceptual loss to penalize mismatch between the network output and ground-truth images. Experiments show that our network significantly outperforms state-of-the-art destriping approaches in real-captured noise images of many imaging fields. Our code is available online at https://github.com/Kuangxd/DDL-SR.

1. Introduction

Destriping is the vital first stage of most multidetector imaging system pipelines such as the atomic force microscope (AFM) [1], passive millimeter-wave (PMMW) radio [2], moderate resolution imaging spectroradiometer (MODIS) [3], and infrared focal plane arrays (IFPA) [4]. It is representatively an ill-posed issue that existing information is corrupted with stripes, which are caused by inconsistent response of different detectors.

People have long realized that the using regularities in natural images contributes to lifting underdetermination. Previous methods have hard-coded hand-crafted inspirations into filters [5]. They construct a filter by combining wavelet and Fourier transform. However, their filters are still hand-crafted and cannot be used in some special cases such as real stripes do not obey the distribution of predefined models.

In this paper, we tackle destriping employing a data-driven learning stripe removal method (DDL-SR). We train our network on a great deal of ground-truth data to exploit the regularity in natural images. We build on the success of deep learning and convolutional neural networks [6]. While authors in [4] have already proposed a data-driven learning method, it is still very crucial to perform a refined training procedure. Their data-driven learning method is only suitable for handling light stripes and cannot handle heavy stripes. Usually complex stripe characteristics where the proportion and intensity of stripe lines vary greatly make this a challenge. To address this challenge, we propose the use of a UNet-like model [7] to learn the complex mapping. Meanwhile, perceptual loss that depends on high-level features extracted from pretrained networks is employed to preserve details and suppress artifacts. We optimize our network combining the benefits of per-pixel loss and perceptual loss instead of only using per-pixel loss relying on low-level pixel information. Our processing algorithm benefits from data-driven learning. Given training images (images with and without stripes), our model can learn the mapping between the stripe-free and striped images. Once the training is done, we can use the network directly to process the stripe images without optimizing the parameters, which have already been obtained through learning. So our network takes a striped image as the input and outputs a stripe-free image that will be used as the input data in the further processing.

Overall, we make the following contributions:

- We propose a data-driven learning method based on convolutional neural networks to achieve best quality on stripe noise removal.
- We demonstrate that our DDL-SR method is able to address a wide range of stripe levels in different imaging fields.
- We confirm that our DDL-SR method achieves higher-quality results than previous work in simulated and real data experiments.

* Corresponding author.
E-mail address: sxbhandsome@163.com (X. Sui).

https://doi.org/10.1016/j.infrared.2018.09.015
Received 17 May 2018; Received in revised form 19 July 2018; Accepted 17 September 2018
Available online 18 September 2018

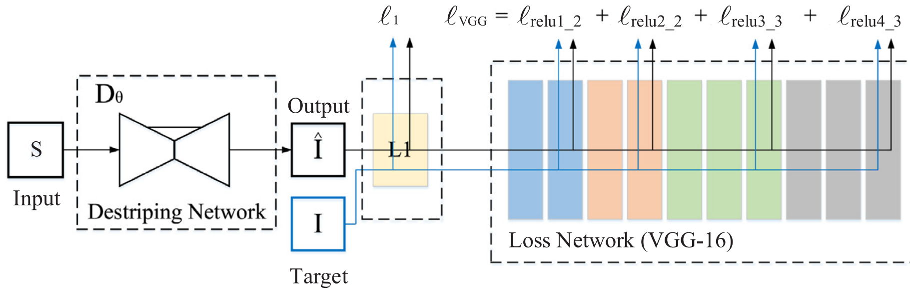

Fig. 1. System overview.

2. Related work

2.1. Previous destriping methods

In the past decades, many destriping approaches have been presented under different frameworks, which can be roughly divided into three categories: transformation-based approaches, statistics-based approaches, and optimization-based approaches. Transformation-based approaches mainly use filters in transform domain to suppress stripes [5,8,9]. Statistics-based approaches generally depend on some statistical properties of numerical values in every detector [3,10]. Optimization-based approaches treat the destriping process as inverse issues, minimizing energy functions under regularized constraints [11–14].

However, these approaches do not seem to achieve completely satisfactory results. We recognize that the above methods may achieve best-quality results in some different cases. However, any existing filtering-based method has a common problem. Specifically, the stripes that should have been removed are not completely removed. Instead, some details that should be preserved are seriously damaged. The main reason is that people's ability to visually distinguish unwanted details often exceeds preferences of filters by far.

A similar data-driven destriping method that trains a simple three-layer convolutional network has recently been proposed to address stripes on infrared images [4]. In weak stripe noise removal experiments, it has achieved a great advantage over other destriping methods, perfectly retaining details and completely removing stripes. However, this method has a fatal flaw where the trained network cannot remove more serious stripe noise even with a larger training set. The main reason for this problem is that the network is too simple to map more serious stripe distributions.

2.2. Neural networks for image processing

Convolutional neural networks (CNN) have recently shown an explosive popularity since its revolution in classification tasks [15]. They are also quickly becoming dominated methods in image processing tasks like image colorization [16,17], super resolution [18,19], style transfer [20,21], image deblurring [22], image de-raining [23], image synthesis [24,25], image completion [26,27], image denoising [28], and image-to-image translation [29].

2.3. Design of convolutional neural networks

Many well-known computer vision tasks are successfully addressed by designing specific network architectures. They show that a deeper network architecture helps greatly improve the network's performance as it allows modeling a more complexity mapping [30,31]. Ledig et al. [19] proposes to use a very deep network with residual blocks [32] to realize photo-realistic image super-resolution. Pathak et al. [33] uses encoder-decoder networks to solve image inpainting problems. Meanwhile, Johnson et al. combines residual blocks and encoder-decoder architectures to design a refined network for style transfer and super-resolution [20]. Recently an encoder-decoder network with skip connections [34] achieves remarkable results on a variety of image-to-image translation tasks [29].

2.4. Loss functions

Per-pixel loss functions like mean squared error (MSE) aim to address the indeterminacy immanent in regaining missing data. However, using MSE tends to find per-pixel averages of scores with maximum probability and hence produce poor visual results [35]. To tackle this issue, researchers proposed the use of perceptual losses. Authors in [36] present using feature maps extracted from a pretrained VGG model rather than low-level per-pixel errors. Specifically, they define loss functions based on Euclidean distances between features extracted from the VGG19 model [30]. Visually more persuasive results are shown for super-resolution and style transfer [20].

3. Methods

Destriping has conventionally been processed employing nonlinear filters, combining prior inspirations of intra-column interrelation, and utilizing intra-patch similarity of images. Using convolutional networks appears a natural solution for the issue in this context. First, they can discover natural interrelation in the data. Second, convolutional networks are capable of representing a superset of pipelines performed using existing approaches.

We treat destriping as a supervised learning issue. We train the network on a set of input measurements where the desired output is known. We build a training dataset from hundreds of thousands of RGB images. We then create a convolutional neural network and train it in an end-to-end fashion.

Fig. 1 shows the basic framework of our method. The destriping network D_θ parameterized by weights θ uses a UNet-like architecture to implement a mapping between input and output. Each loss function calculates a scalar value measuring errors between the output image and the ground-truth image. Then the destriping network is optimized by minimizing a weighted combination of loss functions. We first focus on network architecture and then discuss loss functions.

3.1. Mechanism of stripes formation

It is crucial to generate a large number of training images (images with and without stripes) to train our DDL-SR model. For thermal imaging system, microbolometer focal plane arrays are the core devices of which the working principle is that the infrared radiation incoming to the infrared detector is absorbed on a specially prepared absorber surface and then the infrared energy absorbed by the detector is converted to heat. So the thermal image is produced by translating the changes of heat in resistance of the detector into a time-accumulated electrical signal. However, the responsive nonuniformity of detectors in each column often leads to undesirable stripe noise in the raw infrared data. Previously, a simple offset model [4] and a linear correction [37,38] model have been proposed to describe the relationship between the stripe noise and thermal radiation. To easily generate training