ISPRS Journal of Photogrammetry and Remote Sensing xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

One-two-one networks for compression artifacts reduction in remote sensing

Baochang Zhang [a], Jiaxin Gu [a], Chen Chen [c,*], Jungong Han [d], Xiangbo Su [a], Xianbin Cao [b], Jianzhuang Liu [e]

[a] School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
[b] School of Electronics and Information Engineering, Beihang University, Beijing, China
[c] Center for Research in Computer Vision (CRCV), University of Central Florida, Orlando, FL, USA
[d] School of Computing and Communications, Lancaster University, LA1 4YW, UK
[e] Noah's Ark Lab, Huawei Technologies Co. Ltd., China

ARTICLE INFO

Article history:
Received 27 May 2017
Received in revised form 22 November 2017
Accepted 4 January 2018
Available online xxxx

2010 MSC:
00-01
99-00

Keywords:
Compression artifacts reduction
Remote sensing
Deep learning
One-two-one network

ABSTRACT

Compression artifacts reduction (CAR) is a challenging problem in the field of remote sensing. Most recent deep learning based methods have demonstrated superior performance over the previous hand-crafted methods. In this paper, we propose an end-to-end one-two-one (OTO) network, to combine different deep models, i.e., summation and difference models, to solve the CAR problem. Particularly, the difference model motivated by the Laplacian pyramid is designed to obtain the high frequency information, while the summation model aggregates the low frequency information. We provide an in-depth investigation into our OTO architecture based on the Taylor expansion, which shows that these two kinds of information can be fused in a nonlinear scheme to gain more capacity of handling complicated image compression artifacts, especially the blocking effect in compression. Extensive experiments are conducted to demonstrate the superior performance of the OTO networks, as compared to the state-of-the-arts on remote sensing datasets and other benchmark datasets. The source code will be available here: https://github.com/bczhangbczhang/.

1. Introduction

In remote sensing, the satellite- or aircraft-based sensor technologies are used to capture and detect objects on Earth. Thanks to various propagated signals (e.g., electromagnetic radiation), remote sensing makes the data collection from dangerous or inaccessible areas possible, and therefore plays a significant role in many applications including monitoring, military information collection and land-use classification (Chen et al., 2016; Li et al., 2014, 2015; Vosselman et al., 2017). With the technological development of various satellite sensors, the volume of high-resolution remote sensing image data is increasing rapidly. Hence, proper compression of the satellite image becomes essential, which enables information exchange much more efficient, given a limited band width.

Existing compression methods generally fall into two categories: lossless (e.g., PNG) and lossy (e.g., JPEG) (Wang et al., 2002). The lossless methods usually provide better visual experience to users, but lossy methods often achieve higher compression ratios via non-invertible compression functions along with trade-off parameters to balance the data amount and the decompressed quality. Therefore the lossy compression schemes are always preferred by consumer devices in practice due to higher compression rate (Wang et al., 2002). However, high compression rate comes with the cost of having compression artifacts on the decoded image, which is a barrier for many applications, such as image analysis. Therefore, there is a clear need for compression artifact reduction, which is able to gain visual quality of the decompressed image, which can influence the visual effect and low-level vision processing (Yu et al., 2016).

The compression artifacts are in relation to the schemes used for compression. Take JPEG compression as an example, blocking artifacts are caused by discontinuities at the borders when encoding adjacent 8×8 pixel blocks, which are in the form of ringing effects and blurring due to the coarse quantization of the high frequency components. To deal with these compression artifacts, an improved version of JPEG, named JPEG 2000, is proposed, which

* Corresponding author.
E-mail addresses: bczhang@buaa.edu.cn (B. Zhang), chenchen870713@gmail.com (C. Chen).

https://doi.org/10.1016/j.isprsjprs.2018.01.003
0924-2716/

adopts the wavelet transform to avoid blocking artifacts, but still undergoes ringing effects and blurring. As an excellent alternative, SPIHT (Said et al., 1996) showed that using simple uniform scalar quantization, rather than complicated vector quantization, also yields superior results. Due to its simplicity, SPIHT has been successful on natural (portraits, landscape, weddings, etc.) and medical (X-ray, CT, etc.) images. Furthermore, its embedded encoding process has proved to be effective in a broad range of reconstruction qualities. For instance, it can code fair-quality portraits and high-quality medical images equally well (as compared with other methods in the same conditions). However, in the field of remote sensing, the images usually suffer from severe artifacts after compression as shown in Fig. 1, which poses challenges to many high-level vision tasks, such as object detection (Cheng and Han, 2016; Xiao et al., 2016), classification (Chen et al., 2016; Bian et al., 2017), and anomaly detection (Chang and Chiang, 2002).

To cope with various compression artifacts, many conventional approaches have been proposed, such as filtering approaches (List et al., 2003; Reeve and Lim, 1984; Wang et al., 2013), specific priors (e.g., the quantization table in DSC (Liu et al., 2015)), and thresholding techniques (Liew and Yan, 2004; Foi et al., 2007). Inspired by the great success of deep learning technology in many image processing applications, researchers start to exploit this powerful tool to reduce the compression artifact. Specifically, the Super-Resolution Convolutional Neural Network (SRCNN) (Dong et al., 2014) exhibits great potential of an end-to-end learning in image super-resolution. It is also pointed out that conventional sparse-coding-based image restoration model can be equally seen as a deep model. However, if we directly apply SRCNN to the compression artifact reduction task, the features extracted by its first layer are noisy, which will cause undesirable noisy patterns in reconstruction. Thus the three-layer SRCNN is not suitable for compressed image restoration, especially when dealing with complex artifacts. Thanks to transfer learning, ARCNN (Yu et al., 2016) has been successfully applied to image restoration tasks. However, without exploiting the multi-scale information, ARCNNs fail to solve more complicated compression artifact problems. Although many deep models with different architectures have been explored (e.g., Dong et al., 2014; Yu et al., 2016; Cavigelli et al., 2017) to solve the artifact reduction problem, there is little work incorporating different models in a unified framework to inherit their respective advantages.

In this paper, a generic fusion network, dubbed as one-two-one (OTO) network, is developed for complex compression artifacts reduction. The general framework of the proposed OTO network is presented in Fig. 2. Specifically, it consists of three subnetworks: a normal-scale network, a small-scale network with max pooling to increase the network receptive field, and a fusion network to perform principled fusion of the outputs from the summation and difference models. The summation model aggregates the low frequency information captured from different network scales, while the difference model is motivated by the Laplacian pyramid which is able to describe the high frequency information, such as detailed information. By combining the summation and difference models, both low and high frequency information of the image can be better characterized. This is motivated by the fact

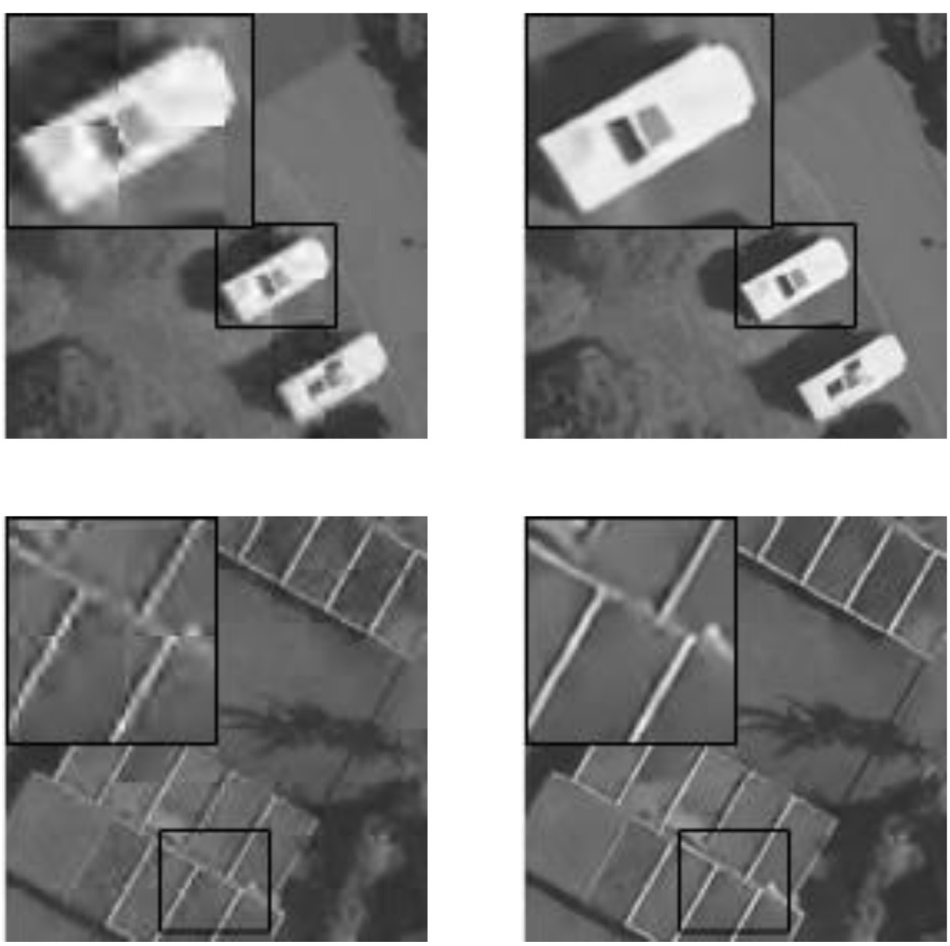

Fig. 1. Left: the SPIHT-compressed remotely sensed images with obvious blocking artifacts. Right: the restored images by our OTO network, where lines are sharp and blurring is removed.

Please cite this article in press as: Zhang, B., et al. One-two-one networks for compression artifacts reduction in remote sensing. ISPRS J. Photogram. Remote Sensing (2018), https://doi.org/10.1016/j.isprsjprs.2018.01.003