Group members:
1. Ike Osafo Nkansah
2. Guillermo de Leon Archila

Deep-Demosaicing using CNNs

Motivation

Digital cameras are readily available all around the world. This was made possible by the development of efficient technologies that improve each step in the image processing pipeline. Many digital cameras use a color filter array (CFA) pattern when recording an image (most commonly the Bayer pattern). The raw image captured by the camera sensor, through the CFA, contains information of only one color per pixel. Here the necessity of an efficient interpolation algorithm for demosaicing arises as a fundamental part of the image processing pipeline.

Traditionally, the demosaicing task has been done by the implementation of linear and nonlinear filters that seek to achieve the best interpolation paradigm. In [1] we can find a comparative study in the performance of nine different demosaicing algorithms. In all nine cases, color artifacts can be observed in some of the processed images. These methods have been outperformed by data-driven approaches, such as the application of convolutional neural networks (CNNs) to solve the demosaicing problem [2]. In this work we seek to evaluate the performance of our own implementation of a convolutional neural network and compare it to traditional demosaicing methods and other data-driven techniques.

An interesting conundrum here is that any training dataset of raw images to colored images would have relied on some specific demosaicing scheme to generate the colored images. This means in principle that with data-driven approaches, we are constrained in performance by whatever approach was used to generate the training dataset. Considering the constraint above, our approach seeks to incorporate priors and regularizers that can mitigate the effect of the prior demosaicing algorithms used to generate the training dataset images. Gharbi Michael et. al. in their paper "Deep Joint Demosaicking and Denoising" [7] note that applying neural machine learning naively to image datasets produce poor performance. In their paper they devise different quality metrics that capture different aspects of what a good reconstruction should yield. It is noted that using machine learning naively struggles just as much as linear and non-linear schemes on image patches with high spatial frequencies. Our work will be to experiment with different neural network architectures, using prior work [2], [3], [4] and [5] as foundations to design our network and to gauge how well our methods perform.

Project Description and Methodology

Objectives: Design a deep neural network (Deep CNN) that is able to faithfully demosaic and denoise even on images with high spatial frequencies. We will measure this using metrics like PSNR and cross-compare with results from traditional methods and those reported by [2], [3], [4] and [5], where applicable. We will also aim to implement regularizations, and incorporate existing metrics (as used by [3]) and new ones, where necessary, to enhance the demosaicing quality of our models.

Methodology:

1. Dataset: The very first step in the implementation of any data-driven learning algorithm is to get a reliable dataset. There are many readily available data-sets that have been used before for the demosaicing task, such as in [6] and [7]. To reduce the complexity of the problem, ideally, we will use a dataset that concentrates only on the demosaicking task and ignores the denoising of the sensor readings (it is common practice to jointly execute the demosaicking and denoising procedures). The dataset provided in [6] has a subset of images used exclusively for the demosaicing task.
2. Models:
    a. *Shallow CNN*: A baseline shallow CNN will be created that takes as input the raw images and generates the demosaiced images. Here we will be experimenting with both the L1 and L2 loss, with other hyperparameter tuning such as batch normalization, layer biasing etc. The result will be a baseline model that will inform how we evaluate other models.
    b. *Deep CNN*: We focus on adjusting the depth of the CNN network from a. together with the hyperparameter tuning insights to observe how depth of the CNN affects the demosaicing task.
    c. *Autoencoder Neural Network*: We will build an autoencoder network that takes as input the raw images and “encodes" it into the RGB format and decodes it back into raw images. Our loss function will then be based on both the generated raw images and/or the generated RGB images in the encoding layers. The motivation here is that whatever RGB image we generate should be reverse decomposable back into its raw image. We believe this may potentially help (or not) the quality of the demosaicing (in particular by acting as a regularizer to dampen the limitations of whatever demosaicing algorithm was used on the training dataset).
    d. *Transfer learning*: This approach will leverage existing pretrained resnet and other architectures and change the output layers and train again for the task of demosaicing. The motivation here is that the weights should have been fine tuned on a large set of images from the natural distribution of images (with rotational invariance etc.) hence the already trained weights could be viewed as regularization to constrain our search space to hopefully yield better demosaicing results.
3. Performance metrics:
    We will evaluate the performance of the CNN after training by comparing the resulting pSNR on images in the test set against the pSNR obtained by demosaicking methods that we have previously explored in class, such as bilinear interpolation and high quality interpolation. Ideally, the CNN will outperform both interpolation methods.

References

[1] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color filter array interpolation," IEEE Signal Processing Magazine, vol. 22, no. 1, pp. 44–54, Jan. 2005.

[2] N. S. Syu, Y. S. Cheng, Y. Y. Chuang, "Learning Deep Convolutional Networks for Demosaicing", Arxiv: Computer Vision and Pattern Recognition. Feb. 2018.

[3] D. S. Tan, W. Chen and K. Hua, "DeepDemosaicking: Adaptive Image Demosaicking via Multiple Deep Fully Convolutional Networks," in *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2408-2419, May 2018, doi: 10.1109/TIP.2018.2803341.

[4] Y. Wang, "A multilayer neural network for image demosaicking," *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 1852-1856, doi: 10.1109/ICIP.2014.7025371.

[5] K. Cui, Z. Jin and E. Steinbach, "Color Image Demosaicking Using a 3-Stage Convolutional Neural Network Structure," *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 2177-2181, doi: 10.1109/ICIP.2018.8451020.

[6] D. Khashabi, S. Nowozin, J. Jancsary, and A. W. Fitzgibbon, "Joint demosaicing and denoising via learned nonparametric random fields," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 4968–4981, 2014.

[7] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, "Deep joint demosaicking and denoising," *ACM Transactions on Graphics*, vol. 35, no. 6, pp. 1–12, 2016.

.