

DCGAN 을 활용한 디모자이킹에 관한 연구

장영채¹, Marco Anisetti², 전광길¹

¹ 인천대학교 임베디드시스템공학과

²Dipartimento di Informatica (DI), Università degli Studi di Milano

e-mail : jyc4261@inu.ac.kr, marco.anisetti@unimi.it, gjeon@inu.ac.kr

A Study on demosaicking using DCGAN

Young-chaе Jang¹, Marco Anisetti², Gwanggil Jeon¹

¹Department of Embedded Systems Engineering, Incheon National University

²Dipartimento di Informatica (DI), Università degli Studi di Milano

요 약

본 연구에서는 일반적으로 R, G, B 색 평면의 높은 상관관계를 이용하여 컬러 복원을 시도하던 기존의 방법의 문제점을 정의하고, DCGAN 을 활용한 디모자이킹에 관한 연구를 소개한다. 약 2000 장의 256x256 이미지를 학습데이터를 이용하였다. 보다 나은 결과를 위하여 R, G, B 색상 채널에 따라, 각각의 네트워크를 구성하고 학습하였다. 제안 방법은 Intel Core i7-7700 CPU(3.60GHz), 16GB Memory, NVIDIA GeForce GTX1080Ti 구성의 Laptop 에서 진행하였고, 평균 PSNR 22.5dB 정도의 성능을 보인다.

1. 서론

높은 품질을 가지는 Demosaicking Algorithm 은 modern 카메라 하드웨어에서 실제 적용을 위해 필수적이며, 컬러 이미지 획득을 위한 중요한 첫 단계이다. 기술적으로 세 개의 CCD 또는 CMOS 센서를 통하여 Full-color channel 이미지를 얻을 수 있지만, 다중 센서의 사용은 결과적으로 고비용을 야기하기 때문에 센서 타입과는 무관하게 일반적으로 단일 센서와 CFA(Color Filter Array) 조합으로 컬러 영상을 얻어낸다. 각 센서 소자는 CFA 패턴에 따라 R, G 또는 B 의 컬러 채널 중 하나의 세기를 기록 할 수 있다. 따라서, 디모자이킹은 컬러 컴포넌트의 누락된 픽셀의 missing color 를 interpolation 하는 과정이 필요하며, 보통 이웃하고 있는 주변 픽셀의 정보를 통해 보간하는 프로세스를 사용한다. 하지만 이 과정은 항상 완벽할 수는 없고 이미지 shapness 저하, Aliasing 발생, reconstruction error 등의 다양한 결함을 동반할 수 있다.

본 논문에서는 최근에 많이 등장한 학습에 기반한 영상처리 방법 중 Generative Adversarial Network(GAN) 과 Convolution Neural Network(CNN)을 이용한 디모자이킹 방법을 소개한다. CNN 은 영상의 특징을 효과적으로 검출하는 특징으로 인해 초 해상도 기법이나 영상 잡음 제거 기법 등에 적용되고 있는 추세이다.[1][2]

2. 본론

다양한 디모자이킹 알고리즘은 지난 수십년 동안 Bayer pattern 에 기반하여 제안되어왔다. Eric Dubois 가 Bayer CFA signal 의 Frequency-domain 에서의 Power

spectral density 가 다른 CFA 보다 이상적이라는 결과 [3]를 바탕으로 CFA 는 Bayer pattern 으로 선정하였다. Input 으로 사용하는 R channel, G channel, B channel 의 모습은 그림 1 을 통과한 신호의 세기가 되며, 그림 2 에 순서대로 나타내었다.



(그림 1) 7x7 Bayer CFA



(그림 2) Red, Green, Blue Input image, Original image

그림 2 을 보면 입력 영상의 차원과 출력 영상의 차원을 일치시키기 위하여 센서 소자에 수광되어지지 않는 값은 zero padding 수행하였다.

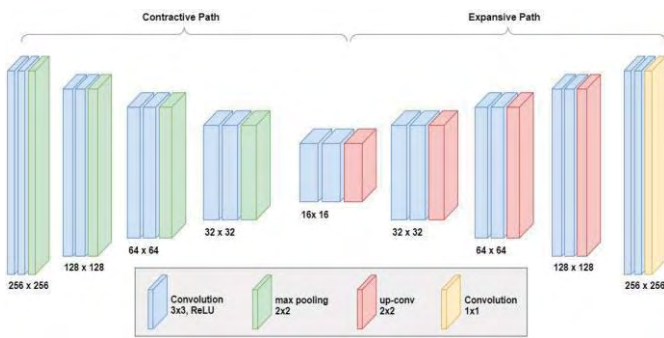
- Architecture guidelines for stable Deep Convolutional GANs
- Replace any pooling layers with strided convolutions (discriminator) and fractional-strided convolutions (generator).
 - Use batchnorm in both the generator and the discriminator.
 - Remove fully connected hidden layers for deeper architectures.
 - Use ReLU activation in generator for all layers except for the output, which uses Tanh.
 - Use LeakyReLU activation in the discriminator for all layers.

(그림 3) Guidelines for stable DCGAN

Architecture guidelines for stable DCGAN[4][5][6]을 따라 DCGAN 모델을 채용하였다.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]. \quad (1)$$

DCGAN 을 수식으로 정리하면, value function V(G, D)에 대한 minimax problem 을 푸는 것과 같다. Goodfellow의 guideline 을 따르면, 위의 value function 에서 log(1-D(G(z)))부분을 Generator 에 대하여 minimize 하는 것보다 log(D(G(z)))부분을 maximize 하도록 학습하는 것이 실용적이라고 소개하고 있다.[4]



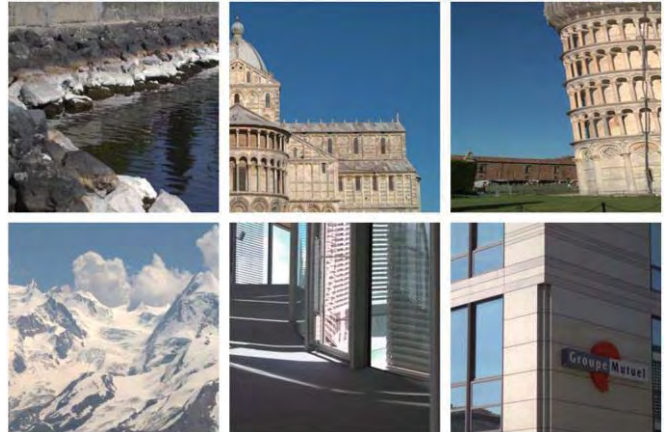
(그림 4) architecture of the model

그림 4 에는 구성한 모델의 architecture 을 나타내었다. Contractive path 는 두 단계의 3 x 3 convolution layer 를 지난 후, 2 x 2 max pooling 과정을 거쳐 다운샘플링을 진행한다. 활성화 함수로서 ReLU(Rectified Linear Unit)을 적용하였다. Expansive path 는 업샘플링 과정으로서 2 x 2 up-convolution layer 로 구성하고, 이는 채널의 수를 반으로 줄이는 과정이다. 마지막 1 x 1 convolution layer 의 output 은 L*a*b* 색 공간을 가지는 3 채널의 결과값이다.

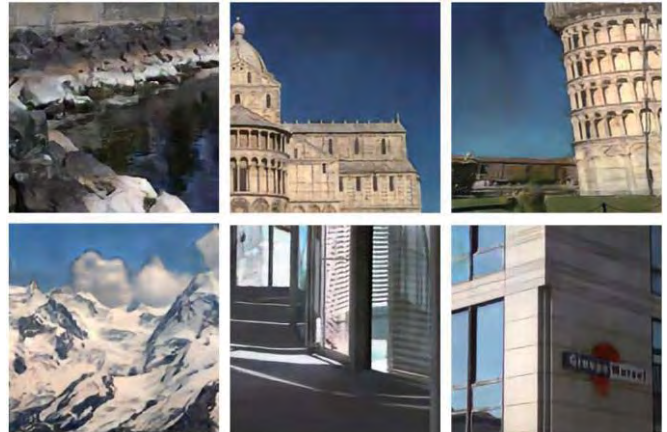
연구는 Intel Core i7-7700 CPU(3.6GHz), 16GB Memory, NVIDIA GeForce GTX1080Ti 구성의 Desktop 에서 진행하였고, MATLAB R2018a 에서 jpg file 을 tiff file 로 변환하여 Dataset 을 생성하였다. jpg file 로 저장시 lossless 로 이미지 저장의 어려움이 있어, resample 시 원본과 PSNR 의 차이를 보였다. 생성한 데이터셋으로 Model training 은 Ubuntu 16.04 에서 진행하였으며, 학습 parameter 는 다음과 같다. Epoch 2000 번, Learning rate 0.001, Batch size 8 으로 테스트 하였다.

3. 결론

본 논문에서는 Convolution 과 Generative Adversarial Network 를 결합한 DCGAN 을 이용한 디모자이킹 연구에 대하여 나타내었다. 표 1 은 연구중인 알고리즘의 객관적 성능에 대한 결과이다. 성능지표로써 원본 신호와 복원된 신호의 Peak signal-to-noise ratio(PSNR) 을 비교하였다. 현 연구의 알고리즘은 평균 PSNR 22.5dB 정도의 성능을 보인다.



(그림 5) 원본 이미지



(그림 6) 디모자이킹 된 이미지

Input Image	1	2	3	4	5	6
PSNR	21.3608	19.7582	22.4494	19.0365	23.8355	28.3208

<표 1> 그림 5 와 그림 6 사이의 PSNR 관계

Acknowledgment

This work was supported in part by the NRF of Korea under Grant NRF-2016K1A3A1A25003543 and by the "Ministero degli Affari Esteri e della Cooperazione Internazionale" of Italy under Grant PGR00217.

참고문헌

[1] X. Li. IEEE Trans. Image Process. 14 (March (3)) 370. (2005)

- [2] L. Zhang, X. Wu. IEEE Trans. Image Process. 14 (December) 2167. (2005)
- [3] E. Dubois, "Frequency-domain methods for demosaicking of Bayer-sampled color image," IEEE Signal Process. Lett, vol.12, no.12, pp847-850, Dec.2005.
- [4] Goodfellow, Ian J., Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron C., and Bengio, Yoshua. Generative adversarial nets. NIPS, 2014.
- [5] Maas, Andrew L, Hannun, Awni Y, and Ng, Andrew Y. Rectifier nonlinearities improve neural network acoustic models. In Proc. ICML, volume 30, 2013.
- [6] Alec Radford & Luke Metz, Unsupervised representation learning with deep convolutional generative adversarial networks. Under review as a conference paper at ICLR 2016.