

컨벌루션 네트워크를 이용한 저조도 환경 카메라 잡음 제거

*박구용 안병용 조남익

서울대학교, 뉴미디어통신공동연구소

*benkay@ispl.snu.ac.kr

Camera noise reduction in the low illumination conditions using convolutional network

*Park, Gu-Yong Ahn, Byeong-Yong Cho, Nam-ik

Seoul National University, INMC

요약

본 논문에서는 카메라 잡음 제거에 딥 러닝 알고리즘을 적용하는 연구를 진행하였다. 합성된 가우시언 잡음에 대하여 좋은 잡음 제거 성능을 보이는 DnCNN(Denoising Convolutional Network)를 이용하여 카메라 잡음을 제거하는 학습과 실험을 진행하였으며, 기존 실험으로는 RGB 색공간의 3채널 모두에 대하여 학습한 신경망(Neural Network)을 사용하였고, 본 논문의 실험에서는 그레이 이미지에 대하여 학습한 신경망을 사용하였다. 신경망의 평가를 위하여 딥 러닝 알고리즘 입력 이미지를 RGB 색공간(RGB Color Space)과 YCbCr 색공간(YCbCr Color Space) 2가지 색공간으로 표현하여 사용하였고, 입력 이미지에 노이즈를 첨가하기 위해 가우시언 노이즈(Gaussian Noise)를 이용하였다. 또한 가우시언 잡음과 다른 성질을 갖는 실제 카메라 잡음에 대해서도 학습과 테스트를 진행하였다.

1. 서론

최근 그래픽카드의 눈부신 성능 향상으로 복잡한 신경망 학습 연산에 소요되는 시간을 줄여주어 다양한 머신 러닝 분야에 딥 러닝이 활용되고 있다. 딥 러닝 알고리즘이 좋은 성능을 보여주는 분야 중 하나가 컴퓨터 비전과 영상처리 분야인데, 최근의 연구 결과들로서 이미지 인식, 이미지 분류, 저해상도 이미지(Low Resolution Image)에서 고해상도 이미지(High Resolution Image)를 추론하는 이미지 초해상 기술과 노이즈 이미지에서 노이즈가 제거된 이미지를 추론하는 이미지 노이즈 제거 등의 분야를 들 수 있다.

그 중 이미지 노이즈 제거 알고리즘은 오래 전부터 연구되어 온 분야이다. 이미지 노이즈 제거 알고리즘의 기본적인 아이디어를 간단하게 설명하면 다음과 같다. 카메라를 통해서 얻은 노이즈 이미지를 y , 노이즈를 v , 노이즈가 제거된 이미지를 x 라고 하면, $y = x + v$ 라는 식이 성립한다. 앞의 식에서 노이즈 모델을 표준편차를 σ 로 가지는 백색 가우시 노이즈(Additive White Gaussian Noise)로 가정하여 우리가 얻고자 하는 잡음이 제거된 이미지 x 를 얻는 것이 목표이다. 베이저안(Bayesian) 관점에서 우도(Likelihood)가 알려진 경우, 사전 이미지 모델링(Image prior modeling)이 굉장히 중요한 역할을 수행하게 된다.

지난 수 년 동안 이미지 노이즈 제거분야에는 수많은 알고리즘이 연구되었다. 사전 이미지 모델링에 NSS(Non-local Self-Similarity) 모델, 희소 모델(Sparse Model), 그라디언트 모델(Gradient Model)과

마르코프 랜덤 필드(Markov Random Field) 등을 사용하는 방법이 연구되고, 특히 BM3D[7], LSSC[8], NCSR[9]과 같은 최첨단 방법에서 NSS 모델을 사용한다. 위에 소개된 방법들은 이미지 제거에 좋은 성능을 보이지만 두 가지 주요 단점을 가지고 있다. 첫 번째 문제점은 테스트 단계에서 복잡한 최적화 문제를 수반하고 있다는 점이다. 따라서 계산의 효율성을 포기하지 않으면서 높은 성능을 기대할 수 없게 된다. 두 번째 문제점은 모델들은 일반적으로 비 볼록형(Non-convex)이고 매개변수(Parameter)를 수동적으로 선택한다는 점이다[6].

본 논문은 이미지 노이즈 제거(Image Denoising)에 기존의 모델들 대신 딥 러닝 알고리즘을 적용하는 연구를 진행하였다. 기본적인 구조는 기존에 가우시언 잡음에 대하여 좋은 성능을 보이는 DnCNN을 이용하였다. 기존 실험으로는 RGB 색공간의 3채널 모두에 대하여 학습이 된 신경망을 사용하였고, 비교실험을 위해서 그레이 이미지에 대해서 신경망 학습을 진행하였다. 이미지는 각각 RGB 색공간과 YCbCr 색공간으로 표현하였고, 딥 러닝 알고리즘의 입력을 생성하기 위해 기존의 이미지에 가우시언 노이즈(Gaussian Noise)를 더하여 노이즈 이미지를 생성하였다. 또한 가우시언 잡음과 다른 성질을 갖는 실제 카메라 잡음에 대해서도 학습과 테스트 실험을 수행하였다.

2. 본론

2.1. YCbCr 색공간(YCbCr Color Space)

디지털 영상을 처리하는 분야에서 가장 널리 쓰이고 있는 색 표현

방식은 RGB 색공간(RGB Color Space)이다. RGB 색공간은 빛의 삼원색인 Red channel, Green channel, Blue channel로 이루어진 색공간이고, 디스플레이 분야와 컴퓨터 기술과 함께 발전되어 사용된 색공간이다. 하나의 색을 표현할 때, RGB 색공간은 가산혼합방식(Additive Color Mixing)을 이용하여 색을 표현한다[1].

YCbCr 색공간은 디지털 영상에서 색을 표현하는 여러 가지 색공간 중 한 가지 색공간이다. YCbCr 색공간은 인간의 눈이 휘도(Luminance)의 차이에 민감하게 반응하고, 색차(Chrominance)의 차이에는 둔감하게 반응한다는 사실에 입각하여 휘도 성분을 이용하기 위해 만든 색공간이다. RGB 색공간에서 좌표계 변환(Coordinate Transform)을 통해 YCbCr 색공간을 정의할 수 있으며, Y channel은 휘도 정보를 나타내고 Cb channel과 Cr channel은 색차 정보를 나타낸다[2]. YCbCr 색공간은 RGB 색공간과 달리 색의 휘도성분을 빠르게 이용할 수 있다는 장점과 RGB 색공간에 비하여 각 색공간간 상관관계(Inter Channel Correlation)가 적다는 장점을 가지고 있다. 그림 1은 RGB 색공간으로 표현된 이미지를 YCbCr 색공간으로 변환하였을 때의 결과를 보여준다.

논문의 실험에서는 이미지들을 각각 RGB 색공간과 YCbCr 색공간으로 표현하여 실험을 진행하였다[3].

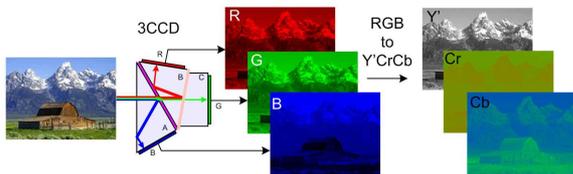


그림 1. RGB 색공간 이미지를 YCbCr 색공간 이미지로 변환한 결과[10]

2.2. 딥 러닝

딥 러닝이란 기계학습(Machine Learning)의 한 분야로서, 컴퓨터가 신경망(Neural Network)를 통해서 데이터를 학습하는 방법을 일컫는다 [4]. 딥 러닝의 개념이 발표된 지는 약 20년 정도 되었지만 이 주목을 받게 시작된 것은 약 5년 전이다. 딥 러닝에 대해 활발하게 연구가 가능하게 된 요인으로는 하드웨어의 급격한 발전과 빅 데이터의 등장을 꼽을 수 있다.

딥 러닝은 심층 신경망(Deep Neural Network, DNN), 합성곱 신경망(Convolutional Neural Network, CNN), 순환 신경망(Recurrent Neural Network, RNN) 등 다양한 신경망들을 사용하여 학습을 진행한다. 그 중 합성곱 신경망 기반의 딥 러닝 알고리즘이 컴퓨터 비전(Computer Vision)과 음성 인식(Speech Recognition) 등의 분야에서 뛰어난 성능을 나타내고 있으며 합성곱 신경망을 이용한 딥 러닝의 기본구조는 그림2와 같다 [5]. 컴퓨터 비전 분야 중 이미지 노이즈 제거(Image Denoising) 분야에 대해서 좋은 성능을 보여주는 딥 러닝 알고리즘들이 많이 발표되고 있다.

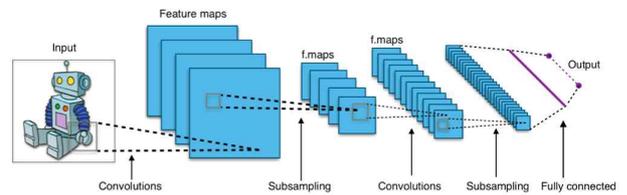


그림 2. 합성곱 신경망을 이용한 딥 러닝의 기본구조[11]

본 논문에서는 가우시언 잡음에 대하여 좋은 성능을 보이는 DnCNN(Denoising Convolutional Neural Network)를 이용하여 실험을 진행하였다. 실험에 사용한 신경망의 구조는 그림 3과 같다. 처음 Conv + ReLU 레이어의 구조는 합성곱 신경망과 ReLU 신경망으로 이루어져 있다. 크기가 $3 \times 3 \times$ 채널수로 이루어진 64개의 필터를 사용한다. 두 번째 Conv + BN + ReLU의 경우 크기가 $3 \times 3 \times 64$ 인 필터 64개를 사용하고 배치 정규화(Batch Normalization, BN)이 추가되었다. 마지막 Conv 레이어의 경우 $3 \times 3 \times 64$ 크기의 필터를 사용하여 출력을 재구성한다 [6].

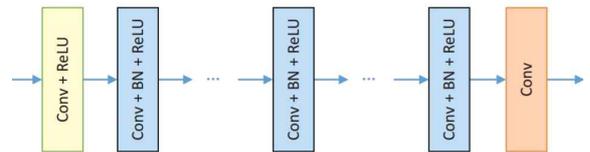


그림 3. DnCNN 신경망 구조[6]

3. 실험 결과

먼저 기준 값을 설정하기 위해 RGB 색공간으로 표현된 이미지 3 채널 모두에 대해서 학습한 신경망을 이용하였다. RGB 색공간으로 표현된 이미지를 테스트 이미지로 설정하여 실험을 진행하였고, 신호 대 잡음비(PSNR)를 측정기준으로 이용하였다. 실험결과로 나온 기준 값은 37.9425dB이다.

또한 본 연구에서는 그레이 이미지에 대해서 학습한 신경망을 이용하여 실험을 진행하였다. 학습에 사용된 이미지 수는 400장이고 학습에 사용된 가우시 노이즈의 σ 의 값은 5, 10, 15, 20을 사용하였다. 학습을 통해서 노이즈를 제거하여 얻은 결과 이미지와 기준 이미지(Reference Image)는 신호 대 잡음비를 이용하여 비교하였고 표1에 정리하였다.

표 1. 실험결과

	기준 값	실험 값
RGB 색공간	37.9425	38.1763
YCbCr 색공간		38.6317

그레이 이미지로 트레이닝을 진행 한 실험의 경우, 기준 실험과는 다르게 이미지의 각 채널을 신경망의 입력으로 사용하여 테스트를 진행하였다. 표 1의 실험결과를 보면 YCbCr 색공간으로 표현된 테스트

이미지가 RGB 색공간으로 표현된 이미지보다 더 좋은 결과를 보임을 알 수 있다.

실험 결과를 각 색공간에 대하여 논의해보면, 그레이 이미지로 학습을 진행한 신경망의 경우가 더 좋은 결과 값을 나타낸다. 특히 YCbCr 색공간의 경우 높은 성능향상을 나타내고 있다. 신경망을 학습할 때 휘도 정보를 가지는 그레이 이미지를 이용하여 진행하였기 때문에, 휘도 정보를 나타내는 Y 채널의 영향임을 추론할 수 있다.

4. 결론 및 향후 연구 방향

이번 실험에서는 DnCNN을 이용하여 이미지 노이즈 제거에 대한 연구를 진행하였다. RGB 색공간으로 표현된 이미지와 그레이 이미지에 대해서 각각에 대하여 트레이닝한 신경망을 이용하였다. 테스트 이미지를 RGB 색공간과 YCbCr 색공간으로 표현한 후 테스트를 진행하였다. 그레이 이미지를 이용하여 트레이닝한 신경망을 이용한 실험 결과를 보면 YCbCr 색공간 이미지가 RGB 색공간 이미지보다 더 좋은 결과를 보이는 것을 알 수 있다. 이를 통해서 향후 연구 방향을 설정할 수 있었다.

실험결과를 토대로 RGB 색공간 외의 다른 색공간들을 이용하여 신경망 학습을 예상해보면, 더 좋은 결과 값이 나올 수 있음을 예상할 수 있다. 실험에서는 그레이 이미지에 대해 신경망을 학습하였는데, 향후 실험에서 YCbCr 색공간의 Y 채널, HSV 색공간의 V채널 등 이미지의 휘도 정보를 나타내는 채널에 대해서 학습을 진행한 후 실험을 진행할 수도 있을 것이다.

5. 참고문헌

- [1] Hasting, G. D., and A. Rubin. "Colour spaces—a review of historic and modern colour models." *African Vision and Eye Health* 71.3 (2012): 133-143.
- [2] Ibraheem, Noor A., et al. "Understanding color models: a review." *ARNP Journal of Science and Technology* 2.3 (2012): 265-275.
- [3] Lucas, Luís FR, et al. "Intra-prediction for color image coding using YUV correlation." *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010.
- [4] Deng, Li, and Dong Yu. "Deep learning: methods and applications." *Foundations and Trends® in Signal Processing* 7.3 - 4 (2014): 197-387.
- [5] Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton. "Speech recognition with deep recurrent neural networks." *Acoustics, speech and signal processing (icassp), 2013 IEEE international conference on*. IEEE, 2013.
- [6] Zhang, Kai, et al. "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising." *IEEE Transactions on Image Processing* (2017).
- [7] Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8), 2080-2095.

[8] Mairal, J., Bach, F., Ponce, J., Sapiro, G., & Zisserman, A. (2009, September). Non-local sparse models for image restoration. In *Computer Vision, 2009 IEEE 12th International Conference on* (pp. 2272-2279). IEEE.

[9] Dong, W., Zhang, L., Shi, G., & Li, X. (2013). Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4), 1620-1630.

[10] 그림1. <https://en.wikipedia.org/wiki/YCbCr>

[11]

그림2.

https://en.wikipedia.org/wiki/Convolutional_neural_network