

ADMM과 깊은 합성곱 신경망 잡음 제거기 이미지 Prior에 기반한 이미지 디블러링

권준형 소재웅 조남익

서울대학교 전기정보공학부 뉴미디어통신공동연구소

{gjh8760, soh90815}@ispl.snu.ac.kr nicho@snu.ac.kr

Image Deblurring Based on ADMM and Deep CNN Denoiser Image Prior

Kwon, Junhyeong Soh, Jae Woong Cho, Nam Ik

Department of ECE, INMC, Seoul National University

요약

오래 전부터 모델 기반 최적화 방법이 이미지 디블러링을 위해 널리 사용되어 왔고, 최근에는 학습 기반 기술이 영상 디블러링에서 좋은 성과를 보이고 있다. 본 논문은 ADMM과 깊은 합성곱 신경망 잡음 제거기 이미지 prior를 이용하여 모델 기반 최적화 방법의 장점과 학습 기반 방법의 장점을 모두 활용할 수 있는 방법을 제안한다. 본 방법을 이용하여 기존 방법보다 더 좋은 디블러링 성능을 얻을 수 있었다.

1. 서론

이미지 디블러링(Image deblurring)은 블러된 관측 이미지로부터 블러가 없는 깨끗한 이미지를 복원하는 과정이다. 블러가 균일하고 위치에 독립적이라면, 이 과정은 다음과 같은 수식으로 표현할 수 있다.

$$y = Hx + n \quad (1)$$

여기서 y 는 블러된 이미지, x 는 깨끗한 원본 이미지, H 는 블러 커널(blur kernel)을 의미하는 연산자, n 은 잡음을 나타낸다.

관측된 블러 이미지 y 에서 원본 이미지 x 를 추정하는 것은 ill-posed 문제이므로, 이미지 디블러링은 일반적으로 어려운 문제이다. 따라서 모델 기반 최적화 방법(model-based optimization method)은 블러 커널이 주어지는 경우인 논블라인드 이미지 디블러링(non-blind image deblurring)을 해결하기 위해 아래 수식과 같이 사후 확률 최대화 기법(MAP, Maximum a posteriori)을 이용한다.

$$\hat{x} = \arg \min_x \frac{1}{2} \|y - Hx\|^2 + \lambda \Phi(x) \quad (2)$$

여기서 $\frac{1}{2} \|y - Hx\|^2$ 은 추정한 이미지의 정확도를 반영하는 데이터 항(fidelity term)이고, $\Phi(x)$ 는 추정한 이미지의 성질을 담고 있는 사전 정보 항(prior term)이다.

모델 기반 최적화 방법은 학습 기반 방법과 달리 훈련에 필요한 데이터셋이 필요하지 않고, 여러 종류의 이미지 디블러링에 유연하게 쓰일 수 있다는 장점이 있다. 하지만 학습 기반 방법에 비해 결과 이미지를 얻기 위한 처리 속도가 느리고, 좋은 성능을 내기 위해 다양하고 복잡한 이미지 prior를 사용해야 한다는 단점이 존재한다. 반면 학습 기반 방법

은 복잡한 이미지 prior를 필요로 하지 않고, 결과 이미지를 매우 빠르게 취득할 수 있는 대신 대량의 학습 데이터를 필요로 한다는 단점이 있다.

IRCNN[1]은 수식 (2)를 변수 분리 기법(variable splitting) 중 한 종류인 HQS (Half Quadratic Splitting)를 사용하여 분리하고, 복잡한 다른 이미지 prior 대신 깊은 합성곱 신경망 잡음 제거기에 기반한 데이터로부터 학습된 이미지 prior를 사용함으로써 효과적으로 이미지 디블러링을 수행하였다.

본 논문에서는 HQS 대신 다른 변수 분리 기법 중 하나인 ADMM (Alternating Direction Method of Multipliers)을 사용하여 논블라인드 영상 디블러링을 수행하였다. 제안된 방법의 성능을 평가하기 위하여, 이미지 디블러링에서 널리 쓰이는 데이터셋인 Set3에 대해 실험한 결과를 비교하였다.

2. HQS 방법

수식 (2)는 보조변수(auxiliary variable) z 를 사용하여 다음과 같이 쓸 수 있다.

$$\hat{x}, \hat{z} = \arg \min_{x, z} \frac{1}{2} \|y - Hx\|^2 + \lambda \Phi(z) \quad s.t. \quad z = x \quad (3)$$

HQS 방법은 다음과 같은 비용 함수를 최소화하게 된다.

$$\mathcal{L}(x, z, \rho) = \frac{1}{2} \|y - Hx\|^2 + \lambda \Phi(z) + \frac{\rho}{2} \|z - x\|^2 \quad (4)$$

(4)는 다음 두 식을 반복적으로 계산함으로써 최소화된다.

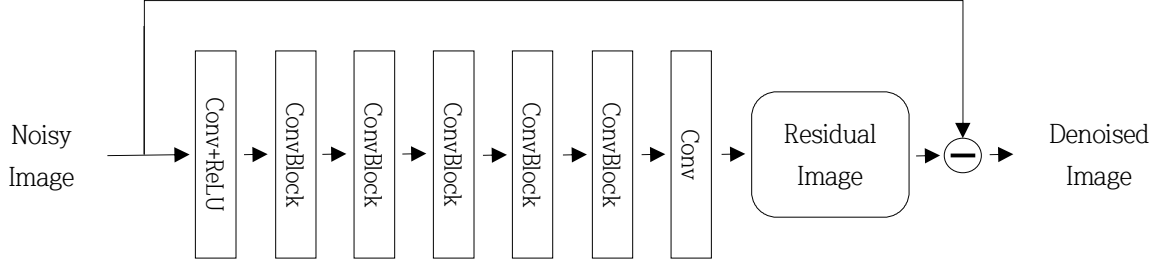


그림 1 사용한 깊은 합성곱 신경망 잡음 제거기의 구조

$$\begin{cases} x_{k+1} = \arg \min_x \|y - Hx\|^2 + \rho \|x - z_k\|^2 \\ z_{k+1} = \arg \min_z \frac{\rho}{2} \|z - x_{k+1}\|^2 + \lambda \Phi(z) \end{cases} \quad (5)$$

이 때 ρ 는 반복 횟수에 대해 단조증가 하도록 설정한다. 반복을 거듭하면서 ρ 가 커지게 되면 수식 (4)의 비용 함수를 최소화하는 문제는 수식 (3)과 같아지게 된다.

수식 (5)에서 x_{k+1} 을 구하는 문제는 x 에 대한 이차 형태 (quadratic form)를 최소화시키는 문제이므로, 이를 x 로 미분하여 정확한 x_{k+1} 을 구할 수 있다. x_{k+1} 은 다음과 같다.

$$x_{k+1} = (H^T H + \rho I)^{-1} (H^T y + \rho z_k). \quad (6)$$

수식 (6)에서 역행렬을 구하는 것은 대략 $O(n^3)$ 의 시간 복잡도를 가지므로, 사용하는 이미지가 커지면 처리 시간이 매우 길어진다. 고속 푸리에 변환(Fast Fourier Transform)을 이용하면 수식 (6)을 효율적으로 계산할 수 있다[2].

수식 (5)에서 z_{k+1} 을 구하는 문제는 아래와 같이 다시 쓸 수 있는데,

$$z_{k+1} = \arg \min_z \frac{1}{2(\sqrt{\lambda/\rho})^2} \|x_{k+1} - z\|^2 + \Phi(z) \quad (7)$$

이는 입력 이미지 x_{k+1} 을 잡음 레벨 $\sqrt{\lambda/\rho}$ 에 해당하는 가우시안 잡음 제거기에 통과시킨 결과와 같다.

$$z_{k+1} = \text{Denoiser}(x_{k+1}, \sqrt{\lambda/\rho}) \quad (8)$$

3. 제안하는 방법

3.1. 제안하는 ADMM 방법

ADMM 방법은 아래와 같은 augmented Lagrangian function이라고 불리는 비용 함수를 최소화한다.

$$\begin{aligned} \mathcal{L}(x, z, \mu, \rho) &= \frac{1}{2} \|y - Hx\|^2 + \lambda \Phi(z) \\ &\quad + \mu^T (x - z) + \frac{\rho}{2} \|x - z\|^2 \end{aligned} \quad (9)$$

수식 (9)는 다음 세 개의 식을 반복적으로 계산함으로써 최소화된다.

$$\begin{cases} x_{k+1} = \arg \min_x \|y - Hx\|^2 + \rho \|x - z_k + \bar{\mu}_k\|^2 \\ z_{k+1} = \arg \min_z \frac{\rho}{2} \|z - x_{k+1} - \bar{\mu}_k\|^2 + \lambda \Phi(z) \\ \bar{\mu}_{k+1} = \bar{\mu}_k + x_{k+1} - z_{k+1} \end{cases} \quad (10)$$

이 때 ρ 는 반복 횟수에 대해 단조증가 하도록 설정한다. 여기서 $\bar{\mu}_k = \frac{1}{\rho} \mu_k$ 이다. 반복을 거듭하면서 ρ 가 커지게 되면 수식 (9)의 비용 함수를 최소화하는 문제는 수식 (3)과 같아지게 된다.

HQS 방법과 같은 과정을 따라가면, 수식 (10)에서 x_{k+1} 과 z_{k+1} 을 아래와 같이 구할 수 있고, 이 중 x_{k+1} 는 고속 푸리에 변환을 이용하여 효율적으로 빠르게 구할 수 있다.

$$x_{k+1} = (H^T H + \rho I)^{-1} (H^T y + \rho(z_k - \bar{\mu}_k)) \quad (11)$$

수식 (10)의 두 번째 식은 아래와 같이 다시 쓸 수 있다.

$$z_{k+1} = \arg \min_z \frac{\rho}{2} \|z - x_{k+1} - \bar{\mu}_k\|^2 + \lambda \Phi(z) \quad (12)$$

이는 입력 이미지 $x_{k+1} + \bar{\mu}_k$ 을 잡음 레벨 $\sqrt{\lambda/\rho}$ 에 해당하는 가우시안 잡음 제거를 수행하는 식이므로, 외부 데이터를 이용하여 학습시킨 잡음 레벨 $\sqrt{\lambda/\rho}$ 에 해당하는 가우시안 잡음 제거기를 통과시키는 것으로 근사할 수 있다. 이를 식으로 나타내면 다음과 같다.

$$z_{k+1} = \text{Denoiser}(x_{k+1} + \bar{\mu}_k, \sqrt{\lambda/\rho}) \quad (13)$$

모델 기반 최적화 방법에 기반한 다른 여러 가지 이미지 디블러링 방법들은 효과적인 디블러링을 위해 TV (total variation) 이미지 prior[3]와 같은 잘 설계된 이미지 prior $\Phi(z)$ 를 사용하였으나, 제안한 방법을 사용하면 대량의 데이터로부터 미리 학습된 가우시안 잡음 제거기가 이러한 이미지 prior를 대신할 수 있다.

3.2. 사용한 잡음 제거기 구조

복잡한 이미지 prior를 대신하기 위해 본 논문에서 사용한 깊은 합성곱 신경망 잡음 제거기의 구조는 IRCNN과 같은 구조를 사용하였다. 일반적인 conv 레이어 대신 dilated conv 레이어[4]를 사용함으로써 효과적으로 receptive field를 확대하였고, 학습 속도를 빠르게 하기 위해 배치 정규화(batch normalization)[5]와 잔차 학습(residual

learning][6]을 도입하였다.

사용한 잡음 제거기는 총 7개의 3×3 dilated conv 레이어로 구성되고, 이 중 첫 번째 dilated conv 레이어는 dilated conv + ReLU로, 마지막 dilated conv 레이어는 뒤에 아무것도 추가하지 않은 dilated conv를 사용하였다. 첫 번째와 마지막 레이어를 제외한 나머지 레이어는 dilated conv +배치 정규화 + ReLU로 이루어진 convblock을 사용하였다. 7개의 dilated conv 레이어의 dilated factor는 첫 번째 레이어부터 마지막 레이어 순서대로 1, 2, 3, 4, 3, 2, 1로 설정하였고, 모든 레이어의 피쳐맵 개수는 64로 고정하였다. 잡음 제거기의 전체적인 구조를 그림 1로 나타내었다.

3.3. 실험 세부 사항

ADMM으로 수식 (9)의 비용 함수를 최소화하기 위해서는 ADMM의 iteration에 따라 ρ 를 점점 증가시켜야 한다. ρ 를 점점 증가시키는 대신, λ 를 고정하고 잡음 레벨인 $\sigma = \sqrt{\lambda/\rho}$ 를 큰 값에서 작은 값으로 감소 시킴으로써 ρ 를 점차 증가시키는 것과 같은 효과를 낼 수 있다. 본 실험에서는 잡음 레벨을 $\sigma = 2:2:50$ 으로 한 총 25가지의 잡음 제거기를 BSDS500 데이터셋[7]을 이용하여 학습시켰고, 잡음 레벨 σ 를 50부터 2까지 2씩 감소시켜 총 25번의 iteration을 ADMM 내에서 반복하였다. 25번의 iteration 동안 λ 값은 0.0001로 고정하였다.

4. 실험 결과

성능 평가에는 영상 디블러링 뿐만 아니라 다양한 이미지 복원 (image restoration) 분야에서 쓰이는 데이터셋인 Set3의 이미지들을 사용하였다. IRCNN과의 결과 비교를 위해 IRCNN에서와 같이 가우시안 블러 커널, Levin et al.[8]에서 사용한 블러 커널 두 가지로 총 세 가지의 블러 커널을 이용하여 Set3의 블러 처리된 이미지를 만들었다. 세 가지의 블러 커널을 그림 2에 시각화하였다. 정량적으로 성능을 평가하기 위해 PSNR을 측정하여 비교하였고, 디블러링을 수행한 실험 결과는 표 1에 정리하였다.

실험 결과, 블러 커널 1, 2에 대해서는 제안한 방법의 PSNR이 IRCNN의 결과보다 대체로 높게 측정되었으나, 가우시안 블러 커널로 블러 처리된 이미지들은 모두 IRCNN보다 PSNR이 낮게 측정되었다. 실험 결과를 정성적으로 확인하기 위해, 제안한 방법으로 Set3에 대해 디블러링을 수행한 결과를 그림 3에 나타내었다. 제안한 방법이 블러 처리된 이미지를 효과적으로 디블러링하는 것을 확인할 수 있다.

5. 결론

본 논문에서는 변수 분리 기법 중 하나인 ADMM을 사용하였고, 이미지 prior를 직접 설계하는 대신 외부 데이터를 이용하여 학습시킨 잡음 제거기에 기반하여 영상 디블러링을 수행하였다. HQS보다 보조 변수를 하나 더 사용하는 ADMM을 사용하여 IRCNN보다 평균적으로 더 좋은 디블러링 결과를 보였다. 본 논문에서 사용한 잡음 제거기는 IRCNN에서 사용한 잡음 제거기와 동일한 구조를 가지고 있는데, 이를 좀 더 개선하거나 잡음 제거기 이미지 prior에 추가적인 이미지 prior를 사용하면 다양한 블러 커널에 대해서도 좀 더 좋은 결과를 낼 수 있을 것으로 기대된다.

표 1 Set3에서의 PSNR 측정 결과. 표의 가장 아래 행은 전체 9개의 이미지에서의 PSNR 평균을 의미한다. 더 높은 수치를 강조하기 위해 빨간색으로 나타내었다.

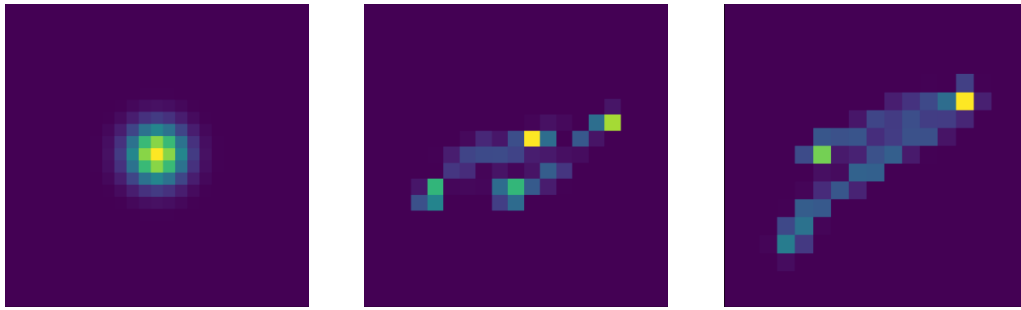
Image	Blur kernel	IRCNN[1]	Ours
Monarch	Gaussian	30.00	29.94
	kernel 1	33.62	33.95
	kernel 2	33.13	33.25
Leaves	Gaussian	29.78	29.74
	kernel 1	33.92	33.95
	kernel 2	33.51	33.69
Parrots	Gaussian	32.07	31.67
	kernel 1	35.49	35.87
	kernel 2	35.17	35.17
		32.97	33.03

감사의 글

이 논문은 2020년도 BK21 플러스 창의정보기술 인재양성사업단에 의하여 지원되었음.

참고문헌

- [1] K. Zhang, W. Zuo, S. Gu and L. Zhang. "Learning deep CNN denoiser prior for image restoration." IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [2] M. E. Helou, F. Dümbgen, R. Achanta and S. Süsstrunk. "Fourier-domain optimization for image processing." arXiv preprint arXiv:1809.04187, 2018.
- [3] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin. "An iterative regularization method for total variation-based image restoration." Multiscale Modeling & Simulation, 2005.
- [4] F. Yu and V. Koltun. "Multi-scale context aggregation by dilated convolutions." arXiv preprint arXiv:2622.07122, 2015.
- [5] S. Ioffe and C. Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." International Conference on Machine Learning. 2015.
- [6] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang. "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising." IEEE Transactions on Image Processing. 2017.
- [7] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li and L. Fei-Fei. "Imagenet: A large-scale hierarchical image database." IEEE Conference on Computer Vision and Pattern Recognition. 2009.
- [8] A. Levin, Y. Weiss, F. Durand and W. T. Freeman. "Understanding and evaluating blind deconvolution algorithms." IEEE Conference on Computer Vision and Pattern Recognition. 2009.



(a) 가우시안 블러 커널 (b) 블러 커널 1 (c) 블러 커널 2

그림 2 실험에 사용된 세 종류의 블러 커널을 시각화한 모습



Blurred image

Groundtruth

Output image

(a) Set3의 “Monarch” 이미지



Blurred image

Groundtruth

Output image

(b) Set3의 “Leaves” 이미지



Blurred image

Groundtruth

Output image

(c) Set3의 “Parrots” 이미지

그림 3 Set3에서의 실험 결과