

Image Demoiréing with Learnable Bandpass Filters

김정현

Vision & Display Systems Lab.

Dept. of Electronic Engineering, Sogang University

Outline

- Background
 - Aliasing & Moire
- Introduction
- Method
 - Moiré Photo Restoration Using Multiresolution Convolutional Neural Networks
 - Image Demoireing with Learnable Bandpass Filters
- Experiments
- Reference

Background

- Aliasing & Moire

- Aliasing: 신호 처리에서 표본화를 할 때 각기 다른 신호를 구별해내지 못하는 효과
 - Sampling 할 때, 주파수의 차이로 인한 현상
 - Sampling 주파수가 2배 이상이어야 aliasing 현상이 생기지 않음
- Moire: 규칙적으로 반복 되는 모양이 여러 번 거듭하여 합쳐졌을 때, 주기의 차이에 따라 생기는 물결무늬의 패턴
 - Moire는 대표적인 aliasing 현상
 - Moire는 주로 모니터를 촬영할 때 많이 발생함

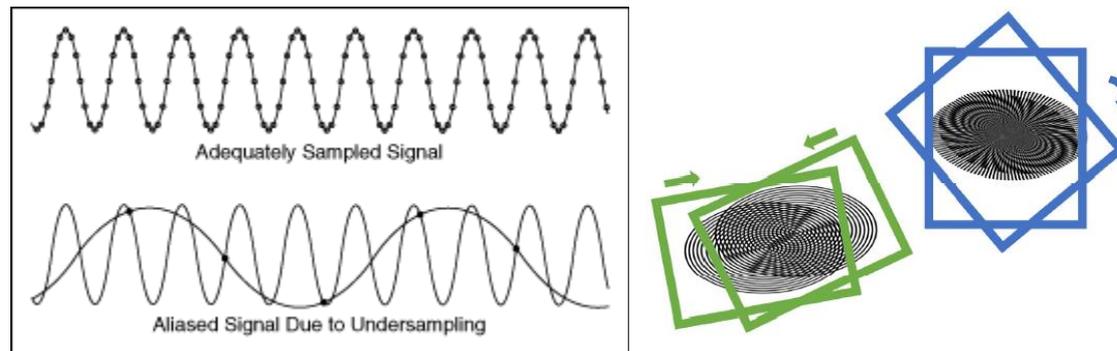
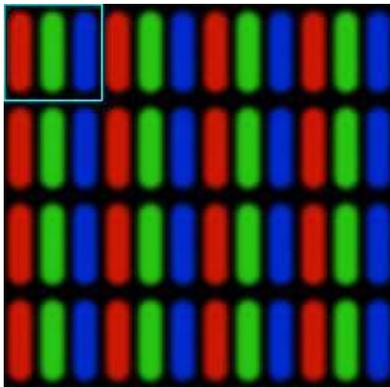


그림1. Aliasing 현상과 moire 현상

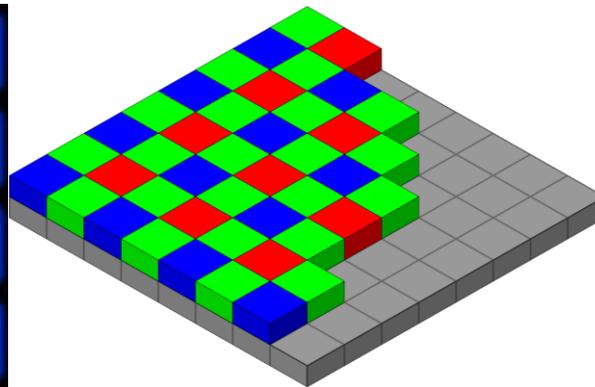
Background

- Aliasing & Moire

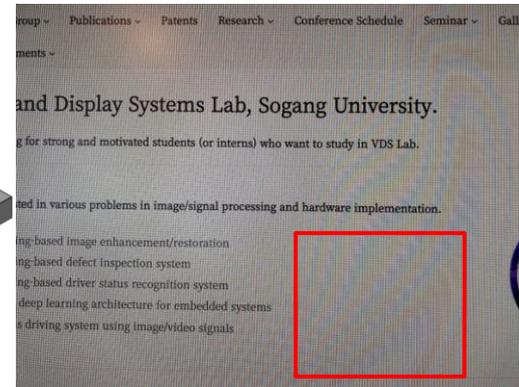
- 모니터의 RGB subpixel과 카메라 image sensor의 color filter array 의 주파수 차이로 인해 Aliasing 현상이 발생
 - Aliasing현상으로 moire가 생성 됨
- 카메라로 주기적인 특성이 있는 피사체를 촬영하면 moire가 발생
 - LCD모니터 위주로 발생하고, 방충망, 줄무늬 옷을 촬영할 때도 발생 함
- Moire는 image에 불규칙, 다양한 크기, 색깔, 주파수의 artifact를 생성
 - Image 품질을 저하시킴



RGB subpixel



camera의 color filter array



moire 예시

Background

- Aliasing & Moire
 - 영상의 품질 개선을 위해 Demoire가 필요
 - 렌즈의 필터를 사용해 전 처리 과정으로 demoire도 가능하지만 사용성이 떨어짐
 - 촬영의 편의성이 낮아져 사용 빈도가 낮음
 - 렌즈필터 또한 고가
 - Anti-aliasing 기능이 탑재된 카메라도 출시
 - 불필요한 anti-aliasing 효과로 image의 quality 저하로 인한 소비자의 선호도 감소
 - 때문에, Demoire 과정은 전처리가 아닌 후처리 위주로 진행



Image Demoireing with Learnable Bandpass Filters

- Introduction

- Demoire는 moire pattern 으로 손상된 image로부터 clean image를 복구하는 task
 - 주파수, 모양, 색상 등의 다양한 이유로, 현재까지 해결이 어려움
- 기존의 방법들은 moire pattern을 모델링을 하지 않음
 - moire pattern을 modeling하는 multi scale bandpass convolutional network (MBCNN) 을 고안
 - Pattern 보다 frequency를 먼저 학습해서 moire pattern 을 modeling
 - 그후 moire image를 texture와 color에 대해서 따로, 복원
- DCT network는 frequency domain에서 moire제거
 - 주파수 영역에서의 학습을 통해 효과적으로 제거 가능
 - Pixel network는 spatial domain 에서 moire를 제거
 - 영상 내에서의 texture와 moire pattern은 구분이 어려움
 - ※ 주파수 영역에서는 계수 분포가 서로 다른 특징을 가짐

Method

- MBCNN

- 카메라에 촬영된 moire 이미지의 modeling

$$I_{moire} = \psi(I_{clean}) + N_{moire} \quad (1)$$

- I_{moire} : Moire image
- I_{clean} : screen 에서 출력하는 clean image
- N_{moire} : Moire texture
- ψ : camera sensor와 screen때문에 생긴 color degradation

$$I_{clean} = \psi^{-1}(I_{moire} - N_{moire}) \quad (2)$$

- ψ^{-1} : ψ 의 inverse function (tone mapping)
- Clean image는 I_{moire} 에서 N_{moire} 를 제거하고, tone mapping (ψ^{-1})을 순서대로 진행

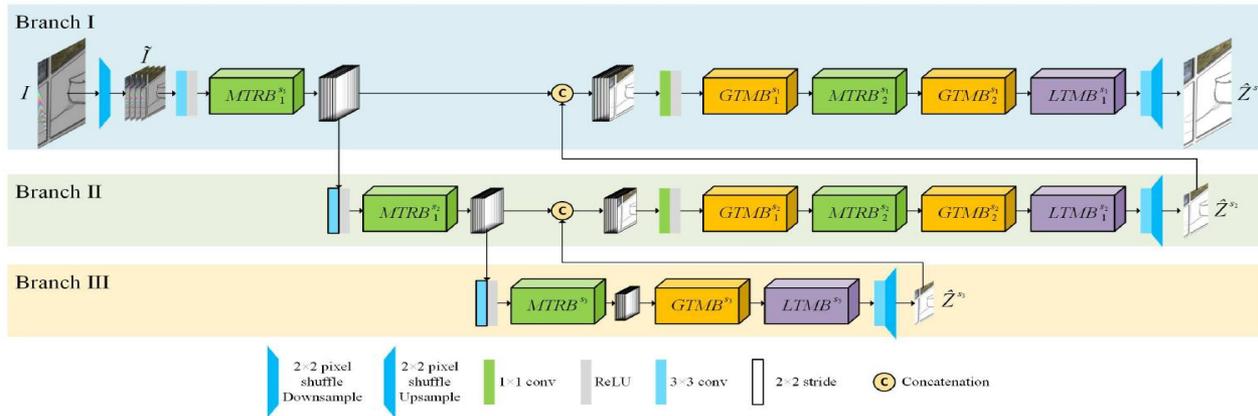
- Image를 modeling 하고, demoireing는 2 step으로 분할

- Moire texture removal
- Tone mapping

Method

- MBCNN

- 3개의 scale과 3가지의 block으로 구성
 - Branch별 각기 다른 3개의 scale
 - Moire texture removal block (MTRB), Global tone mapping block (GTMB), Local tone mapping block (LTMB)
- 각각의 branch 별로 순서대로 moire 제거와 tone mapping을 진행
- 최종 출력단에서 up-scaling
- Up-scaling으로 인한 texture, color error제거를 위해 GTMB,MTRB를 추가



Method

- Moire texture removal

- Moire pattern 에 대해서 scale (s_i) 과 frequency (f_{ij}) 단위로 분석

$$N_{moire} = \sum_i \sum_j N_{f_{ij}}^{s_i} \quad (3)$$

- Frequency problem에 대해서 효과적인 처리를 위해 Block-DCT를 적용

- 각각의 block-DCT domain에 있는 $N_{f_{ij}}^{s_i}$ 의 frequency spectrum을 $FS_{f_{ij}}^{s_i}$ 라고 가정

$$\begin{aligned} N_{moire} &= \sum_i \sum_j D^{-1}(FS_{f_{ij}}^{s_i}) \\ &= D^{-1}\left(\sum_i \sum_j FS_{f_{ij}}^{s_i}\right) \end{aligned} \quad (4)$$

- 각각 color channel에 따른 Moire는 N_P^C , $C \in \{R, G, B\}$

- C: learnable convolution, P는 color image patch라고 할 때

$$C(N_P) = \sum_{C \in \{R, G, B\}} C(N_P^C) \quad (5)$$

Method

- Moire texture removal

- Eq.4 의 Moire를 Eq.5 의 moire에 대입

- $FS^{s_i}|_c$ 는 color channel C의 결합된 frequency spectrum, s_i 는 scale

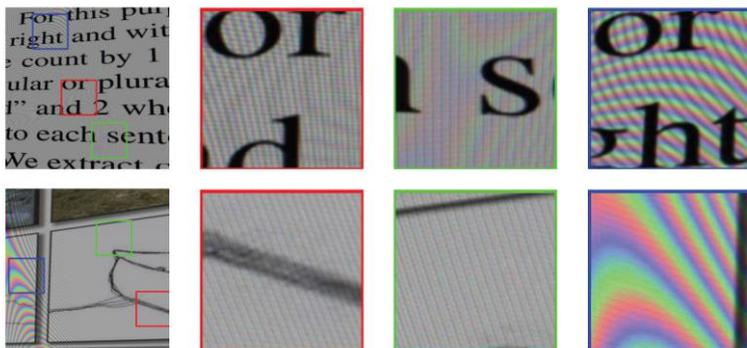
$$\begin{aligned}
 C(N_p) &= \sum_{C \in \{R,G,B\}} C(D^{-1}(\sum_i \sum_j FS_{f_{ij}}^{s_i})) \Big|_c \\
 &= \sum_i C(D^{-1}(\sum_{C \in \{R,G,B\}} \sum_j FS_{f_{ij}}^{s_i} \Big|_c)) \\
 &= \sum_i C(D^{-1}(\sum_{C \in \{R,G,B\}} FS^{s_i} \Big|_c))
 \end{aligned} \tag{6}$$

- $\sum_{C \in \{R,G,B\}} FS^{s_i} \Big|_c = \xi^{s_i}$ 라고 가정하면, ξ^{s_i} 는 Conv연산으로 표현이 가능[4,49]

$$C(N_p) = \sum_i C(D^{-1}(\xi^{s_i})) \tag{7}$$

Method

- Moire texture removal
 - Deep CNN block으로 ξ^{si} 를 추정할 수 있음 [3]
 - Eq. 7은 전부 linear, convolution layer로 modeling이 가능함
 - 특정 주파수를 증폭하거나 감소시키기 위해서 bandpass filter를 사용가능
 - 그러나, moire texture를 modeling 하기전에 frequency spectrum을 얻는 것은 어려움
 - Moire는 scale, frequency,color의 특성이 다르며 서로 영향을 줄 수도 있음
 - 각 frequency마다 learnable weights θ^{si} 를 ξ^{si} 와 곱해서 문제를 해결
 - θ^{si} 는 learnable weights of DCT domain frequencies

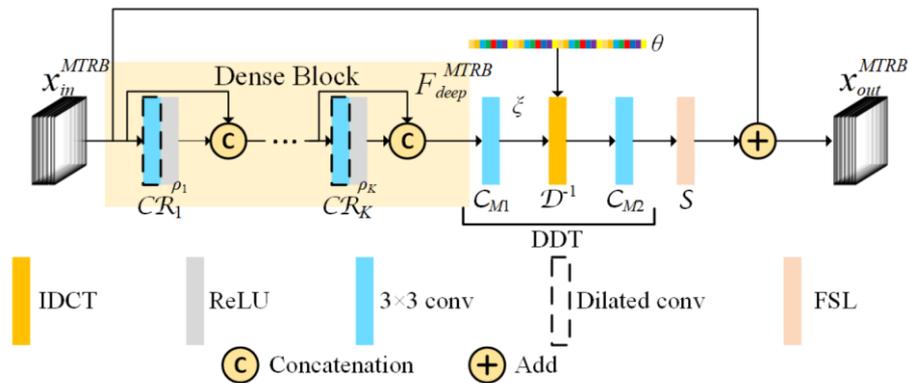


$$C(N_p) = \sum_i C(D^{-1}(\theta^{si} \cdot \xi^{si})) \quad (8)$$

Method

- Moire texture removal
 - Dense block은 F_{deep} 을 생성하는 feature extraction,
 - Dense block은 K개의 $3 \times 3 \times n_D$ 의 dilated convolution (with ReLU) 으로 연결
 - ⋈ Receptive field를 늘리기 위함
 - 3×3 convolution 으로 F_{deep} 에서 ξ 추정 한 뒤, learnable weight θ , block-IDCT D^{-1} , convolution layer C_{M2} 를 순차적으로 연산 (Eq .8)
 - Moire texture를 제거하기위한 residual connection [6]을 도입

$$x_{out}^{MTRB} = x_{in}^{MTRB} + S(C_{M2}(D^{-1}(\theta \cdot \xi)))$$



Method

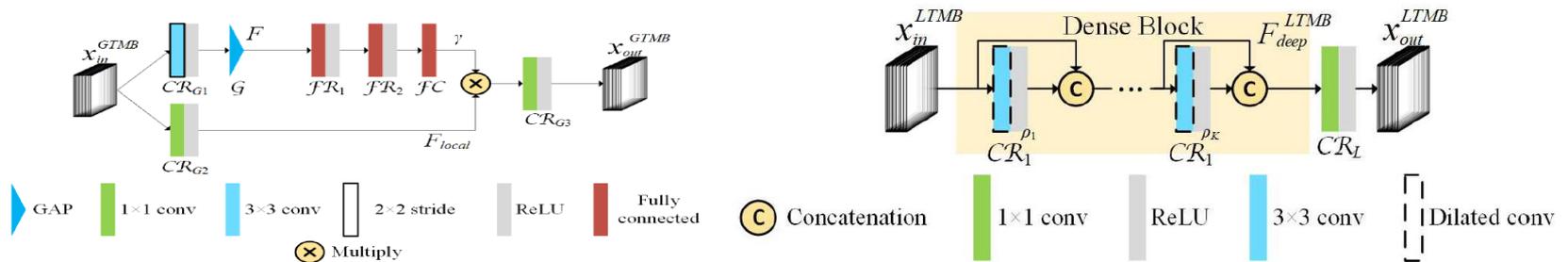
- Tone mapping

- RGB color space는 256^3 개의 색을 표현

- Point-wise tone mapping을 어렵게 하는 요인으로 작용

- GTMB (global tone mapping block) , LTMB (local tone mapping block) 로 이루어진 2-step tone mapping을 제안

- Moire image 와 clean image의 color shift를 비교하기 위함



Method

- Tone mapping (GTMB)

- GTMB는 3×3 Convolution으로 global feature F 를 추출

- Stride = 2, kernel = 3×3 , global average pooling (GAP) 을 사용

- Deep global feature γ 를 추출하기 위해

- 3×3 Convolution으로 global feature F 를 추출

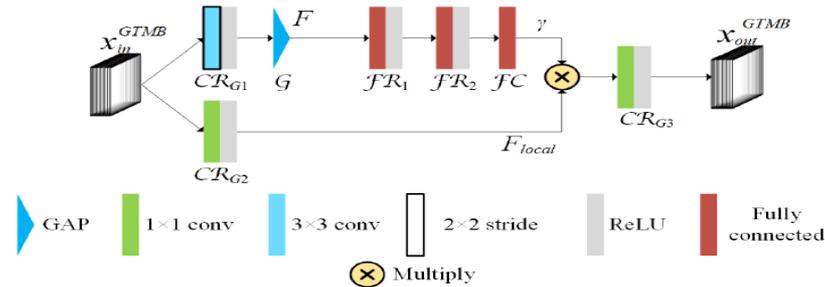
- ※ Stride = 2, kernel = 3×3 , global average pooling (GAP) 을 사용

- 총 3개의 fully connected layers를 통과

- ※ 2 with ReLU, 1 without ReLU

- F_{local} 을 추출하기 위해 1×1 Convolution 을 사용

- 최종 output은 $x_{out}^{GTMB} = CR_{G3}(\gamma \cdot F_{local})$



Method

- Tone mapping vs Channel Attention

- Image restoration, image retrieval 등에서 channel attention이 제안됨

- GTMB가 channel attention block [2,5] 과 유사하게 보임

- 그러나 GTMB는 channel attention과는 다름

- 1. channel attention block은 sigmoid function 만을 사용

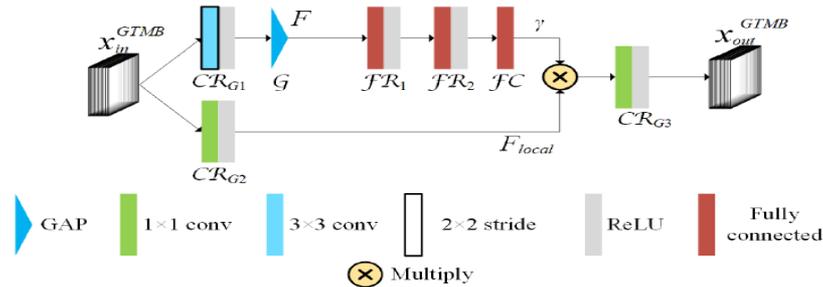
- ※ GTMB의 γ 는 activation function의 제약이 없음 (ReLU를 사용)

- 2. channel attention은 attention input block에 직접적으로 적용

- ※ GTMB의 γ 는 local feature인 F_{local} 에 적용

- 3. channel attention의 목적은 adaptive channel wise feature re-calibration를 만들기 위함

- ※ GTMB의 목적은 global color shift를 만들고, color artifact를 제거하기 위함



Method

- Tone mapping (LTMB)

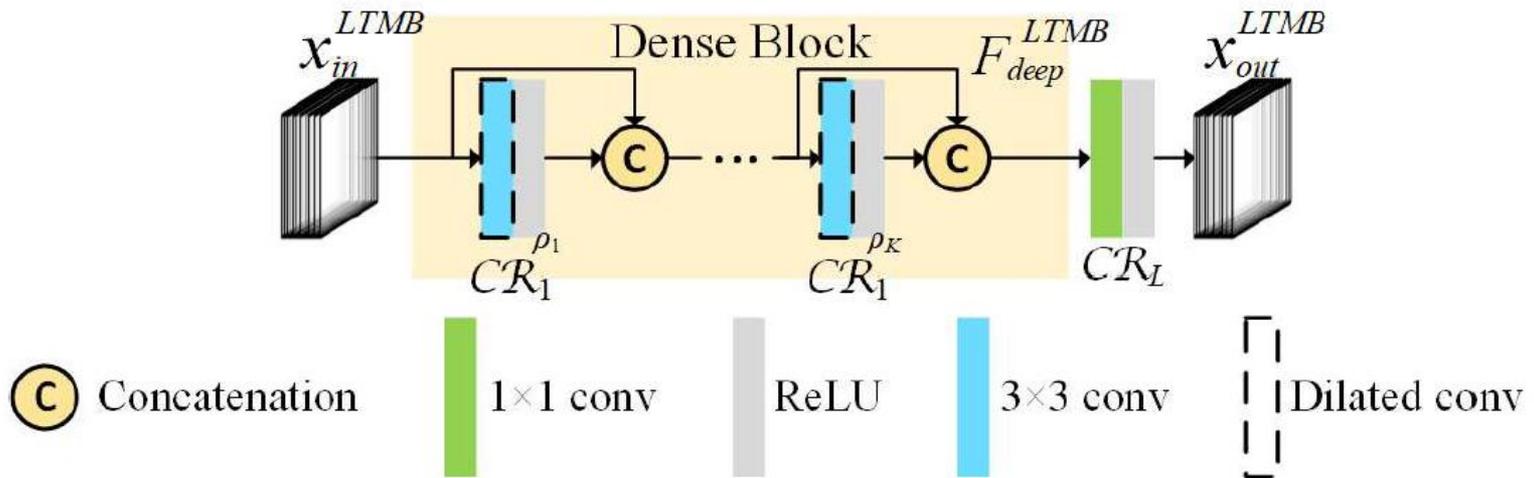
- LTMB는 local fine-grained tone mapping을 fitting하기 위해 제안

- LTMB는 input, F_{in}^{LTMB} 에서 F_{deep}^{LTMB} 를 추출하기 위해 dense block을 통과

- ※ LTMB는 MTRB와 구조가 유사함

- LTMB의 output은 $x_{out}^{LTMB} = CR_L(F_{in}^{MTRB})$

- CR_L 은 1×1 Convolution으로 x_{in}^{LTMB} 의 size와 x_{out}^{LTMB} size는 같음



Method

- Loss function

- Image reconstruction에서 L1 loss function은 L2loss function보다 효율적 [1]

- Moire pattern은 구조적인 artifact이지만 , L1loss 는 구조적 정보를 포함하지 않는 point-wise loss 임

- 구조적인 정보를 포함하기 위한 Advanced Sobel Loss (ASL) 를 제안

$$ASL(\hat{Z}, Z) = \frac{1}{N} \sum |Sobel^*(Z) - Sobel^*(\hat{Z})|$$

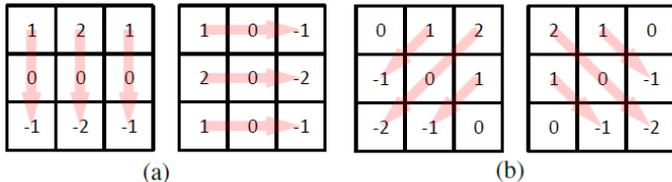
- \hat{Z} = output, Z = ground truth

- Advanced Sobel Loss는 (a)에서 다양한 구조적 정보 포함을 위해 45°방향의 필터를 추가

- 최종 loss는 L1과 ASL을 더함

- λ 는 L1과 ASL의 균형을 위한 hyper parameter

- Multi supervising을 위해 training시에 branch별 loss를



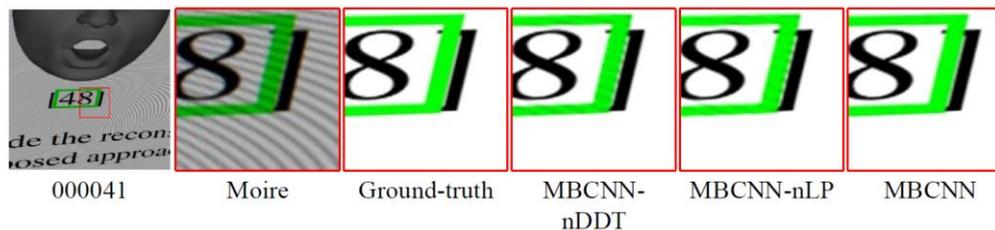
$$Loss(\hat{Z}, Z) = \mathcal{L1}(\hat{Z}, Z) + \lambda \cdot ASL(\hat{Z}, Z)$$

$$loss = Loss(\hat{Z}^{s_1}, Z^{s_1}) + Loss(\hat{Z}^2, Z^{s_2}) + Loss(\hat{Z}^{s_3}, Z^{s_3})$$

Experiments

- Learnable bandpass filter
 - LBF는 DCT domain transform 과 learnable passband(LP)로 구성
 - MBCNN-nDDT: MBCNN without DDT
 - MBCNN-LP: MBCNN without LP
 - MBCNN-nDDT: 1×1 Conv로 변경 (output size 유지가능)
 - MBCNN-nLP: passband의 parameter를 1로 변경
 - MBCNN-nLP는 MBCNN-nDDT보다 PSNR을 0.18dB상승
 - MBCNN은 MBCNN-nLP보다 PSNR 0.95dB상승

Model	MBCNN-nDDT	MBCNN-nLP	MBCNN
PSNR/SSIM	42.91/0.9932	43.09/0.9936	44.04/0.9948



Experiments

- Learnable bandpass filter

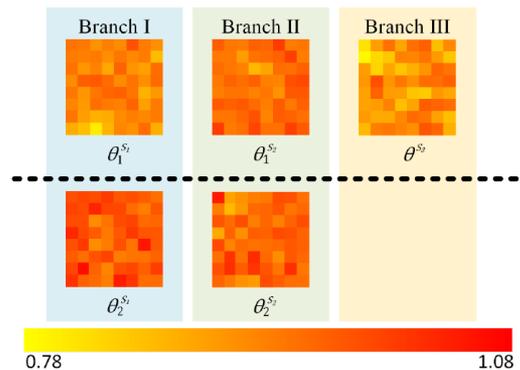
- Block-IDCT size p : DDT parameter에 있어서 중요

- p 가 클수록 LBF가 더 정확히 정확하고, 빠르게 frequency를 learn 할 수 있다

- Block size 는 8×8 일 때, 가장 최고 성능

- ※ p 가 커질수록 frequency 가 learning 하기 어렵고 복잡함

- ※ p 가 크면 receptive field를 지원할 수 없음



Model	MBCNN-6	MBCNN-8	MBCNN-10	MBCNN-12
PSNR/SSIM	43.25/0.9937	44.04/0.9948	43.45/0.9939	43.17/0.9937

Experiments

- Loss function

- Moire 는 구조적 정보가 중요, sobel loss는 구조적인 정보를 제공함

- L1 보다 L1 + Sobel Loss는 +1.81dB의 큰 폭의 성능 향상을 가능하게 함

- Sobel Loss 에서 ASL로의 변화는 +0.40dB의 성능을 향상시킴

- SSIM 과 perceptual loss도 유의미한 성능의 향상

- ※ SSIM loss는 Laplace와 유사한 기능을 하고, perceptual loss는 첫번째보다 0.21dB낮은 2번째로 높은 성능

- ASL이 간단하고 효과적인 Loss function

Loss	λ	PSNR	SSIM
L1	-	41.83	0.9905
L1+Sobel	0.5	43.64	0.9945
L1+Laplace	0.5	42.92	0.9927
L1+SSIM	0.2	43.36	0.9946
L1+perceptual	1.0	43.83	0.9946
L1+ASL	0.25	44.04	0.9948

Experiments

- Comparison on LCD Moire dataset

- CAS-CNN, MWCNN, DMCNN 도 Loss function 은 L1 + ASL을 사용

- ASL의 Loss function 성능이 가장 우수한 것을 증명했으므로

- MBCNN이 다른 Method를 압도

- CAS-CNN 보다 +7.88dB/+0.075 우위

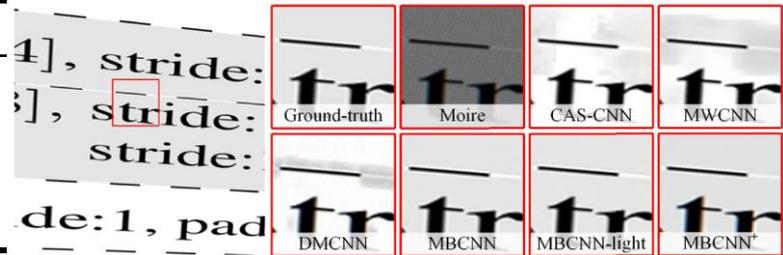
- MBCNN이 moire texture를 제거하고, 대부분의 image details들을 복원 성공

- MBCNN은 사용하는 parameter수가 많음

- $n_G = 64, n_G=32$ 로 줄였을 때를 실험

- MBCNN보다 PSNR-1.46dB이 낮아짐, 하지만 다른 method보다는 성능이 여전히 높음

Loss	CAS-CNN	MWCNN	DMCNN	MBCNN	MBCNN-light
PSNR	36.16	28.93	35.48	44.04	42.81
SSIM	0.9873	0.9698	0.9785	0.9948	0.9940
Time(s)	0.14	0.14	0.10	0.25	0.12



Reference

- [1] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In CVPRW, 2017. 5, 8
- [2] Chenggang Yan, Yunbin Tu, Xingzheng Wang, Yongbing Zhang, Xinhong Hao, Yongdong Zhang, and Qionghai Dai. Stat: spatial-temporal attention mechanism for video captioning. TMM, 2019. 4
- [3] Xi Cheng, Zhenyong Fu, and Jian Yang. Multi-scale dynamic feature encoding network for image demoir'ing. In ICCVW, 2019. 1, 2
- [4] Bolun Zheng, Yaowu Chen, Xiang Tian, Fan Zhou, and Xuesong Liu. Implicit dual-domain convolutional network for robust color image compression artifact reduction. TCSVT, 2019. 2, 3
- [5] Xintao Wang, Kelvin CK Chan, Ke Yu, Chao Dong, and Chen Change Loy. Edvr: Video restoration with enhanced deformable convolutional networks. In CVPRW, 2019. 4
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, 2016. 4
- 그림 1. <https://zone.ni.com/reference/ko-XX/help/371361R-0129/lvanlsconcepts/aliasing/>