

일반논문 (Regular Paper)

방송공학회논문지 제27권 제3호, 2022년 5월 (JBE Vol.27, No.3, May 2022)

<https://doi.org/10.5909/JBE.2022.27.3.351>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

왜곡 정보 모듈을 이용한 이미지 디블러 방법

김 정 환^{a)}, 김 원 준^{b)†}

Distortion-guided Module for Image Deblurring

Jeonghwan Kim^{a)} and Wonjun Kim^{b)†}

요 약

영상 흐려짐은 피사체의 움직임, 카메라의 흔들림 등의 요인으로 발생하는 현상이다. 최근 합성곱 심층신경망(Convolution Neural Network, CNN)을 활용하여 흐려짐 현상을 복원하는 연구가 활발하게 진행되었으며, 원본과 정답 영상의 차이를 이용하여 복원 과정을 가이드하는 방법이 뛰어난 성능을 보였다. 본 논문에서는 왜곡 정보를 기반으로 흐려진 영상 복원 성능을 개선하는 방법을 제안한다. 이를 위해 학습 시, 원본과 정답 영상 차이에 대한 이진화를 수행하여 복원 과정을 가이드 할 수 있도록 하는 트랜스포머(Transformer) 기반 신경망 모듈을 설계하였다. 제안하는 방법은 학습 과정에서 잠재 특징을 기반으로 전역적 추론을 통해 예측한 왜곡 위치 정보 분포를 흐려짐 복원 과정에 반영한다. 다양한 영상 흐려짐 복원 신경망에 제안하는 모듈을 적용하여 복원 성능을 효과적으로 향상시킬 수 있음을 확인하였다.

Abstract

Image blurring is a phenomenon that occurs due to factors such as movement of a subject and shaking of a camera. Recently, the research for image deblurring has been actively conducted based on convolution neural networks. In particular, the method of guiding the restoration process via the difference between blur and sharp images has shown the promising performance. This paper proposes a novel method for improving the deblurring performance based on the distortion information. To this end, the transformer-based neural network module is designed to guide the restoration process. The proposed method efficiently reflects the distorted region, which is predicted through the global inference during the deblurring process. We demonstrate the efficiency and robustness of the proposed module based on experimental results with various deblurring architectures and benchmark datasets.

Keyword : Image Deblurring, Distortion-guided Module, Transformer, Convolution Neural Network

a) 건국대학교 인공지능학과(Department of Artificial Intelligence, Konkuk University)

b) 건국대학교 전기전자공학부(Department of Electrical and Electronics Engineering, Konkuk University)

† Corresponding Author : 김원준(Wonjun Kim)

E-mail: wonjkim@konkuk.ac.kr

Tel: +82-2-450-3396

ORCID: <http://orcid.org/0000-0001-5121-5931>

※ 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2020R1F1A1068080).

※ This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No.2020R1F1A1068080).

· Manuscript March 21, 2022; Revised April 29, 2022; Accepted April 29, 2022.

I. 서론

영상 흐려짐(Image Blurring) 현상은 촬영 도중 카메라의 흔들림, 피사체의 움직임 등으로 인해 발생한다. 이를 해결하기 위해 광학식 손 떨림 보정(Optical Image Stabilization, OIS) 기술과 같은 카메라 센서 기술이 고도화되었음에도, 일상에서 촬영한 영상에서 여전히 흐려짐 현상이 빈번하게 발생한다. 영상 흐려짐이 발생한 경우 재촬영을 통해 선명한 영상을 확보할 수 있지만, 촬영 순간을 재현할 수 없는 경우에는 획득한 영상으로부터 선명한 영상을 복원하는 기술이 필수적이다. 일반적으로 촬영 순간의 주어진 장면에 대한 선명한 영상이 없으므로 기존 영상 흐려짐 복원 연구는 다양한 최적화 기법을 이용하여 복원 과정을 수행하였다. 예를 들어, Fergus^[1] 등은 선명한 영상의 통계적 특성을 기반으로 흐려짐을 발생시키는 커널(Kernel)의 형태를 추론하는 방법을 제안하였다. 이와 비슷한 방법으로 Shan^[2] 등은 확률적 모델을 기반으로 블러(Blur) 커널을 예측하는 반복적 최적화 방법을 제시하였다. Cho와 Lee^[3]는 픽셀 값 대신 그레디언트(Gradient) 정보를 기반으로 잠재 이미지(Latent Image)와 블러 커널을 추정하고 고속 푸리에 변환(Fast Fourier Transform)을 통해 빠르게 선명한 영상을 복원하는 방법을 제안하였다. 그러나 이와 같은 방법들은 객체의 움직임에 의해 지역적으로 불규칙하게 발생하는 흐려짐 문제를 해결하는 데 여전히 많은 어려움을 겪고 있다.

최근 컴퓨터 비전 분야에서는 심층 학습의 성공에 힘입어 합성곱 신경망(Convolution Neural Network, CNN)을 활용한 다양한 알고리즘이 연구되고 있다. 이러한 합성곱 신경망을 기반으로 영상 흐려짐 문제 또한 해결하고자 하는 시도가 최근 활발히 이루어지고 있다. 구체적으로 학습을 위한 대용량 데이터셋이 구축되고 있으며, 압축기-복원기(Encoder-Decoder) 구조를 기반으로 선명한 영상을 추정하는 방법이 제안되고 있다. Nah^[4] 등은 다중 스케일(Multi-Scale) 기반의 흐려짐 복원 신경망을 제안하였으며, 실제 환경 학습을 위한 GoPro 데이터셋을 제공하였다. 해당 방법은 각 스케일에서의 학습을 위해 별도의 신경망을 사용하기 때문에 수행 속도가 느린 단점이 있어 이를 보완하기 위해 Zhang^[5] 등은 다중 패치(Multi-Patch) 기반의 신

경망(DMPHN)을 제안하였다. Lim^[6] 등은 주파수 계수를 기반으로 전역적인 장면 구조를 복원한 후 공간 정보를 이용하여 지역적 상세 정보(텍스처(Texture), 에지(Edge) 등)를 선명하게 복원하는 신경망 구조를 제안하였다. 또한, Cho^[7] 등은 다중 스케일에서 추출된 잠재 특징 간 복합 관계를 구성하는 U자형 신경망 구조(MIMO-UNet)를 제안하여 우수한 흐려짐 복원 성능을 보여주었다. 그 외에도 합성곱 신경망에 순환 신경망(Recursive Neural Network, RNN)을 혼합한 방법^[8]과 적대적 생성 신경망(Generative Adversarial Network, GAN)을 활용한 흐린 영상 복원 방법^[9]이 제안되었다. 가장 최근에는 Purohit^[10] 등이 원본과 정답 영상의 차이에 대한 이진화를 통해 왜곡 위치 정보를 복원 과정에 사용하는 신경망 구조(SPAIR)를 제안하였다. 그러나, 합성곱 신경망의 제한된 수용 영역(Receptive Field)으로 인해 전역적인 임계값으로 이진화 된 왜곡 위치 정보를 효과적으로 반영하지 못하는 문제점이 있다.

이러한 문제를 해결하기 위해 본 논문에서는 트랜스포머(Transformer)^[11]를 활용한 왜곡 정보 가이드 모듈을 제안한다. 제안하는 방법은 학습 과정에서 잠재 특징(Latent Feature)을 기반으로 트랜스포머를 통해 예측한 왜곡 위치 정보 분포를 흐려짐 복원 과정에 반영하도록 한다. 제안하는 방법은 간단한 모듈 형태로 구현 가능하며, 압축기를 통해 도출된 잠재 특징만을 사용하기 때문에 신경망 구조와 관계없이 폭넓게 적용할 수 있다. 영상 흐려짐 복원에 널리 사용되고 있는 GoPro 데이터셋^[4], HIDE 데이터셋^[12]과 RealBlur^[13] 데이터셋을 이용하여 제안하는 왜곡 정보 가이드 모듈의 성능 개선 효과를 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 왜곡 정보 가이드 모듈의 구조에 대해 자세히 설명하며, 3장에서는 다양한 영상 흐려짐 복원 신경망에 적용하여 성능이 향상됨을 확인한다. 또한, 합성곱 신경망과 비교하여 트랜스포머 모듈이 왜곡 위치 정보 학습에 효과적임을 증명한다. 마지막으로 4장에서는 본 논문의 결론을 서술한다.

II. 제안하는 방법

본 장에서는 제안하는 왜곡 정보 가이드 모듈의 전체적

인 구조를 소개한다. 이어서 다양한 영상 흐려짐 복원 신경망에 적용한 방식에 대해 자세히 설명한 후 마지막으로 학습에 사용된 손실 함수에 대해 설명한다.

1. 트랜스포머(Transfomer)를 활용한 왜곡 위치 정보 가이드 모듈

제안하는 방법은 전역적으로 발생하는 흐려짐 왜곡 위치를 효과적으로 학습하기 위해 트랜스포머를 학습 모듈에 적용한다. 자세히 살펴보면, 압축기로부터 도출된 잠재 특징에 대해 트윈 트랜스포머(Twin Transformer)^[14]를 적용하여 흐려짐이 발생한 위치의 분포 특성을 추출하였다. 트윈 트랜스포머는 기존 트랜스포머와 달리 행과 열로 단위로 자기 주의(Self Attention) 값을 계산함으로써 연산량을 크게 줄인 방식이다. 행과 열 단위의 연속적인 자기 주의 계산을 통해 전역적인 수용 영역을 갖게 되며, 계산 복잡도를 $O((H \times W)^2)$ 에서 $O((H + W) \times H \times W)$ 로 줄여 효율적인 메모리 관리 또한 가능하다. 여기서 H 와 W 는 각각 잠재 특징 맵의 높이와 너비 값을 나타낸다. 제안하는 모듈에서 흐려짐이 발생한 위치 지도(즉, 그림 1에서 Blur Mask)를 생성하기 위한 복원기는 업샘플링(Up-sampling)과 ResBlock^[15]을 이용하여 간단하게 구축하였다. 업샘플링 단계에서는 선형 보간과 1×1 합성곱 연산을 수행한다. 해당 복원기는 학습 과정에서만 사용되며 추론(Inference) 시에는 사용되지 않는다. 마지막으로 트랜스포머를 통해 재조정된 특징을 압축기로부터 도출된 잠재 특징과 융합하기 위해 채널 방향으로 결합(Concatenation)한 후 합성곱

연산을 수행한다. 제안하는 왜곡 위치 정보 가이드 모듈의 전체적인 구조는 그림 1과 같다. 학습을 위해 원본 영상과 정답 영상의 차이를 아래와 같이 계산하여 흐려짐이 발생한 위치 지도를 생성한다.

$$G(x) = \begin{cases} 1, & \text{if } (B(x) - S(x))^2 > \mu \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

여기서 B 와 S 는 각각 입력 영상과 정답 영상을 의미한다. x 는 픽셀 위치 인덱스를 나타낸다. μ 는 입력과 정답 영상 차이에 대한 제곱 값의 평균을 의미하며, 이 값을 이용하여 이진화를 수행하여 최종적으로 흐려짐이 발생한 위치 지도 G 를 생성한다.

2. 왜곡 위치 정보 가이드 모듈을 적용한 흐려짐 복원 신경망

제안하는 방법은 압축기를 통해 도출된 잠재 특징만을 사용하기 때문에 신경망 구조와 관계없이 폭넓게 적용될 수 있다. 본 논문에서는 제안하는 방법을 다양한 영상 흐려짐 복원 신경망, 즉, MIMO-UNet^[7], DSSN^[6], 및 Deblur GANv2^[9]에 적용하였다. 이 중 MIMO-UNet에 제안하는 모듈을 적용한 예를 그림 2에 나타내었다. MIMO-UNet은 여러 스케일에서 압축된 잠재 특징을 융합하여 순차적으로 선명한 영상을 복원하는 구조로 특히, 복원기는 앞서 융합한 압축기의 잠재 특징을 참조할 수 있어 효과적으로 선명한 영상 복원을 수행한다. 본 논문에서는 제안하는

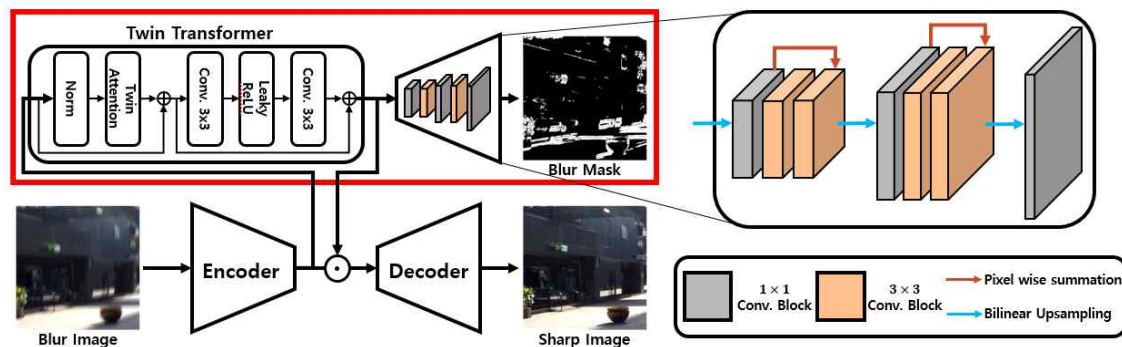


그림 1. 제안하는 왜곡 위치 정보 가이드 모듈(왼쪽)과 복원기의 상세 구조(오른쪽)

Fig. 1. Entire structure of the proposed distortion guided module(left) and details of the decoder(right)

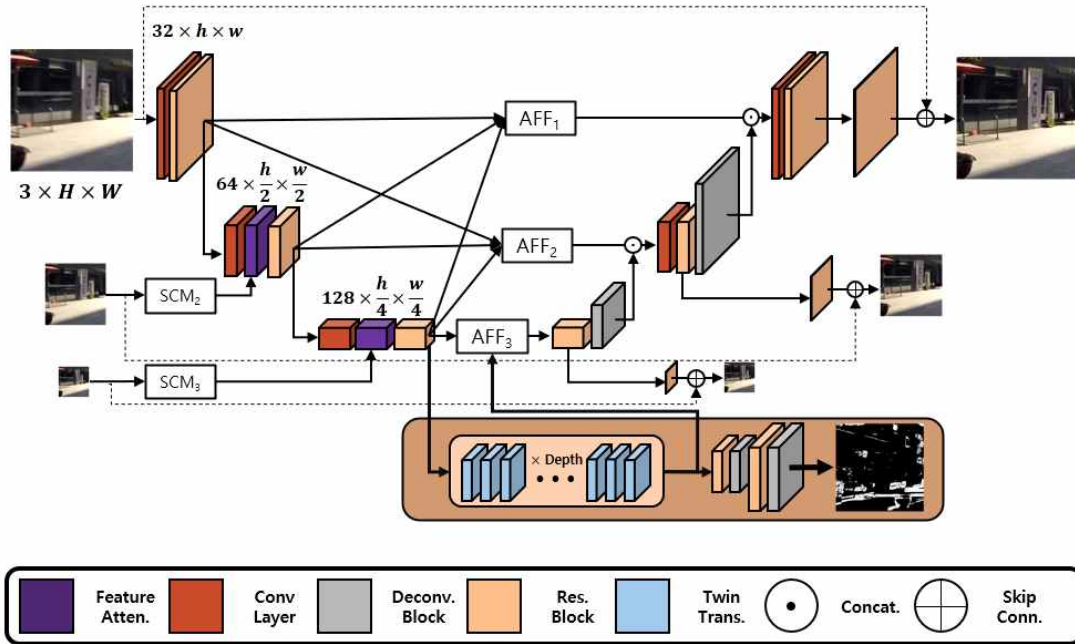


그림 2. 제안하는 왜곡 위치 정보 가이드 모듈을 MIMO-UNet 구조에 적용한 예
Fig. 2. Example of applying the proposed module to MIMO-UNet

모듈을 그림 2와 같이 압축기 단에서 출력된 잠재 특징으로부터 흐려짐이 발생한 위치의 분포 특성을 추출할 수 있도록 삽입하였다. DSSN이나 DeblurGAN에도 이와 비슷한 방법으로 제안하는 왜곡 위치 정보 가이드 모듈을 삽입하여 영상 흐려짐 복원 성능을 효과적으로 향상시킬 수 있다.

3. 손실 함수

본 논문에서는 흐려진 영상 복원을 위해 기존 방법에서 널리 사용되고 있는 손실 함수를 사용하였다. 간단히 살펴보면, 신경망의 최종 예측 값에 적용되는 손실 함수 L 은 영상 복원을 위한 손실 함수 L_{reg} 와 흐려짐이 발생한 위치 지도 복원을 위한 손실 함수 L_{mask} 의 합으로 정의되며 다음과 같이 표현될 수 있다.

$$L = L_{reg} + \lambda L_{mask}, \quad (2)$$

여기서 λ 는 기존 손실 함수와 흐려짐이 발생한 위치 지

도 복원을 위한 손실 함수의 크기를 균일화하기 위한 가중치이다. L_{reg} 를 계산하기 위해 대표적으로 평균 절대 오차 (Mean Absolute Error, MAE)를 사용하고 있으며 다음과 같이 계산할 수 있다^[6,10].

$$L_{reg} = \frac{1}{N} \sum_{x=1}^N |Y(x) - S(x)|, \quad (3)$$

여기서 Y 와 S 는 각각 예측된 영상과 선명한 정답 영상이며, x 는 픽셀 위치 인덱스를 나타낸다. N 은 전체 픽셀 수를 의미한다. 식(1)에서 계산된 흐려짐이 발생한 위치 지도를 학습하기 위해 이진 크로스 엔트로피(Binary Cross Entropy) 손실 함수를 사용하며 L_{mask} 값을 계산한다.

$$L_{mask} = -\frac{1}{N} \sum_{x=1}^N G(x) \log(M(x)) + (1 - G(x)) \log(1 - M(x)) \quad (4)$$

여기서 M 은 제안하는 모듈을 통해 예측된 흐려짐이 발생한 위치 지도이다.

III. 실험 결과 및 분석

본 논문에서는 제안하는 방법의 성능 평가를 위해 세 개의 벤치마크 데이터셋을 사용하였다. 먼저, 영상 디블러 (Image Deblurring) 분야에서 가장 대표적으로 사용되는 GoPro 데이터셋^[4]을 사용하여 성능을 평가하였다. GoPro 데이터셋은 GOPRO4 Hero Black 카메라를 사용하여 240fps 영상을 촬영한 다음 연속 프레임의 평균화를 통해 생성된 흐려진 영상으로 구성되어 있다. 학습을 위해 GoPro 데이터셋은 총 2,103 개의 흐려진 영상과 선명한 영상 묶음이 사용되었고, 1,111 개의 영상 묶음이 검증에 사용되었다. 다음으로 성능 평가에 사용한 HIDE 데이터셋^[12]은 GoPro 데이터셋과 같은 방법으로 촬영되었으며, 영상 속 사람의 근접촬영 여부와 군집 여부에 대한 정보가 포함되어 있다. 총 6,397 개의 학습용 묶음과 2,025 개의 검증용 묶음으로 구성되어 있다. 마지막으로 성능 평가에 함께 사용된 Realblur 데이터셋^[13]은 카메라 흔들림에 대한 시나리오

를 다루며 모션 블러가 가장 빈번하게 발생하는 저조도 환경에서 촬영한 데이터셋이다. 총 3,758 개의 학습용 묶음과 980 개의 검증용 묶음으로 구성되어 있다. 제안하는 방법은 PyTorch^[16] 프레임워크를 기반으로 구현되었다. 본 논문에서는 신경망 가중치를 최적화하기 위한 알고리즘으로 Adam^[17]을 사용하였고, 파워(Power)와 가속도(Momentum) 값은 각각 0.9와 0.999로 설정하였다. 학습 속도(Learning Rate)는 1×10^{-4} 부터 시작하여 500 에포크(Epoch) 지점에서 5×10^{-5} 로 감소하도록 설정하였으며, 총 1000 에포크 동안 학습을 진행하였다. 학습 영상은 원본 영상에서 무작위로 256×256 픽셀 크기로 잘라 생성하였고, 과적합 (Overfitting) 문제를 해결하기 위해 영상을 수평으로 반전하는 데이터 증강 방법을 적용하였다. 학습과 성능 평가에는 Intel(R) Xeon(R) E5-1650v4 @3.60GHz CPU와 NVIDIA GTX Titan XP GPU 4대가 이용되었다. 제안하는 방법의 성능 개선 효과를 검증하기 위해 기존 영상 디블러 방법인 MIMO-UNet^[7], DSSN^[6], 및 DeblurGANv2^[9]에 왜

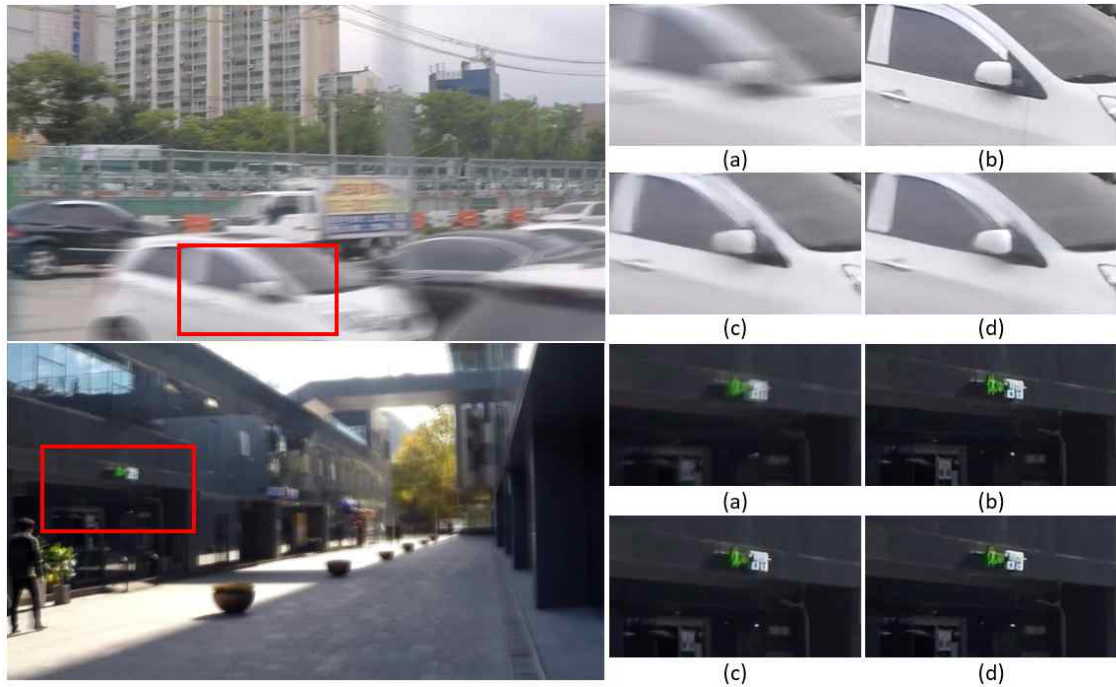


그림 3. GoPro 데이터셋[4]에 대한 흐려진 영상 복원 결과. (a): 입력 영상. (b): 정답 영상 (c): MIMO-UNet의 복원 결과 (d): 제안하는 모듈을 적용한 모듈의 복원 결과

Fig. 3. Deblurring result of GoPro dataset[4]. (a): input image, (b): ground truth, (c): restoration result of MIMO-UNet, (d): restoration result of proposed method

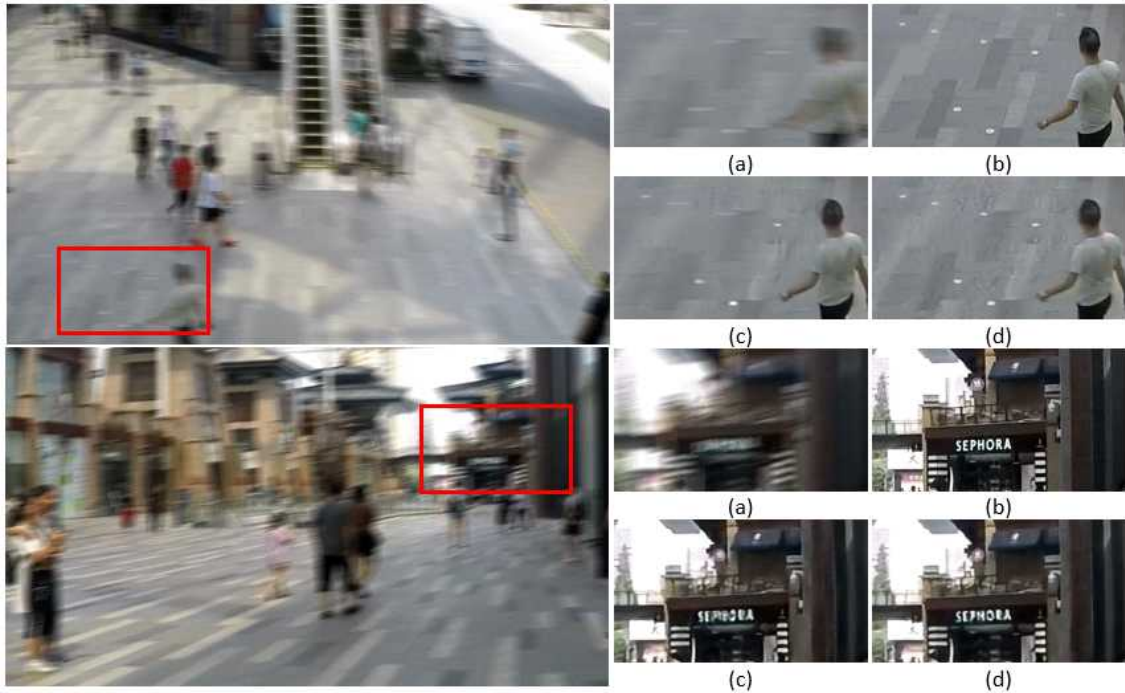


그림 4. HIDE 데이터셋^[12]에 대한 흐려진 영상 복원 결과. (a): 입력 영상. (b): 정답 영상 (c): MIMO-UNet의 복원 결과 (d): 제안하는 모듈을 적용한 모델의 복원 결과

Fig. 4. Deblurring result of HIDE dataset^[12]. (a): input image, (b): ground truth, (c): restoration result of MIMO-UNet, (d): restoration result of proposed method

곡 정보 가이드 모듈을 적용하여 성능을 비교하였다. 정성적인 영상 흐려짐 복원 결과를 그림 3과 4에 나타내었다. 그림에서 볼 수 있듯이, 제안하는 방법을 적용했을 때 흐려진 부분이 더 선명하게 복원된다. 특히, 그림 4에서 흐려짐이 강하게 발생한 복잡한 배경 구조 및 글자 영역이 효과적으로 복원되는 것을 확인할 수 있다. 다음으로는 흐려진 영상 복원 성능의 정량적 평가에 널리 사용되는 최대 신호대 잡음비(Peak Signal-to-Noise Ration, PSNR)와 구조적 유사도(Structural Similarity Index Measure, SSIM)를 활용

하여 성능 비교를 수행하였으며, 해당 결과를 표 1, 2와 3에 나타내었다. 표 1, 2와 3에서 제안하는 방법에 따른 성능 개선 효과를 확인할 수 있다. 따라서, 제안하는 방법은 영상 디블러를 위해 신경망 구조와 관계없이 효과적으로 적용 가능하다. 그러나, 제안하는 방법은 흐려짐 발생 위치 지도를 정확하게 생성하지 못하게 되면 흐려짐 위치 정보를 잘못 참조하여 복원하는 데 어려움이 있다 (예를 들어, 그림 5의 예시).

표 1. GoPro 데이터셋^[4]에서의 정량적 평가 비교

Table 1. Quantitative evaluations on GoPro dataset^[4]

Method	Module	PSNR	SSIM	Module	PSNR	SSIM
MIMO-UNet ^[7]	without DGM	31.18	0.916	with DGM	31.54	0.923
DSSN ^[6]	without DGM	28.34	0.861	with DGM	28.80	0.868
DeblurGANv2 ^[9]	without DGM	23.87	0.733	with DGM	25.77	0.805

Note: DGM denotes distortion guided module

표 2. HIDE 데이터셋^[12]에서의 정량적 평가 비교

Table 2. Quantitative evaluations on HIDE dataset^[12]

Method	Module	PSNR	SSIM	Module	PSNR	SSIM
MIMO-UNet ^[7]	without DGM	28.50	0.884	with DGM	29.39	0.898
DSSN ^[6]	without DGM	27.00	0.843	with DGM	27.46	0.850
DeblurGANv2 ^[9]	without DGM	23.05	0.717	with DGM	24.41	0.770

Note: DGM denotes distortion guided module

표 3. RealBlur-J 데이터셋^[13]에서의 정량적 평가 비교

Table 3. Quantitative evaluations on RealBlur-J dataset^[13]

Method	Module	PSNR	SSIM	Module	PSNR	SSIM
MIMO-UNet ^[7]	without DGM	28.05	0.848	with DGM	28.34	0.854
DSSN ^[6]	without DGM	28.67	0.852	with DGM	28.61	0.852
DeblurGANv2 ^[9]	without DGM	26.31	0.768	with DGM	27.46	0.812

Note: DGM denotes distortion guided module

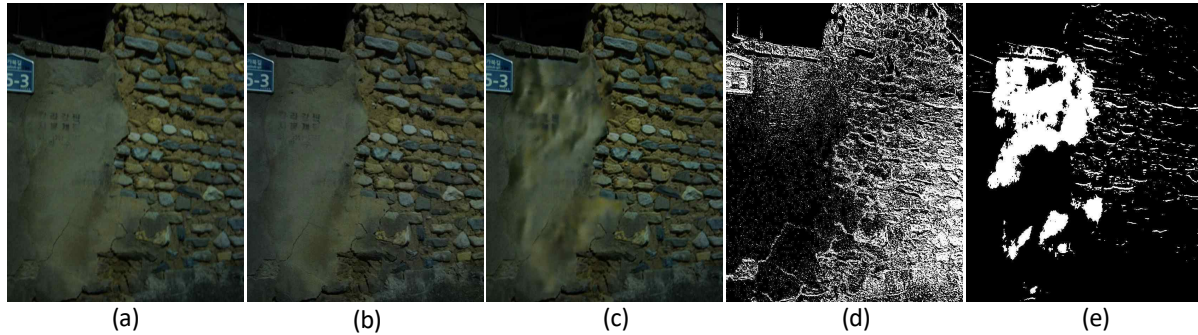


그림 5. 제안하는 방법의 복원 실패 예시. (a): 입력 영상, (b): 정답 영상, (c): 제안하는 모듈을 적용한 모델의 복원 결과, (d): 흐려짐이 발생한 위치 지도, (e): 제안하는 모듈을 적용한 모델에서 출력된 흐려짐이 발생한 위치 지도

Fig. 5. Examples of fail restoration of the proposed method. (a): input image, (b): ground truth, (c): restoration result of proposed method, (d): ground truth of blur mask, (e): blur mask result of proposed method

제안하는 왜곡 정보 가이드 모듈에 사용된 트랜스포머의 효과를 면밀히 검증하기 위해 제안하는 모듈을 합성곱 신경망으로 구성하여 비교 실험을 진행하였으며 그림 6과 표 4에 각각 그 결과를 나타내었다. 그림 6에서 잡재 특징을 시각화하기 위해 채널 방향으로 최대 풀링(Max Pooling)을 적용한 후 반전하였다. 해당 결과에서 볼 수 있듯이 합성곱 신경망보다 트랜스포머를 사용한 경우 잡재 특징이 왜곡 위치 정보를 더 정확하게 반영하는 것을 확인할 수 있다.

더 나아가, 영상 흐려짐 현상이 전역적으로 발생하는 카메라 블러 영상에서의 비교를 진행하여 그림 7에 나타내었다. 그림 7에서 볼 수 있듯이 전역적인 블러가 발생한 영상에서 트랜스포머를 사용한 경우 효과적으로 복원하는 것을 확인할 수 있다. 따라서, 제안하는 트랜스포머 모듈이 전역적 추론을 통해 예측한 왜곡 위치 정보 분포를 추출하는데 효과적임을 확인하였다.

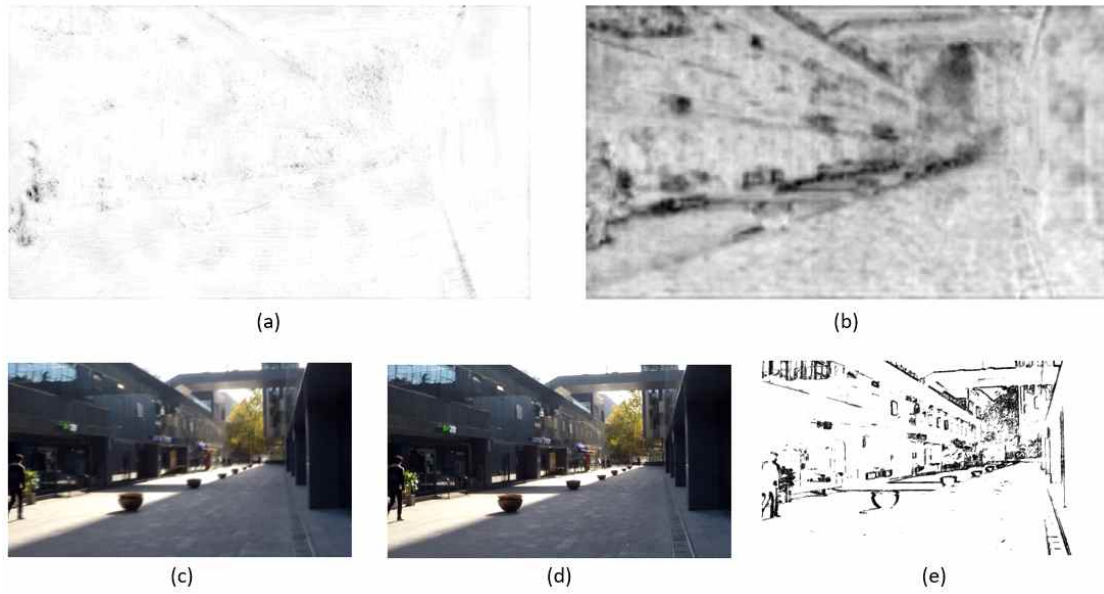


그림 6. (a): 제안하는 모듈에 합성곱 신경망을 적용하여 출력된 잠재 특징, (b): 제안하는 모듈에 트랜스포머를 적용하여 출력된 잠재 특징, (c): 입력 영상, (d): 정답 영상, (e): 흐려짐이 발생한 위치 지도

Fig. 6. (a): latent feature of convolution, (b): latent feature of transformer, (c): input image, (d): ground truth, (e): ground truth of blur mask

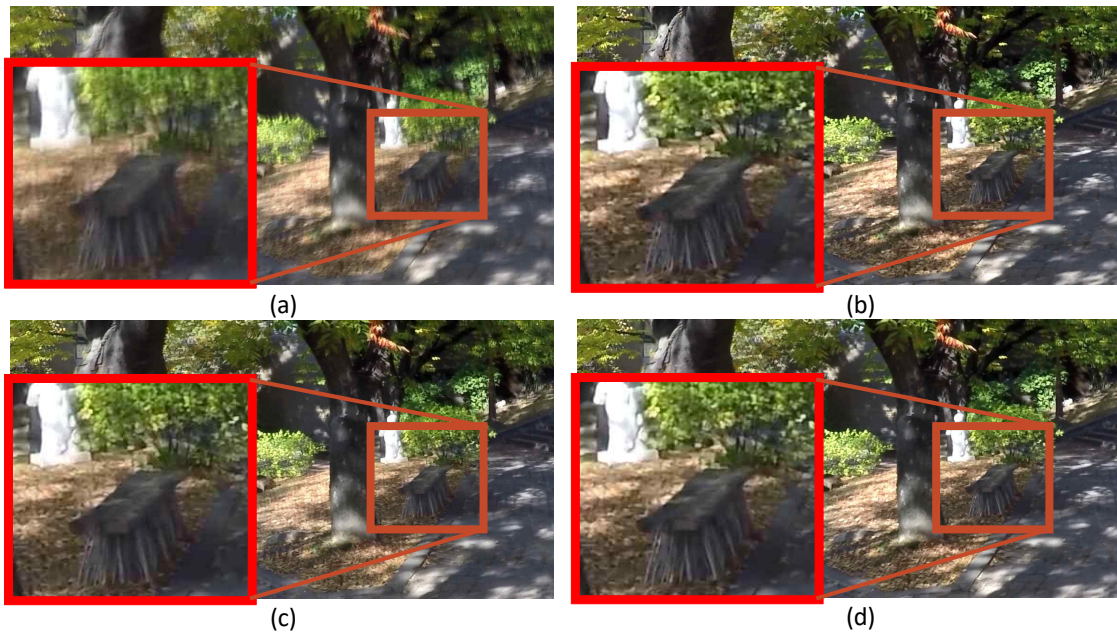


그림 7. (a): 카메라 블러가 발생한 영상 (b): 정답 영상, (c): 제안하는 모듈에 합성곱 신경망을 적용하여 출력된 복원 결과, (d): 제안하는 모듈에 트랜스포머를 적용하여 출력된 복원 결과

Fig. 7. (a): camera shaken blur image, (b): ground truth, (c): restoration result of convolution based module, (d): restoration result of transformer based module

표 4. 제안하는 모듈의 구조 따른 정량적 실험 결과

Table 4. Quantitative experimental results according to the structure

Methods	Dataset	PSNR	SSIM
Convolution Based DGM	GoPro dataset	31.15	0.917
Transformer Based DGM	GoPro dataset	31.54	0.923
Convolution Based DGM	HIDE dataset	28.67	0.888
Transformer Based DGM	HIDE dataset	29.39	0.898

Note: DGM denotes distortion guided module

IV. 결 론

본 논문에서는 영상 디블러를 위해 왜곡 위치 가이드 모듈을 제안하였다. 제안하는 방법은 복원 신경망의 잠재 특징으로부터 왜곡 위치 정보를 추출하고, 복원 과정에서 이를 참조할 수 있도록 한다. 왜곡 위치 정보의 전역적인 특징을 추출하기 위해 트랜스포머 구조를 활용하고, 기존 영상 디블러 방법에 폭넓게 적용할 수 있도록 간단한 모듈 형태로 설계하였다. 다양한 실험을 통해 제안하는 방법이 영상 디블러 성능을 효과적으로 향상시킬 수 있음을 확인하였다.

참 고 문 헌 (References)

- [1] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 787-794, Jul. 2006.
doi: <https://doi.org/10.1145/1141911.1141956>
- [2] Q. Shan, J. Jia, A. Agarwala, "High-quality motion deblurring from a single image," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, pp. 73:1-73:10, Aug. 2008.
doi: <https://doi.org/10.1145/1360612.1360672>
- [3] S. Cho and S. Lee, "Fast motion deblurring," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 5, pp. 145:1-145:8, Dec. 2009.
doi: <https://doi.org/10.1145/1618452.1618491>
- [4] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 3883-3891, Jul. 2017.
doi: <https://doi.org/10.1109/CVPR.2017.35>
- [5] H. Zhang, Y. Dai, H. Li, and P. Koniusz, "Deep stacked hierarchical multi-patch network for image deblurring," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 5978-5986, Jun. 2019.
doi: <https://doi.org/10.1109/CVPR.2019.00613>
- [6] S. Lim, J. Kim, and W. Kim, "Deep spectral-spatial network for single image deblurring," *IEEE Signal Process. Lett.*, vol. 27, no. 1, pp. 835-839, Dec. 2020.
doi: <https://doi.org/10.1109/LSP.2020.2995106>
- [7] S. J. Cho, S. W. Ji, J. P. Hong, S. W. Jung, and S. J. Ko, "Rethinking coarse-to-fine approach in single image deblurring," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 4641-4650, Oct. 2021.
doi: <https://doi.org/10.1109/ICCV48922.2021.00460>
- [8] J. Zhang, J. Pan, J. Ren, Y. Song, L. Bao, R. Lau, and M. Yang, "Dynamic scene deblurring using spatially variant recurrent neural networks," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 2521-2529, Jun. 2018.
doi: <https://doi.org/10.1109/CVPR.2018.00267>
- [9] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "DeblurGAN-v2: deblurring (orders-of-magnitude) faster and better," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 8878-8887, Oct. 2019.
doi: <https://doi.org/10.1109/ICCV.2019.00897>
- [10] K. Purohit, M. Suin, A. N. Rajagopalan, and V. N. Boddeti, "Spatially-adaptive image restoration using distortion-guided networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2309-2319, Oct. 2021.
doi: <https://doi.org/10.1109/ICCV48922.2021.00231>
- [11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Neural Inf. Process. Syst.*, pp. 5998-6008, Dec. 2017.
doi: <https://dl.acm.org/doi/10.5555/3295222.3295349>
- [12] Z. Shen, W. Wang, X. Lu, J. Shen, H. Ling, T. Xu, and L. Shao, "Human-aware motion deblurring," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 5572-5581, Oct. 2019.
doi: <https://doi.org/10.1109/ICCV.2019.00567>
- [13] J. Rim, H. Lee, J. Won, and S. Cho, "Real-world blur dataset for learning and benchmarking deblurring algorithms," in *Proc. Eur. Conf. Comput. Vis.*, pp. 184-201, Aug. 2020.
doi: https://doi.org/10.1007/978-3-030-58595-2_12
- [14] R. Guo, D. Niu, L. Qu, and Z. Li, "SOTR: segmenting objects with transformers," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 7157-7166, Oct. 2021.
doi: <https://doi.org/10.1109/ICCV48922.2021.00707>
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 770-778, Jun. 2016.
doi: <https://doi.org/10.1109/CVPR.2016.90>
- [16] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. Devito, Lin, A. Desmaison, L. Antiga, A. Lerer, "Automatic differentiation in pytorch". in *Proc. Conference and Workshop on Neural Information Processing Systems (NIPS)*, pp. 1-4, 2017.
doi: <https://openreview.net/pdf?id=BJJsrnFCZ>
- [17] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proc. Int. Conf. Learn. Representation*, pp. 1-15, May 2015.
doi: <https://doi.org/10.48550/arXiv.1412.6980>

저 자 소 개



김 정 환

- 2020년 2월 : 건국대학교 학사
- 2021년 8월 ~ 현재 : 건국대학교 인공지능학과 석사과정
- ORCID : <https://orcid.org/0000-0002-5941-8351>
- 주관심분야 : 컴퓨터 비전, 객체 검출, 기계학습, 패턴 인식



김 원 준

- 2012년 8월 : 한국과학기술원(KAIST) 박사
- 2012년 9월 ~ 2016년 2월 : 삼성종합기술원 전문연구원
- 2016년 3월 ~ 2020년 2월 : 건국대학교 전기전자공학부 조교수
- 2020년 3월 ~ 현재 : 건국대학교 전기전자공학부 부교수
- ORCID : <https://orcid.org/0000-0001-5121-5931>
- 주관심분야 : 영상이해, 컴퓨터 비전, 기계학습, 패턴 인식