

적대적 생성 신경망을 이용한 컬러 필터 배열 변환 기법

김성열*, 성치훈*, 김승욱^o

Color Filter Array Mapping Using Generative Adversarial Networks

Seong-Yeol Kim*, Chi-Hun Sung*,
Seung-Wook Kim^o

요약

서로 다른 카메라 센서에서 동일한 장면을 촬영한 정렬된 컬러 필터 배열의 수집이 어렵기 때문에, 관련 데이터셋은 충분한 학습 데이터를 제공하지 못하는 경우가 많다. 본 논문에서는 RGBW와 RGB 컬러 필터 배열 변환을 위해 적대적 생성 신경망을 사용하는 새로운 데이터 증강 및 학습 기법을 제안한다. 제안하는 기법은 추가적인 데이터 수집 작업 없이 향상된 컬러 필터 배열의 변환 성능을 보인다.

Key Words : Color filter array, bayer pattern, raw image mapping

ABSTRACT

The lack of paired data is a critical problem in raw image mapping since it is hard to capture the color filter arrays (CFAs) of the same scene from different cameras. This paper introduces a novel RGBW/RGB CFA data generation method using generative adversarial networks (GANs). The experimental results confirm that the performance of the RGBW-to-RGB CFA mapping can be improved by using the proposed data generation method based on GANs.

1. 서론

컬러 필터 배열(color filter array: CFA)은 픽셀 단위로 특정 파장 대역의 빛을 통과시켜 상 정보를 획득한다. CFA의 일반적인 디자인은 베이어 패턴(Bayer pattern)으로 적색(R), 녹색(G), 청색(B)을 1:2:1 비율로 배치하여 구성한다. 베이어 패턴은 인간 시각 시스템에 적합한 장점을 가지고 있어 많은 영상 센서에서 탑재하여 사용되고 있다¹⁾.

RGB CFA는 투과율이 낮은 필터를 사용하기 때문에 색상의 표현력이 떨어지고, 특히 저조도 환경에서 촬영 시 신호 대 잡음비(signal-to-noise ratio: SNR)가 낮아 영상의 품질이 떨어지는 문제가 있다. 이를 해결하기 위해 백색 가시광선의 대역폭을 이용하는 백색(W) 픽셀을 포함하는 RGBW CFA 패턴이 제안되었다²⁾. RGBW CFA의 W 필터는 가시광선에 대한 투과율이 상대적으로 높기 때문에 영상 센서의 SNR을 크게 증가시켜 저조도 환경에서도 영상의 품질을 향상시키는 장점이 있다. 그러나 RGBW CFA를 사용한 미가공 영상(raw image)은 RGB CFA에 기반하여 설계된 기존의 비전(vision) 알고리즘에서 정상적으로 동작할 수 없는 문제가 있다.

본 논문에서는 RGBW-to-RGB CFA 리모자이크(remosaic) 기법을 제안한다. RGBW CFA는 높은 SNR에 비해 R/G/B 픽셀의 해상도가 낮은 단점이 있다. 또한 영상 촬영의 특성상 동일한 장면에 대해 같은 시점을 가진 정렬된 CFA 쌍을 취득하는 것이 매우 제한적이다. 본 논문에서는 서로 다른 패턴 정보를 가지는 CFA를 효과적으로 변환할 수 있는 신경망 구조를 설계하고, 적대적 생성 신경망(generative adversarial networks: GANs)³⁾을 활용하여 추가적인 RGB 컬러 영상으로부터 제한적인 CFA 데이터셋 규모를 증강할 수 있는 학습 기법을 제안한다.

* 이 논문은 부경대학교 자율창의기술연구비(2023년)에 의하여 연구되었음.

• First Author : (ORCID:0009-0006-5911-5744) Pukyong National University Division of Electronics and Communications Engineering, yeolkim@pukyong.ac.kr, 학생(석사과정), 학생회원

◦ Corresponding Author : (ORCID:0000-0002-6004-4086) Pukyong National University Division of Electronics and Communications Engineering, swkim@pknu.ac.kr, 조교수, 정회원

* Pukyong National University Division of Electronics and Communications Engineering, sch804@pukyong.ac.kr, 학생(학사과정)
 논문번호 : 202312-166-A-LU, Received December 14, 2023; Revised December 28, 2023; Accepted December 29, 2023

II. 본 론

2.1 베이스라인 RGBW-to-RGB 모델

영상 변환을 위해 가장 널리 사용되는 모델은 U-Net^[4] 기반의 신경망이다. 본 논문에서는 NAFNet^[5]을 기반으로 하는 신경망을 사용하여 RGBW-to-RGB 변환을 수행한다. NAFNet은 모델의 활성화된 출력의 개수가 신경망의 연산 속도와 비례한다는 점에 착안하여, 비선형 활성화 함수가 없이 효율적으로 영상 변환을 수행하도록 구현된 신경망 모델이다. 비교를 위한 베이스라인 모델은 RGBW CFA를 입력으로 하고, RGB CFA를 출력으로 하는 신경망이며, 입출력 모두 단일 채널로 이루어진 미가공 영상 쌍을 사용하여 지도 학습을 통해 훈련된다.

2.2 제안하는 RGBW-to-RGB 모델

CFA는 각 픽셀이 고유한 색상 패턴을 가지고 있어, 입출력의 패턴이 다를 경우 극심한 색상 왜곡이 발생할 수 있다. 이를 해결하기 위해 입출력 CFA에 적합한 처리 방법이 필요하다. 제안하는 모델은 CFA 패턴의 색상을 채널별로 분할하고, 보간(interpolation)을 통해 입출력 영상을 다채널 영상으로 변환하여 사용한다. 그림 1은 제안하는 기법에서 사용하는 RGBW 디모자이크(demosaic) 방법을 보여준다. W 픽셀은 색상 정보가 있는 주변 픽셀을 사용하여 색상 정보가 없는 픽셀을 복원한다. R/G/B 픽셀은 CFA 해상도로 인해 이웃 픽셀을 바로 사용하는 것이 어렵다. 따라서 다운 샘플링(down-sampling)을 통해 저해상도 영상을 생성하고 이중선형 보간(bilinear interpolation)을 통해 원래 해상도의 RGB 채널을 복원한다. 제안하는 모델은 가공된 n 번째 4채널의 RGBW 영상 \mathbf{x}_r 를 입력 받아

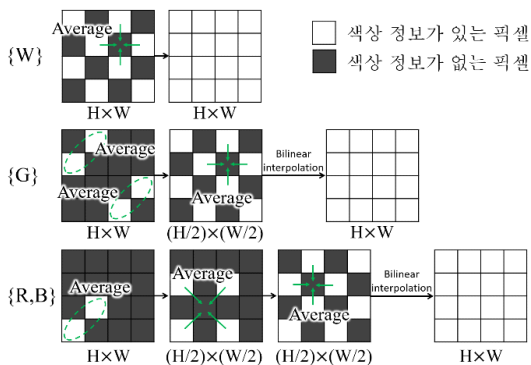


그림 1. RGBW 디모자이크 방법
Fig. 1. RGBW demosaicking method

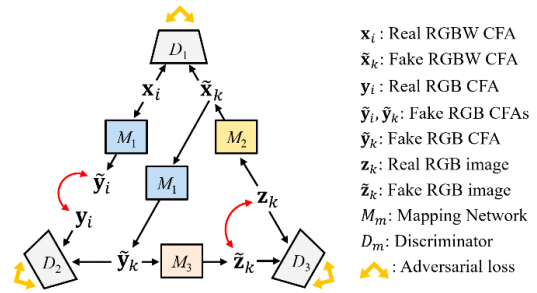


그림 2. 데이터셋 생성 및 학습 프레임워크
Fig. 2. Framework of dataset generation and learning

3채널의 RGB 미가공 영상으로 변환하며, RGB CFA 패턴에 맞게 채널에서 픽셀을 서브샘플링(subsampling)하여 최종적으로 단일 채널 RGB CFA \mathbf{y}_i 에 대한 예측을 수행한다.

2.3 GAN 기반의 데이터 증강 및 학습 기법

CFA 데이터셋 ($\mathbf{x}_i, \mathbf{y}_i$)의 개수 N 이 제한적인 상황을 가정한다. 세 개의 변환 모델이 사용되며, 각각 RGBW CFA를 RGB CFA로 변환하는 M_1 , RGB 영상을 RGBW CFA로 변환하는 M_2 , RGB CFA를 RGB 영상으로 변환하는 M_3 로 나타낸다. 이 중 M_1 이 RGBW CFA 입력 \mathbf{x}_i 를 RGB CFA $\hat{\mathbf{y}}_i$ 로 변환하는 목적 신경망 모델이며, 아래 식과 같이 L_1 -loss $\|\cdot\|_1$ 를 사용하여 학습된다.

$$L_{\text{rec}} = \frac{1}{|B|} \sum_{i \in B} \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|_1 \quad (1)$$

여기서 B 는 CFA 데이터셋의 미니배치(mini-batch)를 나타낸다.

제안하는 기법은 M_2 를 통해 추가적인 P 개의 RGB 영상 \mathbf{z}_k 로부터 RGBW CFA $\tilde{\mathbf{x}}_k = M_2(\mathbf{z}_k)$ 를 생성하는 것을 목표로 한다. 생성된 $\tilde{\mathbf{x}}_k$ 의 품질을 보장하기 위해 순환 생성을 통해 원본 RGB 영상과 유사한 출력을 생성하도록 아래 식과 같은 cyclic consistency loss^[6]를 적용한다.

$$L_{\text{cycle}} = \frac{1}{|C|} \sum_{k \in C} \|\mathbf{z}_k - M_3(M_1(\tilde{\mathbf{x}}_k))\|_1 \quad (2)$$

위 식에서 C 는 RGB 영상 데이터셋에 대한 미니배치를 나타낸다. 추가된 RGB 영상 데이터셋에 사용된 카메라 영상신호처리(ISP) 기법을 알 수 없기 때문에, 변환된 RGB CFA $\hat{\mathbf{y}}_k = M_1(\mathbf{x}_k)$ 를 RGB 영상 $\tilde{\mathbf{z}}_k$ 로 변환하는 모델 M_3 를 추가로 학습한다. 이

과정에서 변환 모델 M_i 이 추가로 생성된 RGBW CFA $\tilde{\mathbf{x}}_k$ 에 대한 변환 함수를 학습할 수 있다.

각각의 변환 모델들의 입출력은 서로 짝을 이루지 않는 실제 데이터가 존재한다. 각 변환 결과를 적절하게 생성하여 데이터를 증강하기 위해 RGBW CFA, RGB CFA, RGB 영상의 진위 여부를 각각 구분하는 판별기 D_1, D_2, D_3 를 도입하고, 변환 모델 M_1, M_2, M_3 와 경쟁적으로 학습한다. GAN 학습을 위한 손실 함수는 다음 식과 같다.

$$L_{adv} = \frac{1}{|B|} \sum_{i \in B} (\log D_1(\mathbf{x}_i) + \log D_2(\mathbf{y}_i)) + \frac{1}{|C|} \sum_{k \in C} (\log(1 - D_1(\tilde{\mathbf{x}}_k)) + \log(1 - D_2(\tilde{\mathbf{y}}_k)) + \log D_3(\mathbf{z}_k) + \log(1 - D_3(\tilde{\mathbf{z}}_k))) \quad (3)$$

안정적인 학습을 위해 적대적 손실 함수에 대한 규제(regularization)^[6] 기법을 적용한다.

전체 학습 과정은 아래의 수식과 같이 경쟁적으로 수행된다.

$$\min_{D_1, D_2, D_3} \max_{M_1, M_2, M_3} L_{adv} + \lambda_{rec} L_{rec} + \lambda_{cycle} L_{cycle} \quad (4)$$

위 식에서 총 손실 함수 가중치는 $\lambda_{rec} = 40, \lambda_{cycle} = 10$ 이 사용되었다.

III. 실험 결과

본 논문에서는 RGBW CFA와 RGB CFA 쌍이 포함된 MIPI Challenge 데이터셋^[7]을 사용하였다. 데이터 셋은 학습 데이터셋 70장, 평가 데이터셋 15장으로 구성되어 있다. 데이터 증강을 위한 RGB 영상은 Adobe5k^[8] 영상 중 500장을 임의로 선택하여 사용했다. 학습 및 평가에 사용된 세부 구현 내용은 표 1과 같다.

표 2는 제안하는 기법과 기존 변환 모델의 RGB CFA에 대한 PSNR, RGB 영상에 대한 PSNR, RGB 영상에 대한 SSIM 성능을 각각 보여준다. 제안하는 RGBW-to-RGB 모델(Proposed)이 데이터 증강 없이 학습했을 때 CFA와 RGB 영상 모두에서 3dB 이상의 성능 향상을 보였으며, 제안하는 데이터 증강을 통한 학습 기법을 적용한 모델(Proposed++)은 추가적인 성능 향상을 보였다. SSIM 측정에서도 동일한 성능 향상을 확인할 수 있다. 실험 결과를 통해 제안하는 변환 모델 및 학

표 1. 구현 세부 정보
Table 1. Implementation details

	구현 내용
변환 모델	NAFNet[5]
분류 모델	ResNet 기반 모델[9]
Optimizer	Adam[10] ($\beta_1 = 0.0, \beta_2 = 0.99$)
학습률	0.001
Weight decay	0.0001
# epochs	200
학습 영상 크기	128×128

표 2. RGB CFA 변환 성능 평가
Table 2. Performance of RGB CFA mapping

	CFA PSNR	RGB PSNR	RGB SSIM
U-Net[4]	32.3	28.5	0.855
NAFNet[5]	33.8	29.1	0.876
Proposed	38.8	32.2	0.923
Proposed++	39.4	33.0	0.946

습 기법이 기존 기법 대비 효과적임을 확인할 수 있다.

IV. 결론

본 논문에서는 CFA의 입출력 패턴 형태에 무관하게 학습 가능한 CFA 변환 모델을 구현하고GAN 기반의 CFA 데이터셋 증강 및 학습 기법을 제안하였다. 실험 결과를 통해 제안하는 기술의 성능을 검증하였다.

References

- [1] B. E. Bayer, *Color imaging array*(1976), U.S. Patent 3971065, Jul. 1976.
- [2] I. Hirota, *Solid-state imaging device, method for processing signal of solid-state imaging device, and imaging apparatus*(2013), U.S. Patent 8436925, Oct. 2013.
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Neural Inf. Process. Syst. (NeurIPS)*, pp. 2672-2680, Montreal, Canada, Dec. 2014.
- [4] O. Ronneberger, P. Fischer, and T. Brox,

- “U-net: convolutional networks for biomedical image segmentation,” in *Proc. Int. Conf. Med. Imag. Comput. Comput.-Assisted Intervention (MICCAI)*, pp. 234-241, Munich, Germany, Oct. 2015.
(https://doi.org/10.1007/978-3-319-24574-4_28)
- [5] L. Chen, X. Chu, X. Zhang, and J. Sun, “Simple baselines for image restoration,” in *Proc. European Conf. Comput. Vis. (ECCV)*, pp. 17-33, Tel Aviv, Israel, Oct. 2022.
(https://doi.org/10.1007/978-3-031-20071-7_2)
- [6] L. Mescheder, S. Nowozin, and A. Geiger, “Which training methods for GANs do actually converge?” in *Proc. Int. Conf. Mach. Learning (ICML)*, pp. 3481-3490, Stockholm, Sweden, Jul. 2018.
- [7] Q. Sun, et al., “MIPI 2023 challenge on RGBW remosaic: Methods and results,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshop (CVPRW)*, pp. 2878-2885, Vancouver, Canada, Jun. 2023.
(<https://doi.org/10.1109/cvprw59228.2023.00289>)
- [8] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, “Learning photographic global tonal adjustment with a database of input/output image pairs,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 97-104, Colorado Springs, USA, Jun. 2011.
(<https://doi.org/10.1109/cvpr.2011.5995413>)
- [9] Y.-J. Choi, Y.-J. Yh, J.-J. Yoo, and J.-W. Ha, “Stargan v2: Diverse image synthesis for multiple domains,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 8188- 8197, Virtual, Jun. 2020.
(<https://doi.org/10.1109/cvpr42600.2020.00821>)
- [10] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proc. Int. Conf. Learning Representation (ICLR)*, San Diego, USA, May 2015.