

적대적 생성 신경망을 활용한 과다 노출 영상 복원 방법

김태하, 양성엽, 강병근, *이의진

서울과학기술대학교

amvision@seoultech.ac.kr, syyang@seoultech.ac.kr,

byeongkeun.kang@seoultech.ac.kr, *yeejinlee@seoultech.ac.kr

Overexposure Image Restoration Using a Generative Adversarial Network

Taeha Kim, Seongyeop Yang, Byeongkeun Kang, *Yeejin Lee

Seoul National University of Science and Technology

요 약

본 논문에서는 과다 노출된 영상을 영상 간 변환(Image-to-Image Translation)을 위해 설계된 적대적 생성 신경망(Generative Adversarial Network)을 활용하여 복원하는 연구를 수행한다. 과다 노출 복원을 위한 기존의 연구에서 과다 노출 영역 판별, 밝기 회복, 색상 보정 과정을 거치는데, 영상 내 과다 노출 영역을 판별하는 과정에서 임의로 결정하는 파라미터에 의해 복원된 영상 결과가 달라지는 한계점을 극복하기 위해 종단간(End-to-End) 신경망을 학습시켜 과다 노출 영역을 별도의 파라미터 선택과 분할된 과정 없이 한 번에 복원하는 방법을 제안한다. 영상 간 변환 신경망 학습에 필요한 과다 노출 여부로 도메인이 분할된 데이터셋은 게임 소프트웨어를 활용하여 만들어 사용하였다. 본 연구에서는 신경망이 생성한 영상이 실제로 과다 노출 영역을 탐지하여 복원하는 것을 확인하였다. 그리고 과다 노출 영역을 탐지하여 복원하는 과정을 학습 단계별로 확인함으로써 신경망이 실제로 과다 노출 복원 과정을 학습함을 보였다.

1. 서론

실제 환경에서 촬영된 영상의 빛의 밝기(Brightness)의 범위는 매우 넓다. 예를 들어, 화창한 날에 야외에서 촬영한 사진의 최대 빛의 밝기는 최소 빛의 밝기와 비교하면 약 400,000 배 정도이다. 실내에서 촬영한 사진의 상대적인 빛의 밝기도 약 1,500 배 정도지만, 일반적인 상용 디지털 카메라의 동적 범위(Dynamic Range)는 1,000:1 정도이다[1]. 과다 노출은 영상 내 밝은 영역의 정보를 자세하게 표현하지 못하는 현상으로, 낮은 동적 범위(Low Dynamic Range, LDR)를 가지는 카메라 센서가 실제 높은 동적 범위(High Dynamic Range, HDR)의 빛의 밝기를 충분히 표현하지 못하여 발생한다.

위와 같은 과다 노출 현상으로 인하여 정보가 손실된 영상을 복원하기 위하여, 영상 내 과다 노출 영역을 판별하고 밝기 회복(Lightness Recovery) 과정과 색상 보정(Color Correction) 과정을 거쳐서 복원하는 기술들이 연구되어 왔다. 일반적으로 입력 영상 내 과다 노출 영역을 판별하는 방법은 영상을 RGB 공간에서 다른 색 공간(Color Space)으로 투영하고 임의로 설정한 경계값을 기준으로 그보다 높으면 과다 노출 영역으로 결정하는 방식이다. 이는 과다 노출된 영역은 RGB 색 공간 내에서 각 픽셀이 가질 수 있는 최대 값에 고정(Clip)되어, RGB 색 공간의 값 만으로는 소실된 정보를 알 수가 없으므로, 밝기와 색상 정보를 분리할 수 있는 다른 색 공간에서 정보 소실 여부를 판단하기 위함이다. 과다 노출 영역을 판별하기 위한 색 공간은 주로 CIELAB 색 공간[1,2]과 YCbCr 색 공간[3]이 활용된다.

그러나 과다 노출된 영역을 임의로 설정한 경계값에 의존하여 결정하는 방식은 임의로 설정한 경계값에 따라서 과다 노출 영역의 판별 결과가 달라지고 복원된 영상도 달라지게 된다. 따라서, 본 연구에서는 기존의 방법과 달리 과다 노출 영역의 판별 과정 없이 과다 노출 영상을 복원하기 위하여 영상간 변환(Image-to-Image Translation) 신경망을 학습하는 새로운 방법을 제안한다. 본 연구에서 제안하는 과다 노출 복원 신경망은 과다 노출 영역 판별, 밝기 회복, 색상 보정 과정을 거치지 않고 과다 노출된 영상을 입력 받아 복원한 영상을 반환한다. 기존의 방법은 과다 노출 영역 판별을 위하여 임의로 결정해줘야 하는 파라미터들이 존재하여 정해진 파라미터 값에 의하여 결과가 달라지나, 본 연구는 휴리스틱(Heuristic)한 파라미터의 결정 과정 없이 종단간(End-to-End) 신경망을 학습하여 과다 노출 영역을 복원한다.

2. 제안하는 방법

종단간 신경망은 목적에 따라서 다양한 종류가 있지만, 본 연구에서는 각 영상의 특성을 대표하는 도메인(Domain)을 학습하여 다른 도메인으로 영상을 변환하는 적대적 생성 신경망(Generative Adversarial Network, GAN)[4]을 사용하여 과다 노출된 영상을 과다 노출에 의한 정보 손실이 보완된 영상으로 복원한다. CycleGAN[5]은 영상 간 변환을 목표로 설계된 신경망 중 하나로 각 영상의 특성을 대표하는 도메인을 변환하는 대표적인 신경망이다.

CycleGAN 을 학습시키기 위한 데이터셋은 도메인을 분명하게 구별할 수 있을수록 이상적이다. 하지만 객체 탐지나 영상 분할 등의 연구 목적과 달리 과다 노출 복원 목적을 위해 과다 노출 유무로 도메인이 명확히 분리된 사용 가능한 데이터셋은 존재하지 않는다. 따라서, 본 연구에서는 데이터셋 부족 문제를 해결하기 위하여 영상 내 의미 분할(Semantic Segmentation)에 필요한 데이터셋을 증강시킬 목적으로 게임 내 영상을 이용한 연구[6]와 유사하게 광원 위치의 조절이 가능한

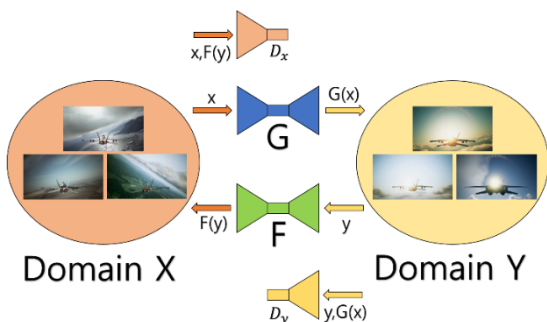


그림 1 게임 소프트웨어로부터 촬영한 영상을 이용하여 구현한 데이터셋의 도메인 분할과 수식을 나타낸 도식도

컴퓨터 게임 소프트웨어[7]를 사용하여 데이터셋을 구현하였다.

그림 1은 구현한 데이터셋을 활용하여 분할한 도메인과 학습 과정을 보여준다. 본 논문에서는 도메인 X 를 과다 노출이 일어나지 않은 영상의 도메인, Y 를 과다 노출이 일어난 영상의 도메인으로 정의하며, 사용한 도메인별 영상의 총량은 아래 표 1과 같다.

표 1 실험에 사용된 데이터셋의 도메인별 영상의 총량

	Domain X	Domain Y
Training	3,864	3,520
Testing	1,212	780

제안하는 과다 노출 복원 적대적 신경망은 2 개의 생성 신경망(Generator Network) G , F 와 2 개의 판별 신경망(Discriminator Network) D_x , D_y 로 구성된다. 생성 신경망은 도메인 변환된 합성 영상을 생성하는 신경망으로, 도메인 X 의 영상 x 를 도메인 Y 의 영상으로 생성 신경망 G 를 통해 합성한 영상은 $G(x)$, 도메인 Y 의 영상 y 를 도메인 X 의 영상으로 생성 신경망 F 를 통해 변환한 영상은 $F(y)$ 로 표현한다. 판별 신경망 D_x 는 생성 신경망 F 가 생성한 과다 노출된 영상 $F(y)$ 와 실제로 과다 노출된 영상 x 를 입력 받아 입력 받은 영상이 실제 영상인지 생성된 영상인지를 판단하는 역할을 하며, 판별 신경망 D_y 는 생성 신경망 G 가 생성한 과다 노출이 없는 영상 $G(x)$ 와 실제 영상 y 를 입력 받아 실제 영상 여부를 판단하는 역할을 한다.

제안하는 신경망 G , F , D_x , D_y 를 학습하기 위한 손실 함수는 적대적 생성 신경망 각각을 학습하기 위한 \mathcal{L}_{GAN} 과 순환 일관성 손실(Cycle Consistency Loss) \mathcal{L}_{Cycle} 의 가중합으로 정의하며, 식 (1)과 같다.

$$\mathcal{L}(G, F, D_x, D_y) = \mathcal{L}_{GAN}(G, D_y, X, Y) + \mathcal{L}_{GAN}(F, D_x, Y, X) + \lambda \mathcal{L}_{Cycle}(G, F) \quad (1)$$

적대적 생성 신경망의 손실함수 \mathcal{L}_{GAN} 은 도메인 변환 과정에서 생성 신경망과 판별 신경망을 학습시키는 손실함수로 도메인 X 에서 도메인 Y 로 변환되는 과정에서 신경망 G 와 D_y 를 학습시키는 \mathcal{L}_{GAN} 은 다음과 같이 정의한다.

$$\mathcal{L}_{GAN}(G, D_y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)} [\log D_y(y)] + \mathbb{E}_{x \sim p_{data}(x)} [\log(1 - D_y(G(x)))] \quad (2)$$

식 (2)에서 $\mathbb{E}_{x \sim p_{data}(x)}$ 는 데이터 x 가 실제 데이터 x 의 확률 밀도 함수를 따를 기댓값을 나타낸다. 판별 신경망 D_y 는 주어진 영상이 실제 도메인 Y 에서 온 영상인지 원래 도메인 X 에서 G 를 통해 변환된 이미지인지를 판별하므로 이 손실함수를 최대화(maximize)하려고 하며 생성 신경망 G 는 반대로 이

손실함수를 최소화(minimize)하여 판별 신경망 D_Y 가 생성된 영상을 실제 영상으로 판별하도록 한다. 반대로 도메인 Y 에서 도메인 X 로 변환되는 과정에서 신경망 F 와 D_X 를 학습시키는 \mathcal{L}_{GAN} 은 다음과 같이 정의한다.

$$\mathcal{L}_{GAN}(F, D_X, Y, X) = \mathbb{E}_{x \sim p_{data}(x)} [\log D_X(x)] + \mathbb{E}_{y \sim p_{data}(y)} [\log (1 - D_X(F(y)))] \quad (3)$$

순환 일관성 손실은 영상을 목표 도메인으로 변환하고, 다시 원래의 도메인으로 변환한 영상과 원래 영상과의 차이를 손실함수로 사용한다. 순환 일관성 손실은 생성 신경망 G, F 의 학습에만 사용하며 손실함수 \mathcal{L}_{Cycle} 는 다음과 같이 정의된다.

$$\mathcal{L}_{Cycle}(G, F) = \mathbb{E}_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1] \quad (4)$$

생성 신경망 G 와 F 는 각 도메인의 실제 영상 x, y 와 도메인 변화를 거쳐서 원래의 도메인으로 돌아온 영상 $F(G(x)), G(F(y))$ 의 차의 L1-norm 의 기댓값을 나타내는 순환 일관성 손실을 최소화하는 방향으로 학습한다.

3. 실험 및 성능 평가

제안하는 과다 노출 복원 네트워크의 학습은 구현한 데이터셋을 사용하여 2 장에서 제안한 적대적 생성 신경망 구조를 적용하여 수행하였다. 입력 영상과 출력 영상의 해상도는 256×256 이며 200 Epoch 동안 Batch size 는 4 개로 학습하였으며, Adam Optimizer 를 사용하여 Momentum rate 는 0.5, Learning rate 는 0.0002 로 설정하고, 100 Epoch 이후엔 선형적으로 감소하게 설정하였다. CPU 는 Intel I7 8700, GPU 는 NVIDIA RTX 2080Ti 11GB, RAM 64GB 의 컴퓨터 환경에서 50 시간 동안

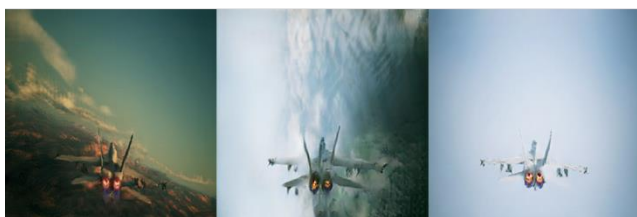
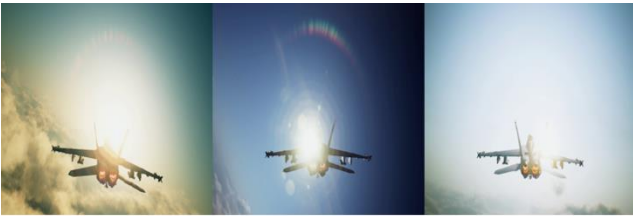


그림 2 구현한 데이터셋으로 학습한 신경망을 이용하여 얻은 출력 영상들의 예시. 위쪽 영상들은 도메인 X 의 입력 영상, 아래쪽 영상들은 도메인 Y 의 출력 영상.

학습하였다.

그림 2 는 위의 실험환경으로 학습한 신경망을 사용하여 얻은 출력 영상의 예시이다. 그림 2 의 위쪽 영상은 도메인 X 에서 나온 영상으로 영상 내 태양에 의해 비행기의 일부가 과다 노출되어 정보가 소실된 영상이다. 그림 2 의 아래쪽 영상은 이를 신경망을 이용해 복원한 영상으로 위쪽 그림에서 나타나지 않은 비행기의 부분을 복원한 모습을 확인할 수 있었다. 결과를 바탕으로 신경망이 도메인 변환을 학습하는 과정에서 비행기의 동체를 학습하여 과다 노출이 일어난 부분을 찾아 비행기의 동체를 복원하였음을 확인하였다.



그림 3 신경망의 학습 과정에서 생성한 영상들의 예시. 왼쪽부터 1, 5, 10, 14, 18 Epoch 에서 출력한 영상.

그림 3 은 CycleGAN 이 영상 내 과다 노출된 부분을 찾아 복원하는 것을 학습하는 과정을 Epoch 별로 나타낸 예시이다. 왼쪽부터 1, 5, 10, 14, 18 Epoch 에서 출력한 영상들이며, 학습 초기에는 태양에 의해 과다 노출된 영역을 복원하지 못한 결과를 출력하였으나 학습이 진행되며 과다 노출된 영역을 복원하고 영역 내 비행기 동체 외의 부분을 주변 배경에 맞게 생성하는 것을 확인할 수 있었다.

4. 결론

본 연구는 적대적 생성 신경망을 활용하여 영상 내 과다 노출된 부분을 찾는 과정 없이 영상 내 과다 노출 영역에 손실된 정보를 복원하였다. 그러나 신경망 학습을 위한 과다 노출 유무로 도메인이 명확히 분리된 사용 가능한 데이터셋은 극히 제한적이므로, 광원 조절이 가능한 소프트웨어를 활용하여 데이터셋을 구축하였다. 또한, 신경망이 학습하는 과정을 통해서 과다 노출된 영역을 스스로 찾아서 복원하는 과정을 학습할 수 있음을 실험 결과를 통해 증명하였다.

Acknowledgement

이 연구는 2021 년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0008473, 2021 년 산업혁신인재성장지원사업).

참고문헌

- [1] Y. Shen, R. Mo, Y. Zhu, L. Wei, W. Gao and Z. Peng, "Over-exposure image correction with automatic texture synthesis," *2011 4th International Congress on Image and Signal Processing*, 2011, pp. 794-797, doi: 10.1109/CISP.2011.6100306.
- [2] D. Guo, Y. Cheng, S. Zhuo and T. Sim, "Correcting over-exposure in photographs," *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 515-521, doi: 10.1109/CVPR.2010.5540170.
- [3] D. Lee, Y. Yoon, S. Kang and S. Ko, "Correction of the overexposed region in digital color image," in *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 173-178, May 2014, doi: 10.1109/TCE.2014.6851990.
- [4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. "Generative adversarial nets." in *Neural Information Processing Systems*, 2014.
- [5] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks" *IEEE International Conference Computer Vision*, 2017.
- [6] Richter, Stephan R., et al. "Playing for data: Ground truth from computer games." *European Conference on Computer Vision*, Springer, Cham, 2016.
- [7] (2021) Ace Combat 7 Official Website. [Online]. Available : <https://www.bandainamcoent.com/games/ace-combat-7>