

자연스러운 저조도 영상 개선을 위한 비지도 학습

이헌상[†], 손광훈^{**}, 민동보^{***}

Unsupervised Learning with Natural Low-light Image Enhancement

Hunsang Lee[†], Kwanghoon Sohn^{**}, Dongbo Min^{***}

ABSTRACT

Recently, deep-learning based methods for low-light image enhancement accomplish great success through supervised learning. However, they still suffer from the lack of sufficient training data due to difficulty of obtaining a large amount of low-/normal-light image pairs in real environments. In this paper, we propose an unsupervised learning approach for single low-light image enhancement using the bright channel prior (BCP), which gives the constraint that the brightest pixel in a small patch is likely to be close to 1. With this prior, pseudo ground-truth is first generated to establish an unsupervised loss function. The proposed enhancement network is then trained using the proposed unsupervised loss function. To the best of our knowledge, this is the first attempt that performs a low-light image enhancement through unsupervised learning. In addition, we introduce a self-attention map for preserving image details and naturalness in the enhanced result. We validate the proposed method on various public datasets, demonstrating that our method achieves competitive performance over state-of-the-arts.

Key words: Unsupervised Learning, Low-light Enhancement, Bright Channel Prior

1. 서 론

저조도 환경에서 획득된 영상은 적은 광자수와 낮은 신호 대 잡음비로 인하여 낮은 가시성과 대비, 심각한 노이즈에 의해 열화될 수 있다. 노출 시간을 늘림으로서 보다 높은 화질의 영상을 획득할 수 있으나, 이 방법은 카메라나 물체가 움직이는 상황에서 블러 문제를 초래한다. 이러한 열화현상은 객체 인식 및 추적 [1, 2, 28], 영역 분할 [3] 등과 같은 다양한 컴퓨터비전 기술들의 성능에 심각한 영향을 끼쳐, 이를 해결하기 위해 지난 수십 년간 다양한 방식들이 제안되어 왔다.

히스토그램 기반 [4, 5] 방식은 관측되는 영상의 동적 범위(dynamic range)를 늘려줌으로써 저조도 영역의 밝기를 증폭시켜 줄 수 있으나, 단순히 대비를 향상시키는 것이 목적이기에 개선된 결과 영상이 자연스럽지 못하며 영상을 지나치게 혹은 덜 향상시킬 수 있다. 다른 접근 방법으로 영상을 반사도와 조명 성분으로 분해하는 Retinex 이론에 기반한 방식들 [6-8]이 존재한다. 이러한 방법들은 추정된 반사도 영상을 개선된 영상으로 가정하거나 이를 조명 영상과 곱하여 개선된 영상을 얻는다. 하지만 어두운 영역에서의 분해가 잘 이루어지지 않으며 과한 개선으로 인하여 영상에 부자연스러움이 존재한다.

※ Corresponding Author : Dongbo Min, Address: (03760) 332, Jinseonmi building, Ewha Womans University, 52 Ewhayeodaegil, Seodaemun-gu, Seoul, Korea 03760, TEL : +82-2-3277-6892, E-mail : dbmin@ewha.ac.kr

Receipt date : Nov. 25, 2019, Revision date : Jan. 8, 2020
Approval date : Jan. 14, 2020

[†] School of Electrical & Electronic Engineering, Yonsei University (E-mail : hslee91@yonsei.ac.kr)

^{**} School of Electrical & Electronic Engineering, Yonsei University (E-mail : khsohn@yonsei.ac.kr)

^{***} Department of Computer Science & Engineering, Ewha Womans University

※ This work was supported by research fund of Chungnam National University. This work was performed when Dongbo Min was with Chungnam National University.

앞서 설명한 수작업 방식들과 달리, 최근 딥러닝 기반 방식들이 영상복원 분야에서 매우 뛰어난 성능을 보임에 따라, 저조도 영상 개선을 위해 딥러닝을 접목한 방식 [9-13, 22] 등이 활발히 연구되고 있다. 딥러닝의 뛰어난 특징 표현 능력으로 인해 수작업 방식들보다 뛰어난 성능을 보이지만, 몇 가지 문제점을 가지고 있다. 우선, 이들 대부분은 지도 학습에 기반하기 때문에 저조도 영상을 입력으로, 일반 조도 영상을 정답으로 사용하는 많은 양의 데이터 셋을 필요로 한다. 하지만, 실제 환경에서 그러한 정답이 포함된 대규모 학습 데이터를 확보하는 것은 현실적으로 매우 어렵다. 예를 들어, 동일한 장면에서 노출 시간을 바꿔가며 영상들을 얻고 이를 통해 high dynamic range (HDR) 영상을 획득하여 정답으로 사용하기는 비교적 쉬우나, 이는 정적이며 낮에 획득한 장면에 대해서만 가능하다. 야간에 촬영되는 저조도 영상과 낮에 촬영되는 정답 영상은 정확히 같은 위치에서 시간 차이를 두고 획득해야 하기에 학습 데이터 획득에 제한이 발생할 수밖에 없다. Lore [9]는 감마 보정을 통해 인위적으로 개선된 저조도 영상에 가우시안 노이즈를 추가한 학습 영상을 사용하기에, 실제 저조도 영상에 적용할 경우 큰 도메인 차이로 인해 낮은 성능을 보인다. Wei [10]는 동일한 카메라로 ISO와 노출 시간을 조절하여 같은 장소를 여러 장 촬영함으로써 저조도와 일반조도 영상 쌍을 획득하였으나, 이는 카메라 설정으로 인하여 얻은 영상으로 실제 저조도 환경을 반영한다고 보기 어렵다. Wang [13]은 저조도 영상 개선을 위한 새로운 데이터 셋을 공개하였다. 이들은 카메라 촬영이나 Flickr [14]를 통해 일반 조도 영상을 획득한 뒤, MIT-Adobe FiveK [15] 데이터 셋과 유사한 방식으로 세 명의 전문가가

Adobe Lightroom으로 영상을 수정하여 저조도 영상을 얻었다. 하지만 이는 많은 시간과 인적 자원을 필요로 한다.

또한, 몇몇 CNN 기반의 방식들은 시각적으로 매우 부자연스러운 결과를 제공한다 [10-12]. 예를 들면, Fig. 1에서 Wei [10]는 비지도 기반의 분해 방식의 한계로 인해 조도가 낮은 영역에서 잘못된 반사도 맵을 예측하고, 대응되는 조명 값을 과도하게 증가시킴으로써, 매우 부자연스러운 영상이 출력된다. Li [11]는 조명 맵을 예측하고 이를 감마 보정을 통해 개선하기 때문에 영상에 따라 결과가 여전히 어둡게 보이는 경향이 있으며, 영상의 구조적인 정보를 고려하지 않고 오버-스무딩된 조명 맵으로 인해 경계 주위에서 후광 효과(halo artifacts)가 존재한다. 또한, Ren [12]은 이러한 과도한 개선 문제와 후광 효과를 모두 가지고 있다.

이러한 문제를 해결하기 위하여, 본 논문에서는 비지도 학습 기반의 저조도 영상 개선 방식을 제안한다. 이를 위해 안개 제거 연구 [16, 17]에서 사용되는 dark channel prior (DCP)에 착안하여, bright channel prior (BCP)에 기반한 비지도 딥러닝 네트워크를 설계하였다. BCP는 저조도가 개선된 영상의 패치 내에서 가장 밝은 값을 가지는 픽셀의 값이 1에 가까워야 한다는 제약 조건으로, 본 논문에서는 BCP를 통해 초기 조명 맵을 예측하고, 이를 유사 정답으로 사용해 네트워크를 학습한다. 우리가 아는 한도 내에서, 제안 방식은 비지도 학습을 통해 저조도를 개선하는 제거하는 첫 번째 연구로 정답이 포함된 대규모 학습데이터가 필요하지 않다. 또한, 대비가 개선된 결과 영상에서 자연스러움을 보존하기 위하여 자기-주목(self-attention) 맵을 사용하였다. 이는 제안하

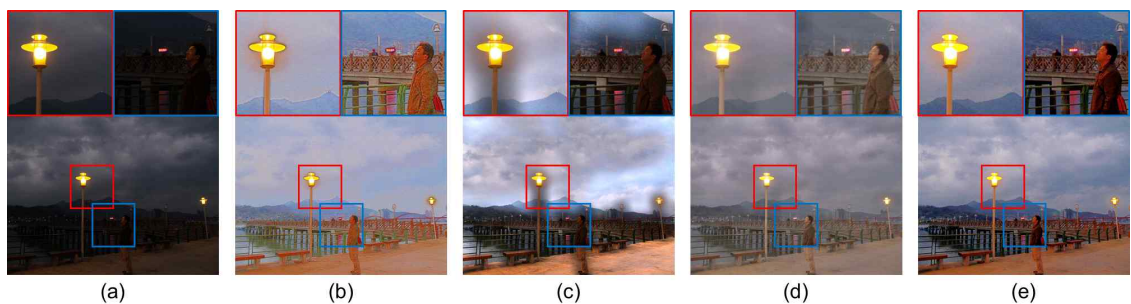


Fig. 1. Unnatural artifacts of existing methods. Each image represents (a) input image, (b) RetinexNet [10], (c) LightenNet [11], (d) DHN [12] and (e) the proposed method.

는 네트워크가 과도한 개선을 하지 않도록 제어해주며, 어두운 영역에 존재하는 물체의 디테일을 보존할 수 있게 만든다.

요약하자면 본 논문의 기여는 다음과 같다.

- 데이터 셋 획득에 따른 어려움을 해결하기 위해 비지도 학습 기반의 저조도 영상 개선 네트워크를 제안한다.
- 자기-주목(self-attention) 맵을 통해 영상의 자연스러움과 디테일을 보존함과 동시에 저조도를 개선한다.

2. 관련 연구

2.1 수작업 기반 저조도 개선 연구

저조도 영상 개선은 컴퓨터 비전에서 가장 중요한 문제 중 하나로 오랫동안 많은 방법이 연구됐다. 가장 기초적인 방법으로 히스토그램에 기반한 연구들이 있는데, 대부분 히스토그램 평활화 (Histogram Equalization: HE)를 통해 입력 영상의 히스토그램을 적절히 조정함으로써 저조도 영상을 개선한다. 대비 제한 적응 히스토그램 평활화(Contrast Limited Adaptive Histogram Equalization: CLAHE) [4]는 영상을 작은 블록으로 나누고, 해당 블록 별로 히스토그램 평활화를 적용하며, 대비 제한 기법을 사용하여 경계 부분의 노이즈를 감소시킨다.

다른 연구로 반전된 저조도 영상을 안개 영상으로 간주하여 처리하는 방식들 [18, 19]이 있다. 반전된 영상에 기존의 안개 제거 방법 [16]을 적용하여 영상을 개선한 후 다시 반전시킴으로써 개선된 영상을 얻을 수 있다. Zhang [18]은 저조도 영상을 개선한 후 joint-bilateral filter [20]를 적용하여 노이즈를 억제하였으며, Li [19]는 결과 영상의 품질을 향상시키기 위하여 영상을 슈퍼픽셀로 나눈 뒤, 각 분할된 슈퍼 픽셀들에 BM3D [21]를 수행하여 노이즈를 제거하였다.

이외에도 영상을 반사도와 조명 성분으로 분해하는 Retinex 모델을 사용하여 저조도 영상을 개선하려는 시도가 많이 이루어졌는데 [6-8], Wang [6]은 조악한 조명 맵에 경계 보존 필터링을 적용하여 개선된 영상을 획득하였다. Fu [7]는 Retinex 모델에서 많이 사용되는 로그 변환에 대한 문제점을 지적하고, 이를 해결하기 위해 weighted variational model을

제안하여 정규화 함수를 개선하였다. Guo [8]은 조명 맵을 예측하고 구조적 정보를 보존할 수 있는 스무딩 기법을 적용하였고, 이를 저조도 영상과 나누어 획득한 반사도 영상을 개선된 영상으로 간주하였다.

2.2 딥러닝 기반 저조도 개선 연구

앞서 언급한 전통적인 방법들과는 다르게, 최근에는 딥러닝을 사용한 학습 기반 방법들이 저조도 영상 개선을 위해 활발히 사용되고 있다. Lore [9]는 stacked auto-encoder를 사용하여 패치 단위로 저조도 영상 개선과 노이즈 제거를 차례대로 수행하고, Li [11]은 저조도 영상과 일반조도 영상의 조명 맵 간의 맵핑 관계를 찾고자 하였다. Wang [13] 또한 낮은 해상도 도메인에서 두 조명 맵 간의 맵핑 관계를 찾는 방식으로 저조도 영상을 개선하고자 하였으며, bilateral grid [22]을 사용하여 해상도를 증가시켰다. Wei [10]는 Retinex 모델을 사용하여 영상을 반사도 성분과 조명 성분으로 나누었다. 조명 성분의 값을 증가시키고 노이즈 제거 기법인 BM3D [21]를 반사도 영상에 적용한 후 두 성분을 다시 결합함으로써 개선된 영상을 획득하였다. Ren [12]은 전역적 특징을 예측함으로써 가시성을 개선하였으며, 공간 적응적 recurrent neural networks (RNN)을 사용하여 영상의 디테일을 보정하였다. 앞서 언급한 방법들은 모두 정답 데이터를 통한 지도 학습에 기반하며, 실제 환경이 고려되지 않은 인위적인 데이터 셋에 의존하여 학습된다.

3. 제안하는 방식

이 장에서는 비지도 학습 기반 저조도 개선 방식을 제안한다. 전체 네트워크 구조는 Fig. 2에서 볼 수 있듯이 간단한 encoder-decoder로 구성되어 있다. 먼저 저조도 영상을 입력으로 받아 BCP [23]에 기반하여 초기 조명 맵 \tilde{t} 를 계산하고, 이를 통해 비지도 손실 함수 L_{BCP} 를 정의하여 최종적으로 대비 개선을 위한 조명 맵 t 를 예측한다. 초기 조명 맵에 존재하는 블록 효과를 억제하기 위해서 soft-matting [24]을 사용한 정규화 함수를 추가하였으며, 자기-주목(self-attention) 맵은 과도한 개선을 막으면서 개선된 영상의 디테일을 보존함과 동시에 더욱 자연스러운 결과를 제공한다. 본 논문에서 제안하는 방법은 다른 방법들 [10-13]과 달리 복잡한 구조의 네트워크

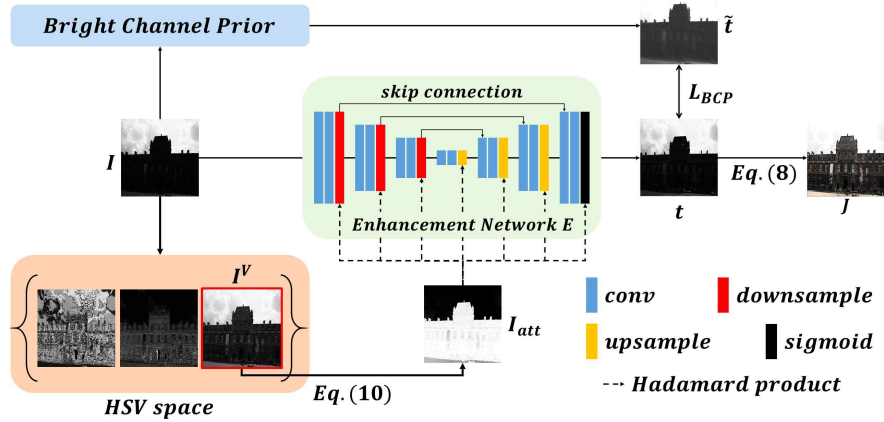


Fig. 2. Overall architecture of the proposed method. An initial illumination map \tilde{t} and self-attention map I_{att} are computed from input image. Then, the enhancement network E estimates a final illumination map.

를 사용하지 않고 매우 간단한 encoder-decoder 구조만을 사용하면서, 비지도 학습임에도 불구하고 최근의 지도 학습 기반 방법들 [10-13]과 유사하거나 더 나은 성능을 보인다.

3.1 초기 조명 맵 계산

BCP [23]은 안개 영상의 투과도 맵을 예측하여 안개를 제거하기 위해 사용된 dark channel prior (DCP) [16]의 변형으로 영상의 지역적인 노출값을 조정하기 위하여 처음 제안되었다. 본 논문에서는 비지도 손실 함수를 정의하는데 사용되는 초기 조명 맵을 예측하기 위하여 BCP를 사용한다. 저조도 개선을 위한 목적 함수는 다음과 같이 정의하였다.

$$I_p = t_p J_p + (1 - t_p) A, \quad (1)$$

여기서 $I_p \in \mathbb{R}^3$ 는 관측되는 저조도 영상이며, $J_p \in \mathbb{R}^3$ 는 저조도가 개선된 영상을 의미한다. $t_p \in \mathbb{R}$ 과 $A \in \mathbb{R}^3$ 은 각각 투과도 맵과 환경 조명을 의미한다. 픽셀 p 를 중심으로 한 패치 $\Omega(p)$ 내의 가장 밝은 값은 다음과 같은 수식으로 나타낼 수 있다.

$$J_p^{bright} = \max_{c \in r, g, b} (\max_{q \in \Omega(p)} J_q^c) \quad (2)$$

c 는 영상의 r, g, b 채널을 의미한다. 여기서 A 는 이미 주어지고, 패치 내의 조명 맵 \tilde{t}_p 가 일정하다고 가정하면, 식 (1)의 양쪽에 최대 연산자를 적용하여 다음과 같이 정리할 수 있다.

$$\max_{c, q} J_q^c = \tilde{t}_p J_p^{bright} + (1 - \tilde{t}_p) A^c \quad (3)$$

수식의 간편화를 위하여

$\max_{c, y} (k) = \max_{c \in r, g, b} (\max_{y \in \Omega(x)} (k))$ 로 정의하였으며, 영상 I 와 J 는 0부터 1 사이의 값으로 정규화되어 있다. BCP [28]에 의하면 패치 내의 가장 밝은 값이 1이 되어야 하므로

$$J_p^{bright} \rightarrow 1 \quad (4)$$

이며, 식 (4)를 식 (3)에 대입한 뒤 조명 맵에 대해서 아래와 같이 정리할 수 있다.

$$\tilde{t}_p = 1 - \max_{c, q} \left(\frac{1 - I_q^c}{1 - A^c} \right) \quad (5)$$

하지만, 위 식을 통해 계산된 초기 조명 맵으로 저조도 개선을 수행하면 그림자 혹은 검은 물체와 같이 원래 어둡게 보여야 하는 영역까지 밝아질 수 있으므로 영상이 부자연스럽게 보일 수 있다. 이러한 영향을 줄이기 위해 추가적인 매개 변수 $\omega (0 < \omega < 1)$ 를 추가하였으며, 본 논문에서는 ω 를 0.85로 고정해 놓고 조명 맵을 예측하였다.

$$\tilde{t}_p = 1 - \omega \cdot \max_{c, q} \left(\frac{1 - I_q^c}{1 - A^c} \right) \quad (6)$$

식 (6)은 A 가 이미 주어지고 있다고 가정한 뒤 도출한 것이므로, A 를 먼저 계산해주는 과정이 필요하다. A 를 가장 어두운 픽셀의 값으로 정의할 수도 있으나, 이는 하나의 픽셀만으로 전체 영상의 환경 조명 값을 정의하는 것으로 이로 인해 잘못된 조명 맵을 계산할 수 있다. 그러므로, DCP [16]과 유사하게 저조도 영상에서 bright channel ($= \max_{c \in r, g, b} I^c$)을 계산한 뒤,

bright channel 내에서 가장 어두운 0.1% 픽셀들의 인덱스(=K)를 통해 아래와 같이 계산한다.

$$A = \frac{1}{|K|} \sum_{p \in K} I_p \quad (7)$$

3.2 비지도 기반 조명 맵 예측

3.2.1 Bright Channel Prior 손실 함수

제안 방법에서는 식 (7)을 식 (6)에 대입하여 계산된 초기 조명 맵을 유사 정답 데이터로 이용하여 네트워크를 학습한다. Fig. 2에서 저조도 개선 네트워크는 저조도 영상을 입력으로 받아 \tilde{t} 와의 지도 학습을 통해 조명 맵 t 를 예측한다. 예측된 조명 맵 t 는 다음 식을 통해 개선된 영상을 예측하는 데 사용된다.

$$J_p = \frac{I_p - A}{t_p} + A \quad (8)$$

하지만 \tilde{t} 가 패치 내에서 일정하다고 가정하기 때문에, Fig. 3의 (e)와 같이 블록 효과가 존재한다. 그러므로, \tilde{t} 를 조명 맵으로 사용할 경우, J 또한 블록 효과가 존재할 수 있기 때문에, 추가적인 개선이 필요하다. 식 (1)은 출력 영상을 전경과 배경의 블록한 선형 조합으로 보는 영상 매칭의 수식과 유사하다. 따라서 본 논문에서는 soft matting [24] 알고리즘을 추가적인 정규화 함수로 사용한 비지도 기반 손실

함수를 사용하여 블록 효과를 억제한다. 정규화 함수가 추가된 BCP 기반 손실 함수는 다음과 같이 정의된다.

$$L_{BCP} = \frac{1}{N} \sum_p \left\{ (t_p - \tilde{t}_p)^2 + \lambda \sum_{i,j \in \Psi(p)} w_{ij} (t_i - t_j)^2 \right\} \quad (9)$$

여기서 w_{ij} 는 I_i 와 I_j 간의 유사도를 계산하는 matting Laplacian 행렬의 가중치를 의미하며, $\Psi(p)$ 는 중심점 p 주변의 3×3 패치를 의미한다. N 은 픽셀의 총 개수이며, λ 는 양쪽의 손실 함수 간의 비중을 조절하는 가중치 상수이다.

3.2.2 자기-주목(self-attention) 기반 저조도 개선

식 (7)에서 A 는 어두운 픽셀들의 평균으로 계산되기에 일부 영역에서만 도출되지만, 영상 전체에 대해서 일정하다. 따라서 Fig. 4와 같이 영상 내의 밝기의 편차가 큰 영상에 대해서 개선을 수행하면, 하늘이나 눈처럼 수행 전에서도 밝은 영역은 개선 후에 과포화(over-saturation)가 발생할 수 있다. 환경 조명 A 를 지역적으로 변화시키며 학습함으로써 과포화를 해결할 수도 있으나, 비지도 개선 방식에서 이는 어려운 과제이다. 따라서 본 논문에서는 이를 해결하기 위하여 자기-주목(self-attention) 맵을 통한 저조도 영상 개선을 수행한다. 이는 다음과 같이 저조도 입력 영상을 HSV 컬러 공간으로 변환한 뒤 계산한다.

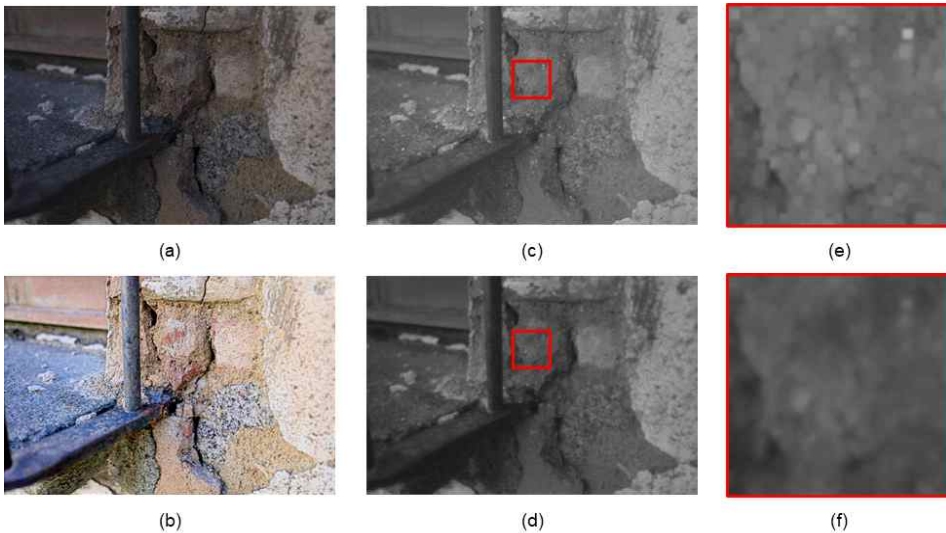


Fig. 3. Comparison with initial illumination map and predicted illumination map. (a) low-light image, (b) enhanced image, (c) initial illumination map, (d) predicted illumination map, (e) partial image of (c), (f) partial image of (d).



Fig. 4. An example of saturation, (a) input image, (b) enhanced image with saturation.

$$I_{att} = (1 - I^V)^\gamma \quad (10)$$

입력 영상의 밝기만을 고려하기 위해 HSV에서 V 채널만 사용하며, $\gamma \geq 1$ 은 자기-주목(self-attention) 맵의 곡률을 계산하기 위한 상수이다. 위와 같이 자기-주목(self-attention) 맵은 매우 간단한 연산을 통해 구할 수 있어 효율적이며, 입력 영상만을 사용하기 때문에 추가적인 데이터가 필요하지 않다. I_{att} 는 Fig. 2에서 볼 수 있듯이 저조도 개선 네트워크의 모든 컨볼루션 계층의 출력과 곱해짐으로써 입력 영상의 밝기에 따라 네트워크의 출력 값을 적절하게 조절해준다. 따라서 밝은 영역일수록 낮은 가중치를 두어 개선을 억제하여 과포화(over-saturation) 문제를 감소시키고 어두운 영역일수록 높은 가중치를 두어 더 많은 개선을 수행하도록 함으로써 개선된 영상이 시각적으로 더욱 자연스러워지도록 만들어준다.

4. 실험

이 장에서는 제안한 모델의 저조도 향상 성능을 검증한다. 먼저 실험 환경에 대해서 설명하고, 최신 기법들과의 정량적, 정성적 비교를 통해 제안한 모델의 성능을 평가한다.

4.1 학습 상세

제안된 방법은 비지도 학습을 기반으로 하기 때문에, 학습에 어떠한 저조도 영상도 사용할 수 있다. 본 논문에서는 LOL [10]과 SICE [25]과 같은 공개된 데이터 셋으로부터 학습 데이터를 수집하였다. LOL

데이터 셋 [10]은 1,500개의 저조도와 일반조도 영상 쌍을 제공하며, SICE 데이터 셋 [25]은 589개 장면의 다중 노출 연속 영상과 그에 해당하는 HDR 영상을 정답으로 제공한다. 학습/검증 데이터 분할을 위해서 LOL에서 1485장의 영상을, SICE에서 534장의 영상을 학습에 사용하였으며, 나머지 영상들을 검증에 사용하였다. 최신 방법인 UPE [13]을 따라서, 다음과 같이 500장의 MIT-Adobe FiveK 데이터 셋 [15] 영상을 평가에 사용하였으며, 추가적으로 저조도 영상 개선 연구에서 많이 사용되는 LIME [8]과 NPE [6] 데이터 셋 또한 평가를 위해 사용하였다.

제안 방식은 Python기반의 Tensorflow를 사용하여 구현되었으며, NVidia TITAN RTX를 사용하여 학습되었다. 학습 시에는 Adam [26]을 사용하여 최적화하였으며, 학습률은 10^{-3} 에서 시작하여 10^{-5} 까지 로그 도메인에서 점차적으로 감소시켰고, 32의 배치 크기로 50 epoch까지 학습하였다. 네트워크의 가중치는 평균이 0이고 분산이 0.1이 되도록 랜덤하게 초기화하였으며, 데이터 증가를 위해 다음과 같이 3가지의 방법을 차례대로 적용하였다. 1) 임의의 위치에서 영상을 192×192 크기로 잘라 패치를 획득하고, 2) 패치를 임의의 각도(90° , 180° , 270°)로 회전시킨 뒤, 3) 회전시킨 패치의 상하좌우를 임의로 반전시켰다. 위의 3가지 단계를 적용하여 얻어진 영상을 네트워크의 입력으로 사용하였다. 또한, BCP를 계산할 때의 패치 Ω 의 크기는 15×15 로 설정하였으며, 정규화 함수를 위한 패치 Ψ 의 크기는 3×3 을 사용하였다. 손실 함수의 비중을 위한 λ 는 0.1, 자기-주목(self-at-

tention) 맵의 곡률을 결정하는 γ 는 2.5로 설정하였다.

4.2 정량적 결과

본 논문에서는 제안하는 방법은 비지도 학습 기반이기 때문에 정량적 결과를 비교하기 위하여 다음과 같은 두 가지의 평가 방법을 사용하였다. 첫 번째로는 대비 개선 결과 영상의 컬러 왜곡을 측정하는 lightness order error (LOE) [6]를 사용하였고, 두 번째로는 이미지의 전반적인 자연스러움을 평가하는 모델로 성능을 측정하는 naturalness image quality evaluator (NIQE) [27]를 사용하였다. 일반적으로 LOE와 NIQE 모두 값이 낮을수록 개선된 영상이 자연스럽다는 것을 의미한다. 먼저, Table 1은 자기-주목(self-attention) 맵의 성능을 검증하기 위하여 해당 맵의 사용 여부에 따른 성능을 비교하였다. 비교 데이터셋으로는 FiveK [15], LIME [8], NPE [6]를 사용하였다. 여기서 L^{BCP} 는 Bright Channel Prior 손실 함수만을 사용한 결과이고, $L^{BCP}+I_{att}$ 는 손실 함수에 자기-주목 맵을 함께 적용한 결과이다. 굵게 표시된 숫자가 가장 좋은 성능을 나타내며, 표에서 확인할 수 있듯이 자기-주목(self-attention) 맵이 과포화(over-saturation) 현상과 같은 영상의 부자연스러움을 유발하는 요인들을 잘 억제하여 자연스러운 영상을 생성한다는 것을 확인할 수 있다. Table 2는

FiveK [15], LIME [8], NPE [6] 데이터 셋에서의 정량적 비교 결과를 나타낸 것이다. 여기서 4개의 방법 [10-13]은 지도 학습에 기반한 방법들이며, LIME [8]은 수작업 방식의 저조도 개선 방법이다. 굵게, 빨간색, 파란색으로 표시된 결과는 첫 번째, 두 번째, 세 번째로 성능이 좋다는 것을 의미한다. 제안 방법은 NIQE에서는 최근 방법인 DHN [12], UPE [13]과 유사한 성능을 보이며, LOE에서는 제안 방법이 다른 방법들보다 뛰어난 것을 확인할 수 있다.

4.3 정성적 결과

다음으로 제안 방법과 다른 저조도 개선 방법을 시각적으로 비교하였다. Fig. 5는 FiveK [15] 데이터 셋에서의 결과이며, Fig. 6은 LIME [8]과 NPE [6] 데이터 셋에서의 결과이다. Fig. 5와 6에서 볼 수 있듯이, RetinexNet [10]는 영상을 과도하게 개선하는 경향이 있으며, 이로 인하여 매우 부자연스러운 영상을 생성한다. 또한, LightenNet [11]은 조명 왜곡으로 인해 후광 효과와 과포화(over-saturation) 문제가 발생하는 것을 확인할 수 있다. DHN [12]은 가시성은 향상되나 대비의 개선이 잘 이루어지지 않는다. 따라서, 자연스럽고 생생한 컬러를 보여주지 못한다. UPE [13]는 제안 방법보다 더욱 가시성이 높고 자연스러운 영상을 출력하지만, 몇몇 영상에서 여전히 컬러

Table 1. The performance validation of self-attention map. Avg. is the mean value of the results across all datasets

Method	Dataset							
	FiveK [15]		LIME [8]		NPE [6]		Avg.	
	NIQE	LOE	NIQE	LOE	NIQE	LOE	NIQE	LOE
L^{BCP}	4.05	751	4.50	506	5.27	1329	4.22	823
$L^{BCP}+I_{att}$	3.41	417	3.83	321	3.53	332	3.55	404

Table 2. Quantitative evaluation with state-of-the-art methods. The lower is the better

Method	Dataset							
	FiveK [15]		LIME [8]		NPE [6]		Avg.	
	NIQE	LOE	NIQE	LOE	NIQE	LOE	NIQE	LOE
RetinexNet [10]	4.26	2125	4.75	1518	4.39	1675	4.34	1907
LightenNet [11]	3.40	2378	3.49	1914	3.22	1403	3.33	1952
LIME [8]	3.50	1203	4.25	1134	3.53	1241	3.55	1215
DHN [12]	3.29	369	3.59	374	3.16	511	3.25	428
UPE [13]	3.23	1003	3.92	1237	2.98	796	3.16	930
Ours	3.33	228	3.78	209	3.18	327	3.30	305

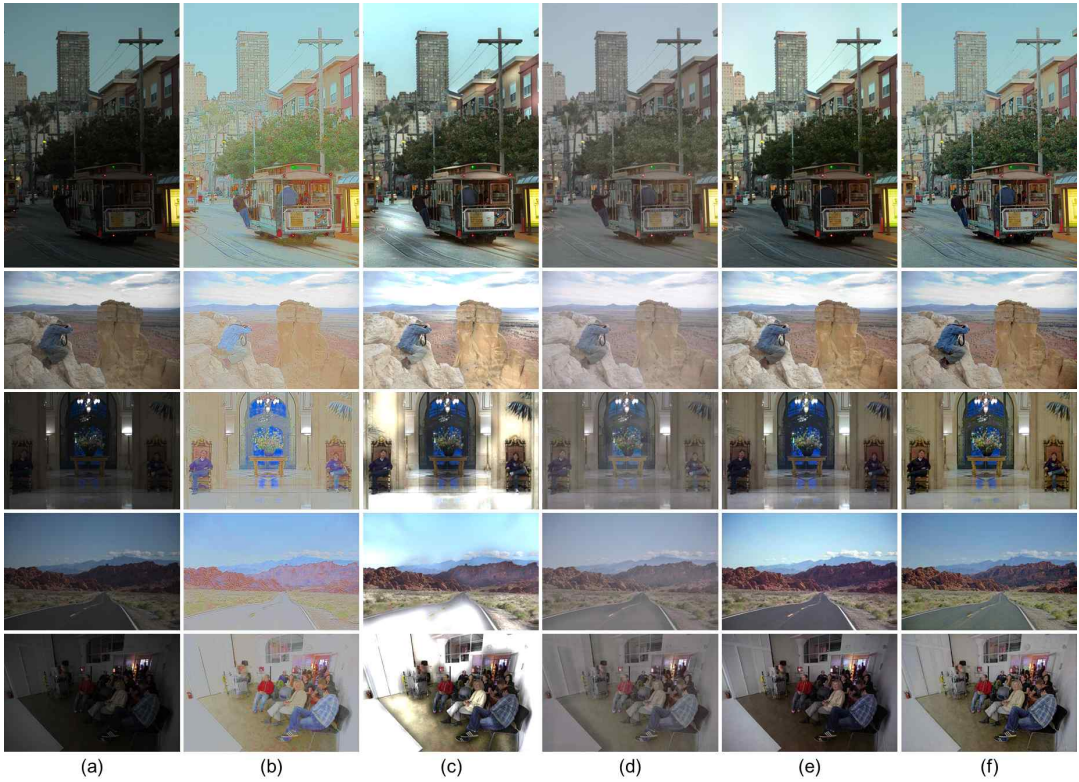


Fig. 5. Qualitative evaluation of FiveK [15] dataset. (a) input image, (b) RetinexNet [10], (c) LightenNet [11], (d) DHN [12], (e) UPE [13], (f) the proposed method.

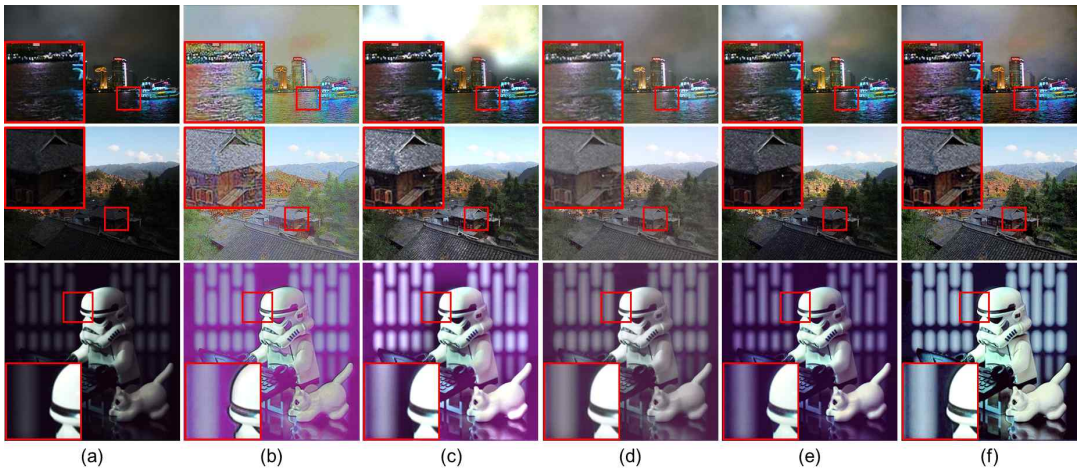


Fig. 6. Qualitative evaluation of LIME [8] and NPE [6] dataset. (a) input image, (b) RetinexNet [10], (c) LightenNet [11], (d) DHN [12], (e) UPE [13], (f) the proposed method.

러 불일치 문제가 존재하는 것을 확인할 수 있다. 예를 들어, Fig. 6의 네 번째 줄의 영상에서 다른 방법들과 마찬가지로 배경의 조명이 검은색에서 보라색으

로 변형되었다. 반면에, 제안한 방법은 저조도 영상의 색감을 비슷하게 유지하고 어두운 영역의 가시성을 향상시키면서 디테일을 보존하여 더욱 자연스러

운 결과를 생성한다. 따라서, 정량적 결과 비교에서도 확인할 수 있었듯이, 제안 방법의 결과가 기존 방법들과 비교하여 더욱 자연스러운 영상을 생성하며 디테일을 보존한 저조도 개선을 이루었다고 볼 수 있다.

5. 결 론

본 논문에서는 저조도 영상을 개선하기 위한 새로운 방법을 제안하였다. 기존 방식들이 실제 저조도 환경에서의 정답 데이터가 부족하다는 이유로 인위적인 데이터 셋을 만들어 학습하던 문제를 해결하기 위하여 BCP를 사용한 비지도 학습을 통해 저조도 개선을 수행하였다. 또한, 자기-주목(self-attention) 맵을 사용하여 과포화(over-saturation) 문제와 같은 현상을 억제함과 동시에 자연스러운 영상을 추정할 수 있었다. 정량적, 정성적 비교를 통하여 제안한 비지도 기반 학습 방식이 최신의 지도 학습 기반 방법들과 비교했을 때 경쟁력 있는 성능을 낸다는 것이 입증되었다. 후속 연구로, 저조도 영상에는 노이즈가 포함되어 있으므로 개선된 영상에서의 증폭된 노이즈를 함께 고려할 수 있는 비지도 기반 네트워크를 연구하고자 한다.

REFERENCE

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-time Object Detection," *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.
- [2] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and L.F. Fei, "Imagenet: A Large-scale Hierarchical Image Database," *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 248-255, 2009.
- [3] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "Deeplab: Semantic Image Segmentation with Deep Convolutional nets, Atrous Convolution, and Fully Connected Crfs," *Transaction on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 4, pp. 834-848, 2017.
- [4] S.M. Pizer, R.E. Johnston, J.P. Ericksen, B.C. Yankaskas, and K.E. Muller, "Contrast-limited Adaptive Histogram Equalization: Speed and Effectiveness," *Proceeding of Conference on Visualization in Biomedical Computing*, pp. 337-345, 1990.
- [5] M.A.A. Wadud, M.H. Kabir, M.A.A. Dewan, and O. Chae, "A Dynamic Histogram Equalization for Image Contrast Enhancement," *Transaction on Consumer Electronics*, Vol. 53, No. 2, pp. 593-600, 2007.
- [6] S. Wang, J. Zheng, H.M. Hu, and B. Li, "Naturalness Preserved Enhancement Algorithm for Non-uniform Illumination Images," *Transaction on Image Processing*, Vol. 22, No. 9, pp. 3538-3548, 2013.
- [7] X. Fu, D. Zeng, Y. Huang, X.P. Zhang, and X. Ding, "A Weighted Variational Model for Simultaneous Reflectance and Illumination Estimation," *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 2782-2790, 2016.
- [8] X. Guo, Y. Li, and H. Ling, "Lime: Low-light Image Enhancement Via Illumination Map Estimation," *Transaction on Image Processing*, Vol. 26, No. 2, pp. 982-993, 2016.
- [9] K.G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement," *Pattern Recognition*, Vol. 61, pp. 650-662, 2017.
- [10] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep Retinex Decomposition for Low-light Enhancement," *Proceeding of British Machine Vision Conference*, pp. 1-12, 2018.
- [11] C. Li, J. Guo, F. Porikli, and Y. Pang, "Lighthnet: A Convolutional Neural Network for Weakly Illuminated Image Enhancement," *Pattern Recognition Letters*, Vol. 104, pp. 15-22, 2018.
- [12] W. Ren, S. Liu, L. Ma, Q. Xu, X. Xu, X. Cao, et al., "Low-light Image Enhancement Via a Deep Hybrid Network," *Transaction on Image*

- Processing*, Vol. 28, No. 9, pp. 4364–4375, 2019.
- [13] R. Wang, Q. Zhang, C.W. Fu, X. Shen, W.S. Zheng, and J. Jia, “Underexposed Photo Enhancement Using Deep Illumination Estimation,” *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 6849–6857, 2019.
- [14] Flickr, <https://www.flickr.com/> (accessed November 22, 2019).
- [15] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, “Learning Photographic Global Tonal Adjustment with a Database of Input/Output Image Pairs,” *Proceeding of Conference on Computer Vision and Pattern Recognition*, pp. 97–104, 2011.
- [16] K. He, J. Sun, and X. Tang, “Single Image Haze Removal Using Dark Channel Prior,” *Transaction on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 12, pp. 2341–2353, 2010.
- [17] A. Golts, D. Freedman, and M. Elad, “Unsupervised Single Image Dehazing Using Dark Channel Prior Loss,” *arXiv Preprint ArXiv:1812.07051*, 2018.
- [18] X. Zhang, P. Shen, L. Luo, L. Zhang, and J. Song, “Enhancement and Noise Reduction of Very Low Light Level Images,” *Proceeding of International Conference on Pattern Recognition*, pp. 2034–2037, 2012.
- [19] L. Li, R. Wang, W. Wang, and W. Gao, “A Low-light Image Enhancement Method for Both Denoising and Contrast Enlarging,” *Proceeding of International Conference on Image Processing*, pp. 3730–3734, 2015.
- [20] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital Photography with Flash and No-flash Image Pairs,” *Transaction on Graphics*, Vol. 23, No. 3, pp. 664–672, 2004.
- [21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image Denoising by Sparse 3-D Transform-domain Collaborative Filtering,” *Transaction on Image Processing*, Vol. 16, No. 8, pp. 2080–2095, 2007.
- [22] M. Gharbi, J. Chen, J.T. Barron, S.W. Hasinoff, and F. Durand, “Deep Bilateral Learning for Real-time Image Enhancement,” *Transaction on Graphics*, Vol. 36, No. 4, pp. 118, 2017.
- [23] Y. Wang, S. Zhuo, D. Tao, J. Bu, and N. Li, “Automatic Local Exposure Correction Using Bright Channel Prior for Under-exposed Images,” *Signal Processing*, Vol. 93, No. 11, pp. 3227–3238, 2013.
- [24] A. Levin, D. Lischinski, and Y. Weiss, “A Closed-form Solution to Natural Image Matting,” *Transaction on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 2, pp. 228–242, 2008.
- [25] J. Cai, S. Gu, and L. Zhang, “Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images,” *Transaction on Image Processing*, Vol. 27, No. 4, pp. 2049–2062, 2018.
- [26] D.P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *Proceeding of International Conference on Learning Representations*, 2015.
- [27] A. Mittal, R. Soundararajan, and A.C. Bovik, “Making a “Completely Blind” Image Quality Analyzer,” *Signal Processing Letters*, Vol. 20, No. 3, pp. 209–212, 2012.
- [28] S. Jung, J. Kim, D. Lee, S. Lee, and G. Kim, “Intelligent Video Event Detection System Used by Image Object Identification Technique,” *Journal of Korea Multimedia Society*, Vol. 13, No. 2, pp. 171–178, 2010.



이 현 상

2016년 충남대학교 컴퓨터공학과
학사 졸업
2018년 충남대학교 컴퓨터공학과
석사 졸업
2018년~현재 연세대학교 전기전
자공학부 박사 과정

관심분야: 컴퓨터비전, 영상처리, 딥러닝



민 동 보

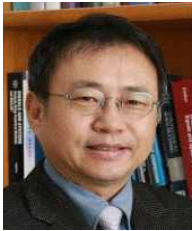
2003년 연세대학교 전기전자공학
부 학사 졸업
2005년 연세대학교 전기전자공학
부 석사 졸업
2009년 연세대학교 전기전자공학
부 박사 졸업

2009년~2010년 Mitsubishi Electric Research Lab
(미국) 박사후연구원

2010년~2015년 Advanced Digital Sciences Center
(싱가포르) 연구과학자

2015년~2018년 충남대학교 컴퓨터공학과 조교수
2018년~현재 이화여자대학교 컴퓨터공학과 조교수

관심분야: 컴퓨터비전, 영상처리, 딥러닝



손 광 훈

1983년 연세대학교 전자공학과
학사
1985년 미국 미네소타대학교 전
자공학과 석사
1991년 미국 노스캐롤라이나 주
립대학교 전기·컴퓨터공
학과 박사

1988년~1992년 North Carolina A&T State University,
Image Proc. & Computer Vision Lab., Research
Associate

1992년~1993년 한국전자통신연구소 위성통신기술연구
단, 선임연구원

1994년 Georgetown University MRI Center, Post
Doctoral Fellow

1995년~현재 연세대학교 전기전자공학부 교수

2002년~2003년 Nanyang Technological University,
School of Computer Engineering, Visiting
Professor

2009년~2010년 국방과학연구소 전문위원

관심분야: 3차원 영상 신호 처리, 컴퓨터 비전