

일반논문 (Regular Paper)

방송공학회논문지 제24권 제1호, 2019년 1월 (JBE Vol. 24, No. 1, January 2019)

<https://doi.org/10.5909/JBE.2019.24.1.87>

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

심층 신경망을 이용한 저조도 영상에서 Retinex 기반 반사 영상 생성

김원회^{a)}, 황인철^{b)}, 김만배^{a)†}

Generating a Retinex-based Reflectance Image from a Low-Light Image Using Deep Neural Network

Wonhoi Kim^{a)}, In-Chul Hwang^{b)}, and Manbae Kim^{a)†}

요약

저조도 영상의 개선에 관한 연구는 대부분 대비 개선을 목적으로 한다. 저조도 영상에서 밝기 개선, 대조 개선, 및 조명 성분 감쇠 등의 다양한 연구가 진행됐다. 최근에 인공신경망으로 상기 방법들을 대체하는 연구가 진행 중이다. 본 논문에서는 Retinex 이론에 기반하여 조명 광원이 존재하는 저조도 영상으로부터 조명 성분을 감쇠하고, 반사 성분만을 생성하는 기법을 심층신경망으로 대체하는 방법을 제안한다. 실험에서는 102장의 저조도 영상으로 학습시킨 인공신경망으로 반사 영상을 생성하였는데, PSNR=30.8682(db), SSIM=0.4345를 얻었다.

Abstract

Improvement of low-light image mainly focuses on the contrast enhancement. Many researches have been carried out for brightness enhancement, contrast improvement and illumination reduction. Recently, the aforementioned approaches have been replaced by artificial neural networks. This paper proposes a methodology that can replace the Retinex-based reflectance image acquisition by deep neural network. Experiments carried out on 102 low-light images validated the feasibility of the replacement by producing PSNR=30.8682(db) and SSIM=0.4345.

Keywords: Low-light image, Retinex, Deep neural network, Reflectance, Illumination

a) 강원대학교 컴퓨터정보통신공학부(Department of Computer and Communications Eng., Kangwon National University)

b) 강원대학교 전기전자공학부(Department of Electrical and Electronics Eng., Kangwon National University)

† Corresponding Author : 김만배(Manbae Kim)

E-mail: manbae@kangwon.ac.kr

Tel: +82-33-250-6395

ORCID: <https://orcid.org/0000-0002-4702-8276>

※ 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터지원사업의 연구결과 (IITP-2018-0-01433) 및 이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2017R1D1A3 B03028806).

※ This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2018-0-01433) supervised by the IITP(Institute for Information & communications Technology Promotion) and This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1D1A3B03028806).

· Manuscript received September 2, 2018; Revised October 26, 2018; Accepted October 26, 2018.

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

고해상도 영상장비 시장에서 요구되는 장치들의 기대 수준은 빠른 속도로 성장해 왔다. 현재의 영상장비들은 물체 감지를 통한 영상 보안 및 인식의 영역까지 확대되었다. 상기 영역을 넘어선 물체 인식 기능들을 효과적으로 수행 할 수 있는 많은 연구가 진행되고 있다.

이와 같은 연구의 효율성을 증대하기 위해서는 다양한 외부 환경을 자율적으로 제어할 수 있는 기능을 필요로 한다. 다양한 외부 환경 중 광원은 영상 화질에 큰 영향을 주게 된다. 광원량이 과다하거나 부족한 경우 영상 표현 능력을 저하시키게 되어 영상에 대한 판별 능력을 어렵게 한다. 또한, 어두운 환경에서 존재하는 광원의 영향은 전체 영상의 인지에 큰 영향을 준다. 상기 영상의 품질 저하를 보완하기 위해, 히스토그램 평활화(histogram equalization)^[1-3], 감마보정(gamma correction)^[4] 등의 기능들이 사용되어 왔다.

Land 등은 실제 인간의 시각에 들어오는 빛은 물체의 반사를 통해서 들어오는 반사(reflectance) 성분과 광원을 통하여 들어오는 조명(illumination) 성분으로 나눌 수 있음을 실험적으로 입증하였다^[5]. Jobson 등은 Land의 시각모델과 인간의 시각모델인 Werber-Fachner Law를 바탕으로 Single Scale Retinex(SSR) 알고리즘을 제안하였다^[6]. 상기 방법은 영상의 조명 성분을 최대한 배제하여 물체가 가지는 고유한 색을 유도할 수 있다. SSR(Single Scale Retinex) 방식은 적절한 크기의 저주파 필터로 필터링된 결과를 영상의 조명 성분으로 추정하며, 조명성분을 감쇠함으로써 반사성분을 얻는다.

위에서 언급한 것처럼 어두운 환경에서 촬영한 저조도(low-light) 영상은 주변에 존재하는 조명에 의해서 일부 영역은 빛에 의해 높은 밝기를 가지는 반면에, 일부 영역은

제한된 밝기 영역으로 인해서, 어둡게 보이는 현상이 종종 발생한다. 그림 1의 네 장의 영상을 관찰하면, 주변의 조명 또는 모니터의 밝기 때문에, 일부 영역은 원래 밝기의 대조를 손실하고, 낮은 대조로 물체의 구조를 시각적으로 이해하는 것이 어렵다. 이러한 조명 광원의 영향을 줄이고, 반사성분을 얻는 것이 Retinex 이론이다.

최근에 기존의 hand-crafted 특징을 사용하지 않고, 심층신경망(Deep Neural Network: DNN)을 이용하여 저조도 영상의 화질을 개선하는 방법들이 소개되고 있다^[7-9]. DNN 기반 방법은 다른 신경망 모델과 마찬가지로 학습 데이터를 활용한다. 입력영상과 실측(ground-truth) 영상을 학습하면서 실측 영상과 유사한 영상을 생성할 수 있도록 신경망을 모델링한다. Lore 등은 저조도 영상에 stacked denoising 자기부호화기(autoencoder)를 사용하였다^[7]. Shen 등은 합성곱 신경망(Convolutional Neural Network)을 적용하여 저조도 영상을 향상하였다^[8]. 또한 Park 등은 자기부호화기 및 CNN의 듀얼 신경망 모델을 제안하였다^[9]. 저조도 영상의 밝기 향상 기법들은 대부분 고조도(high-light) 영상에 수작업으로 감마(gamma)변환을 적용하여 어둡게 만든 후에, 신경망으로 학습하는 과정을 거친다. 이 방법은 실험에서 사용한 감마변환에 잘 맞지 않으면, 일반 영상은 현실적으로 성능이 저하되는 단점이 발생한다.

본 연구는 기존의 저조도 영상의 향상 알고리즘을 심층신경망으로 대체하는 새로운 개념의 이론을 제안하고 타당성(feasibility)을 검증한다. 이를 위해서 먼저 Retinex 이론 및 특정 반사영상 생성 알고리즘을 소개한다. 이 알고리즘으로 얻은 입력영상과 출력영상을 DNN으로 학습을 하게 한 후에, DNN의 출력과 알고리즘의 결과영상을 비교하면서 성능을 조사한다.

그림 2는 기존 저조도 영상 향상 방법과 제안방법의 개념

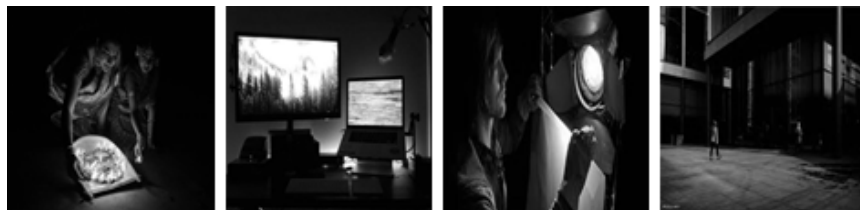


그림 1. 광원 조명의 영향을 받는 저조도 영상
Fig. 1. Low-light images affected by light source

의 차이를 보여준다. 기존 생성방법은 그림 2(a)에서 보듯이, 다량의 학습영상과 실측 타겟영상으로 신경망을 학습시킨 후에, 예측(predicted)영상을 얻는 방식이다. 실험에서는 타겟영상이 주어지면, 감마변환으로 영상을 어둡게 한다. 어두운 영상을 입력으로, 밝은 영상을 출력으로 하여 신경망 모델을 학습시킨다. 이와는 달리 제안방법은 그림 2(b)에서 보는 것처럼 기계작된 방법으로 반사영상을 생성한 후에, DNN으로 입력영상과 출력영상을 학습한다. 최종적으로 DNN 모델을 기계작 반사영상 생성 알고리즘을 대체하고자 하는 것이다.

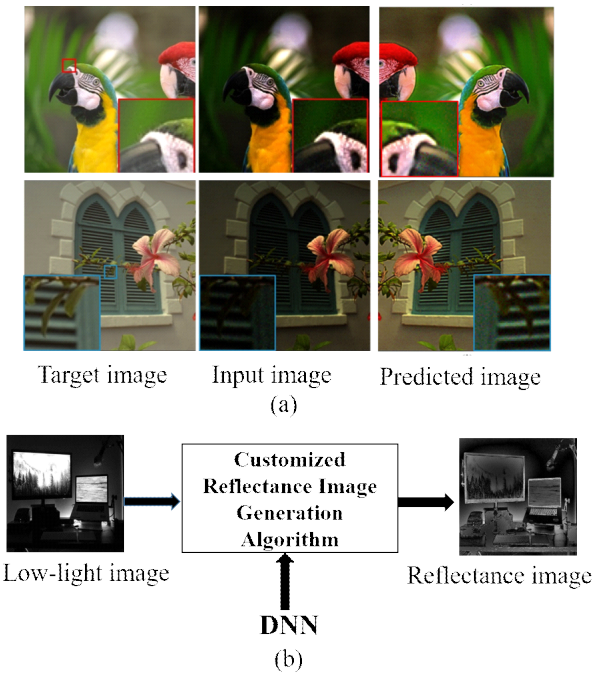


그림 2. 기존 저조도 영상 향상과 제안 방법의 차이. (a) 일반적인 DNN 기반 영상 생성[5] 및 (b) 기계작된 반사 영상 생성 알고리즘을 DNN으로 대체하는 구조도

Fig. 2. Difference of conventional reflectance image methods and the proposed method. (a) The approach of the former and (b) diagram replacing a customized reflectance map generation algorithm with DNN

저조도 영상 향상 또는 반사성분 추출 방법의 성능은 본 연구의 목적이 아니고, 성능에 관계없이 어떠한 반사 성분 생성 알고리즘이라도 DNN으로 대체가 가능한지에 대한 타당성을 검증하는 것이 주 목적이다. 이를 위해 기존에 소개된 반사성분 추출 알고리즘을 활용하여 입력영상으로부

터 반사영상을 생성하고, DNN을 학습시킨다. DNN이 예측한 출력 반사영상과 기존 알고리즘의 출력 반사영상의 오차를 검증한다. 더욱이 높은 복잡도를 가지는 반사영상 생성 기법들을 낮은 복잡도를 가지는 DNN으로 대체할 수 있으면, 상당한 활용이 기대될 수 있다. 결과적으로 어떠한 반사영상 생성 기법을 연구하더라도 이를 대체할 수 있는 장점이 있다.

본 논문의 구성은 다음과 같다. 다음 장에서는 Retinex 이론을 설명하고 반사영상을 생성하는 과정을 소개한다. III 장에서는 제안방법의 전체 흐름도를 설명한다. IV 장에서는 실험 결과를 분석한다. 마지막으로 V 장에서는 결론 및 향후 연구를 정리한다.

II. Retinex 이론

Land 등은 실험적으로 인간 시각 기관이 인지하는 물체의 색이 광원과 물체의 반사성분의 곱으로 나타낼 수 있음을 입증하였으며^[5], 수식으로 표현하면 밝기(brightness) 영상 $f(x,y)$ 는 두 개의 성분인 광원 조명(illumination) 성분 $i(x,y)$ 와 반사(reflectance) 성분 $r(x,y)$ 의 곱으로 다음과 같다.

$$f(x,y) = i(x,y) \cdot r(x,y) \quad (1)$$

여기서 (x,y) 는 픽셀의 좌표이다.

상기 식은 조명성분을 추정할 수 있다면 산술적인 방법을 통하여 물체 고유의 색인 반사성분을 알아낼 수 있음을 의미한다. 조명성분은 천천히 변화하는 저주파 성분을 가지고 있고, 반사성분은 고주파 성분을 가지고 있다. 따라서 적절한 필터링을 활용하면, 조명성분의 영향을 줄일 수 있다.

Weber-Fechner 법칙에 따르면 인간의 감각기관이 인지하는 감각의 차이들이 로그(log)스케일을 갖고 있는데, i 와 r 를 분리하기 위해서 식 (1)에 로그 변환을 적용하면 다음과 같이 얻어진다.

$$z(x,y) = \log[f(x,y)] = \log[i(x,y)] + \log[r(x,y)] \quad (2)$$

로그 영상 $z(x,y)$ 에 푸리에 변환 F 을 적용하면 다음과 같다.

$$Z(u,v) = I(u,v) + R(u,v) \quad (3)$$

여기서 $I(u,v) = F \log [i(x,y)]$, $R(u,v) = F \log [r(x,y)]$, 및 $Z(u,v) = F[z(x,y)]$ 이다.

식 (3)에 고역통과필터 $H(u,v)$ 을 곱해준다. I 는 저주파 성분이므로 $H \cdot I \approx 0$ 이다.

$$H \cdot Z = H \cdot I + H \cdot R \approx H \cdot R \quad (4)$$

고역통과필터는 저주파를 감쇠하고 고주파 성분을 통과시킨다. 본 실험에서는 감쇠 계수 0.5인 버터워스(Butterworth) 필터를 이용하였고, 식은 식(5)에서 보여진다.

$$H_{high} = 1 - \frac{1}{1 + \left(\frac{D(u,v)}{D_0}\right)^n} \quad (5)$$

여기서 $D(u,v) = D(r) = \sqrt{u^2 + v^2}$, D_0 는 차단주파수이다.

$H \cdot R$ 에 역푸리에 변환 F^{-1} 을 사용하여 공간 좌표의 반사 영상, $r_R(x,y)$ 을 복원한다.

$$r_R(x,y) = F^{-1}[H(u,v) \cdot R(u,v)] \quad (6)$$

그림 3은 상기 Retinex 이론에 기반하여 생성된 반사영상을 보여준다. 그림 3(a)은 입력영상 f 을 보여주는데 조명성분이 뚜렷이 관측되고 있다. 이 조명성분을 감쇠시킨 반사영상 r_R 은 그림 3(b)에서 보여진다. 조명성분이 감쇠된 반사성분을 관측할 수 있다. 신경망 학습과정에서 그림 3(a)은 신경망 네트워크의 입력데이터이고, 그림 3(b)은 실측 출력영상으로 사용된다. 신경망 모델을 학습한 후에, 새로운 저저도 영상이 입력되면 신경망은 예측된 반사영상을 출력하게 된다.

III. 제안 방법

제안 방법은 그림 4에서 보여진다. 먼저 입력 그레이스케일 저조도 영상에 저주파 필터를 적용한다. 입력영상과 필터링된 영상의 차영상을 저장한다. 필터링된 영상은 심층 신경망의 입력 데이터이다. 타겟 반사영상에도 저주파필터를 적용시켜 필터링된 영상을 만든다. 저주파 필터로는 가우시안 필터를 이용하였고, 필터의 크기는 3x3이며 $\sigma = 3$ 이다. 저주파 필터가 적용된 영상들은 DNN에서 입력과 출력으로 사용된다. 즉, 저주파 필터를 적용시킨 영상들로 DNN을 학습시키면 예측된 반사 영상이 생성된다.

신경망 학습에서는 입력영상과 출력 타겟영상을 사용하는데, 완벽하게 동일한 출력영상을 생성하는 것은 어렵다. 일반적으로 예측영상들은 부드러워진, 즉 저주파 필터가 적용된 것과 유사하다. 예측영상은 부드럽게 생성되기 때



그림 3. 결과 그림. (a) 입력영상, f 및 (b) 반사 영상, r_R
 Fig. 3. Resulting images. (a) input images, f and (b) reflectance images, r_R

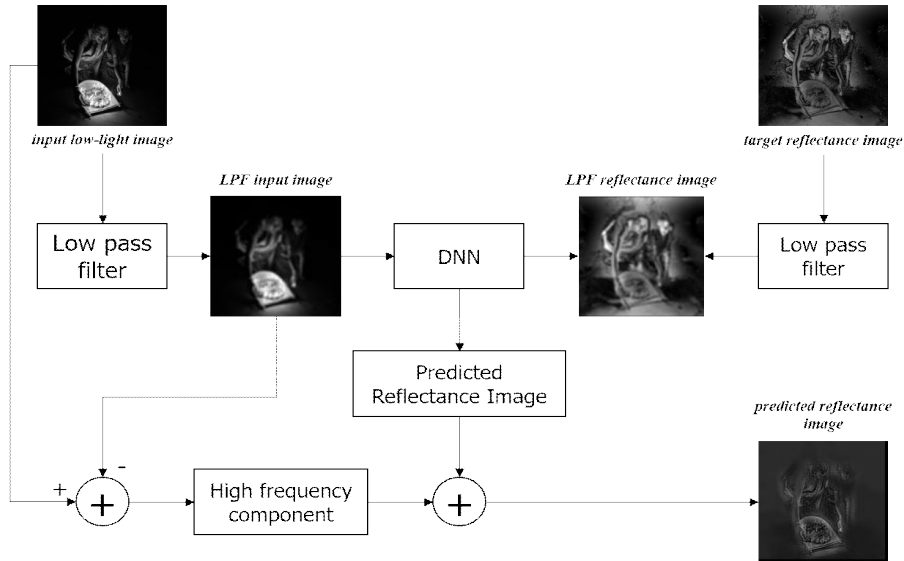


그림 4. 제안 방법의 전체 흐름도
 Fig. 4. The overall block diagram of the proposed method

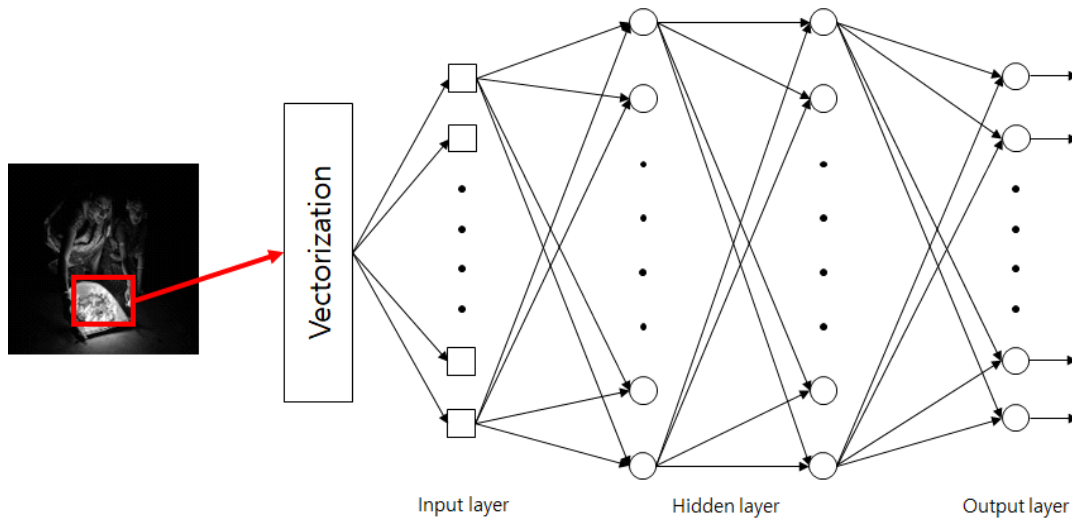


그림 5. 심층신경망의 네트워크 구조
 Fig. 5. The network structure of the deep neural network

문에 신경망 입출력에 저주파 필터를 적용한다. 그러므로 고주파 성분이 감쇄된 영상을 입출력으로 사용함으로써, 오차를 줄일 수 있다. 또한 고주파 성분은 경계 등의 시각적 성분을 포함하므로, 고주파 성분을 가지고 있는 차영상은 출력에서 생성된 예측 반사 영상에 합해지게 된다. 그림 5의 심층신경망은 입력층과 출력층 그리고 여러 개의 은닉층(hidden layer)들로 구성된다. 본 논문에서 사용된 신경망

은 일반적인 구조를 가지며 역전파 알고리즘을 통해 심층신경망의 학습을 수행한다. 신경망은 입력 저조도 영상으로부터 조명성분을 추산하기 때문에 저조도와 고조도 영상의 대조값을 얻기 어렵다. 그러므로 이 논문에서는 입력층에는 영상 패치(patch)를 이용하여 학습 데이터를 생성한다. $N \times N$ 패치는 $N^2 \times 1$ 벡터로 변환하므로, 입력층에서 노드의 개수는 N^2 이다. 그리고 출력층에는 입력층과 같은

노드 개수 N^2 가 있다. 전체적인 심층신경망의 구조는 그림 5와 같다. 은닉층의 개수는 1~3개로 설정한다.

IV. 실험 결과

실험에서 사용한 저조도 실험영상은 그림 6에서 보여진다. 인터넷에서 수집한 영상들의 일부를 보여준다. 영상들은 어두운 환경에서 광원의 영향을 받고 있어서, 제안 방법이 효율적으로 적용될 수 있다. 광원의 영향을 받는 곳은 밝고, 다른 영역은 상대적으로 대조가 낮아 내부의 구조를 자세히 알 수 없다. 영상 해상도는 128x128이고 신경망 학습과 테스트는 운영체제 WINDOWS 10에서 MATLAB R2017a를 사용하였다. 학습데이터로 사용한 이미지는 총 102장이다. 102장에서 82장은 학습에 사용하고, 나머지 20장은 테스트에 사용하여 성능을 검증한다. 20장은 랜덤하게 선택하였다.

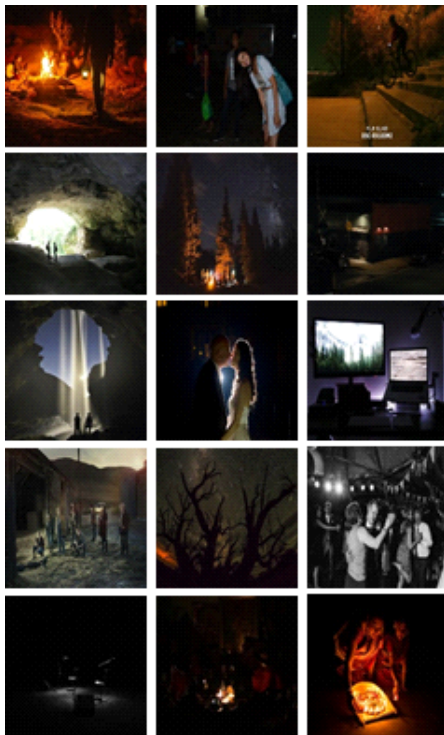


그림 6. 실험에 사용된 광원이 존재하는 저조도 영상
Fig. 6. Low-light images under illumination used in experiments

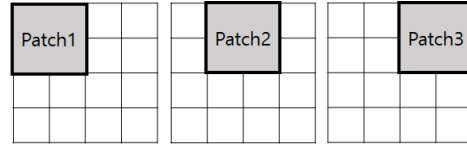


그림 7. 스트라이드=1로 중첩 패치 생성
Fig. 7. Generating overlapped patches with stride=1

영상 데이터는 [0,1]로 정규화한다. 학습데이터는 패치 단위로 처리된다. 이미지를 학습하기 위해 한 장의 사진이 들어올 때마다 중첩된 패치를 생성한다. 패치 생성 방법은 두 가지로 구분되는데, 첫 번째는 영상에서 패치를 랜덤하게 생성한 후에, 합성은 중첩된 패치들의 평균값으로 영상을 복원한다⁷⁾. 다른 방법은 그림 7처럼, 스트라이드(stride) 값을 조절하면서 중첩된 패치를 생성하고, 복원은 첫 번째와 같은 방법을 사용한다^{8,9)}. 실험에서는 두 번째 방법을 선택한다. 패치의 스트라이드가 3이고 크기가 16x16이면 해상도 128x128에서 패치의 개수는 1,444개이다. 패치의 크기는 8x8과 16x16 두 가지로 실험하였다.

신경망의 학습률 α 는 0.001로 설정하였고 신경망 은닉층의 개수와 은닉 노드의 개수를 조절하며 실험하였다. 1개의 은닉층일 때 은닉층 노드의 개수는 128, 2~3개의 은닉층에는 모두 64개로 설정하였다. 마지막으로 출력층 노드의 개수는 입력층 노드의 개수와 같다. 모든 층은 시그모이드(sigmoid) 활성화함수를 사용하였다. 신경망의 가중치 초기화는 Xavier 초기화¹⁰⁾을 사용하고, 가중치 갱신 방식은 확률적 경사 하강법(stochastic gradient descent)을 사용한다. 최적화 기법은 Adam optimization¹¹⁾을 사용하였다. ρ_1 과 ρ_2 는 각각 0.9, 0.999이다. 그리고 δ 는 10^{-8} 이다.

신경망 학습의 에폭(epoch)은 500으로 고정하였으며 에폭에 따른 픽셀의 로스(loss)값은 그림 8의 그래프와 같다. y축은 로스를 나타내며 x축은 에폭을 나타낸다. 그림 8(a)와 (b)는 패치의 크기가 8x8이며 각 각의 은닉층 개수는 2개와 3개이다. 8x8 패치의 크기를 가진 신경망의 로스값은 0.4~0.8사이의 값을 가지며 에폭이 낮을 때 0.8에 가까운 값을 보여주지만 에폭이 높아질수록 0.4에 수렴하고 있는 것을 볼 수 있다. 그림 8(c)와 (d)는 패치의 크기가 16x16이며 은닉층의 개수는 각각 2개와 3개이다. 16x16 패치의 크기를 가진 신경망의 로스값은 1.2~3.2사이의 값을 가지며

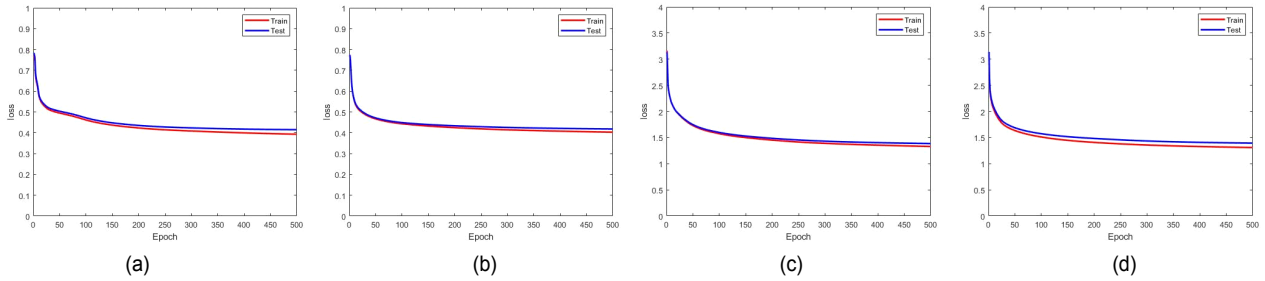


그림 8. 패치의 크기와 은닉층의 개수에 따른 로스의 그래프. (a) 8x8 patch 및 은닉층 개수 2개, (b) 8x8 patch 및 은닉층 개수 3개, (c) 16x16 patch 및 은닉층 개수 2개 및 (d) 16x16 patch 및 은닉층 개수 3개.

Fig. 8. Graph of loss with respect to the size of patches and the number of hidden layers. (a) 8x8 patch and two hidden layers, (b) 8x8 patch and three hidden layers, (c) 16x16 patch and two hidden layers, and (d) 16x16 patch and three hidden layers.

에폭이 낮을 때 3.2에 가까운 값을 보여주며 에폭이 높아질수록 1.2에 수렴하고 있다. 그림 (a)와 (b)보다 (c)와 (d)의 로스값의 변화폭이 더 크다는 것을 볼 수 있다.

주관적 평가를 위해서 그림 9는 입력영상, 실측 출력



그림 9. (a) 입력 저조도 영상, (b) 출력 타겟 영상, 및 (c) 예측 영상
 Fig. 9. (a) input low-light images, (b) output target images, and (c) predicted images

영상, 및 예측영상을 보여준다. 그림 9(a)는 그림 4에서 *input low-light image*에 해당하는 영상으로 저역통과 필터를 적용하기 전인 그레이스케일 영상이다. 조명으로 인한 밝은 부분과 어두운 부분이 대조적으로 나타나는 영상으로 이루어져 있다. 그림 9(b)는 그림 4에서 *target reflectance image*에 해당하는 영상이다. 이 영상은 Retinex 이론에 기반하여 생성된 영상으로 신경망의 타겟영상에 해당된다. 그림 9(c)는 그림 4에서 *predicted reflectance image*이다. 신경망에서 학습되어 출력된 영상에서 *input low-light image*와 *LPF input image*의 차성분을 더해서 만들어진 최종 출력 영상이다.

제한한 심층신경망의 객관적 성능을 검증하기 위해서 저조도영상 향상 성능 검증에서 사용하는 다음 3가지의 객관적 메트릭을 계산한다; 1) Root Mean Squared Error (RMSE), 2) Peak Signal-to-Noise Ratio (PSNR), 3) Structural Similarity Index (SSIM). 아래 식에서 y_i 는 예측된 값이고, d_i 는 실측 픽셀값이다. 신경망에서는 y, d 모두 $[0, 1]$ 의 정규화 값을 가지므로, 255를 곱하여 $[0, 255]$ 로 계산한다.

1) RMSE

$$SE = \sqrt{\frac{1}{K} \sum_{i=1}^N (y_i - d_i)^2} \quad (7)$$

K는 영상 픽셀 개수이다.

2) PSNR

$$PSNR = 10 \log_{10} \frac{255^2}{(y_i - d_i)^2} \text{ (dB)} \quad (8)$$

3) Structural Similarity Index (SSIM)^[12]

SSIM은 전체적인 관점에서 영상 안에서의 구조적 정보인 휘도(luminance), 명암비(contrast), 구조(structure)들을 추출하여 구조적 유사도를 구한 다음 영상의 품질을 측정할 때 사용된다. 1에 가까울수록 유사도가 높으며 반대로 0에 가까울수록 유사도가 낮다.

$$SSIM(d, y) = \frac{(2\mu_d\mu_y + C_1)(2\sigma_{dy} + C_2)}{(\mu_d^2 + \mu_y^2 + C_1)(\sigma_d^2 + \sigma_y^2 + C_2)} \quad (9)$$

μ_d 는 d 의 평균값, μ_y 는 y 의 평균값, σ_{dy} 는 x 와 y 의 공분산이다.

평균 RMSE, PSNR 및 SSIM의 성능은 표 1에서 보여진

다. 표 1은 그림 4에서 *LPF input image*와 *LPF reflectance image*의 성능 결과를 보여준다. 8x8, 16x16 두 개의 패치가 있다. 각각 은닉층의 1~3개이고, 은닉층의 노드 개수는 각각, 128, 64-64, 64-64-64이다. RMSE는 16x16 패치가 8x8 패치보다 작고 PSNR는 패치의 크기와 상관없이 비슷한 값을 보여준다. 그리고 SSIM은 16x16 패치가 8x8 패치보다 작다.

표 2는 그림 4에서 *target reflectance image*와 *predicted reflectance image*의 성능 결과를 보여준다. 실험조건은 표 1과 같다. RMSE는 8x8 패치가 16x16 패치보다 작고 특히 학습보다 테스트에서 차이값이 크게 나타난다. PSNR은 패치의 크기와 상관없이 비슷한 값을 보여주며 SSIM은 패치가 16x16 patch보다 작다.

표 2의 PSNR, SSIM값은 제안방법과 비교할 수 있는 논문이 없기 때문에, 간접적으로 [9]의 결과를 참고하여 설명한다. [9]에서 PSNR의 범위는 13~16 dB이고, SSIM의 범위는 0.38~0.65이다. 제안방법의 PSNR은 [Train, Test]=[30.8682, 31.0892]이고, SSIM은 [0.4345, 0.3684]이다.

표 1. *LPF input image*와 *LPF reflectance image*의 RMSE, PSNR, 및 SSIM의 객관적 성능 평가. Bold는 가장 좋은 성능을 표기함
Table 1. Performance evaluation of RMSE, PSNR and SSIM of *LPF input image* and *LPF reflectance image*. Bold numbers represent the best scores

| Patch size | No. of hidden layers | No. of nodes | Avg. RMSE (in pixels) | | Avg. PSNR (in dB) | | Avg. SSIM [0,1] | |
|------------|----------------------|--------------|-----------------------|--------|-------------------|---------|-----------------|--------|
| | | | Train | Test | Train | Test | Train | Test |
| 8x8 | 1 | 128 | 7.8659 | 7.7851 | 31.6261 | 32.1545 | 0.4504 | 0.4183 |
| | 2 | 64-64 | 7.6329 | 7.6607 | 31.8720 | 32.3097 | 0.4273 | 0.4363 |
| | 3 | 64-64-64 | 7.5511 | 7.6151 | 31.9500 | 32.4587 | 0.4081 | 0.4090 |
| 16x16 | 1 | 128 | 7.5748 | 7.5536 | 31.7840 | 32.1449 | 0.4293 | 0.4307 |
| | 2 | 64-64 | 7.4680 | 7.5314 | 31.8653 | 32.1130 | 0.4308 | 0.4273 |
| | 3 | 64-64-64 | 7.4830 | 7.5744 | 31.7973 | 31.9753 | 0.3858 | 0.3848 |

표 2. *target reflectance image*와 *predicted reflectance image*의 RMSE, PSNR, 및 SSIM의 객관적 성능 평가. Bold는 가장 좋은 성능을 표기함
Table 2. Performance evaluation of RMSE, PSNR and SSIM of *target reflectance image* and *predicted reflectance image*. Bold numbers represent the best scores

| Patch size | No. of hidden layers | No. of nodes | Avg. RMSE (in pixels) | | Avg. PSNR (in dB) | | Avg. SSIM [0,1] | |
|------------|----------------------|--------------|-----------------------|--------|-------------------|---------|-----------------|--------|
| | | | Train | Test | Train | Test | Train | Test |
| 8x8 | 1 | 128 | 8.6450 | 8.5357 | 30.2463 | 30.4431 | 0.3874 | 0.3331 |
| | 2 | 64-64 | 8.4554 | 8.4619 | 30.3840 | 30.5047 | 0.4083 | 0.3496 |
| | 3 | 64-64-64 | 8.3979 | 8.4145 | 30.4659 | 30.5955 | 0.3889 | 0.3323 |
| 16x16 | 1 | 128 | 8.0506 | 7.9831 | 30.8173 | 31.0892 | 0.4301 | 0.3672 |
| | 2 | 64-64 | 7.9783 | 7.9654 | 30.8682 | 31.0647 | 0.4345 | 0.3684 |
| | 3 | 64-64-64 | 7.9794 | 7.9915 | 30.8613 | 31.0191 | 0.4044 | 0.3440 |

[9]의 결과와 비교하면 제안방법이 만족스러운 성능을 보여준다고 판단할 수 있다. 16x16 패치에서 2개의 은닉층, 64-64개의 은닉층의 노드 수에서 RMSE=7.9654, PSNR=31.0647, SSIM=0.3684로 다른 모델보다 좋은 성능을 보여주고 있다.

V. 결론

저조도 영상에서는 빛의 밝기가 부족하여 영상에서 객체를 인식하거나 추적, 감지 등이 힘들어진다. 이를 극복하기 위한 기본 방법은 Retinex 이론을 이용한 조명성분의 감쇠 및 반사성분의 생성으로 다양한 기법들이 제안되었다. 본 논문에서는 이러한 Retinex 이론에 기반한 반사성분 생성을 심층신경망으로 대체하는 방법을 제안하였고, 가능성을 조사하였다. 실험결과에 따르면 RMSE는 7.9654, PSNR은 31.0647(dB), SSIM은 0.3684로 성능 측면에서는 만족스러운 결과를 얻을 수 있었다. 본 연구는 기존의 hand-crafted 방법을 신경망으로 대체하는 예를 보였주었으며, 향후 다양한 분야에서 활용이 가능하다.

참고 문헌 (References)

[1] P. Trahaninas and A. Venetsanopoulos, "Color image enhancement

through 3-d histogram equalization," *11th IAPR international Conference*, pp. 545-548, 1992.

[2] H. Cheng and X. Shi, "A simple and effective histogram equalization approach to image enhancement," *Digital Signal Processing*, pp. 158-170, 2004.

[3] S. Pizer, E. Amburn, J. Austin, and R. Cromartie, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, Vol. 39, pp. 355-368, Jan. 1987.

[4] C. Poynton, "The Rehabilitation of Gamma," *Human Vision and Electronic Imaging III, Proceedings of SPIE*, vol. 3299, p. 232-249, 1998.

[5] E. Land and J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1 - 11, 1971.

[6] D. Jobson, Z. Rahman, and G. Woodell, "Properties and Performance of a Center/Surround Retinex," *IEEE Transactions on Image Processing*, Vol. 6, no. 3, pp. 451-462, Mar. 1997.

[7] K. 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.

[8] L. Shen, Z. Yue, F. Feng, and Q. Chen, "MSR-net: Low-light image enhancement using deep convolutional network," *arXiv:1711.02488*.

[9] S. Park, S. Yu, M. Kim, K. Park, and J. Paik, "Dual Autoencoder Network for Retinex-based Low-Light Image Enhancement", *IEEE Access*, Vol. 6, Mar. 2018.

[10] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep forward neural networks", *Int' Conf. Artificial Intelligence and Statistics, Society for Artificial Intelligence and Statistics*, 2010.

[11] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*

[12] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity". *IEEE Transactions on Image Processing*, 13(4), 600-612, 2004.

저자 소개



김원희

- 2017년 2월 : 강원대학교 컴퓨터정보통신공학과 학사
- 2017년 3월 ~ 현재 : 강원대학교 IT대학 컴퓨터정보통신공학과 석사
- 주관심분야 : 컴퓨터비전, 컴퓨터그래픽스, 기계학습응용

저 자 소 개

황 인 철



- 1993년 : 고려대학교 전자공학과 학사
- 1995년 : 고려대학교 대학원 전자공학과 석사
- 2000년 : 고려대학교 대학원 전자공학과 박사
- 2000년 ~ 2001년 : University of Illinois, Urbana-Champaign, postdoctoral RA
- 2001년 ~ 2007 : 삼성전자 system LSI사업부, 책임연구원
- 2007년 ~ 현재 : 강원대학교 전기전자공학과 교수
- ORCID : <http://orcid.org/0000-0002-4702-8276>
- 주관심분야 : 고속/저전력 SoC설계, 지능반도체설계

김 만 배



- 1983년 : 한양대학교 전자공학과 학사
- 1986년 : University of Washington, Seattle 전기공학과 공학석사
- 1992년 : University of Washington, Seattle 전기공학과 공학박사
- 1992년 ~ 1998년 : 삼성종합기술원 수석연구원
- 1998년 ~ 현재 : 강원대학교 컴퓨터정보통신공학과 교수
- 2016년 ~ 현재 : 강원대학교 정보통신연구소 소장
- ORCID : <http://orcid.org/0000-0002-4702-8276>
- 주관심분야 : 3D영상처리, 비전점유센서, 객체인식