

뉴럴 네트워크를 이용한 Retinex 기반 반사 영상 생성

김원희 김만배

강원대학교 컴퓨터정보통신공학과

*wonhoi@kangwon.ac.kr manbae@kangwon.ac.kr

Generating Retinex-based Reflectance Image Using Neural Network

Wonhoi Kim and Manbae Kim

Computer & Communications Engineering, Kangwon National University

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.

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 연구에서의 광원은 영상 화질에 큰 영향을 주게 된다. 광원량이 과도하거나 부족한 경우 영상 표현 능력을 저하시키게 되어 영상에 대한 판별 능력을 어렵게 한다. 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. 이러한 조명광원의 영향을 줄이고, 반사 성분을 얻는 것이 Retinex 이론이다.

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

2. Retinex 이론

Land 등은 실험적으로 인간 시각 기관이 인지하는 물체의 색이 광원과 물체의 반사 성분의 곱으로 나타낼 수 있음을 입증하였으며[1], 수식으로 표현하면 밝기영상 $f(x,y)$ 는 두 개의 성분인 광원 조명성분 $i(x,y)$ 와 반사성분 $r(x,y)$ 의 곱으로 다음과 같다.

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

여기서 (x,y) 는 픽셀의 좌표이다.

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 를 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 1은 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 1(a)은 신경망 네트워크의 입력데이터이고, 그림 1(b)은 실제 출력영상으로 사용된다.



그림 1. 결과 그림. (a) 입력 영상, f 및 (b) 반사 영상, r_R

3. 신경망의 구조

심층신경망은 입력층과 출력층 그리고 여러 개의 은닉층들로 구성된다. 본 논문에서 사용된 신경망은 일반적인 구조를 가지며 역전파 알고리즘을 통해 심층신경망의 학습을 수행한다. 신경망을 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 이고 은닉층의 개수는 1~3개로 설정한다.

4. 실험 결과

실험에서 사용한 저조도 실험영상은 그림 2에서 보여진다. 영상 해상도는 128x128이고 학습데이터로 사용한 이미지는 총 102장이다. 102장에서 82장은 학습에 사용하고, 나머지 20장은 테스트에 사용하여 성능을 검증한다. 20장은 랜덤하게 선택하였다. 영상 데이터는 [0,1]로 정규화한다. 학습데이터는 패치 단위로 처리하며 이미지를 학습하기 위해 한 장의 사진이 들어올 때마다 중첩된 패치를 생성한다. 패치의 스트라이드가 3이고 크기가 16x16이면 패치의 개수는 1,444개이다. 8x8과 16x16 두 가지 패치의 크기를 설정하여 실험하였다. 그림 2(b)는 입력 저조도영상, 출력영상, 및 예측된 영상을 보여준다.

신경망 학습의 에폭(epoch)은 500으로 고정하였다. 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. 모든 층은 시그모이드(sigmoid) 활성화함수를 사용하였고 신경망의 가중치 초기화는 Xavier 초기화[2]을 사용하였다. 가중치 갱신 방식은 확률적 경사 하강법(stochastic gradient descent)을 사용한다. 최적화 기법은 Adam optimization[3]을 사용하였다. ρ_1 과 ρ_2 는 각각 0.9, 0.999이고 δ 은 10^{-8} 이다.



그림 2. (a) 실험에 사용된 광원이 존재하는 저조도 영상. (b) 입력 저조도 영상, 출력 타겟 영상, 예측 영상

제한한 심층신경망의 객관적 성능을 검증하기 위해서 저조도영상 향상 성능 검증에서 사용하는 3가지의 객관적 메트릭 1) RMSE, 2) PSNR, 3) SSIM을 사용하여 계산한다.

평균 RMSE, PSNR 및 SSIM의 성능은 표1에서 보여진다. LPF를 거쳐 나온 결과와 DNN에서 나온 결과값이다.

표 1. PSNR, 및 SSIM의 객관적 성능 평가.

Patch size	No. of hidden layers	No. of nodes	Avg. PSNR (in dB)		Avg. SSIM [0,1]	
			Training	Test	Training	Test
8x8	1	128	31.6261	32.1545	0.4504	0.4183
	2	64-64	31.8720	32.3097	0.4273	0.4363
	3	64-64-64	31.9500	32.4587	0.4081	0.4090
16x16	1	128	31.7840	32.1449	0.4293	0.4307
	2	64-64	31.8653	32.1130	0.4308	0.4273
	3	64-64-64	31.7973	31.9753	0.3858	0.3848

5. 결론

본 논문에서는 Retinex 이론에 기반한 반사성분 생성을 심층신경망으로 대체하는 방법을 제안하였고, 가능성을 조사하였으며 성능 측면에서는 만족스러운 결과를 얻을 수 있었다.

감사의 글

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2017R1D1A3 B03028806). 본연구는 과학기술정보통신부 및 정보통신 기술진흥센터의 대학ICT연구센터 지원사업의 연구결과임 (IITP- 2018-0-01433).

참고 문헌 (References)

- [1] E. Land and J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1-11, 1971.
- [2] 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.
- [3] D. Kingma and J. Ba, "Adam: A method for stochastic optimization", 2014, arXiv:1412.6980