

역광 및 저조도 환경에서의 얼굴 인식 성능 향상 기법

김민수, *이성주

반도체시스템공학과 및 지능형드론융합전공, 세종대학교

*전자정보통신공학과 및 지능형드론융합전공, 세종대학교

zmzmtls@itsoc.sejong.ac.kr, *seongjoo@sejong.ac.kr

Improvement Techniques for Facial Recognition in Backlit and Low-Light Environments

Kim min su, *Lee Seong joo

Dept. of Semiconductor System Engineering and Dept. of Convergence Engineering for Intelligent Drone

*Dept. of Electrical Engineering and Dept. of Convergence Engineering for Intelligent Drone
Sejong University

요약

본 논문은 역광과 저조도 환경에서의 물체 식별이 카메라 센서 한계로 어렵기에 CNN을 활용한 향상 기법을 제안한다. 그중 Retinex Theory을 활용하였으며, 하드웨어로 구현하는데 있어 Computing Power를 줄이고자 Self-Attention 기술을 활용하여 성능을 유지하면서 복잡성을 줄인 방법을 통해 이미지 향상을 수행함과 동시에 얼굴 인식에 있어서도 비슷한 성능을 나타냈다.

I. 서론

최근 무인 이동 장치가 다양한 분야에서 적용됨에 따라 주로 사용하는 영역인 영상 처리 영역의 연구 또한 활발하게 이루어지고 있다. 영상 처리 영역에서 사용되는 카메라 센서는 대비비를 통해 물체를 인식하는 인간의 눈과 다르게 빛의 밝기와 방향에 대해 영향을 받기에 중요한 정보의 소실 가능성이 있다. 이로 인해 역광 혹은 저조도 환경에서 적응하지 못하여 물체 분별하는데 있어 어려움이 발생한다.

이러한 문제를 해결하고자 따라서 Retinex Theory[1] - [4]와 같은 방법을 통해 사람의 눈과 같이 시간의 변화에 따라 영상에도 적응 가능하도록 구현하였다. 해당 방법은 Illumination map과 Reflectance map으로 이루어진 두 map을 이용하는 것으로 [1]는 3가지의 구조인 Decom, Denoise, Relight 로 분할하여 저조도 이미지를 향상시킨다. 하지만 [1]에서 사용한 방식은 하드웨어로 구현하는데 있어 Complex 수준이 높아 성능 저하, 자원 소모, 열 발생, Real-time 처리와 같은 현실적인 실현 가능성에 어려움이 발생할 수 있다. 그러므로 최근 각광받고 있는 Self-Attention 기술 [5]을 접목하여Complex를 낮추고 동시에 결과 성능은 비슷하게 유지하고자 Channel-wise Attention[6]과 Global Context Block[7]를 사용한 [8]를 응용하는 방식을 제안한다.

이를 통해 저조도 혹은 역광 이미지에 있어서의 향상 수준은 비슷하지만 Complex는 기존 R2Rnet 논문 대비 42% 낮아지는 효과를 얻었으며, MIRNet_v2대비 8% 낮아지는 효과를 얻었다. 또한, Yolov4를 이용한 얼굴인식에 있어서도 10% 수준의 오차범위를 갖는 결과물을 획득하였다.

II. 본론

2-1) Network Architecture

본 논문의 알고리즘은 기존 논문의 결과[1]를 토대로 구현하였다. 해당 알고리즘의 Contrast & Detail Enhancement 부분을 살펴보면 반복적인

FFT 와 IFFT, 그리고 Residual Module로 인해 Complex가 높아져 하드웨어로 구현하는데 있어 어려움이 생기게 된다.

따라서, 해당 부분의 알고리즘을 Self-Attention 기반의 구조를 사용한 논문[8]으로 바꿔주는 것을 제안한다.

해당 논문에서 사용된 Self-Attention은 Channel-wise Attention[6] (SKFF)과 Global Context block[7](RCB)으로 이루어져 있다.

이러한 두 Self-Attention 모듈을 하나의 구조로 엮은[8]의 방식에서 연산량을 낮추기 위한 새로운 구조인 그림1로 개선하였다.

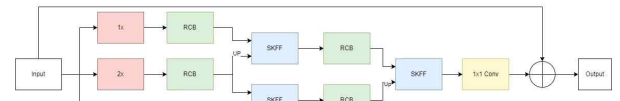


그림 1 MRB(Multi-scale Residual Block)

그림 1인 MRB는 그림 2에서와 같이 하나의 block으로 사용되며 이를 이용하여 Top module을 Residual 형태로 연결하였다.

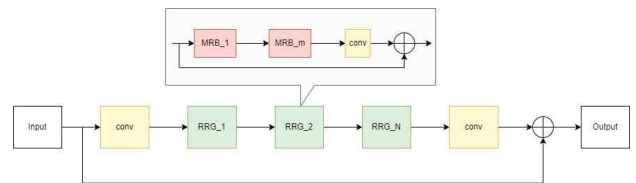


그림 2 RRG module

2-2) Loss function

본 논문에서 사용하는 Loss function은 RRG module에 대한 L1-loss인 ζ_{RRG} 가 존재하며, 이는 content loss와 perceptual loss로 나뉜다 [9].

* 교신저자

$$\zeta_{RRG-con} = \frac{1}{N} \sum_{i=1}^N |S_{low}^i - S_{nor}^i| \quad (1)$$

$$\zeta_{RRG-per} = \frac{1}{C_i H_i W_i} \|\varphi_j(\widehat{S}_{low}^i) - \varphi_j(S_{nor}^i)\|_2^2 \quad (2)$$

다음과 같은 수식을 통해 RRG module에 대한 content와 perceptual에 대한 loss를 구한다. 여기서 \widehat{S}_{low}^i 는 모듈을 통과하고 나온 i번째 저조도 이미지이며 S_{nor}^i 는 학습을 위해 pair로 들어가는 i번째 정상 이미지이다.

두 수식을 연결하여 최종적으로 사용되는 RRG module의 loss는 다음과 같으며, 여기서 사용된 $\alpha=0.1$ 로 설정하였다.

$$\zeta_{RRG} = \zeta_{RRG-con} + \alpha * \zeta_{RRG-per} \quad (3)$$

III. 실험 및 결과

본 논문에서 사용된 Train Dataset은 Low/Normal-Light로 이루어진 LSRW를 사용하였다. 이를 이용해 본 네트워크 구조를 학습 및 추론한 결과에 대한 성능 향상 결과는 표 1과 표 2로 확인할 수 있다.

표 1 네트워크 구조에 대한 FLOPS & Parameters (Input size = 3x256x256)

Net name	Computation Complexity [GMac]			Number of Parameters [M]	
Relight-Net [4]	total	276.65	100.00%	10.5	100.00%
	CE	147.75	53.407%	7.59	72.286%
	DE	124.04	44.836%	2.84	27.048%
MIRNet_v2 [8]	total	140.7	50.858%	5.86	55.809%
proposed	total	130.34	47.114%	5.86	55.809%

표 1의 내용을 살펴보면, [4]의 네트워크 대비 사용하는 Computation Complexity가 47%, 그리고 Parameters는 55% 수준의 향상 효과를 보이며 표 2를 보면 성능 향상 대비 PSNR과 SSIM의 성능 저하는 크게 일어나지 않았음을 보여준다.

표 2 저조도 이미지 향상 결과에 대한 PSNR과 SSIM

methods	PSNR	SSIM
R2RNet	18.2918	0.5531
R2R + MIR	17.9828	0.5365
Proposed	17.9073	0.5328



그림 3 각 Network 구조에 대한 얼굴 인식 성능 비교

또한, 위의 그림 3은 앞서 나타내었던 각 논문의 Network 구조에 대한 Yolov4를 이용한 얼굴인식 성능에 대한 비교를 나타낸 그림이다.

먼저 [4]는 90%의 성능을 보이고 [4] + [8]는 89%의 성능 그리고 제안한 proposed Algorithm의 경우 86%의 성능이 나타났다.

이를 통해 Self-Attention을 사용하여 연산량의 축소를 얻음과 동시에 이미지 향상 성능도 일정 수준 유지함을 확인하였다. 따라서 추후 연구를 통해 구조 변화로 인해 발생하는 성능 저하 현상에 대한 부분을 억제하는 알고리즘과 전체적인 Complex 성능을 향상하는 알고리즘에 대한 연구를 진행하고자 한다.

ACKNOWLEDGMENT

이 성과는 정부(교육부)의 재원으로 한국연구재단의 이공분야 대학중점연구소지원사업의 지원(No. 2020R1A6A1A03038540) 및 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송혁신인재양성(메타버스융합대학원)사업(IITP-2023-RS-2023-00254529)의 지원을 받아 수행한 연구 결과이며, 검증을 위한 EDA관련 툴은 IDEC의 지원을 받았음.

참고 문헌

- [1] X. Ren, W. Yang, W. -H. Cheng and J. Liu, "LR3M: Robust Low-Light Enhancement via Low-Rank Regularized Retinex Model," in IEEE Transactions on Image Processing, vol. 29, pp. 5862-5876, 2020, doi: 10.1109/TIP.2020.2984098.
- [2] W. Yang, W. Wang, H. Huang, S. Wang and J. Liu, "Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement," in IEEE Transactions on Image Processing, vol. 30, pp. 2072-2086, 2021, doi: 10.1109/TIP.2021.3050850.
- [3] C. Wei, W. Wang and W. Yang, "Deep Retinex decomposition for low-light enhancement", Proc. Brit. Mach. Vis. Conf., pp. 1-12, 2018.
- [4] Jiang Hai, Zhu Xuan, Ren Yang, Yutong Hao, Fengzhu Zou, Fang Lin, Songchen Han, R2RNet: Low-light image enhancement via Real-low to Real-normal Network, Journal of Visual Communication and Image Representation, Volume 90, 2023, 103712, ISSN 1047-3203, https://doi.org/10.1016/j.jvcir.2022.103712.
- [5] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).
- [6] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [7] Cao, Yue, et al. "Global context networks." IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).
- [8] Zamir, Syed Waqas, et al. "Learning enriched features for fast image restoration and enhancement." IEEE transactions on pattern analysis and machine intelligence 45.2 (2022): 1934-1948.
- [9] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in: European Conference on Computer Vision, Springer, 2016, pp. 694 - 711.
- [10] Dai, Yuekun, et al. "Flare7k: A phenomenological nighttime flare removal dataset." Advances in Neural Information Processing Systems 35 (2022): 3926-3937.