

조도 변화에 강인한 얼굴 이미지 디블러링 네트워크

권혁민, 김용록, 고현석*

한양대학교 전자공학과

hspeedkwon@hanyang.ac.kr, helloyr12@hanyang.ac.kr, *hyunsuk@hanyang.ac.kr

Face Image Deblurring Network Robust to Illumination Changes

Hyukmin Kwon, Yongrok Kim, Hyunsuk Ko*

Department of Electrical and Electronic Engineering
Hanyang Univ.

요약

얼굴 이미지 내의 블러는 영상의 품질을 저하시키며 얼굴 인식 성능을 감소시킨다. 따라서 블러를 제거하는 디블러링 기술이 필요하다. 하지만 조도 변화가 있는 환경에서 디블러링의 성능이 저하되는 문제점이 있다. 본 논문에서는 디블러링과 조도 개선을 통해 블러와 조도 변화가 있는 얼굴 이미지를 선명한 이미지로 복원하는 CNN 기반 네트워크를 제안한다. 제안하는 네트워크는 얼굴의 시멘틱(semantic) 마스크를 활용해 디블러링 성능을 높였다. 그리고 원본 이미지와 조도 변화 블러 이미지 쌍을 학습에 활용하여 디블러링 뿐만 아니라 저조도와 고조도 이미지로부터 조도를 개선한다.

I. 서론

이미지 내의 블러는 카메라의 흔들림, 피사체의 이동 그리고 아웃포커싱(out-focusing) 등에 의해 발생한다. 특히 조도가 낮은 환경에서는 카메라의 노출시간이 길어 모션 블러(motion blur)가 빈번하게 발생한다. 이러한 블러는 영상의 품질을 저하시키는 주요한 원인 중 하나이다. 특히, 얼굴 이미지 내의 블러는 사람을 식별하는데 어려움을 야기한다.

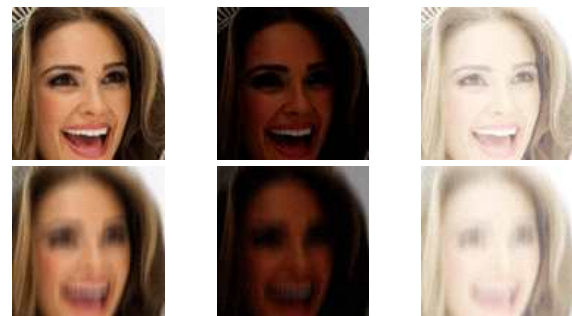
디블러링은 블러를 제거하여 선명한 이미지로 복원하는 기술이다. 얼굴 이미지의 디블러링은 선명한 이미지를 복원하고 얼굴 인식 성능을 높이는 데에 도움을 준다. 얼굴 이미지 내의 눈, 코, 입 등은 사람을 식별하는데 중요한 정보를 제공하기 때문에 시멘틱 마스크를 활용 시 얼굴 이미지의 디블러링 성능 향상과 얼굴 식별에 도움을 준다 [1, 2].

본 논문에서는 시멘틱 마스크를 활용하여 블러가 발생한 얼굴 이미지를 선명한 이미지로 복원하고, 저조도와 고조도 환경에서 조도를 개선할 수 있는 CNN 기반 디블러링 네트워크를 제안한다.

II. 데이터셋

본 연구를 위해 기존의 얼굴 이미지 데이터셋인 CelebA-HQ [3]를 활용해 원본 이미지, 조도 변화 이미지 그리고 각각의 블러 이미지를 생성했다. CelebA-HQ의 1024x1024 고화질 얼굴 이미지 30,000장을 128x128 사이즈로 다운샘플링하여 원본 이미지로 사용했다. 조도 변화 이미지 생성을 위해 원본 이미지에 감마 보정을 적용했으며, 감마 계수를 조절하여 3단계의 고조도 이미지와 3단계의 저조도 이미지를 각각 생성했다. 블러 이미지 생성을 위해 영상 취득 시 카메라가 흔들리는 궤적을 모델링하여 생성한 블러 커널을 활용했다 [4]. 수식 (1)과 같이 $13 \times 13 \sim 31 \times 31$ 의 10개 사이즈 블러 커널 K 를 선명한 원본 이미지 S 와 컨볼루션 연산 \otimes 을 통해 블러 이미지 B 를 생성했다. 이미지를 획득 및 처리하는 과정에서 발생하는 노이즈를 고려하여 블러 이미지에 가우시안 노이즈 N 를 추가했다.

$$B = S \otimes K + N \quad (1)$$



원본 이미지 저조도 이미지 고조도 이미지

그림 1. 생성한 데이터셋: 1행-원본 이미지, 2행-블러된 이미지

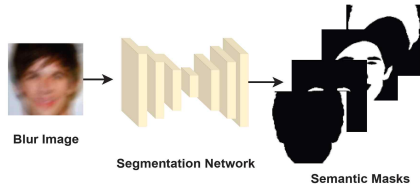
조도 변화 블러 이미지는 원본 이미지에 적용된 블러 커널을 조도 변화 이미지에 동일하게 적용하여 생성했으며 해당 데이터셋을 그림 1에 나타냈다.

III. 네트워크 구조

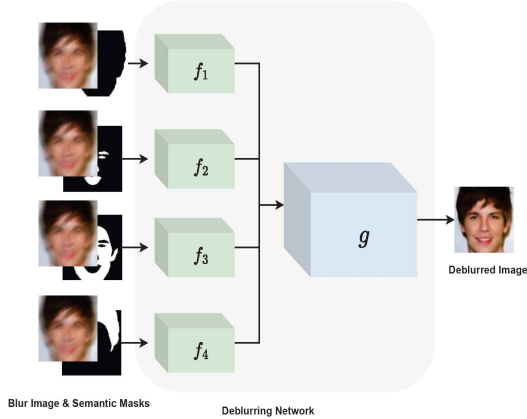
네트워크 구조는 세그멘테이션 네트워크와 디블러링 네트워크로 구성되며 이를 그림 2에 나타냈다. 세그멘테이션 네트워크에서는 얼굴의 구조적인 특징을 보다 잘 반영하기 위해 입력 블러 이미지로부터 4개의 시멘틱 마스크(배경, 얼굴 피부, 얼굴 내부 요소, 머리카락)를 추출한다. 추출한 4개의 시멘틱 마스크는 블러 이미지와 함께 디블러링 네트워크의 입력으로 사용한다. 디블러링 네트워크는 블러 이미지로부터 각 시멘틱 마스크에 해당하는 특성맵을 추출하는 4개의 $f(\cdot)$ 네트워크와 추출된 특성맵으로 선명한 이미지를 복원하는 $g(\cdot)$ 네트워크가 이어진 2단계 구조로 구성된다.

IV. 학습 방법 및 실험 결과

디블러링 네트워크를 학습시키기 위해 네 가지 손실함수로 Content



(a) Segmentation network



(b) Deblurring network

그림 2. 네트워크 구조

loss (L_c), Perceptual loss (L_p), Landmark loss (L_l), Wavelet loss (L_w)를 사용했다. Content loss는 원본 이미지와 복원 이미지 사이의 MSE (Mean Squared Error)로 계산된다. 이는 복원 이미지가 원본 이미지에 가깝도록 학습이 이루어진다. Perceptual loss는 사전 학습된 VGG-Face Net [5]의 Relu2_2 층에서 원본과 복원 이미지의 특성맵 사이의 MSE로 계산된다. 이는 복원 이미지의 인지 화질을 높이도록 학습이 이루어진다. Landmark loss는 사전 학습된 MobileFace Net [6]을 이용해 원본과 복원 이미지의 랜드마크를 추출하고 평균 픽셀 거리로 계산한다. 이는 원본 이미지의 랜드마크 정보를 보존하도록 학습이 이루어진다. Wavelet loss는 Wavelet 변환이 적용된 원본과 복원 이미지 사이의 MAE (Mean Absolute Error)로 계산한다. 이는 원본 이미지의 디테일한 텍스처 정보를 보존하도록 학습이 이루어진다. 학습에 사용한 총 손실함수 (L_t)는 식 (2)와 같다.

$$L_t = 60 * L_c + 0.01 * L_p + 0.1 * L_l + 5 * L_w \quad (2)$$

조도 개선을 위해 제안하는 네트워크와 UMSN, MSPL은 정상 밝기 데이터셋으로 사전 학습하고, 원본 이미지와 조도 변화 블러 이미지 쌍과 정상 밝기 데이터셋을 사용하여 추가 학습했다. 학습한 네트워크에 조도 변화 블러 이미지와 정상 밝기 블러 이미지로 실험한 결과를 그림 3과 표 1에 나타냈다. 복원된 이미지의 화질을 평가하기 위해 원본 이미지와 복원 이미지 사이의 PSNR, SSIM을 계산하고, 얼굴 인식 성능을 평가하기 위해 원본 얼굴 랜드마크와 복원 얼굴 랜드마크의 차이를 NME (Normalized Mean Error)로 계산했다. 평균 PSNR, SSIM, NME 결과에서 제안한 네트워크의 성능이 가장 좋은 모습을 보였다. 제안 네트워크는 UMSN 및 MSPL 대비 PSNR 기준 약 1.8dB, 12.2dB의 우수한 복원 성능을 보인다. 또한, 그림 3에서 볼 수 있듯이 제안 네트워크는 정성적인 화질 복원 측면에서도 가장 우수한 성능을 보였다.

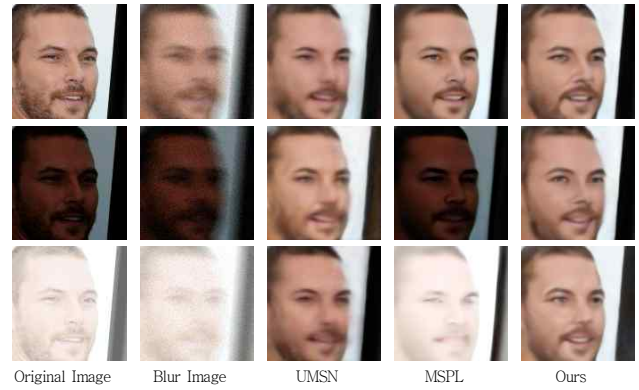


그림 3. 제안 방법과 기존 방법의 정성적 비교

표 1. 제안 방법과 기존 방법의 정량적 비교

Methods	Our Dataset		
	PSNR (↑)	SSIM (↑)	NME5 (↓)
UMSN [1]	21.80	0.804	7.90
MSPL [2]	11.37	0.473	8.10
Ours	23.61	0.835	5.96

V. 결론

본 논문에서는 얼굴 이미지 내의 블러를 제거하고 조도를 개선하여 선명한 이미지로 복원하는 CNN 기반 네트워크를 제안했다. 제안하는 네트워크는 기존 네트워크보다 디블러링 성능과 조도 개선 성능이 높았다. 또한, 얼굴 랜드마크에서 더 우수한 성능을 보여 기존 네트워크에 비해 더 원본에 가까운 얼굴 이미지로 복원한 것을 확인할 수 있었다.

참고 문헌

- [1] R. Yasarla, F. Perazzi and V. M. Patel, "Deblurring Face Images Using Uncertainty Guided Multi-Stream Semantic Networks," in IEEE Transactions on Image Processing, vol. 29, pp. 6251–6263, 2020.
- [2] T. B. Lee, S. H. Jung and Y. S. Heo, "Progressive Semantic Face Deblurring," in IEEE Access, vol. 8, pp. 223548–223561, 2020.
- [3] T. Karras, T. Aila, S. Laine, J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," ICLR, 2018.
- [4] G. Boracchi and A. Foi, "Modeling the Performance of Image Restoration From Motion Blur," in IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3502–3517, Aug. 2012.
- [5] O. M. Parkhi, A. Vedaldi and A. Zisserman, "Deep face recognition," In Proc. BMVC., 2015.
- [6] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices," In Chinese Conference on Biometric Recognition, pp. 428–438, 2018.