

Auto White Balance for multi-illuminant scene

jyp

팀원: 박윤정 유수민 조윤수

지도교수: 김선주

조교: 김동영

목차

1. 연구 주제
2. 연구의 필요성
 - 2.1 기존 연구의 한계점
 - 2.2 Large Scale Multi-Illuminant Dataset
3. 연구 내용
 - 3.1 연구 주제 및 목표 선정
 - 3.2 기존 알고리즘 분석
 - 3.3 U-net3+
 - 3.4 U-net3+로 LSMI dataset 학습
4. 현재 진행 상황 및 향후 계획
5. 일정 및 역할 배분

1. 연구 주제

다중 조명 상황에서의 Auto White Balance 구현:

Auto White Balance(이하 AWB)는 사진에서 조명의 영향을 제거해 물체의 고유 색상이 잘 표현되도록 하는 기술이다. 장면 내에 여러 개의 조명이 존재하는 경우 적용 가능한 AWB 알고리즘에 대하여 연구하고자 한다.

2. 연구의 필요성

2.1 기존 연구의 한계점

기존의 AWB 알고리즘은 카메라 센서가 장면의 광원 및 조도를 추정한 후 이를 기반으로 보정을 하는 방식으로, 추정에 있어서 단일 광원을 가정한다. 단일 광원 AWB 알고리즘으로는 Gray-world methods, white-patch hypothesis와 같은 Statistical methods와 gamut-based methods와 같은 Learning-based methods가 있다. 그러나 현실 상황에서는 하나의 장면 내에 다양한 광원이 존재하기에, 이러한 기존 AWB 알고리즘을 적용할 경우 빨강거나 파랗게 물드는 등 실제 색과 전혀 다른 결과물을 낼 가능성이 매우 높아진다.

2.2 Large Scale Multi-Illuminant Dataset

관련 연구로 제시된 ‘Large Scale Multi-Illuminant (LSMI) Dataset for Developing White Balance Algorithm under Mixed Illumination’은 dataset의 각 이미지에 대하여, 조명의 종류, 광원의 색도, 광원의 혼합 비율을 픽셀 단위로 제시한다. 더불어, LSMI dataset에 CNN 기반의 pixel-level White Balance 알고리즘을 적용하여 pixel-level이 patch-based method보다 성능이 뛰어남을 보여주었다. CNN 기반 모델로는 U-net과 HDRnet이라는 두 기본적인 모델을 사용하였는데, 해당 연구에서는 사용한 CNN 기반 모델에 있어 여전히 개선의 여지가 있음을 시사하였다.

3. 연구 내용

3.1 연구 주제 및 목표 선정

LSMI dataset을 사용하며 개선된 AWB 모델을 찾고자 한다. 기존 모델을 수정하여 U-net3+ 모델을 적용시켜볼 뿐만 아니라, transformer를 활용하여 심화 모델을 구현하는 것이 목표이다.

3.2 기존 알고리즘 분석

LSMI dataset 연구의 dataset과 모델 구조를 이해하기 위해, 해당 연구에서 제시된 코드를 활용하였다. U-net 모델을 학습, 테스트하며 모델의 구조 및 코드의 구성을 이해하였다. (관련 링크: <https://github.com/DY112/LSMI-dataset>)

3.3 U-net3+

U-net3+는 원래 의료 영상을 다루기 위해 개발된 encoder-decoder 구조의 모델이다. image segmentation 문제에서 우수한 성능을 보이며, 오늘날 image processing, image-to-image translation과 같은 pixel-level의 task에서 널리 사용된다.


U-net3+가 기존 u-net과 다른 점은 다음과 같다.

- Full-scale skip connection: low scale과 high scale의 feature를 합쳐 더 적은 parameter로 더 정확한 segmentation을 얻는다.
- Full-scale deep supervision: 각각의 decoder에서의 결과를 deep supervision으로 사용하여 성능을 개선한다.
- Classification Guided Module: CGM을 사용하여 over-segmentation을 방지하고 더 정확한 결과를 낸다.

3.4 U-net3+로 LSMI dataset 학습

LSMI dataset에서 사용하였던 기존 U-net 모델을 U-net3+ 모델로 수정하기 위하여 official U-net3+의 class UNet_3Plus를 사용하였다. U-net3+는 본래 classification이 목적인 모델이기 때문에 코드의 구조를 일부 수정하였다. (관련 링크: https://github.com/ZJUGiveLab/UNet-Version/blob/8ecd06740d18cf508e5c8b2cc2bf38a026aa0e9b/models/UNet_3Plus.py)

모델 크기가 24GB 정도로 커서 GPU 3개를 사용하여 학습하였음에도 batch size를 줄여야 하는 문제가 있었다.

```
 from torchsummary import summary
model = UNet_3Plus().cuda()
summary(model, (3, 256, 256), batch_size = 8)
```

```

=====
Total params: 26,969,538
Trainable params: 26,969,538
Non-trainable params: 0
-----
Input size (MB): 6.00
Forward/backward pass size (MB): 23940.00
Params size (MB): 102.88
Estimated Total Size (MB): 24048.88
=====

```

이에 batch size를 32에서 8로 줄여 학습을 진행하였고, 학습 시간이 오래 걸려 현재까지 진행된 결과를 아래에 첨부하였다.










총 2000 epoch 중 1300 epoch까지 학습된 모델의 validation 결과로, MAE는 Mean Angular Error를 의미한다. validation MAE가 처음에는 약 6도에서 시작하여 epoch가 진행될수록 감소하는 추세로 보아 학습이 잘 진행되고 있는 것으로 보인다.

4. 향후 계획

- 테스트 데이터를 사용하여 학습된 U-net3+ 모델의 성능 평가.
- 기존 U-net 모델의 batch size를 32에서 8로 수정한 후 다시 학습, 테스트.
- U-net과 U-net3+ 두 모델의 성능 비교
- U-net3+ 모델의 성능 개선을 확인한 후, transformer를 활용한 더 심화된 모델 구현.
- 심화 모델에 있어 parameter를 변경하는 등 성능을 테스트할 수 있는 다양한 방법을 적용하며 실험해 볼 예정

5. 일정 및 역할 배분

(주차)

	7	8	9	10	11	12	13	14	15
U-net3+ 학습 및 테스트									
U-net 재학습 및 U-net3+과 성능 비교									
중간 발표									
transformer 모델 구현 및 성능 실험									
최종 발표									
하이라이트 영상 상영회									
온라인 전시회									

박윤정: 연구 자료 조사, 개발, 보고서 작성, 최종 발표

유수민: 연구 자료 조사, 개발, 보고서 작성, 중간 발표

조운수: 연구 자료 조사, 개발, 보고서 작성, 연구제안 발표