

그림자 정보를 활용한 실외 환경 광원 추적에 대한 연구

Outdoor Illumination Estimation using Shadow Information

Team

소프트웨어학부 3학년 이나혁, 이하윤

About Our Team

팀원 소개



이나혁 | Nahyuk Lee

소프트웨어학부 3학년

Research Interest : 3D Vision

Github : [@NahyukLEE](https://github.com/NahyukLEE)



이하윤 | Hayun Lee

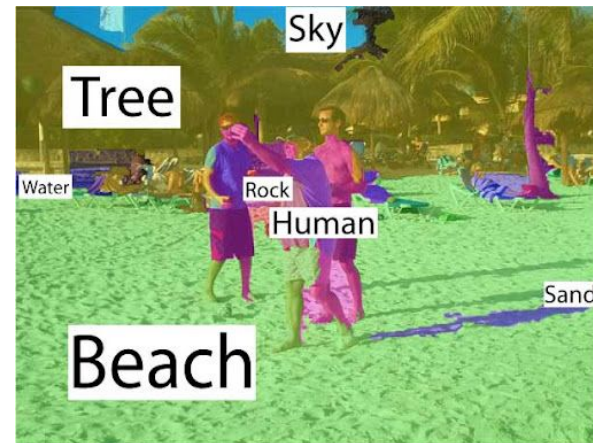
소프트웨어학부 3학년

Research Interest : Robot Learning

Github : [@sksmslhy](https://github.com/sksmslhy)

Introduction

문제 제기 및 필요성 (1)



광원에 대한 해석은 컴퓨터 비전 내 **Scene Understanding** 분야에 있어 중요한 문제 중 하나

Introduction

문제 제기 및 필요성 (2)

가상공간서 모든 경제활동... '메이드 인 메타버스' 시대 열린다

매경TV `2021혁신성장포럼`

일상침투해 경제적 가치 창출
경험중심의 3D공간으로 대체
공상과학 같은 세상 펼쳐질것

이어령 "메타버스가 종착점"

정부, 내년 메타버스 5개 핵심 주제 R&D에 259억원 투입

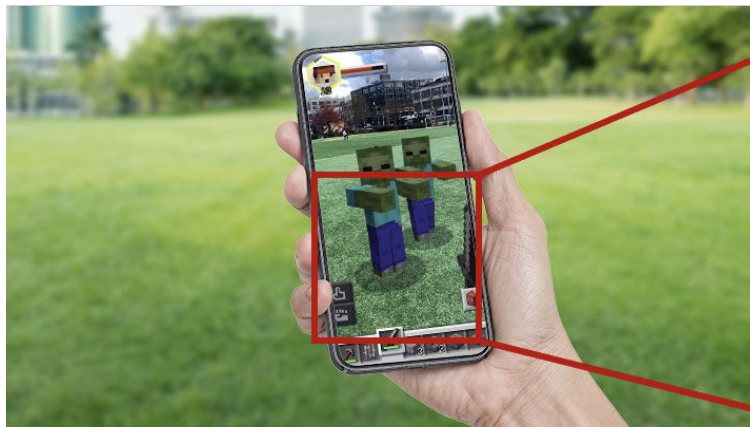
I ITP 행사서 내년도 과제 일부 공개...내년 초 확정 공고 계획

지난달 마이크로소프트(MS)와 엔비디아는 '메타버스(Metaverse)'를 차세대 핵심 사업으로 선언하고, 메타버스 관련 신기술을 소개했다. 젠슨 황 엔비디아 최고경영자(CEO)는 "메타버스가 현재 경제보다 더 큰 새로운 경제가 될 것이라 확신한다"고 했다. 이에 앞서 지난 10월 세계 최대 소셜미디어 그룹 페이스북은 창립 17년 만에 회사명을 '메타(Meta)'로 바꿨다. 마크 저커버그 메타 CEO는 "메타버스는 인터넷의 다음 단계"라며 "시간이 지날수록 우리가 메타버스 회사로 인식되길 바란다"고 했다.

Metaverse(Meta+Universe)가 급부상하며 장면에 대한 이해는 더욱 중요해짐

Introduction

문제 제기 및 필요성 (3)



증강현실(**AR: Augmented Reality**) 등의
어플리케이션에서 가상의 3차원 물체에게
생동감을 부여하기 위해선 **조명 환경에 맞는
그림자가 필수적임**



실제와 같은 그림자를
렌더링하기 위해선 **광원의
위치를 알아야 함**

Baseline

베이스라인 연구 (1)

Deep Outdoor Illumination Estimation^[1]

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde, CVPR 2017(Oral)

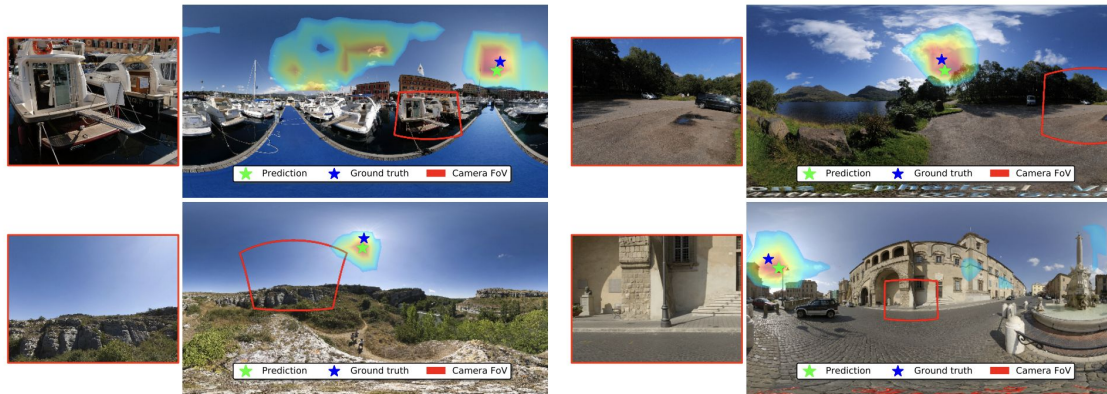


Figure 6. Examples of sun position estimation from a single outdoor image. For each example, the input image is shown on the left, and its corresponding location in the panorama is shown with a red outline. The color overlay displays the probability distribution of the sun position output by the neural network. A green star marks the most likely sun position estimated by the neural network, while a blue star marks the ground truth position.

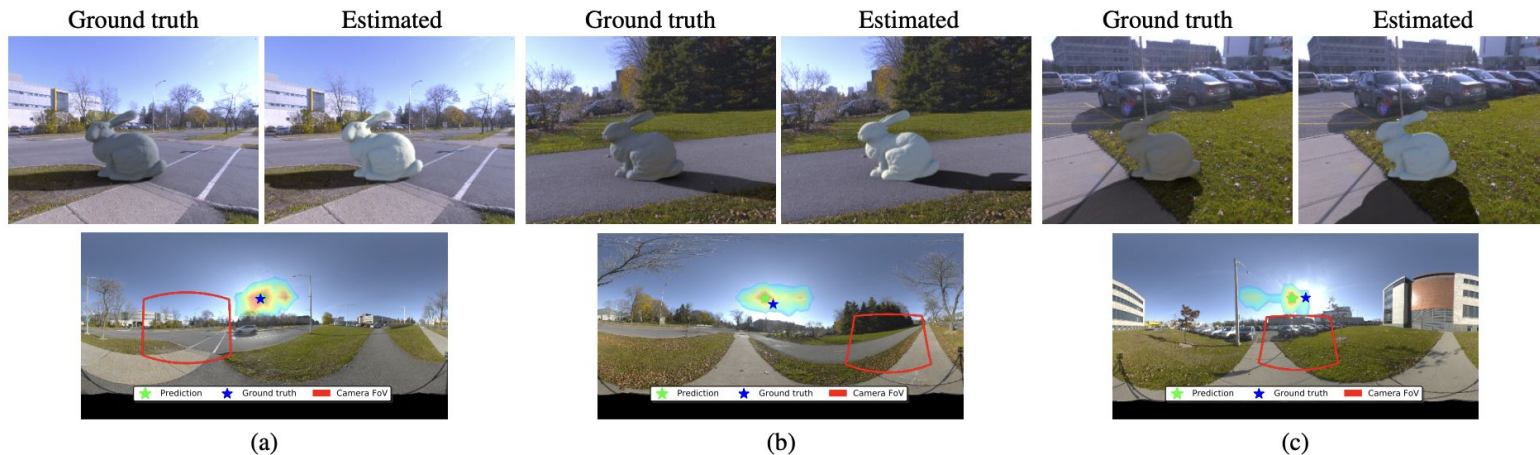
한 장의 실외 파노라마 이미지로부터 광원(태양)의 위치와 Camera, Sky Parameter를 추정

Baseline

베이스라인 연구 (2)

Deep Outdoor Illumination Estimation^[1]

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde, CVPR 2017(Oral)



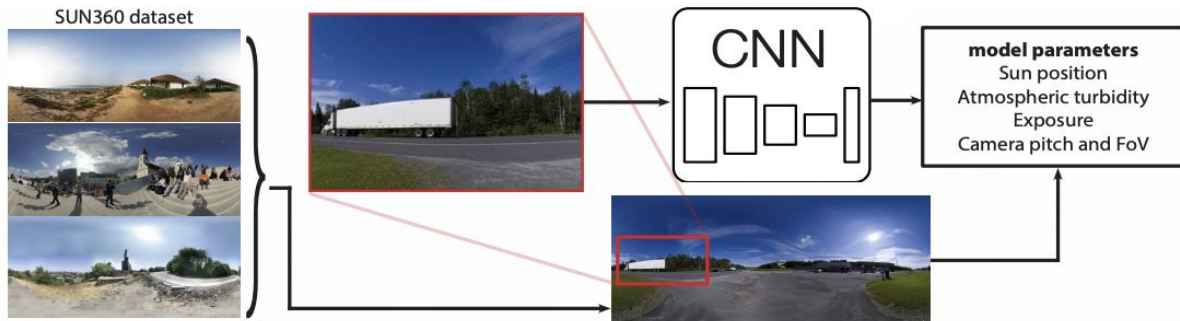
Project URL: <http://vision.gel.ulaval.ca/~jflalonde/publications/projects/deepOutdoorLight/index.html>

Baseline

베이스라인 연구 (3)

Deep Outdoor Illumination Estimation^[1]

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde, CVPR 2017(Oral)



Model Architecture

Layer	Stride	Resolution
Input		320×240
conv7-64	2	160×120
conv5-128	2	80×60
conv3-256	2	40×30
conv3-256	1	40×30
conv3-256	2	20×15
conv3-256	1	20×15
conv3-256	2	10×8
FC-2048		
FC-160 LogSoftMax		
FC-5 Linear		
Output: sun position distribution s		Output: sky and camera parameters q

Baseline

베이스라인 연구 (4)

Deep Outdoor Illumination Estimation^[1]

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde, CVPR 2017(Oral)

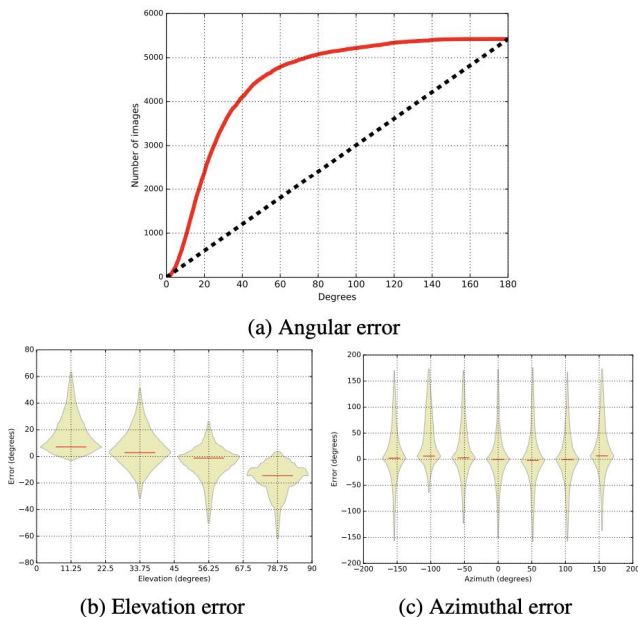


Figure 4. Quantitative evaluation of sun position estimation on all 8659 images in the SUN360 test set. (a) The cumulative distribution function of the angular error on the sun position. The estimation error as function of the sun elevation (b) and (c) azimuth relative to the camera (0° means the sun is in front of the camera). The last two figures are displayed as “box-percentile plots” [10], where the envelope of each bin represents the percentile and the median is shown as a red bar.

Result

Sky Model

실외 환경 광원 추적

태양에 관한 몇 가지 정보

- **Light Direction:** Directional Light

- **Sun Position:**

구 좌표계 상에서 카메라 위치를 중심으로 방위각과
고도를 통해 2차원 값으로 표현할 수 있음

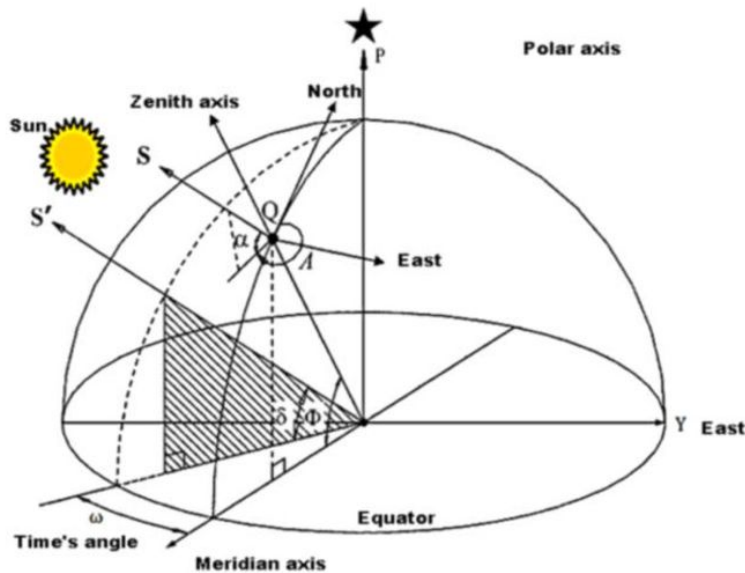
- **LM(Lalonde-Matthews) Sky Model [2]:**

$$f_{LM}(\mathbf{l}; \mathbf{q}_{LM}) = f_{sun}(\mathbf{l}; \mathbf{q}_{sun}, \mathbf{l}_{sun}) + f_{sky}(\mathbf{l}; \mathbf{q}_{sky}, \mathbf{l}_{sun})$$

$$\mathbf{l}_{sun} = [\theta_{sun}, \varphi_{sun}]$$

$$\mathbf{q}_{LM} = \{ \mathbf{w}_{sky}, t, \mathbf{w}_{sun}, \beta, \kappa, \mathbf{l}_{sun} \}$$

다음의 11개의 파라미터로 실외 조명 환경 모델링



구 좌표계

Dataset

데이터셋(1)

- **The Laval Outdoor HDR Dataset**[\[3\]](#) (Hold-Geoffroy, Yannick, et al. (CVPR 2019))
 - LM Sky Model을 기반으로 하여 하늘에 대한 정보가 **Labeling** 되어 있음
 - 205장의 360도 파노라마 이미지와 태양 위치를 포함한 Sky Parameter
- Sky model parameter: [Link](#)



Dataset

데이터셋(2)

- **Tone Mapping**

- 가지고 있는 64bit-float 형의 HDR 파노라마 이미지를 8bit-int형의 LDR 이미지로 Tone Mapping 하여 사용

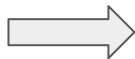


Dataset

데이터셋(3)

- **Panorama Crop**

- 파노라마 사진 1장 당 7장의 이미지를 320*240 크기로 crop하여 데이터셋 제작
- 총 $205 * 7 = 1,435$ 장

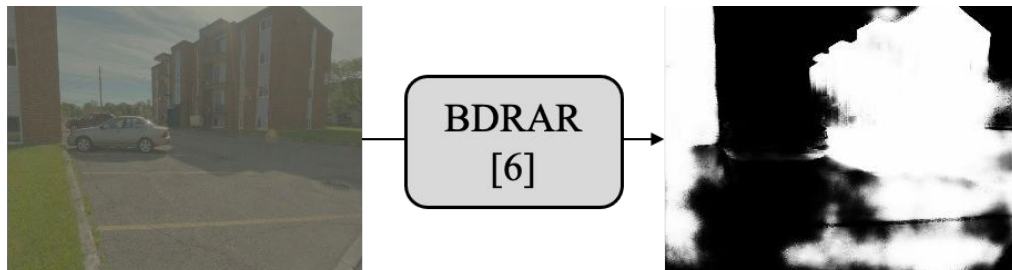


Dataset

데이터셋(4)

- **Shadow Information**

- BDRAR[4](Bidirectional Feature Pyramid Network with Recurrent Attention Residual Modules for Shadow Detection) 이용
- SBU Dataset을 이용하여 BDRAR 모델 학습 후 crop된 우리의 LDR 이미지로 inference 진행해 그림자 정보 획득

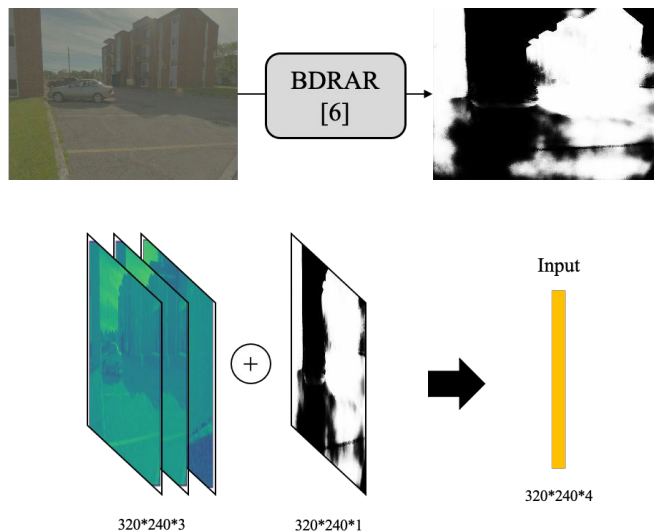


Model Input

학습 이미지

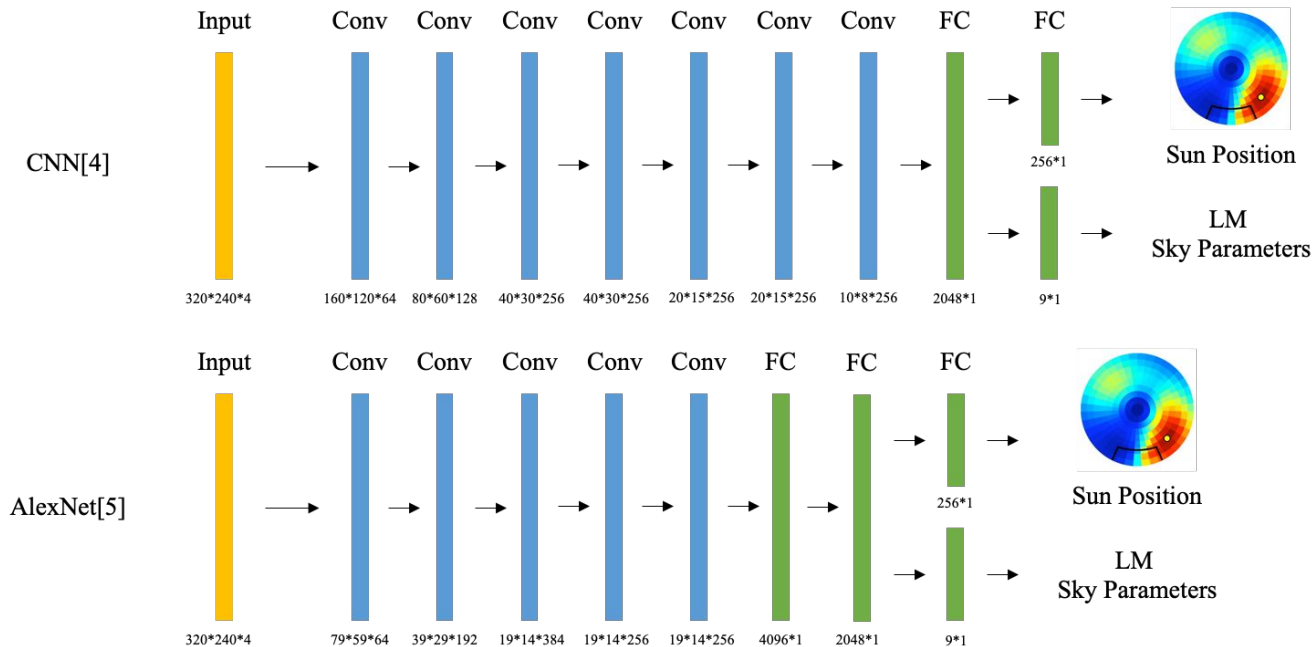
- **Input Data**

- 320*240 크기로 crop된 LDR 이미지와 해당 이미지로부터 검출된 Gray-Scale 그림자 이미지가 합쳐진 4 Channel image



Deep Learning Network

딥러닝 네트워크



CNN기반의 기본 네트워크

Loss Function

손실 함수

- LM(Lalonde-Matthews) Sky Model [2]:

$$f_{\text{LM}}(\mathbf{l}; \mathbf{q}_{\text{LM}}) = f_{\text{sun}}(\mathbf{l}; \mathbf{q}_{\text{sun}}, \mathbf{l}_{\text{sun}}) + f_{\text{sky}}(\mathbf{l}; \mathbf{q}_{\text{sky}}, \mathbf{l}_{\text{sun}})$$

$$\mathbf{l}_{\text{sun}} = [\theta_{\text{sun}}, \varphi_{\text{sun}}]$$

$$\mathbf{q}_{\text{LM}} = \{ \mathbf{w}_{\text{sky}}, \quad t, \quad \mathbf{w}_{\text{sun}}, \quad \beta, \quad \kappa, \quad \mathbf{l}_{\text{sun}} \}$$

- Loss Function:

$$\mathcal{L}(\mathbf{s}^*, \mathbf{q}^*, \mathbf{s}, \mathbf{q}) = \underbrace{\mathcal{L}(\mathbf{s}^*, \mathbf{s})}_{\text{MSE Loss}} + \beta \underbrace{\mathcal{L}(\mathbf{q}^*, \mathbf{q})}_{\text{KL Divergence Loss}}, \quad (5)$$

Training

학습

- 실험은 총 8개로 나뉘어 진행
 - 기준 : 네트워크 구조, 그림자 정보의 포함 여부, 타겟 파라미터의 종류
- 네트워크 구조
 - Hold-Geoffroy et al[1]. 에서 사용된 7층 CNN 구조 혹은 Shang et al[5]. 에서 사용된 AlexNet 구조
- 그림자 정보의 포함 여부
 - 4채널 : 그림자 정보 포함 RGB+Shadow
 - 3채널 : 그림자 정보 미포함, RGB
- 예측 타겟
 - 11개 파라미터 예측 : 태양의 위치(2) + 하늘 파라미터(9) 추정
 - 2개 파라미터 예측 : 태양의 위치(2)만 추정
- 학습 세부사항
 - Train : Test = 7: 3
 - Batch Size: 64, 500 Epoch, Learning rate : 0.01 (15 epoch마다 0.5 곱함)

Result

결과

Model Architecture	CNN[4]	AlexNet[5]	CNN[4]	AlexNet[5]	CNN[4] (Y. Hold-Geoffroy et al.)	AlexNet[5] (Zhang et al.)	CNN[4]	AlexNet[5]
Shadow	with Shadow (Proposed)				without Shadow			
Target	Sky Parameter + Sun Position		Sun Position		Sky Parameter + Sun Position		Sun Position	
MAE	57.707	63.642	47.265	45.304	40.527	50.775	51.493	68.912
MAE (mean)	60.675		46.285		45.651		60.203	

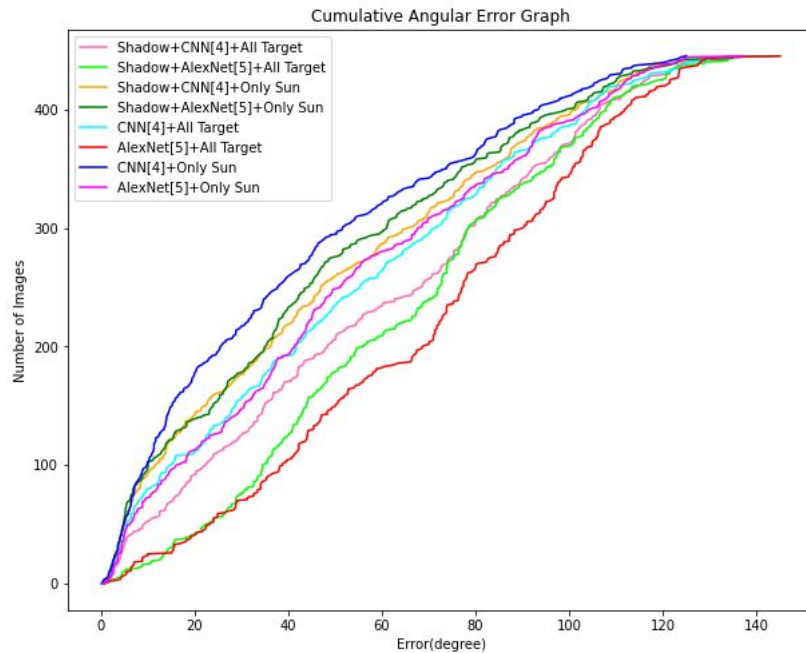
각 실험 별 MAE를 나타낸 도표

사용 네트워크 구조, 그림자의 입력 포함 여부, 타겟 파라미터 등에 따라 실험을 구분하여 표기.

5, 6번째 실험은 각각 Y.Hold-Geoffroy et al., Zhang et al.에서 제시된 실험과 동일한 환경으로 구성됨

Result

결과



각 실험별 sun position의 angular error를 누적하여 나타낸 그래프

Discussion

결론

- 성능
 - Sky Parameter를 반드시 필요로 하였던 기존의 방법과 비교하여 Sky Parameter에 대한 추정 없이도 Shadow 정보를 이용하면 충분히 그에 버금가는 추적 성능을 달성할 수 있음.
- 발전 가능성
 - 그림자 검출의 성능을 향상시키면 성능 향상에도 도움이 될 것임.
 - 그림자를 딥러닝 모델에 포함시키는 방법을 개선할 수 있음.
 - 데이터셋의 양을 늘리고 더 견고히 하여 성능 향상을 기대할 수 있음.

Reference

참고 문헌

- [1] Hold-Geoffroy, Y., Sunkavalli, K., Hadap, S., Gambaretto, E., & Lalonde, J. F. (2017). Deep outdoor illumination estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7312-7321).
- [2] Lalonde, J. F., & Matthews, I. (2014, December). Lighting estimation in outdoor image collections. In *2014 2nd International Conference on 3D Vision* (Vol. 1, pp. 131-138). IEEE.
- [3] Hold-Geoffroy, Y., Athawale, A., & Lalonde, J. F. (2019). Deep sky modeling for single image outdoor lighting estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6927-6935).
- [4] Zhu, L., Deng, Z., Hu, X., Fu, C. W., Xu, X., Qin, J., & Heng, P. A. (2018). Bidirectional feature pyramid network with recurrent attention residual modules for shadow detection. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 121-136).
- [5] Zhang, J., Sunkavalli, K., Hold-Geoffroy, Y., Hadap, S., Eisenman, J., & Lalonde, J. F. (2019). All-weather deep outdoor lighting estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10158-10166).