

그림자 정보를 활용한 실외 환경 광원 추적에 대한 연구

이나혁*, 이하윤*

Outdoor Illumination Estimation using Shadow Information

Nahyuk Lee*, Hayun Lee*

*School of Computer Science & Engineering, Chung-Ang University

요 약

실외 환경에 대한 광원을 추적하는 것은 Scene Understanding, AR 등에 있어 굉장히 중요한 문제이다. 현실 세계에서 그림자에 대한 정보는 광원의 위치를 짐작할 수 있도록 하는 중요한 정보인데, 본 보고서는 이를 기반으로 한 실험을 통해 CNN 기반의 딥 러닝 네트워크에 그림자 정보를 입력으로 추가함으로써 LDR Photo로부터의 태양의 위치 추적 성능이 향상됨을 보인다. 또한, 딥 러닝 네트워크의 추적을 Sky Parameter를 통해 보조했을 때와 그렇지 않을 때의 성능에 대하여 비교, 분석하여 그 성능을 평가하였다.

Abstract

Estimating outdoor illumination is a very important problem for scene understanding, AR, etc. In real world, shadow information is important information that allows us to guess the position of illumination, and we shows that performance of light estimation from LDR photo is improved by adding shadow information as input for CNN-based deep learning network. In addition, we proved the performance of deep learning network, which evaluated by comparing whether network are assisted through sky parameter or not.

Key words

Scene Understanding, Outdoor Illumination Estimation, Deep Learning, Shadow

1. 서 론

1.1 주제 선정 개요

컴퓨터 비전 내 장면 이해(Scene Understanding) 영역에서 광원에 대한 해석은 굉장히 중요한 문제로 여겨진다. 광원은 장면의 모양을 결정하며 광원을 추적하는 것은 장면을 재구성하고 편집하는 데 매우 중요하다는 것이 그 이유이다.

특히 최근 들어 메타버스(Metaverse)와 증강 현실(Augmented Reality)와 같이 현실 세계의 3차원 환경과 가상의 환경을 연결 짓는 기술에 대한 수요가 급부상하며 장면에 대한 이해는 더욱 중요해졌다. 가상공간에서 3차원 물체에 자연스러움과 생동감을 부여하기 위해서는 조명 환경에 맞는 렌더링과 그림자가 필수적이다. 따라서 실외 환경에서 실제와 같은 그림자를 렌더링하기 위해 태양의 위치를 파악하는 것은 중요한 과제이다.

장면에서 광원에 대한 정보는 빛의 경향성, 하늘의 색상, 물체의 밝기 정보, 그림자 등을 통해 얻을 수 있다[1]. 그 중, 그림자는 광원에 대해 종속성을 가지며 매우 직관적인 정보로 광원 추적에서 중요한 단서로 작용할 수 있다. 따라서 기존 실외 환경에서의 광원 추적에 대한 연구를 분석하고 이에 그림자 정보에 대한 입력을 추가한 딥 러닝 기반의 광원 추적 모델을 제안한다.

1.2 선행 연구 분석

기존 실외 환경에서의 광원 추적에 대한 연구는 크게 Parametric model based method와 Eschewed parametric model based method로 구분된다. Parametric model based method는 Hosek-Wilkie sky model[3], Lalonde-Matthews sky model[2] 등의 HDR sky parameter의 추정을 기반으로 태양의 위치를 추적하는 방법으로 Y. Hold-Geoffroy et al.[4]는 CNN 네트워크를 통해 LDR Photo로부터 Hosek-Wilkie HDR sky parameter[3]를 추정하며 태양의 위치를 추적한다. 한 장의 LDR Photo로 HDR sky parameter를 추정하는 딥 러닝 기반의 최초의 시도라는 의미를 가지지만 사실적인 그림자 렌더링을 하지 못하며 다양한 날씨 환경에 대응하지 못한다. 한편, Zhang et al.[5]은 모든 날씨 환경에 대응할 수 있는 Sky parameter 추정을 위해 Parameter 레이블링을 위한 생성 모델과 AlexNet[10] 기반의 LM Sky parameter[2] 추정 네트워크를 제안하였다. 다만 앞서 언급한 두 논문은 태양 추적에 대한 정확도가 낮기 때문에 여전히 추적에 대한 정확도는 풀어야 할 과제로 남아있다.

한편, Eschewed parametric model은 이러한 Sky parameter에 대한 추적 없이 Encoder를 통해 LDR 이미지의 특징을 추출하고 이를 해석하여 태양의 위치를 추적한다[11, 12, 13]. 이는 Sky parameter에 대한 추적이 없기 때문에 그 추적에 대한 신뢰성이 떨어진다는 단점을 가지나, 따로 Sky parameter에 대한 레이블링이 필요하지 않다는 장점을 가진다.

우리는 태양 위치 추적에 있어 Parametric model이 반드시 태양의 추적에 효과적일까라는 의문에



그림 1. LDR Tone Mapping 결과 예시

답하고, 그림자 정보가 태양 위치 추적에 효과적이라는 선행 연구[1]에 따라 이를 딥 러닝 모델을 통해 검증하는 실험을 진행하였다.

II. 데이터셋 구축

2.1 Sky model

데이터셋은 실외 환경에 대한 360 HDR 파노라마 이미지와 사전 계산된 Sky Parameter가 포함되어 있는 LAVAL HDR Dataset[14]을 활용하였다. 이 데이터는 Lalonde-Matthews (LM) sky model[2]을 기반으로 하여 하늘에 대한 정보가 레이블링 되어 있다. 아래 서술할 LM Sky model[2]에 대한 설명은 Zhang et al.[5]을 인용하였다. LM Sky Parameter[2]는 아래 수식 (1)로써 정의된다:

$$f_{LM}(l; q_{LM}) = f_{sun}(l; q_{sun}, l_{sun}) + f_{sky}(l; q_{sky}, l_{sun}), \quad (1)$$

이 때, $l_{sun} = [\theta_{sun}, \phi_{sun}]$ 은 실외 환경의 외부 광원인 태양의 좌표를 구 좌표계를 통해 방위각과 고도로 표현한 것을 의미한다. q_* 는 아래 정의되는 파라미터를 의미한다.

수식 (1)에서 Sky 요소 $f_{sky}(l)$ 은 Preetham sky model[8]에서의 $f_P(\cdot)$ 를 의미하고, w_{sky} 는 3차원으로 표현되는 Sky의 평균 색상을 의미한다:

$$f_{sky}(l; q_{sky}, l_{sun}) = w_{sky} f_P(\theta_{sun}, \gamma_{sun}, t), \quad (2)$$

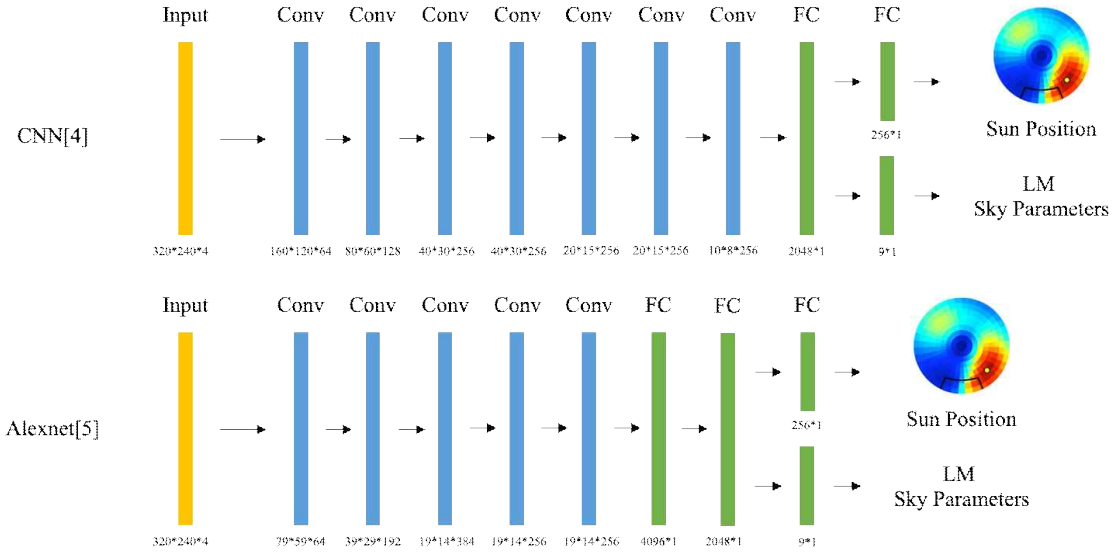


그림 2. 딥 러닝 네트워크 구조

이 때, γ_{sun} 는 Sky 요소 l 과 태양의 위치 l_{sun} 사이의 각도를 의미하며 t 는 Sky Turbidity(연무율) 을 의미한다. 마찬가지로 수식 (1)에서 정의된 Sun 요소 $f_{sun}(l)$ 는 아래와 같이 수식 (3)으로 정의된다.

$$f_{sun}(l; q_{sun}, l_{sun}) = w_{sun} e^{-\beta e^{(-x/\gamma_{sun})}}, \quad (3)$$

이 때, β 와 x 는 태양의 형태를 조절하는 두 개의 파라미터이고, w_{sun} 은 3차원으로 표현되는 태양의 평균 색상을 의미한다. 따라서, LM sky model[2] 은 다음의 11개의 파라미터로 표현되는 것이다.

$$q_{LM} = \{w_{sky}, t, w_{sun}, \beta, x, l_{sun}\}. \quad (4)$$

우리는 각 HDR 파노라마 이미지마다 11개의 파라미터가 레이블링 되어있는 데이터셋[14]을 이용하였다. 우리는 실외 환경에서의 태양의 위치를 추정하는데 Sky Model의 Parameter가 미치는 영향에 대하여 파악하기 위해 Sky Parameter와 태양의 위치를 함께 딥 러닝으로 추론하는 기존의 방법[4, 5]과 오로지 태양의 위치만을 추정하는 방법을 비교 분석하였다.

2.2 Tone Mapping

우리는 단일 LDR 이미지를 통해 HDR 파노라마 환경에서 계산된 Sky Parameter와 태양의 위치를 추정한다. 따라서 가지고 있는 64bit-float형의 HDR 파노라마 이미지를 8bit-int형의 LDR 이미지로 Tone Mapping하여 사용하였다. Tone Mapping 기법으로는 Drago Tone mapping[7]을 활용하였으며 γ 의 경우 [2.5, 3.5, 4.5, 5.5, 6.5]의 각각의 값에 대응되는 LDR 파노라마 이미지를 생성한 후 직접 가장 적합한 밝기를 가졌다고 판단되는 이미지를 택하여 사용하였다. 그림 1은 Tone Mapping의 예시이다.

2.3 Shadow Information

우리는 딥 러닝 모델에 추가적인 그림자 정보가 주어지면 더 뛰어난 학습이 진행된다는 가정 하에 그림자 정보를 새롭게 획득하기 위해 딥 러닝 기반의 그림자 추출 기법을 활용하였다. 사용한 신경망은 BDRAR(Bidirectional Feature Pyramid Network with Recurrent Attention Residual Modules for Shadow Detection)[6]이며, SBU Dataset[9]을 이용하여 모델을 학습시킨 뒤, Crop된 우리의 LDR photo로 Inference를 진행하여 그림자 정보를 획득하였다.

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				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	

표 1. 실험 결과: 각 실험 별 MAE(Mean Angular Error)를 나타낸 도표로, 사용한 네트워크 구조, 그림자의 입력 포함 여부, 타겟 파라미터 등에 따라 실험을 구분하여 표기하였음. 5, 6번째 실험은 각각 Y. Hold-Geoffroy et al.[4], Zhang et al.[5]에서 제시된 실험과 동일한 환경으로 구성됨.

학습 세부 사항은 다음과 같다. 우리는 Train Batch Size를 8로 설정하여 3,000번의 Epoch로 반복 학습하였고, Optimizer는 SGD를 0.9의 momentum, 그리고 0.0005의 weight decay로 설정하여 학습을 진행하였다. 학습은 Tesla V100을 기준으로 약 30분이 소요되었으며 기타 학습 환경의 경우 모두 BDRAR[6]의 기본 설정과 동일하게 진행하였다.

III. 광원 추적

3.1 데이터셋 구성

공개되어있는 데이터셋인 LAVAL HDR Dataset[14]를 이용하였으며, 총 205장의 HDR Panorama 이미지를 사용하였다. HDR Panorama 이미지를 LDR로 Tone Mapping한 후, LDR Panorama 이미지 한 장 당 7장의 이미지를 320*240 크기로 Random crop 하였다. Cropping을 통해 생성된 LDR 이미지는 총 1,435장으로 이 이미지 각각을 사전 훈련된 BDRAR[6]의 입력으로 사용하여 그림자를 추정 및 획득하였다. 그림 3은 네트워크의 입력 데이터를 구성한 방법을 도식화한 것으로, 320*240 크기로 Crop된 LDR 이미지와 해당 이미지로부터 검출된 Gray-Scale 그림자 이미지가 합쳐진 4채널 이미지를 네트워크의 입력으로 가져가게 된다.

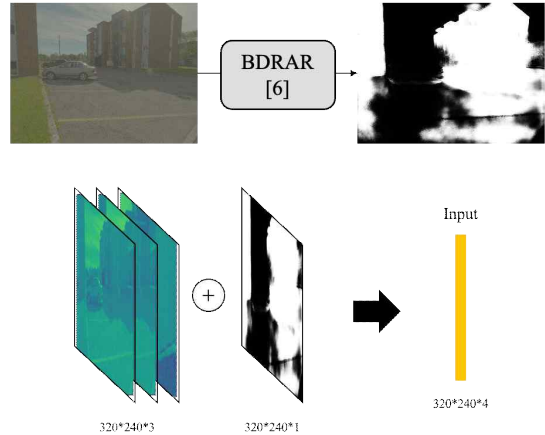


그림 3. BDRAR[6]를 이용한 그림자 검출과 입력 데이터 구성

3.2 딥 러닝 네트워크 구조

우리는 CNN 기반의 2개의 딥 러닝 네트워크를 이용하여 학습을 진행하였다. 그림 2는 해당 네트워크의 구조를 도식화 한 것이다. CNN[4] 네트워크의 경우, Y. Hold-Geoffroy et al.[4]에서 제시된 간단한 합성곱 신경망으로 7개의 컨볼루션 Layer와 2개의 Fully-Connected Layer를 거쳐 태양 위치의 분포와 9개의 Sky Parameter를 획득한다. AlexNet[5] 네트워크는 Zhang et al.[5]에서 제시된 네트워크인데, AlexNet[index]의 구조에서 마지막 부분이 수정된 형

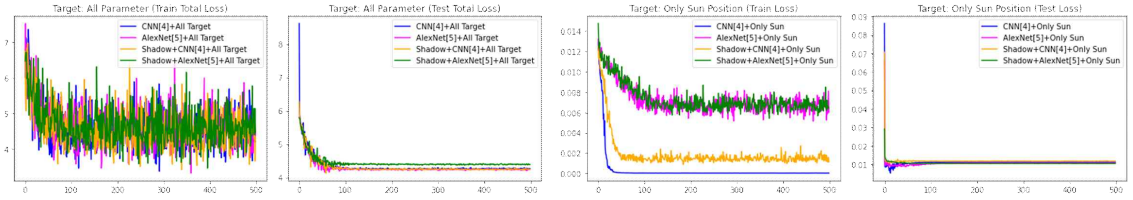


그림 4. 좌측 2개의 그래프는 Sky Parameter와 Sun Position을 모두 Target으로 한 모델의 Train/Test Loss 그래프이며, 우측 2개의 그래프는 Sun Position만을 Target으로 한 모델의 Train/Test Loss 그래프임.

태로 CNN[4] 네트워크와 마찬가지로 태양 위치의 분포와 9개의 Sky Parameter를 획득한다.

우리는 학습을 위해 다음의 Loss Function(수식 (5))을 사용하였다. 해당 Loss Function에 대한 서술은 CNN[4]를 참조하였다.

$$L(s^*, q^*, s, q) = L(s^*, s) + \beta L(q^*, q), \quad (5)$$

이 때, $\beta = 160$ 으로 s 의 bins의 수를 곱하여 값을 보정하였다. Ground Truth 태양의 위치 s^* 는 각각의 bin s_j 에 대해 아래의 수식과 같이 계산된다.

$$s_j^* = e^{xI_s^* T_{I_j}}, \quad (6)$$

또한 $\sum_j s_j^* = 1$ 이 되도록 정규화를 수행하였다.

수식 (6)에서 우리는 $x = 80$ 으로 정의하였으며, 수식 (5)의 q^* 는 Ground Truth Sky Parameter를 의미한다.

우리는 $L(q^*, q)$ 의 계산을 위해 MSE Loss를 활용하였으며, 태양의 위치에 대한 $L(s^*, s)$ 를 계산하기 위해 KL Divergence Loss를 활용하였다. 이 때의 KL Loss는 태양의 위치에 대한 분포를 학습하도록 하여 학습 성능 향상을 기대하였다.

3.3 학습 세부 사항

실험은 총 8개로 나뉘어 진행되었으며, ‘네트워크 구조’, ‘그림자 정보의 입력 포함 여부’, ‘타겟 파라미터의 종류’의 총 3가지 기준에 따라 나뉘어 진행되었다. 네트워크 구조의 경우 Hold-Geoffory et

al.[4]에서 사용된 7층 CNN[4] 구조 혹은 Zhang et al.[5]에서 사용된 AlexNet[5] 구조를 사용하였으며, 그림자 정보의 입력이 포함된 경우는 4채널 입력, 포함되지 않은 경우는 그림자 1채널을 제외한 RGB 3채널을 입력으로 하였다. 예측 타겟의 경우 Sky Parameter와 Sun Position을 모두 예측하여 총 11개의 파라미터를 예측하는 경우와 Sky parameter를 예측하지 않고 오로지 Sun Position 파라미터 2개만을 예측하는 경우로 진행되었다.

Train 데이터셋과 Test 데이터셋은 7:3 비율로 나누었으며, 모든 실험은 동일한 데이터셋 환경에서 진행되었다. 각 실험은 500 epoch만큼 진행되었으며 Train Batch Sizes는 64로 설정하였다. 초기 Learning Rate는 0.01로 설정하여 15 epoch마다 0.5를 곱해주었다.

IV. 평가

4.1 Loss Graph

그림 4는 모든 실험에 대한 Train Loss와 Test Loss를 시각화한 그래프이다(그림 5의 경우 All Parameter를 타겟으로 한 모델에 대해서만 Sun Loss를 따로 시각화). Train Loss에 비해 Test Loss가 안정적으로 줄어드는 모습을 볼 수 있으며, 특히 Sun Position만을 타겟으로 하는 모델보다 Sky Parameter와 Sun Position을 모두 타겟으로 하는 모델이 더 불안정적으로 Loss가 감소하거나 수렴하지 못하는 모습을 확인할 수 있다.

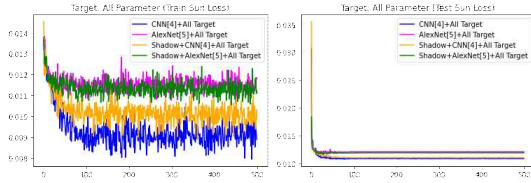


그림 5. Sky Parameter와 Sun Position을 모두 타겟으로 하는 모델의 Sun Loss 그래프

4.2 Mean Angular Error (MAE)

표 1은 모든 실험에 대한 Mean Angular Error를 계산하여 이를 표기한 결과를 나타낸다. 가장 성능이 좋은 경우는 Shadow 정보 없이 Sky Parameter와 Sun Position을 모두 타겟으로 하여 CNN[4] 신경망으로 학습을 진행한 경우였으며, 그 뒤의 경우 Shadow 정보와 함께, 그리고 Sun Position만을 타겟으로 하여 딥 러닝 학습을 진행한 경우임을 확인할 수 있다. 한 가지 발견할 수 있는 사실은 오로지 Sun Position만을 타겟으로 하는 딥 러닝 모델에서는 그림자 정보가 광원 추적에 크게 효과적이라는 사실이다. 그림 6은 누적 각도 오차 그래프로서 각 실험에 대한 성능을 직관적으로 확인할 수 있도록 한다.

V. 결론

본 보고서에서는 실외 환경에서의 광원 추적을 위해 제안된 딥 러닝 네트워크를 소개하고, 또 다른 방법의 실외 환경 광원 추적 방법론을 제안한다. Sky Parameter를 반드시 필요로 하였던 기존의 방법과 비교하여 Sky Parameter에 대한 추정 없이도 Shadow 정보를 이용하면 충분히 그에 버금가는 추적 성능을 달성할 수 있다는 사실을 보인다.

이 연구는 추후 발전 가능성이 굉장히 높다. 그 근거는 다음과 같다. 첫째, 그림자 검출의 성능을 향상시키면 성능 향상에도 도움이 될 것이다. 그림자 정보를 이용하면 태양 위치 추적 성능 향상에 도움이 된다는 사실은 실험을 통해 증명하였다. 둘째, 그림자를 딥 러닝 모델에 주는 방식을 개선할 수 있다. 지금은 단순히 3채널 입력값에 채널을 더

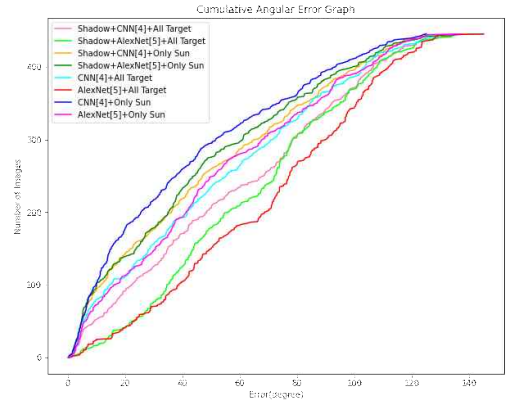


그림 6. Cumulative Angular Error Graph

해 4채널 입력을 수행하지만, 더 나은 방식으로 딥 러닝 모델 안에서 처리될 수 있다. 셋째, 데이터셋의 양을 늘리고 더 견고히 하여 성능 향상을 기대할 수 있을 것이다.

그럼에도 불구하고 이 보고서를 통해 충분히 우리가 제시하는 방법이 유의미하다는 점과 향후 발전 가능성이 높다는 사실을 보일 수 있다는 점에 이 보고서는 큰 의의를 가진다.

VI. 부록

Github 주소: <https://github.com/sksmslhy/Outdoor-Illumination-Estimation-using-Shadow-Information>

LAVAL Outdoor HDR Dataset: <http://outdoor.hdrdb.com/>

SBU Dataset: https://www3.cs.stonybrook.edu/~cvl/projects/shadow_noisy_label/index.html

참 고 문 헌

- [1] Lalonde, J. F., Efros, A. A., & Narasimhan, S. G. (2009, September). Estimating natural illumination from a single outdoor image. In 2009 IEEE 12th International Conference on Computer Vision (pp. 183-190). IEEE.
- [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] Hosek, L., & Wilkie, A. (2012). An analytic model for full spectral sky-dome radiance. *ACM Transactions on Graphics (TOG)*, 31(4), 1-9.
 - [4] 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).
 - [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).
 - [6] 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).
 - [7] Drago, F., Myszkowski, K., Annen, T., & Chiba, N. (2003, September). Adaptive logarithmic mapping for displaying high contrast scenes. In *Computer graphics forum* (Vol. 22, No. 3, pp. 419-426). Oxford, UK: Blackwell Publishing, Inc.
 - [8] Preetham, A. J., Shirley, P., & Smits, B. (1999, July). A practical analytic model for daylight. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (pp. 91-100).
 - [9] Vicente, T. F. Y., Hou, L., Yu, C. P., Hoai, M., & Samaras, D. (2016, October). Large-scale training of shadow detectors with noisily-annotated shadow examples. In *European Conference on Computer Vision* (pp. 816-832). Springer, Cham.
 - [10] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
 - [11] Zhang, J., & Lalonde, J. F. (2017). Learning high dynamic range from outdoor panoramas. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4519-4528).
 - [12] Calian, D. A., Lalonde, J. F., Gotardo, P., Simon, T., Matthews, I., & Mitchell, K. (2018, May). From faces to outdoor light probes. In *Computer Graphics Forum* (Vol. 37, No. 2, pp. 51-61).
 - [13] 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).
 - [14] 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).