

## 노말맵과 알베도맵 추정을 통한 영상 기반 재조명

안흥현 이영현 김영서 \*강동완

서울과학기술대학교

\*dongwann@seoultech.ac.kr

## Image-based relighting using normal map and albedo map prediction

Honghyun Ahn Younghyeon Lee Youngseo Kim \*Dongwann Kang

Seoul National University of Science and Technology

## 요약

영상에 새로운 광원을 추가하거나 기존의 광원을 변경하여 영상 내 오브젝트들에 적용된 조명을 변경하는 것을 영상 기반 재조명이라 한다. 하지만, 영상에는 재조명을 위해 필요한 광원과 오브젝트들의 3차원 기하 정보가 부재하다는 문제가 있다. 이를 해결하기 위해, 본 연구에서는 영상으로부터 재조명에 필요한 요소들을 추정하는 접근법을 취한다. 오브젝트 표면의 노말과 알베도는 조명의 주 요소이지만 광원에는 독립적이므로 새로운 광원에 대한 재조명을 가능케 한다. 따라서 본 연구는 영상으로부터 노말맵과 알베도맵을 추정한 뒤, 이를 이용하여 영상 기반 렌더링하는 영상 재조명 방법을 제안한다. 조건부 적대적 생성망을 다양한 조명 환경에서 렌더링된 3차원 오브젝트 영상들과 그에 대응하는 노말맵, 알베도맵을 이용해 학습함으로써, 임의의 영상에 대한 노말맵과 알베도맵 추정기를 생성한다. 이를 통해 추정된 노말맵과 알베도맵은 3차원 공간상에서 새로운 광원에 대해 렌더링됨으로써 재조명 영상을 생성한다. 마지막으로, 영상 기반으로 재조명된 영상과 ground truth와의 비교 실험을 통해 본 연구에서 제안한 방법이 유효함을 확인한다.

## 1. 서론

영상 기반 재조명(image-based relighting)은 2차원 영상내 오브젝트들에 적용된 조명을 새로운 광원을 추가하거나 기존 광원의 속성을 수정함으로써 변경하는 것을 의미한다. 디지털 영상 미디어 및 영상 기반 소셜 서비스 등의 대중화, 그리고 가상/증강현실 등 새로운 영상 기술의 등장으로 인해 디지털 영상의 가공에 대한 수요가 증가하고 있는데, 디지털 영상들을 서로 합성하거나 수정하는 가공 과정에서 각각의 영상에 사용된 조명의 불일치는 시각적으로 어색한 영상을 야기한다. 이를 해결하기 위한 영상의 재조명은 일반적으로 숙련된 디자이너의 기술에 의존하는데, 이는 컴퓨터에 의해 자동화된 영상 재조명 기술 개발의 필요성을 증대시킨다. 이에 본 연구에서는 영상에 기반한 재조명 기술을 제안한다.

현실에는 조명하의 오브젝트의 표면 색상에 연관된 다양한 요소들이 존재하지만, 이를 보다 단순하게 근사화하는 3차원 컴퓨터그래픽스의 지역 조명 모델의 관점에서는 오브젝트 표면의 색상, 조명과 뷰잉 조건을 주된 요소로 꼽을 수 있다. 따라서 이들이 주어진다면 조명 환경을 바꿔 오브젝트를 다시 렌더링, 즉 재조명하는 것이 가능하다. 하지만, 2차원 영상에는 이와 같은 3차원 기하 정보들이 부재하다는 문제를 갖는다. 이를 해결하기 위해, 본 연구는 지역 조명 모델에 필요한 요소들을 주어진 2차원 영상으로부터 추정한 뒤, 조명을 바꾸어 렌더링함으로써 재조명하는 접근 방법을 취한다.

본 연구에서 영상으로부터 추정하는 지역 조명 모델의 요소들은 오브젝트 표면의 노말벡터와 알베도이다. 오브젝트 표면의 노말벡터와 알

베도는 광원에 독립적인 지역 조명 모델의 주 요소들이므로, 영상으로부터 이들의 추정은 새로운 광원에 대한 재조명을 가능케 한다. 2차원 영상으로부터 노말맵과 알베도맵을 추정하기 위해, 본 연구에서는 조건부 적대적 생성망을 사용하는 방법을 제안한다. 조건부 적대적 생성망을 이용해 조명이 적용된 오브젝트의 영상과 이에 대응되는 노말맵과 알베도맵과의 상관관계를 학습함으로써 임의의 영상에 대응하는 노말맵과 알베도맵을 추정할 수 있다. 추정된 노말맵과 알베도맵은 스크린 공간 기반 렌더링을 통해 새로운 조명 환경에서의 오브젝트 렌더링의 근사를 제공한다.

본 논문의 구성은 다음과 같다. 2절에서는 영상 기반 재조명과 관련된 연구들을 살펴본다. 3절에서는 본 연구에서 제안하는 영상 기반 재조명을 노말맵 및 알베도맵 추정과 이를 통한 재조명 렌더링의 관점에서 자세히 설명한다. 4절에서는 본 연구의 결과를 5절에서는 결론과 향후 연구를 다룬다.

## 2. 관련연구

일반적으로 영상 기반 재조명은 영상으로부터 얻은 빛 전달 행렬과 새로운 빛 벡터를 곱하는 것으로 수행된다.[1] 따라서 영상 기반 재조명 연구들은 영상으로부터 빛 전달 행렬을 계산하는 효과적인 방법을 제안한다. 영상 기반 재조명 분야에서는 주어진 장면에 대해 다양한 조명하에서의 영상을 획득하여 빛 전달 행렬을 구하는 연구들이 수행되어왔다.[2-5] 이 연구들은 영상 획득을 위해 특별하게 고안된 장비를 필요로 한다는 점에서 일반적인 영상을 대상으로 재조명을 하고자 하는 본 연구

의 목표와 거리가 있다. 또한, 임의의 영상에서 동작할 수 있는 재조명을 목표로 하는 본 연구와 달리, 주어진 하나의 장면에 대한 다양한 조명의 영상셋을 통해 정교한 재조명을 수행하며, 그 재조명은 다른 장면에 적용될 수 없다는 점에서 차이가 있다.

영상 기반 재조명은 안면 영상을 대상으로도 연구되어졌다.[6] 단일 안면 영상에 대해 재조명을 수행하기 위해 심층 합성곱 신경망을 사용한 연구에서는 Stacked hourglass network이 안면 기하와 조명을 추정하도록 학습되어 사용되었다. 따라서 학습된 모델은 임의의 단일 안면 영상에 대해 주어진 조명하에서 재조명된 영상을 생성한다. 이 연구는 임의의 영상에 대해 재조명을 수행하고자 하는 본 연구와 유사한 목표를 공유한다. 하지만, 비교적 제한된 기하구조를 갖는 안면 영상에 비해, 본 연구에서 다루고자 하는 오브젝트 영상들은 매우 복잡하고 다양한 기하구조를 갖는다는 점에서 주된 차이점이 있다.

### 3. 노말맵과 알베도맵 추정을 통한 재조명

본 연구에서 제안하는 영상 기반 재조명 방법은 조건부 적대적 생성망을 이용하여 노말맵과 알베도맵을 추정하는 부분과 이를 이용하여 영상을 재조명하는 부분으로 구성된다. 다음 절에서 각각의 상세에 대해 다룬다.

#### 3.1 조건부 적대적 생성망을 이용한 노말맵과 알베도맵 추정

입력 영상으로부터 노말맵과 알베도맵을 추정하기 위해 본 연구에서는 조건부 적대적 생성망을 사용한다. 조건부 적대적 생성망은 적대적 생성망의 변형된 모델로서 주어진 조건에서의 데이터 확률 분포를 추정한다. 이때, 영상을 조건을 사용할 경우, 입력 영상과 조건 영상의 상관관계를 학습할 수 있다. 따라서, 조건부 적대적 생성망은 주어진 오브젝트 영상에 대해 그에 대응하는 노말맵과 알베도맵을 생성하는 것을 목표로 하는 본 연구에 적합하다. 본 연구에서는 pix2pix 조건부 적대적 생성망을 이용하여 오브젝트 영상과 노말맵 및 알베도맵 간의 상관관계를 학습한다.[7]

본 연구에서의 학습을 위해 오브젝트 영상과 그에 대한 ground truth 노말맵과 알베도맵이 요구된다. 하지만 실사 영상을 학습에 사용하고자 할 경우, ground truth 노말맵과 알베도맵을 획득하는 것이 용이하지 못하다. 본 연구에서는 학습 데이터셋 구축 용이성을 고려하여, 3차원 오브젝트 모델들로부터 오브젝트 영상과 노말맵, 알베도맵을 생성하여 데이터셋으로 활용하는 접근방법을 취한다.

먼저, 주어진 조명 환경 하에서 3차원 모델을 다양한 시점에서 렌더링을 수행한 장면들을 2차원 영상들로 저장한다. 이 때, 조명 환경을 바꿔가며 이 과정을 반복수행함으로써 실사 영상에서 나타날 수 있는 다양한 조명에 대응할 수 있도록 한다. 다음으로, 앞서 사용된 시점들의 장면들에 대해 오브젝트 표면의 노말벡터를 렌더링함으로써 노말맵 데이터셋을 얻고, 조명효과를 제거한 오브젝트 표면 고유의 색상을 렌더링함으로써 알베도맵 데이터셋을 얻는다(그림 1). 본 연구에서는 다양한 범주의 오브젝트들에 대한 3차원 모델 대용량 데이터셋인 ShapeNet을 사용하여 데이터셋을 구축한다.[8]

그림 2는 pix2pix 조건부 적대적 생성망을 이용하여 생성한 노말맵

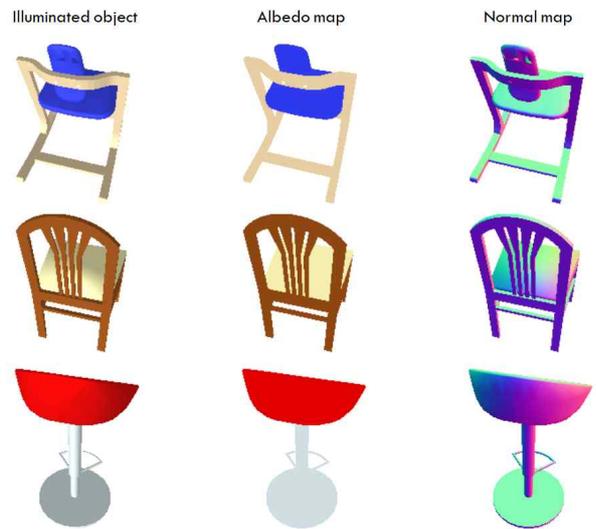


그림 1. 3차원 모델로부터 생성한 학습 데이터셋: 오브젝트 영상(왼쪽), 알베도맵(중앙), 노말맵(오른쪽). 노말맵의 x,y,z 성분은 r,g,b 색상으로 정규화되어 표현되었다.

과 알베도맵 생성기를 도식화한 것이다. 3차원 모델에 대한 렌더링 영상과 이에 대응하는 노말맵과 알베도맵 간의 상관관계를 학습함으로써, 임의의 렌더링 영상에 대응하는 노말맵과 알베도맵을 생성 가능하다.

#### 3.2 노말맵과 알베도맵을 이용한 영상 재조명

Phong의 조명 모델에 의하면, 3차원 오브젝트에 대한 조명(illumination)에 기여하는 것은 오브젝트 표면의 색상과 재질, 노말벡터, 그리고 시점(view point)과 광원이다.[9] 실사 영상에서는 시점에 따른 투영의 차이가 발생하지만, 평행 투영(parallel projection)되었다 간주할 경우 조명에서 시점 요소를 제거하는 근사가 가능하다. 비슷한 방식으로, 광원의 종류를 태양광과 같이 방향성을 가진 광원으로 간주하면 조명에서 광원의 위치를 제거할 수 있다. 본 연구에서는 평행 투영과 임의의 방향성을 갖는 빛을 이용한 렌더링 상황을 가정하여, 3.1절에서 생성한 오브젝트 표면의 노말벡터(노말맵)와 색상(알베도맵)으로부터 조명이 적용된 오브젝트 렌더링을 수행한다.

노말벡터와 알베도맵은 2차원 스크린에 투영된 3차원 오브젝트 표면의 정보(노말벡터와 색상)를 알려줄 뿐, 3차원 기하정보를 제공하지는 못한다. 따라서 이를 통한 일반적인 3차원 오브젝트 모델에 대한 렌더링은 불가능하다. 이를 해결하기 위해, 본 연구에서는 스크린 공간에서의 렌더링 방식을 통해 3차원 오브젝트 모델의 렌더링을 수행한다. 먼저, 스크린 공간의 크기에 대응하는 3차원 공간상의 평면 기하를 생성한다. 다음으로, 노말맵과 알베도맵을 그 평면의 텍스처로 설정한 뒤, 래스터화된 평면의 fragment들에 대해 각 텍스처로부터 노말벡터와 색상을 취한다. 마지막으로, 노말벡터와 오브젝트 표면 색상을 이용하여, 주어진 광원에 대해 Phong 조명 모델의 난반사요소와 주변광 요소를 구함으로써 렌더링을 수행하고, 스크린 상의 장면을 영상으로 저장한다(그림 3).

3.1절에서 생성한 노말맵은 단위 벡터로 저장된 ground truth 노말맵과 달리 그 크기가 제각각이다. 따라서 본 연구에서는 생성된 노말맵의 벡터들을 단위벡터로 정규화한 뒤 사용한다.

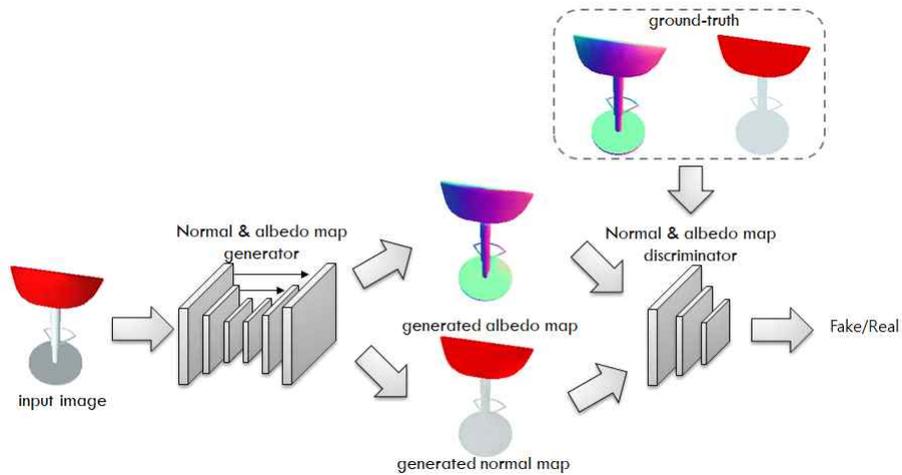


그림 2. 노말맵과 알베도맵 생성을 위한 조건부 적대적 생성망

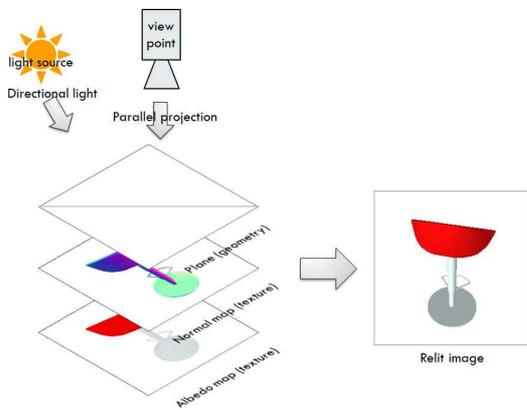


그림 3. 재조명을 위한 노말맵과 알베도맵 이용 렌더러

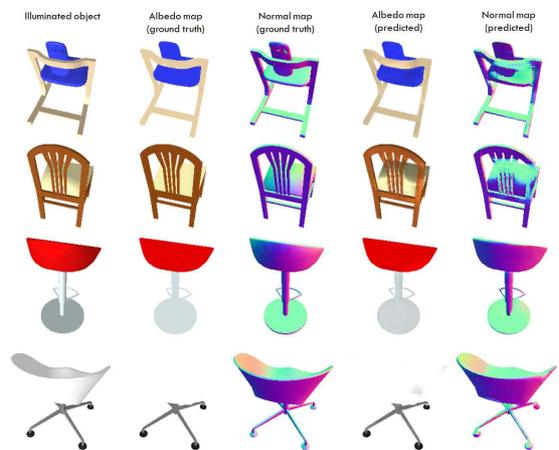


그림 4. 노말맵과 알베도맵 추정 결과

#### 4. 결과

본 연구에서는 ShapeNet의 다양한 오브젝트들 중에서 chair 범주에 속하는 모델 6,777종에 대해 데이터셋을 생성했다. 하나의 오브젝트에 대해 총 12개의 시점과 6개의 광원 방향, 3가지의 조명 환경을 조합하여 216장의 오브젝트 렌더링 영상을 생성하였고, 그에 대응하는 노말맵과 알베도맵을 생성하였다. 따라서 노말맵과 알베도맵 학습에 사용된 데이터셋은 오브젝트 영상만 약 140만장에 달한다. 데이터셋 생성을 위한 렌더러의 뷰포트 크기는 가로, 세로 모두 256픽셀로 설정되었고, 동일한 크기의 영상들이 렌더링되어 저장되었다. 조건부 적대적 생성망의 학습에는 Geforce RTX 3090 4개로 구성된 시스템이 사용되었고, 학습에 약 120시간이 소요되었다.

그림 4는 주어진 오브젝트 영상(그림의 맨 왼쪽열)에 대한 노말맵과 알베도맵의 추정 결과를 보여준다. 알베도맵의 경우, 대체로 ground truth에 근접한 결과를 추정함을 볼 수 있으나, 다소 복잡한 기하구조를 가진 오브젝트에서는 노말맵의 추정에 오차가 있음을 확인할 수 있다. 그림 5는 추정된 노말맵과 알베도맵을 이용해 입력 오브젝트 영상에 조명을 적용한 결과, 즉 생성된 재조명 영상을 보여준다. 입력 오브젝트 영상을 렌더링할 때와 동일한 조명 조건을 적용하여 ground truth 렌더링 결과와 비교함으로써 본 연구에서 제안한 재조명 방법의 성능을 확인하였다. 본 연구에서 사용된 데이터셋에 대해 ground truth 렌더링 결

과와 재조명된 영상 간에 측정된 평균 제곱근 오차(RMSE) 평균값은 약 0.084로, 그림에서 보여주듯 크게 두드러지는 시각적 차이가 발생하지는 않았다. 하지만, 노말맵과 알베도맵의 추정에서 오차가 발생한 일부 지점들에서 재조명 오차가 발생했음을 볼 수 있다. 또한, 조건부 적대적 생성망을 이용해 생성한 노말맵과 알베도맵에서 나타나는 artefact들이 재조명에 영향을 주고 있음을 발견할 수 있다.

#### 5. 결론

본 연구에서는 영상 기반 재조명 방법을 제안하였다. 2차원 영상에 부재한 기하정보를 얻기 위해, 조건부 적대적 생성망을 이용해 입력 오브젝트 영상에 대응하는 노말맵과 알베도 맵을 생성하였다. 이 과정에서 조건부 적대적 생성망의 학습을 위해 3차원 오브젝트들을 다양한 환경에서 렌더링한 데이터셋을 사용하였다. 그리고 스크린 공간 기반 렌더러를 이용해 노말맵과 알베도맵을 새로운 조명 환경 하에서 렌더링 함으로써 재조명된 영상을 얻었다.

본 연구에서 사용한 조건부 적대적 생성망은 artefact를 갖는 영상을 생성한다는 한계를 갖고, 이는 재조명 성능에 악영향을 미친다. 향후 연구에서는 망의 개선을 통해 artefact를 억제하고 노말맵과 알베도맵 추정 성능을 높이는 것을 연구하고자 한다.



그림 5. Ground truth와 재조명된 영상간의 비교

## Acknowledgment

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00994, 이용 환경을 반영하는 자율적 VR·AR 콘텐츠 생성 기술개발)

## 참고문헌

- [1] Ng, Ren, Ravi Ramamoorthi, and Pat Hanrahan. "All-frequency shadows using non-linear wavelet lighting approximation." ACM SIGGRAPH 2003 Papers. 2003. 376-381.
- [2] Debevec, Paul, et al. "Acquiring the reflectance field of a human face." Proceedings of the 27th annual conference on Computer graphics and interactive techniques. 2000.
- [3] Wenger, Andreas, et al. "Performance relighting and reflectance transformation with time-multiplexed illumination." ACM Transactions on Graphics (TOG) 24.3 (2005): 756-764.
- [4] Wang, Jiaping, et al. "Kernel Nystrom method for light transport." ACM SIGGRAPH 2009 papers. 2009. 1-10.
- [5] Ren, Peiran, et al. "Image based relighting using neural networks." ACM Transactions on Graphics (ToG) 34.4 (2015): 1-12.
- [6] Zhou, Hao, et al. "Deep single-image portrait relighting." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.
- [7] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [8] Chang, Angel X., et al. "Shapenet: An information-rich 3d model repository." arXiv preprint arXiv:1512.03012 (2015).
- [9] Phong, Bui Tuong. "Illumination for computer generated pictures." Communications of the ACM 18.6 (1975): 311-317.