

불균일 안개 영상 합성을 이용한 딥러닝 기반 안개 영상 깊이 추정

최영철^{1O} 백지현¹ 주광진² 이동건² 황경하³ 이승용^{2*}

¹포스코ICT ²포항공과대학교 ³영남대학교

¹{ycchoi, jeehyun100}@poscoict.com ²{gwangjin, dalelee, leesy}@postech.ac.kr ³ghhwang@yu.ac.kr

**Non-Homogeneous Haze Synthesis for
Hazy Image Depth Estimation Using Deep Learning**
Yeongcheol Choi^{1O} Jeehyun Paik¹ Gwangjin Ju² Donggun Lee²
Gyeongha Hwang³ Seungyong Lee^{2*}
¹POSCO ICT ²POSTECH ³Yeungnam University

요 약

영상의 깊이를 추정하는 다양한 영상 분석의 기반이 되는 기술이다. 딥러닝 모델을 활용한 분석 방법이 대두되면서, 영상의 깊이 추정 분야 또한 딥러닝을 활용하는 연구가 활발하게 이루어지고 있다. 현재 대부분의 딥러닝 영상 깊이 추정 모델들은 깨끗하고 이상적인 환경에서 학습되고 있다. 하지만 연무, 안개가 낀 열악한 환경에서도 깊이 추정 기술이 잘 동작할 수 있으면 이러한 환경의 데이터를 포함하여야 한다. 하지만 열악한 환경의 영상을 충분히 확보하는 것이 어려운 실정이며, 불균일한 안개 데이터를 얻는 것은 특히 어려운 문제이다. 이를 해결하기 위해, 본 연구에서는 불균일 안개 영상 합성 방법과 이를 활용한 단안 기반의 깊이 추정 딥러닝 모델의 학습을 제안한다. 안개가 주로 실외에서 발생하는 것을 고려하여, 실외 위주의 데이터 세트를 구축한다. 그리고 실험을 통해 제안된 방법으로 학습된 모델이 합성 데이터와 실제 데이터에서 깊이를 잘 추정하는 것을 보인다.

Abstract

Image depth estimation is a technology that is the basis of various image analysis. As analysis methods using deep learning models emerge, studies using deep learning in image depth estimation are being actively conducted. Currently, most deep learning-based depth estimation models are being trained with clean and ideal images. However, due to the lack of data on adverse conditions such as haze or fog, the depth estimation may not work well in such an environment. It is hard to sufficiently secure an image in these environments, and in particular, obtaining non-homogeneous haze data is a very difficult problem. In order to solve this problem, in this study, we propose a method of synthesizing non-homogeneous haze images and a learning method for a monocular depth estimation deep learning model using this method. Considering that haze mainly occurs outdoors, datasets mainly containing outdoor images are constructed. Experiment results show that the model with the proposed method is good at estimating depth in both synthesized and real haze data.

키워드: 깊이 추정, 불균일 안개, 안개 합성, 딥러닝

Keywords: depth estimation, non-homogeneous haze, haze synthesis, deep learning

1. 서론

영상 깊이 추정 기술은 영상 정보를 활용하여 영상 내 사물들의 거리를 추정하는 기술이다. 영상 깊이 추정 기술은 다양한 영상

분석 기술의 기반이 되는 기술이며 로봇, 자율주행, AR/VR 등의 산업이 발전함에 따라 중요성이 대두되고 있다. 특히 로봇, 자율주행과 같은 산업의 경우 다양한 외부 환경에서도 잘 작동할 수 있는 기술을 요구한다. 이에 따라 악조건의 환경에서도 잘 작동하는 깊이 추정 기술과 악조건의 환경을 개선하기 위한

*corresponding author: Seungyong Lee/POSTECH(leesy@postech.ac.kr)

기술이 중요해지고 있다[1]. 여러 악조건의 환경 중에서도 특히 안개는 영상에서 깊이를 추정할 수 있는 정보를 많이 손실시키기 때문에 깊이 추정을 어렵게 한다.

최근에는 딥러닝 기술이 발전함에 따라 이를 활용한 영상 깊이 추정 연구가 활발히 이루어지고 있어, 안개 상황에서도 깊이를 추정할 수 있는 딥러닝 모델에 대한 연구가 필요한 실정이다. 딥러닝 모델 학습을 위한 영상 깊이 추정 데이터 세트는 맑은 날씨에서 촬영된 것이 대부분이다. 이 때문에 악조건의 날씨에서는 영상의 깊이를 추정할 때 성능이 떨어질 수 있다는 단점을 가지고 있다. Fig. 1 과 같이 깊이 추정 모델이 NH-HAZE [2]의 동일한 장면에 대하여 깨끗한 영상(clean image)에 대해서는 깊이(depth)를 잘 추정하지만, 안개 영상(hazy image)에서는 영상의 깊이를 제대로 추정하지 못하는 것을 볼 수 있다. 이러한 문제를 해결하기 위한 데이터 활용 접근 방법으로, 안개 영상이 포함된 데이터 세트를 구축하여 모델을 학습하는 방식을 많이 사용한다.

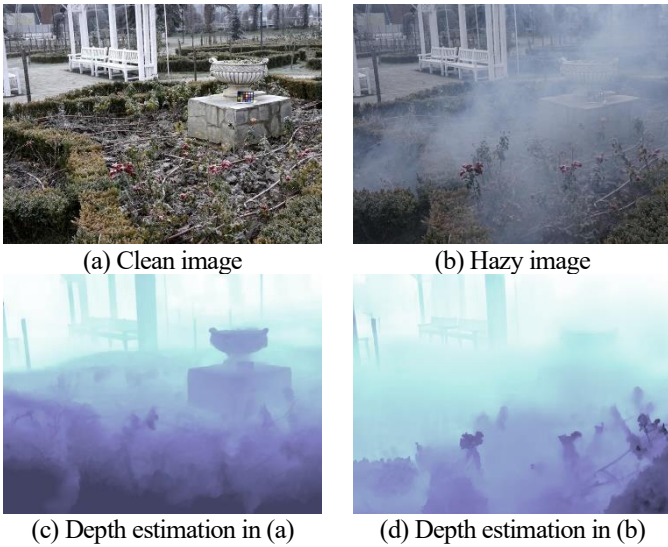


Figure 1: An example of the depth estimation result of the Boosting [3] model based on LeRes [4]. For haze images, the model does not estimate the depth well.

안개 영상이 포함된 데이터 세트를 구축하기 위해선 실제 안개 현상이 발생하는 장소와 시점을 찾아 영상을 촬영하고, 해당 영상에서의 ground truth 깊이를 얻어야 한다. 하지만 실제로 이런 방법으로 데이터 세트를 구축하는 것은 매우 어렵다. 이 때문에 안개 영상이 포함된 데이터 세트를 구축할 때, 주로 대기산란 공식 [5]을 활용하여 기존의 데이터 세트에 안개를 합성하는 방식이 많이 이용된다[6, 7]. 하지만 이런 합성 방식은 균일한 안개 영상(homogeneous hazy image)을 합성하므로, 불균일한 안개 영상(non-homogeneous hazy image)을 합성할 수 없다.

본 연구에서는 대기산란 공식 [5]과 dark channel prior(DCP) [8]를 기반으로 불균일한 안개 영상을 합성하는 방법을 제안한다. 본 연구에서는 제안하는 합성 방법으로 기존 데이터를 합성 및 증폭하여 불균일한 안개 영상에서도 깊이를 추정할 수 있도록 모델을 개선한다. 그리고 실제 불균일 영상에서도 깊이를 추정할 수 있도록 한다. 이를 위해 깊이 추정용 합성 데이터 세트(synthetic haze dataset)와 실제 데이터 세트(real haze dataset)를 구축한다. 실제 데이터 세트는 NH-HAZE [2], O-HAZE [9] 데이터 세트를 활용하여 구축한다. 합성 데이터 세트의 경우 Make3D [10, 11] 데이터 세트를 활용하여 구축한다.

본 연구의 기여는 다음과 같다.

- 본 연구에서는 대기산란 공식 [4]과 DCP [8]를 활용하여 불균일한 안개 영상을 합성하는 방법을 제안한다.
- 본 연구에서는 제안하는 합성 방식을 활용하여, 깊이 추정 모델 학습을 위한 불균일 안개 데이터 세트를 구축한다.
- 본 연구에서는 구축한 데이터 세트로 기존의 깊이 추정 모델을 학습하여 성능이 개선됨을 확인한다.

2. 관련 연구

2.1 대기산란 공식(Atmospheric Scattering Model)

안개 영상은 대기 중의 입자로부터 산란된 빛과 물체로부터 반사된 빛에 의해 취득된다. 카메라로부터 거리가 멀고 대기 중에 퍼져 있는 입자의 밀도가 클수록 빛의 산란이 강해지고 영상의 대조도(contrast)와 채도(saturation)가 감소한다. 안개 영상을 표현하기 위한 가장 일반적인 방법으로 대기산란 공식 [5]이 사용된다[8, 12, 13].

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (1)$$

이때, x 는 2 차원 벡터로 높이 H , 너비 W 의 픽셀 크기를 가지는 영상의 각 픽셀(pixel) 위치를 의미한다. $I(x)$ 는 안개 영상을 의미하고, $J(x)$ 는 $I(x)$ 에서 안개가 제거된 깨끗한 영상을 의미한다. $J(x)$ 와 $I(x)$ 는 각 픽셀 별로 RGB 값을 가지는 $H \times W \times 3$ 형태의 3 차원 배열이다. A 는 대기 산란광(global atmospheric light)을 의미하며 각 RGB에 대응하는 3 가지 값이 표현된 벡터이다. $t(x)$ 는 대기 투과율(transmission map)을 의미하며 $H \times W$ 형태를 가지는 2 차원 배열이다. 대기산란 공식 [5]은 안개 영상을 깨끗한 영상과 대기 산란광의 대기 투과율에 따른

요소별 비율의 합으로 나타낸 식이다. 즉, $t(x)$ 의 값이 작을수록 대기 산란광이 많은 영향을 미쳐 심한 안개 영상이 만들어진다.

$t(x)$ 는 $t(x) = e^{-\beta d(x)}$ 로 표현한다. 여기서 $d(x)$ 는 영상의 깊이를 나타내며 HxW 형태를 가지는 2 차원 배열이다. β 는 대기산란 공식의 계수인데, 안개를 유발하는 입자의 대기 중 밀도를 나타낸다고 할 수 있다. β 값은 각 RGB 영역에 대응하는 3 가지 값이 표현된 벡터이다. 따라서 모든 영역에 β 균일하게 적용된다.

2.2 Dark Channel Prior

DCP [8]는 안개가 없는 깨끗한 영상의 경우 각 픽셀은 R, G, B 값 중 하나가 매우 낮은 경향을 보인다는 영상의 특징이다. 이를 Eq. (2)와 같은 식으로 나타낼 수 있다.

$$J^{dark}(x) = \min_{c \in \{r, g, b\}} (\min_{y \in \Omega(x)} J^c(y)) \quad (2)$$

이때, x 는 2차원 벡터로 각 픽셀의 위치를 의미한다. $J^{dark}(x)$ 는 깨끗한 영상의 dark channel, $J^c(y)$ 는 픽셀 x 의 R, G, B 중 하나를 의미하며 $\Omega(x)$ 는 x 를 중심으로 한 일정 범위 안의 픽셀 집합을 의미한다. DCP [8]을 Eq. (1)에 적용하면 Eq. (3)과 같이 안개 영상으로부터 각 픽셀의 대기 투과율을 구할 수 있다.

$$t(x) = 1 - \min_{c \in \{r, g, b\}} (\min_{y \in \Omega(x)} \frac{I^c(y)}{A^c}) \quad (3)$$

2.3 영상 깊이 추정(Depth Estimation)

깊이 추정 기술은 크게 양안(stereo) 방식과 단안(mono) 방식으로 나뉜다. 양안 방식은 사람이 보는 것처럼 동일한 장면에 대하여 두 개의 촬영 영상을 사용하여 영상의 깊이를 추정한다 [14, 15, 16]. 단안 방식은 단일 영상에서 단일 영상에서 깊이를 추정하는 방식이다. 일반적으로 양안 방식이 많이 사용되어 왔으나, 딥러닝 기술의 활용으로 단안 방식도 양안 방식만큼의 성능을 보인다 [17, 18, 19, 20, 21]. 그리고 단안 방식도 크게 영상 깊이 ground truth 값을 활용한 지도 학습(supervised learning) 방식과 준 지도 학습(semi-supervised learning) 방식으로 나뉜다.

일반적으로 딥러닝을 활용한 영상의 깊이 추론 모델은 안개 영상에서 성능이 좋지 않다. 이는 안개 환경의 데이터가 부족하기 때문이다. 본 연구에서는 안개를 영상에 합성하는 방법을 제안하여 이를 해결하고자 한다.

2.4 불균일 안개 영상

균일 안개와 불균일 안개는 안개를 유발하는 입자의 밀도가 영상 전체에 균일한지의 여부에 따라 구분할 수 있다. 이 밀도는 Eq. (1)에서 β 값으로 표현되는 부분이다. Eq. (1)에서는 β 값을 각 RGB 영역에 대응하는 3 가지 값이 합쳐진 벡터로 표현하는데, 이럴 경우 영상의 모든 영역에 입자의 밀도가 균일하게 적용되므로 균일한 안개가 표현된다. 그리고 $t(x) = e^{-\beta d(x)}$ 에 따라, 영상의 깊이가 큰 영역일수록 RESDIE [22]의 데이터인 Fig. 2의 (a)처럼 짙은 안개가 생성되게 된다. 즉, 균일한 안개 영상에서는 깊이에 따라 안개의 정도가 달라진다.

반면, 불균일한 안개 영상은 β 값이 영상의 픽셀 별로 적용되어 입자의 밀도가 영상 전체에 대해 불균일하다. 영상이 높이 H, 너비 W의 픽셀 크기를 가진다고 할 때, 불균일 안개 영상에서는 β 를 $\beta(x)$ 와 같이 HxW 형태를 가지는 2 차원 배열로 표현한다. 이때, x 는 2 차원 벡터로 높이 H, 너비 W의 픽셀 크기 영상의 각 픽셀 위치를 의미한다. 영상의 깊이가 작더라도 해당 픽셀의 β 값에 따라 NH-HAZE [2]의 데이터인 Fig. 2의 (b)처럼 짙은 안개가 생성될 수 있다. 즉, 불균일한 안개 영상에서는 깊이와 β 에 따라 안개의 정도가 달라진다.



(a) Homogeneous hazy image

(b) Non-homogeneous hazy image

Figure 2: Example images of a homogeneous hazy image and a non-homogeneous hazy image. In (b), thick haze is formed due to β even at close distance

기준에 불균일한 안개 영상을 생성하기 위한 다양한 방법이 시도되었다. NH-HAZE [2]에서는 안개 생성기를 활용하여 야외에 안개를 인위적으로 만들어 준 후, 해당 장면을 촬영하여 데이터 세트를 구축했다. 실제 안개를 만들었다는 점에서 가장 현실적인 불균일 안개 영상을 생성했지만, 방대한 양의 불균일 안개 영상을 확보하는 것이 어렵다는 한계가 있다.

이외에도 대기산란 공식 [5]과 슈퍼픽셀(superpixel) 활용하여 불균일한 안개를 합성한 연구가 있다[23]. 이 연구에서는 슈퍼픽셀을 활용하여 깨끗한 영상의 영역을 나뉘인다. 그리고 각 영역에 적용하기 위한 무작위 $t(x)$ 값들과 영상 전체에 사용될 임의의 A 값을 추출하여 대기산란 공식 [5]에 따라 안개를 합성해준다. 이를 통해 깊이 값에 따라 안개의 정도가 결정되지

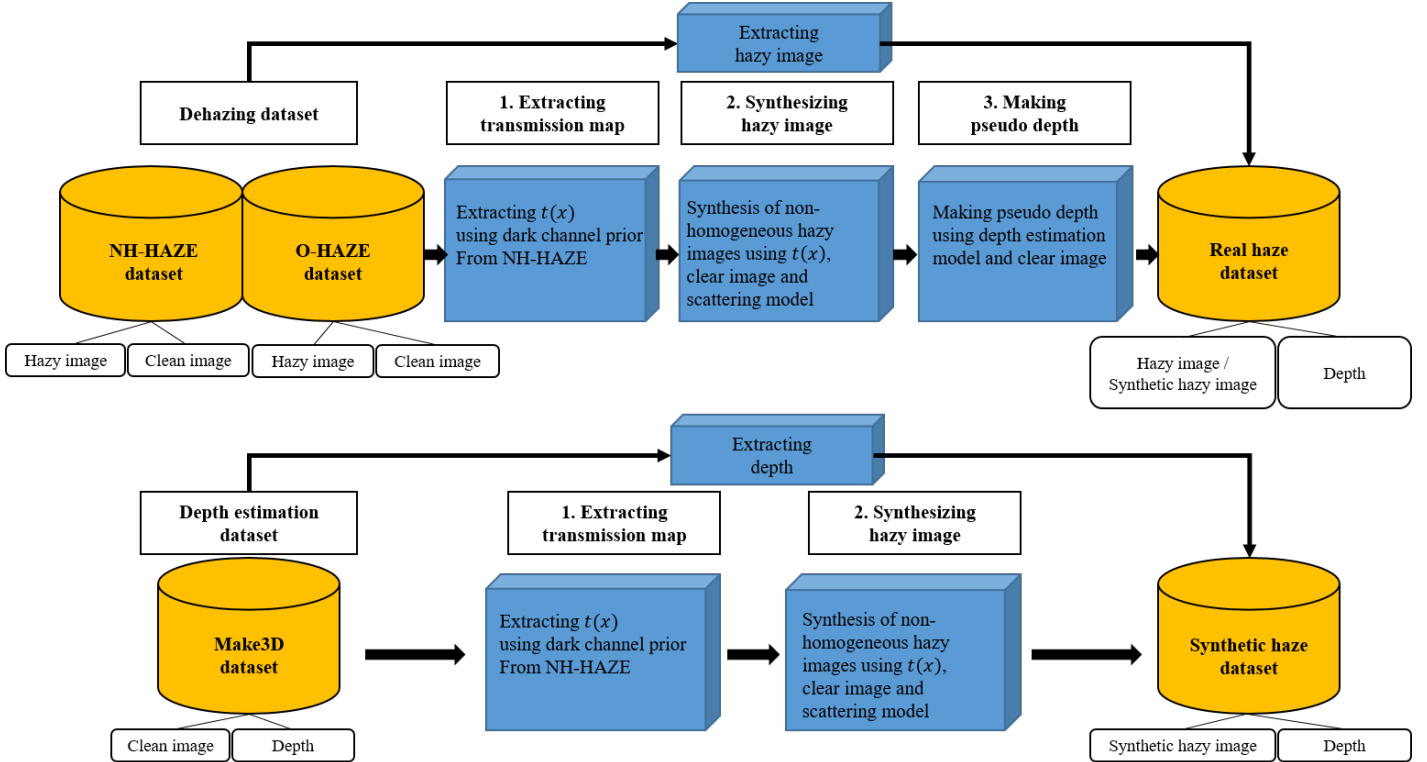


Figure 3: Haze data generation process for depth estimation. Create pseudo depth and synthetic haze in the dehazing dataset and depth estimation dataset.

않는 불균일한 안개를 생성할 수 있다. 하지만 안개의 형상이 슈퍼픽셀에 따라 결정되기 때문에, 실제 안개의 형상과 다를 수 있다는 한계가 있다.

본 연구에서는 기존 방법의 한계를 개선하여 실제 안개의 형상을 고려하고, 방대한 양의 불균일 안개 데이터를 확보할 수 있는 불균일 안개 합성 방법을 제안한다.

3. 깊이 추정용 안개 데이터 세트

본 연구에서는 불균일 안개 합성 방법을 제안하고 이를 통해 기존의 깊이 추정 모델을 개선하고자 한다. 이를 위해서 본 연구에서는 DCP [8]를 활용한 합성 방법을 사용하여 깊이 추정 모델 학습을 위한 안개 데이터 세트를 구축한다.

기존의 깊이 추정 과제를 위한 벤치마크 데이터 세트(depth estimation dataset)는 깨끗한 영상과 깊이 값을 한 쌍으로 제공한다. 하지만 안개 영상 데이터는 제공하지 않는다. 반면 안개 제거 과제를 위한 벤치마크 데이터 세트(dehazing dataset)는 안개 영상과 깨끗한 영상을 제공할 뿐 깊이 값을 제공하지 않는다. 그래서 DCP [8]를 활용하여 추출한 실제 안개 정보를 사용하여 깊이 추정 과제 데이터 세트에 안개 영상을 합성한다. 이렇게 합성된 안개 영상만을 가지는 깊이 추정용 합성 안개 데이터 세트를 만든다.

그리고 안개 제거 과제 데이터 세트에서는 깊이 값을 제공하지 않으므로 단안 영상 깊이 추정 네트워크를 사용하여 의사(pseudo) 깊이 값을 추가한다. 그리고 DCP [8]를 활용하여 추출한 실제 안개 정보로 안개 영상을 추가로 합성한다. 이렇게 실제 안개와 합성 안개가 포함된 깊이 추정용 실제 데이터 세트를 만든다. 데이터 세트를 만드는 방법을 도식화하면 Fig. 3과 같다.

3.1 불균일 안개 합성 방법

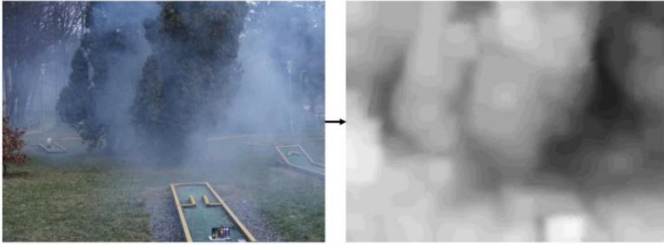
일반적으로 균일한 안개를 합성하는 방식으로 대기산란 공식 [5] Eq. (1)를 사용한다. 이때, $t(x)$ 는 $e^{-\beta d(x)}$ 로 구할 수 있기 때문에, 영상의 깊이 값을 활용하면 β 값을 조정하여 균일한 안개를 깨끗한 이미지에 투영할 수 있는 $t(x)$ 를 구할 수 있다. 그리고 안개 영역은 일반적으로 다른 지역보다 픽셀의 최소 채널 값이 크기 때문에 이를 고려하여 대기 산란광 값을 설정해두면 Eq. (1)을 활용하여 균일한 안개를 생성할 수 있다. 불균일 안개 영상의 경우 균일 안개 영상과 다르게 픽셀별로 β 값이 다르게 적용된다. 따라서 영상의 깊이 값을 알고 있어도 β 를 설정해주는 것이 매우 어렵다. 이 때문에 이와 같은 방식으로 안개를 합성하면 균일한 안개밖에 합성할 수 없고, 영상이 깊이 값을 반드시 가지고 있어야 한다는 단점이 있다.

반면 영상의 깊이 값을 가지고 있지 않아도 DCP [8]를 활용하면 안개 영상으로부터 근사적으로 $t(x)$ 를 구할 수 있다.

DCP [8]는 Eq. (3)에 따라 안개의 심한 정도를 판별하여 픽셀 별로 대기 투과율을 계산한다. 이 때문에 β 값이 픽셀 별로 자연스럽게 다르게 적용된 근사적 $t(x)$ 를 곧바로 얻을 수 있다.

만약 특정 불균일 안개 영상에서 DCP [8]로 $t(x)$ 를 구한 뒤, $t(x)$ 에 포함된 영상의 깊이 정보를 최대한 배제해준다면, 다른 영상 합성에 활용할 수 있는 불균일한 안개의 특성만을 가진 $t(x)$ 를 추출할 수 있다. 이때 깊이 정보는 $t(x)$ 에 풍경, 사물 등의 윤곽 형태로 남게 된다. 이렇게 추출된 $t(x)$ 값을 다른 깨끗한 영상에 활용해준다면 Eq. (1)에 따라 불균일한 안개 영상을 합성할 수 있다.

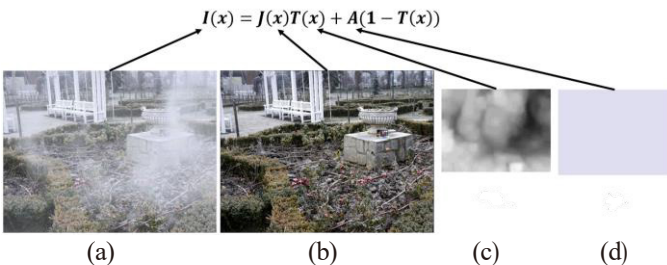
본 연구에서는 DCP [8]를 사용하여 실제 불균일 안개 영상에서 합성에 활용 가능한 $t(x)$ 를 구했다. 그리고 박스 블러 필터를 사용하여 이 $t(x)$ 에 사물, 풍경 등의 윤곽으로 남아있는 깊이 정보를 최대한 제거하여 Fig. 4와 같이 최종적인 합성용 $t(x)$ 를 얻었다.



(a) Non-homogeneous hazy image (b) $t(x)$ for synthesis extracted in (a)

Figure 4: Example of extracting $t(x)$ for synthesis from a real hazy image using the scattering model and dark channel prior

이렇게 만들어진 $t(x)$ 로 Eq. (1)을 사용하여 Fig. 5와 같이 깨끗한 영상에 불균일한 안개를 합성한다. 대기 산란광의 경우, 깨끗한 영상의 dark channel을 기준으로 밝기가 상위 0.1% 안에 들어가는 픽셀들의 R, G, B값을 각각 평균을 내어 구한다.



(a) (b) (c) (d)

Figure 5: Example of synthesizing a non-homogeneous hazy image using the scattering model [5]. (a) is a synthetic hazy image, and (b) is a clean image used for the synthesis. (c) and (d) are $t(x)$ for synthesis and randomly selected global atmospheric light, respectively.

3.2 깊이 추정용 실제 안개 데이터 세트 구축

깊이 추정 과제를 위한 벤치마크 세트 중, 실제 불균일 안개 영상과 영상 깊이 값을 한 쌍으로 제공하는 데이터 세트는 존재하지 않는다. 이 때문에 안개 제거 과제용 벤치마크 데이터 세트를 활용한다. 이 데이터 세트는 안개 영상과 깨끗한 영상을 한 쌍으로 제공한다. 이 데이터 세트에는 영상의 깊이 값이 존재하지 않는다. 하지만 이미 학습된 영상 깊이 추정 모델을 활용하면 깨끗한 영상에서 의사 깊이 값을 뽑아낼 수 있다. 이 같은 방법으로 영상의 의사 깊이 값을 추출해 데이터 세트를 사용 가능하도록 보완한다. 이때, 이미 학습된 영상의 깊이 추정 모델은 LeRes [4]기반의 Boosting [3]모델을 사용한다.

안개 제거 과제용 벤치마크 데이터 세트는 야외 안개 데이터를 위주로 선정한다. 실제 안개가 많이 발생하는 장소는 야외이기 때문이다. 이를 고려하여 NH-HAZE 2020 [2]과 O-HAZE [9] 데이터 세트를 사용한다. 특히 불균일 안개 영상 데이터 세트는 NH-HAZE 2020 [2] 데이터 세트가 많이 사용된다. 본 연구에서는 이 데이터 세트를 활용하여 불균일 안개 합성에 사용할 수 있는 합성용 $t(x)$ 28개를 추출한다. 이렇게 추출한 합성용 $t(x)$ 를 활용하여 각 데이터 세트의 학습용(training)과 검증용(validation) 데이터 세트를 증폭시킨다. 최종적으로 NH-HAZE 2020 [2]과 O-HAZE [9]를 합쳐서 학습용 데이터 1,600개, 검증용 데이터 200개, 평가용(test) 데이터 10개인 실제 안개 데이터 세트를 구축한다.

3.3 깊이 추정용 합성 안개 데이터 세트 구축

깊이 추정 과제를 위한 벤치마크 데이터 세트는 깨끗한 영상과 영상 깊이 값을 한 쌍으로 제공한다. 하지만 대부분 안개 영상 데이터를 제공하지 않는다. 따라서 본 연구에서 제안하는 불균일 안개 합성을 사용하여 안개 영상을 생성한다.

NH-HAZE 2020 [2] 데이터 세트에서 추출한 합성용 $t(x)$ 를 활용하여 안개 합성을 진행한다. 자연스러운 합성을 위해 야외 영상이면서 최대한 NH-HAZE 2020 [2]과 비슷한 환경을 가지고 있는 데이터 세트를 고려하여, Make3D [10, 11] 데이터 세트를 선택한다. 이는 합성용 $t(x)$ 가 야외의 자연환경 및 풍경의 영상에서 추출되었기 때문이다. 추가적으로 Make3D [10, 11]에서 이런 조건을 만족하지 않는 데이터를 정제한다. 이 과정을 통해서 학습용 데이터는 360 개에서 249 개, 검증용 데이터는 40 에서 24 개, 평가용 데이터는 134 개에서 91 개로 줄어든다. 그리고 최종적으로 불균일 안개 영상 합성 방법을 활용하여 학습용 데이터 7,221 개, 검증용 데이터 696 개, 평가용 데이터 2,639 개인 합성 안개 데이터 세트를 구축한다.

4. 단안 영상 깊이 추정 네트워크

본 연구에서는 안개 상황에서 제안하는 합성 방법을 기반으로 학습된 모델이 그렇지 않은 동일 모델보다 성능이 개선될 것이라고 기대한다. 이를 확인하기 위하여 실험을 위한 적절한 영상 깊이 추정 딥러닝 모델을 선정한다. 본 연구에서 구축한 데이터 세트를 활용하기 위해서는 단안 깊이 추정 방식의 지도 학습 모델이 필요하다. 그리고 주로 넓은 야외 환경에서 잘 작동하는 모델이 필요하다.

깊이 추정 과용 벤치마크 데이터 세트 중 야외 환경의 영상을 위주로 하는 데이터 세트는 Make3D [10, 11]와 KITTI [24]이다. 이 두 가지 데이터 세트에서 가장 좋은 성능을 보이는 단안 방식 기반의 딥러닝 모델을 확인했을 때, SharinGan [19], Monodepth2 [21], DPT-Hybrid [20] 그리고 LapDepth [18] 모델이 이를 만족한다. 최종적으로 지도 학습 기반의 모델 중 성능이 더 뛰어난 모델을 고려하여, Table. 1 에서 LapDepth [18] 모델을 본 연구에서 선택하고 사용한다.

Table 1: Comparison of SOTA models on benchmark datasets containing outdoor images.

Model	Datasets	RMSE	Type of learning	Method
SharinGan [19]	Make3D [10, 11]	8.388	unsupervised	mono
Monodepth2 [21]	Make3D [10, 11]	7.417	self-supervised	mono
DPT-Hybrid [20]	KITTI [24]	2.573	supervised	mono
LapDepth [18]	KITTI [24]	2.446	supervised	mono

5. 실험 결과 및 분석

5.1 실험 상세

불균일 안개 영상에서 모델의 깊이를 추정 성능을 객관적으로 검증하기 위해서는 실제 불균일 안개 영상과 영상의 깊이 값이 한 쌍으로 존재하는 데이터 세트가 필요하다. 하지만 실제로 이런 데이터 세트가 존재하지 않기 때문에 본 연구에서 구축한

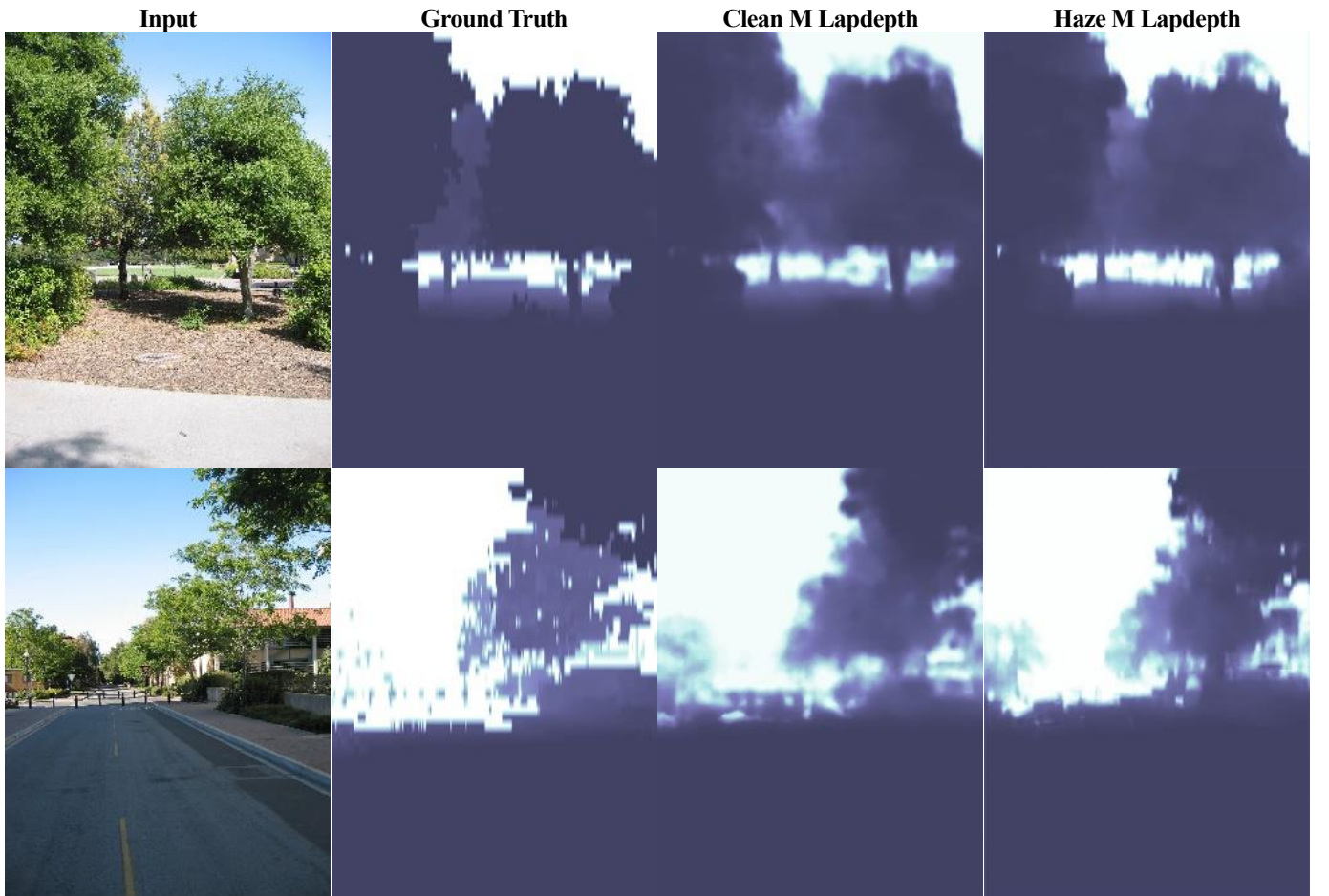


Figure 6: Qualitative results of Make3D dataset experiments. The results of the Haze M Lapdepth model and the Clean M Lapdepth model are similar.

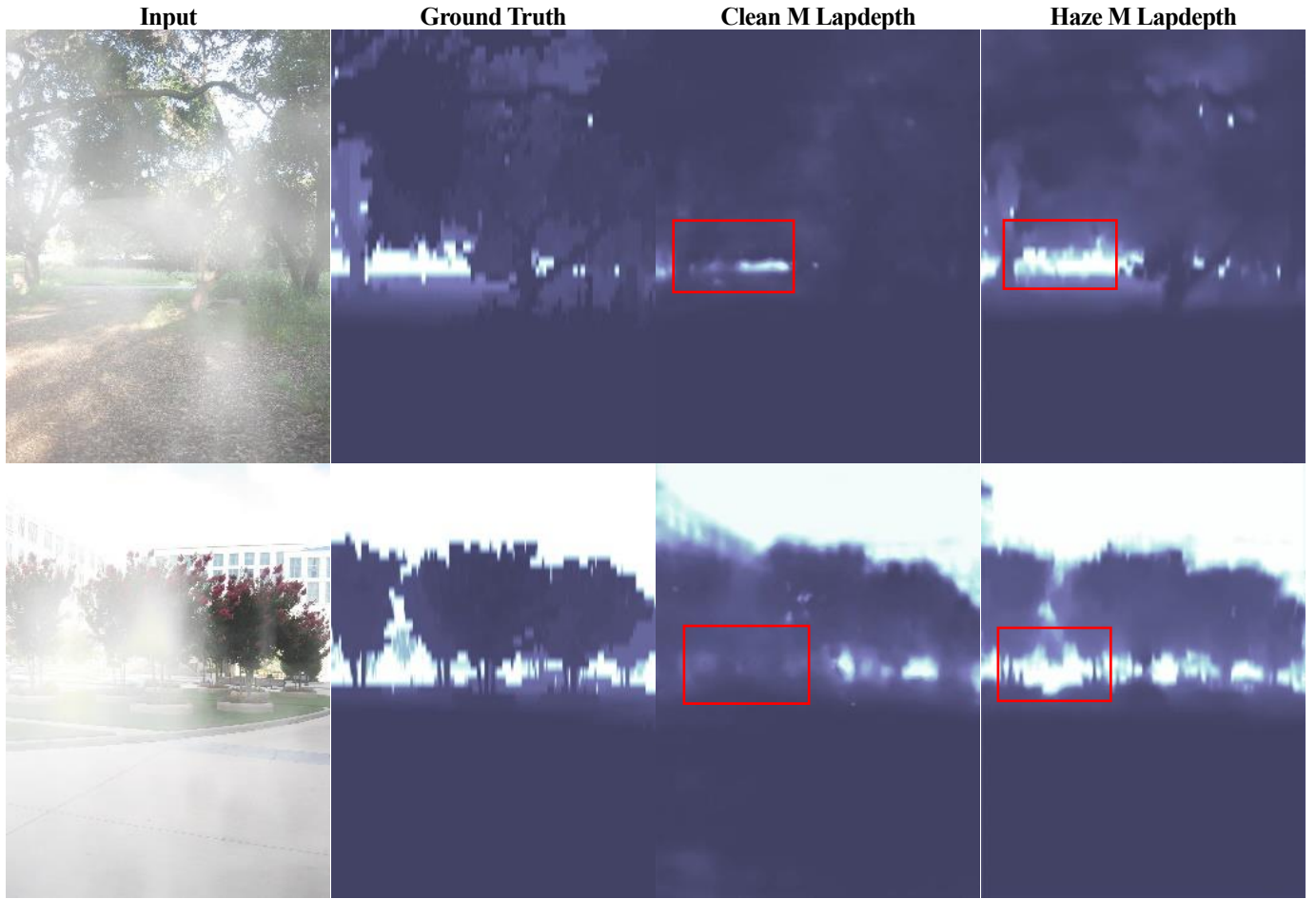


Figure 7: Qualitative results of synthetic haze dataset experiments. The area surrounded by the red box is an area where the output of the Haze M Lapdepth model is visually more similar to the ground truth than the output of the Clean M Lapdepth model.

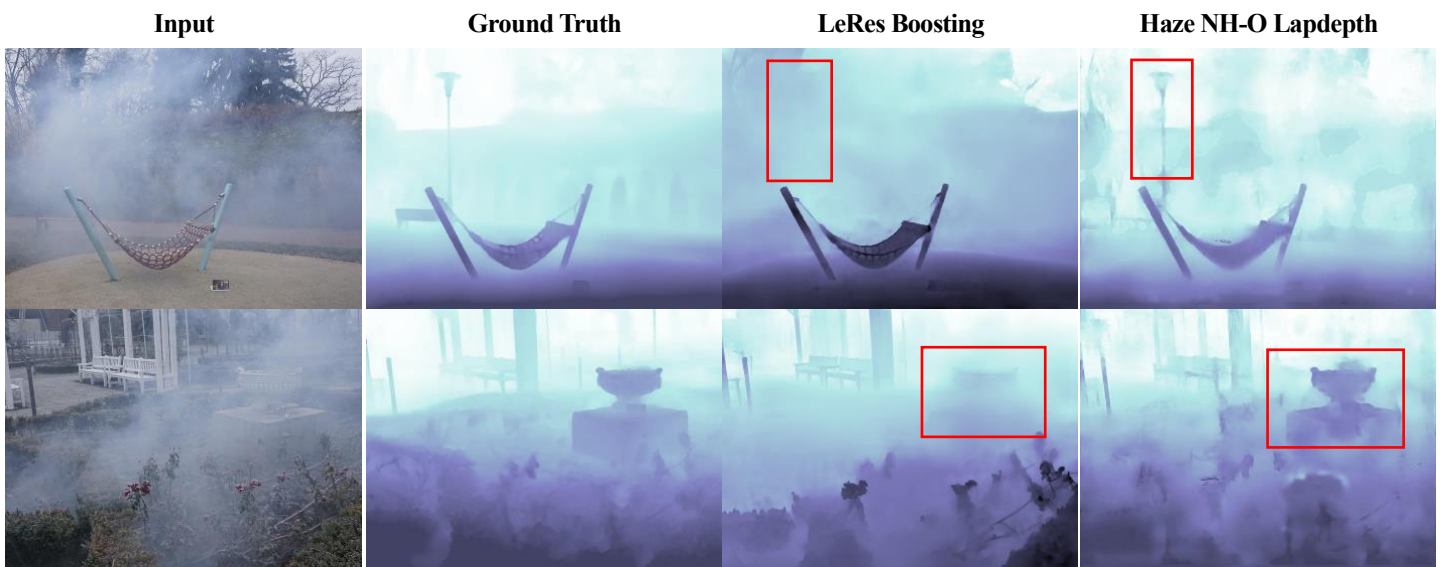


Figure 8: Qualitative results of real haze dataset experiments. The area surrounded by the red box is an area where the output of the Haze NH-O Lapdepth model is visually more similar to the ground truth than the output of the LeRes [4] Boosting [3] model

합성 안개 데이터 세트와 실제 안개 데이터 세트를 사용하여 두 가지 실험을 진행한다.

첫 번째 실험은 합성 안개 데이터 세트를 활용한 실험이다. 합성 안개 데이터 세트에는 영상의 실제 깊이 값이 있으므로, 이 데이터 세트로 학습된 Lapdepth [18] 모델이 안개 상황에서 영상의 실제 깊이 값을 잘 추정하는지를 중점적으로 보는 실험을 진행한다.

두 번째 실험은 실제 안개 데이터 세트를 활용한 실험이다. 실제 안개 데이터 세트에는 실제 불균일한 안개 영상이 있으므로, 이 데이터 세트로 학습된 Lapdepth [18] 모델이 실제 불균일한 영상에서도 깊이 값을 추정할 수 있는지를 중점적으로 보기 위한 실험을 진행한다.

5.2 합성 안개 데이터 세트 실험 결과

합성 안개 데이터 세트는 Make3D [10, 11] 데이터 세트에 불균일 안개를 합성하여 증폭시킨 데이터 세트이다. 이 데이터 세트를 가지고, 학습된 Lapdepth [18] 모델을 본 연구에서는 Haze M Lapdepth [18] 모델 그리고 Make3D [10, 11] 데이터 세트 만을 가지고 학습된 Lapdepth [18]

모델을 본 연구에서는 Clean M Lapdepth 모델이라고 명명한다. 이 두 개의 모델을 합성 안개 데이터의 평가용 세트를 사용하여 성능을 비교한다.

Table. 2에서 정량적 평가 결과를 확인해보면 모든 성능 지표에서 Haze M Lapdepth 모델이 더 나은 성능 수치를 보여주는 것을 확인할 수 있다. 또한, Fig. 7의 정성적인 결과를 보면 Haze M Lapdepth 모델이 시각적으로 ground truth에 더 가까운 결과를 보여준다. 특히 Fig. 7의 빨간색으로 둘러싸인 부분은 안개가 심한 영역에서 Haze M Lapdepth 모델이 잘 추론한 영역이다. 이를 통해서 불균일 안개 영상 합성을 학습에 활용할 경우 안개 환경에서 영상의 실제 깊이를 추론하는 성능이 개선될 수 있음을 확인할 수 있다.

Table 2: Quantitative results of synthetic haze dataset experiments

Model	RMSE	SSIM
Clear M Lapdepth	11.32	0.92
Haze M Lapdepth	9.09	0.94

추가로 Haze M Lapdepth 모델이 깨끗한 영상에서도 성능이 유지되는 지 확인하기 위해서, Make3D [10, 11] 데이터의 평가용 세트를 사용하여 정량적 평가와 정성적 평가를 진행했다. Table. 3에서 정량적 평가 결과를 확인해보면 Clean M Lapdepth 모델과 거의 비슷한 성능 수치를 보이고 있다. 또한 Fig. 6의 정성적인 결과도 비슷한 성능을 보이고 있다. Haze M Lapdepth 모델이 불균

일한 안개 영상으로만 학습된 모델이라는 것을 감안하면, 깨끗한 영상에서도 추정 정확도가 유지된다고 볼 수 있다.

Table 3: Quantitative results of Make3D test dataset experiments

Model	RMSE	SSIM
Clear M Lapdepth	8.76	0.94
Haze M Lapdepth	9.16	0.93

5.3 실제 안개 데이터 세트 실험 결과

실제 안개 데이터 세트는 NH-HAZE 2020 [2]과 O-HAZE [9] 데이터 세트에 불균일 안개를 합성하여 증폭시킨 데이터 세트이다. 이 데이터 세트로 학습된 Lapdepth 모델을 본 연구에서는 Haze NH-O Lapdepth 모델이라 명명한다. 이 실험에서 사용된 ground truth 값은 의사 깊이 값이기 때문에 실제 불균일 안개 영상에서 영상의 깊이를 추론할 수 있는지 정성적인 측면에서 성능을 확인한다. 그리고 의사 깊이 값을 추출할 때 사용한 LeRes [4] 기반의 Boosting [3] 모델의 결과와 비교한다. Fig. 8에서 정성적인 결과를 보면 Haze NH-O Lapdepth 모델의 결과가 시각적으로 ground truth와 더 비슷한 것을 확인할 수 있다. 이를 통해서 실제 안개 영상에서도 제안하는 불균일 안개 영상 합성 방법을 통해 모델이 잘 학습될 수 있음을 확인할 수 있다.

6. 결론 및 한계점

본 연구에서는 대기산란 공식 [5]과 DCP [8]를 활용하여 불균일한 안개 영상을 합성하는 방법을 제안한다. 이전의 안개 영상 합성 방법으로는 불균일한 안개 영상을 생성하지 못한다. 또한, 이전의 방법으로는 영상의 깊이 값을 알고 있어야만 합성할 수 있었지만 본 연구에서는 영상을 합성할 때 깊이 값이 필요 없다는 장점을 가지고 있다. 그리고 제안하는 합성 방법을 활용하여 기존의 딥러닝 영상 깊이 추정 모델이 성능을 개선하는 것을 확인할 수 있다.

본 연구의 한계점은 영상을 합성하기 위해 실제 불균일 안개 영상을 필요로 한다는 것이다. 따라서 합성하기를 원하는 특성 또는 조건을 가지는 안개 영상이 있어야만 합성이 가능하다는 한계점을 가지고 있다.

감사의 글

이 연구는 포스코 ICT 산학협력과제의 지원으로 수행되었습니다.

References

- [1] Lee, Jonghyeop, et al., "Slow Sync Image Synthesis from Short Exposure Flash Smartphone Images," *Journal of the Korea Computer Graphics Society*, 27.3, 1-11, 2021.
- [2] Ancuti, Codruta O., Cosmin Ancuti, and Radu Timofte., "NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020.
- [3] Miangoleh, S.M.H., et al., "Boosting Monocular Depth Estimation Models to High-Resolution via Content-Adaptive Multi-Resolution Merging," *In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9685-9694, 2021.
- [4] Yin, W., et al., "Learning to recover 3d scene shape from a single image," *In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 204-213, 2021.
- [5] Narasimhan, S.G., Nayar, S.K., "Vision and the atmosphere," *International journal of computer vision*, **48**(3), pp. 233-254, 2002.
- [6] Ancuti, C., Ancuti, C.O., De Vleeschouwer, C., "D-hazy: A dataset to evaluate quantitatively dehazing algorithms," *In: 2016 IEEE international conference on image processing (ICIP)*, pp. 2226-2230, 2016.
- [7] Li, B., et al., "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, **28**(1), pp. 492-505, 2018.
- [8] He, K., Sun, J., Tang, X., "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, **33**(12), pp. 2341-2353, 2010.
- [9] Ancuti, Codruta O., et al., "O-haze: a dehazing benchmark with real hazy and haze-free outdoor images," *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018.
- [10] Saxena, Ashutosh, Sung Chung, and Andrew Ng., "Learning depth from single monocular images," *Advances in neural information processing systems*, 18, 2005.
- [11] Saxena, Ashutosh, Min Sun, and Andrew Y. Ng., "Make3d: Learning 3d scene structure from a single still image," *IEEE transactions on pattern analysis and machine intelligence*, 31.5, 824-840, 2008.
- [12] Narasimhan, S.G., Nayar, S.K., "Chromatic framework for vision in bad weather," *In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2000* (Cat. No. PR00662), pp. 598-605, 2000.
- [13] FATTAL, R., "Single image dehazing," *ACM Transactions on Graphics*, vol. 27., 2008.
- [14] Tankovich, V., et al., "Hitnet: Hierarchical iterative tile refinement network for real-time stereo matching," *In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14362-14372, 2021.
- [15] Shamsafar, F., et al., "MobileStereoNet: Towards Lightweight Deep Networks for Stereo Matching," *In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2417-2426, 2022.
- [16] Khamis, S., et al., "Stereonet: Guided hierarchical refinement for real-time edge-aware depth prediction," *In: Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 573-590, 2018.
- [17] Lee, Jonghyeop, et al., "SINGLE PANORAMA DEPTH ESTIMATION USING DOMAIN ADAPTATION," *Journal of the Korea Computer Graphics Society*, 26.3, 61-68, 2020.
- [18] Song, M., Lim, S., Kim, W., "Monocular Depth Estimation Using Laplacian Pyramid-Based Depth Residuals," *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.
- [19] PNVR, K., Zhou, H., Jacobs, D., "Sharingan: Combining synthetic and real data for unsupervised geometry estimation," *In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13974-13983, 2020.
- [20] Ranftl, R., Bochkovskiy, A., Koltun, V., "Vision transformers for dense prediction," *In: Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12179-12188, 2021.
- [21] Godard, C., et al., "Digging into self-supervised monocular depth estimation," *In: Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3828-3838, 2019.
- [22] Li, Boyi, et al., "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, 28.1, 492-505, 2018.
- [23] Liu, Chunxiao, et al., "Non-homogeneous haze data synthesis based real-world image dehazing with enhancement-and-restoration fused CNNs," *Computers & Graphics*, 2022.
- [24] Geiger, Andreas, et al., "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, 32.11, 1231-1237, 2013.

〈 저자 소개 〉



최영철

- 2019년 인천대학교 산업경영공학과(학사)
- 2020년 1월~2020년 10월 포항공과대학교 인공지능연구원 연구원
- 2020년 12월~현재 포스코ICT AI기술그룹 연구원
- 관심분야: 영상 및 비디오 처리, 패턴인식, 머신비전, 딥러닝
- <https://orcid.org/0000-0003-2525-873X>



백지현

- 2007년 경원대학교 소프트웨어학부(학사)
- 2021년 연세대학교 공학대학원 인공지능(석사)
- 2007년~현재 포스코ICT AI기술그룹 책임연구원
- 관심분야: 패턴인식, 머신비전, 딥러닝
- <https://orcid.org/0000-0001-5485-2602>



주광진

- 2020년 포항공과대학교 컴퓨터공학(학사)
- 2020년~현재 포항공과대학교 인공지능대학원 석박통합과정
- 관심분야: 딥러닝, 이미지 프로세싱, 컴퓨터 비전
- <https://orcid.org/0000-0003-1006-8351>



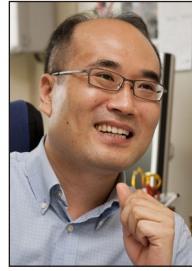
이동진

- 2020년 광주과학기술원 전기전자컴퓨터학과(학사)
- 2021년~현재 포항공과대학교 인공지능대학원 석박통합과정
- 관심분야: 딥러닝, 머신러닝, 컴퓨터 비전
- <https://orcid.org/0000-0001-6482-8892>



황경하

- 2006년 포항공과대학교 컴퓨터공학과(학사)
- 2012년 포항공과대학교 수학과(박사)
- 2012년~2014년 서울대학교 PARC 박사후연구원
- 2014년~2014년 울산과학기술원 박사후연구원
- 2014년~2015년 울산과학기술원 초빙 조교수
- 2015년~2018년 National Center for Theoretical Sciences 박사후연구원
- 2018년~현재 영남대학교 수학과 조교수
- 관심분야: 편미분방정식, 역문제
- <https://orcid.org/0000-0001-9791-8130>



이승용

- 1988년 2월 서울대학교 계산통계학과(학사)
- 1990년 2월 한국과학기술원 전산학과(석사)
- 1995년 2월 한국과학기술원 전산학과(박사)
- 1995년 3월~1996년 9월 미국 City College of New York / CUNY 연구원
- 2003년 8월~2004년 7월 독일 Max-Planck-Institut fur Informatik 방문교수
- 2010년 10월~2011년 9월 미국 Adobe Systems 방문교수
- 1996년 10월~현재 포항공과대학교 컴퓨터공학과 교수
- 관심분야: 영상 및 비디오 처리, 삼차원 복원, 그래픽스 응용시스템
- <https://orcid.org/000-0002-8159-4271>