

Real-time Defog Processing Using Cooperative Networks

Sanghyun Jung*

*Researcher, ENERZAI, Seoul, Korea

[Abstract]

In this paper, we propose a deep learning model and inference pipeline that can process high-resolution fog video in real-time, addressing limitations found in classical defogging algorithms and existing deep learning-based defogging models. The key idea is separating the tasks of inferring fog color and estimating the amount of fog into two distinct models, allowing for a more efficient, lightweight design that improves inference speed. While many deep defogging models perform well on synthetic fog images, they suffer from reduced effectiveness on real-world fog images with diverse fog colors and backgrounds. We solve this problem by introducing a synthetic fog dataset generation method tailored for real-world conditions. Through experiments, we demonstrate the increase in visible distance achieved by proposed model and compare its inference speed and defogging performance against pre-trained models on real-world CCTV fog images.

▶ **Key words:** Defogging, Dehazing, Image restoration, Deep learning, Light-weight model, Real-time

[요약]

본 논문에서는 기존의 고전적인 안개 제거 알고리즘과 딥러닝 기반의 안개 제거 모델들의 문제점을 개선하고, 고해상도 안개 영상을 실시간으로 처리할 수 있는 딥러닝 모델과 추론 방식을 제안한다. 핵심 아이디어는 안개 영상에서 안개의 색상을 추론하는 모델과 안개량을 추론하는 모델을 분리하여 학습시켜 각각의 모델을 경량화함으로써 추론 속도를 향상시키는 것이다. 또한, 합성 안개 영상에 대해서는 잘 작동하지만 다양한 안개 색상과 배경을 갖는 실제 안개 영상에서는 성능이 떨어지는 문제점을 새로운 데이터셋 생성 방식을 이용하여 해결한다. 실험을 통해 우리의 안개 제거 모델로 안개 이미지를 처리한 후의 가시거리 증가량을 측정하고, 실제 CCTV 영상에 대하여 추론 속도와 안개 제거 성능을 사전 학습된 기존의 모델들과 비교한다.

▶ **주제어:** 안개 제거, 연무 제거, 이미지 복원, 딥러닝, 경량화 모델, 실시간

I. Introduction

안개가 없는 맑은 날씨 환경에서 카메라로 촬영하는 경우, 광원의 빛이 물체에 반사되어 카메라의 이미지 센서에 도달하여 이미지가 생성된다. 이때 물체에서 반사되어 카메라로 전달되는 빛을 scene radiance라고 한다. 안개는 미세한 입자들로 구성되어 있으며 그 종류는 미세한 물방울, 먼지, 모래 입자 등 다양하다. 광원이 이러한 안개 입자들에 의해 반사되어 나타나는 빛과 그 색상을 airlight라고 한다. 따라서 airlight는 안개의 색상 정보를 나타낸다. 안개가 끼게 되면 scene radiance는 대기 중의 안개 입자로 인해 흡수 또는 산란되어 카메라까지 온전히 전달될 수 없다. 또한, 안개 입자에 의해 발생한 airlight가 일부 섞여 카메라에 전달되게 된다. 따라서 정확한 색상 정보 및 물체의 경계 정보를 전달할 수 없게 되어 안개가 낀 환경에서 촬영된 영상은 가시거리가 제한되며 상대적으로 낮은 품질을 보인다. 이러한 상황이 발생하면 일부 CCTV 카메라들은 안개에 의해 흐려진 영상을 실시간으로 보정하기 위해 dark channel prior (DCP) [1] 나 이미지의 대비 증가 방식의 실시간 알고리즘을 탑재하기도 하지만 [2], DCP에서 사용되는 통계적 특성이 무채색의 안개 색상에만 적용된다는 한계와 대비 증가 알고리즘을 적용했을 때 영상의 노이즈가 증가하는 문제점이 발생한다. 따라서 기존의 실시간 안개 제거 알고리즘은 심한 노이즈로 인해 확보할 수 있는 가시거리의 증가량이 높지 않거나 황사가 섞인 노란색의 안개나 새벽녘의 푸른 안개와 같이 특정 색상의 안개에는 적용할 수 없는 문제가 발생한다.

안개 제거 처리가 필요한 상황은 CCTV의 관제 시스템과 같이 처리된 영상을 관제자가 직접 확인하거나 탐지 모델의 입력으로 사용하는 경우가 대다수를 차지하며, 특성상 실시간 처리를 필요로 하는 경우가 많다. 하지만 딥러닝 기반의 안개 제거 모델의 경우, 일반적으로 성능과 처리 속도가 반비례 관계에 있기 때문에 적절한 성능과 처리 속도를 동시에 얻는 것은 어려운 일이다. 이러한 문제는 안개 제거 모델뿐만 아니라 대부분의 딥러닝 기반의 모델들에서 발생한다. 또한 이러한 모델들은 대부분 합성된 안개 이미지 데이터셋에 대해서만 학습되어 다양한 색상의 안개와 배경으로 구성된 실제 안개 영상에 대해서는 낮은 안개 제거 성능을 보여주는 문제도 발생한다.

본 논문에서는 경량화된 모델과 추론 방식을 새롭게 제안하고, 이를 이용하여 기존 모델들의 문제점이었던 느린 추론 속도를 향상시켜 실시간 안개 제거 처리가 가능함을 보이며, 거리 표지가 있는 안개 영상을 합성하여 안개 제

거를 통한 가시거리의 증가량을 측정하고 그 효과를 확인한다. 또한, 실제 CCTV로부터 취득한 안개 영상에 대한 모델의 안개 제거 성능을 정성적으로 비교하여 제안하는 방법의 효과를 확인한다.

II. Preliminaries

1. Background

1.1 Fog image formation modeling

안개 이미지가 생성되는 두 가지 요소는 안개의 색상을 의미하는 airlight $A \in [0, 255]^3$ 와 scene radiance가 안개를 투과하여 카메라까지 전달되는 비율을 나타내는 transmission $t \in (0, 1]^{H \times W}$ 이다. 안개가 없는 이미지 I 와 airlight A , transmission t 가 주어졌을 때, 안개 이미지 F 는 Koschmieder visibility formula [3]를 이용하여 각각의 픽셀 좌표 (x, y) 에서 다음과 같이 선형 결합으로 표현된다.

$$F(x, y) = t(x, y)I(x, y) + (1 - t(x, y))A$$

이를 이용하여 반대로 안개 이미지로부터 안개가 없는 scene radiance를 얻기 위해서는 안개 생성의 역과정을 진행해야 하며, 다음과 같이 나타내어진다.

$$I(x, y) = \frac{F(x, y) - A}{t(x, y)} + A$$

이때 물체와 카메라 간의 거리가 멀수록 사이에 존재하는 안개 입자들에 의해 빛이 산란될 확률이 높아지므로 transmission은 지수적으로 낮아지게 된다. 또한, 안개 입자들의 밀도나 특성에 의해서도 transmission은 변하게 되며, 이는 안개 계수 β 로 표현된다. 안개 계수 β 와 카메라와 물체와의 거리를 나타내는 depth map d 가 주어졌을 때, transmission t 는 다음과 같이 나타낼 수 있다.

$$t(x, y) = e^{-\beta d(x, y)}$$

2. Related Works

2.1 Classical methods

기존에 고전 알고리즘을 통한 안개 제거 처리 방법들이 존재했으나, 주어진 안개 이미지로부터 정확한 airlight와 transmission을 계산하는 것은 어려운 문제이다. 고전적인 방식에서는 transmission을 유추하기 위해 주로 DCP를 이용한 알고리즘을 사용한다. DCP는 선명하게 보이는 자연 물체의 픽셀은 RGB 채널 중 하나 이상의 값이 0에 가깝다는 통계적 관찰을 기반으로 한다. 따라서, 픽셀 좌표 (x, y) 에 대해 local window $\Omega(x, y)$ 안에 있는 모든

픽셀들의 R, G, B 값 중 최솟값이 해당 픽셀 좌표의 dark channel J^{dark} 이 되며 다음의 수식으로 표현할 수 있다.

$$J^{dark}(x, y) = \min_{c \in \{R, G, B\}} \min_{(x', y') \in \mathcal{Q}(x, y)} I^c(x', y')$$

DCP에 의해, 선명한 scene을 촬영한 이미지의 경우 dark channel의 값이 0에 가까워야 한다. 하지만 안개의 airlight는 일반적으로 R, G, B 값이 모두 높은 회색 또는 흰색이므로 dark channel의 값을 증가시킨다. 따라서 dark channel의 값이 높은 영역에 많은 안개가 끼었음을 통계적으로 유추할 수 있으며, 이를 통해 transmission을 예측할 수 있다. 또한, dark channel의 값이 상위 0.5%인 영역은 안개량이 가장 많은 영역이므로 해당 영역의 이미지 픽셀값이 안개의 색인 airlight가 된다.

하지만 이러한 방식은 무채색 계열의 안개에만 적용할 수 있으며, 장면에 대한 전체적인 이해가 아닌 국소적인 영역에 대한 통계적인 특성만을 이용하므로 회색의 아스팔트 도로나 건물 등의 물체를 안개가 낀 영역으로 인식하여 검은색에 가깝게 과변형시키는 문제가 발생한다.

2.2 Deep learning-based methods

최근에는 딥러닝을 이용하여, 모델 [4, 5, 6] 에 안개 이미지를 입력하면 안개가 제거된 이미지가 직접적으로 추론되는 end-to-end 방식으로 안개 제거 처리를 시도하는 경우들이 많아졌다. 이러한 방식은 학습의 용이함과 이미지의 분포에 대한 모델의 자체적인 이해를 활용할 수 있어 다양한 이미지 복원 분야에서 사용되고 있다. 하지만 모델이 저품질 이미지들의 분포에서 고품질 이미지들의 분포로의 변환을 이해해야 하므로 모델의 크기가 커져야 하고, 이로 인해 추론 속도가 느려져 실시간 처리가 어려워진다.

또한, 학습 데이터의 분포와 실제 적용해야 하는 데이터의 분포가 다를 경우 성능이 떨어지는 문제가 발생할 수 있다. 특히 합성 데이터셋이 많이 사용되는 안개 제거 분야에서는 현실의 안개 이미지에 대해서 안개가 제대로 제거되지 않는 경우가 발생한다.

III. The Proposed Scheme

1. Cooperative Networks

기존의 딥러닝 기반의 안개 제거 모델은 end-to-end 방식으로, 안개 이미지로부터 깨끗한 이미지를 추론하기 위해 필요한 prior를 잠재적으로 학습하게 된다. 이러한 잠재적인 prior를 학습 가능한 모델을 만들기 위해서는 모델의

parameter들이 많이 필요하다. 하지만 본 논문에서 제시하는 안개 제거 모델은 이미 알려져 있는 안개 이미지 생성 원리를 기본적인 prior로 사용하여 안개의 색상과 관련된 airlight를 예측하는 모델 A 와 안개의 질기와 관련된 transmission을 예측하는 모델 T 로 분리하여 구성하였다. 각각의 모델들은 fog image formation modeling에서 제시하는 기초적인 prior들을 학습하기 때문에 모델의 경량화에 이점이 있다. 안개 이미지 x 가 주어졌을 때, 안개가 제거된 결과 이미지 y 는 airlight 모델 A 와 transmission 모델 T 에 의해 다음의 수식에 의해 얻어진다.

$$y = \frac{x - A(x)}{T(x)} + A(x)$$

하지만 두 모델을 개별적으로 학습하게 되면 airlight와 transmission으로부터 안개가 제거된 이미지를 추론하였을 때, 각 모델에서 예측한 값들의 부조화로 인해 결과 이미지가 부자연스러운 경우가 발생하였다. 주된 원인은 안개 제거 이미지를 얻을 때 transmission으로 나누게 되면서 안개가 심한 부분에서는 transmission의 값이 0에 가까워지므로 약간의 오차가 결과 이미지를 크게 바꾸게 되며, 이러한 오차와는 무관하게 학습된 airlight 모델의 결과와 합해졌을 때 Fig. 1과 같이 색감이 크게 달라져 부정확한 결과를 얻게 된다.



Fig. 1. Defog results when airlight model and transmission model are trained individually and trained with cooperative networks method.

이러한 문제를 해결하기 위해 본 논문에서는 협동적 네트워크를 제안한다. Generative adversarial networks (GANs) [7]의 경우에는 생성 모델과 판별 모델이 번갈아가며 적대적으로 학습되어 두 모델의 성능을 점점 향상시킨다. 제안하는 협동적 네트워크는 airlight 모델과 transmission 모델이 번갈아가며 협동적으로 학습되며, airlight 모델은 현재 transmission 모델이 예측한 transmission에 대응되는 적절한 airlight를 생성하도록 학습되며, transmission 모델은 현재 airlight 모델이 예측한 airlight에 대응되는 적절한 transmission을 생성하

도록 학습된다. 따라서 학습이 진행될수록 모델들이 서로의 오차를 감안하여 학습되므로 최종적으로는 두 모델로부터 추론한 airlight와 transmission을 이용하여 안개를 제거했을 때의 결과 이미지가 자연스러워지게 된다. 따라서 안개 제거를 위한 협동적 네트워크는 학습 시 다음의 cooperative loss를 줄이는 것을 목적으로 한다.

$$\min_A \min_T L_{coop} = \left\| \left\{ \frac{x - A(x)}{T(x)} + A(x) \right\} - y \right\|_2^2$$

하지만 cooperative loss는 예측된 transmission 값으로 나뉘야하므로 학습 초기의 불안정성이 존재한다. 따라서 각각의 모델이 초기 학습에서 적절한 결과를 낼 수 있도록 다음과 같은 reconstruction loss를 추가하였다.

$$\min_{A, T} L_{rec} = \| A(x) - A_{GT} \|_2^2 + \| T(x) - t_{GT} \|_2^2$$

결과적으로 협동적 네트워크를 이용하여 학습하기 위해 사용한 loss function은 다음과 같다.

$$L = L_{coop} + \lambda L_{rec}$$

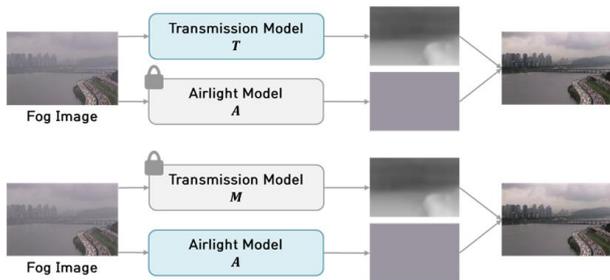


Fig. 2. Pipeline to train cooperative networks. For every iteration, one of cooperative networks is frozen and other network is trained using the frozen network.

2. Model Architectures

Airlight를 예측하기 위해서는 이미지의 각 부분에 어느 정도의 안개가 끼어있는지를 비교하고, 그 정보를 통해 안개의 색상을 찾아야 한다. 따라서 airlight 모델은 이미지를 각각의 패치로 잘라 패치 간의 attention 정보를 통합하여 결과를 예측하는 Vision Transformer (ViT) [8]의 구조를 기반으로 한다. 이미지 분류 모델인 ViT를 경량화하여 안개 이미지에 대한 airlight 색상을 R, G, B의 세 가지 채널 값에 대한 regression 모델로 변형하였다.

Transmission 예측 모델은 이미지 복원 분야에서 뛰어난 성능을 나타내는 image-to-image 모델인 NAFNet [9]의 구조를 사용하였다. Transmission을 예측하기 위해서는 depth map을 기반으로 예측해야 하므로, 전체적인 이미지의 구조를 이해하기 위해 여러 scale의 feature들을 종합하여 결과 이미지를 추론하는 U-Net [10] 구조

의 모델이 필요하며, 그 중 성능과 속도면에서 이점이 있는 NAFNet을 선택하였다. 이 모델을 원하는 성능을 유지하면서 추론 속도가 빨라질 수 있도록 경량화하였고, RGB 3채널의 안개 이미지를 입력하면 1채널의 transmission 이미지를 출력하는 모델로 구조를 변형하였다.

3. Intermittent Inference

안개 제거 처리는 일반적으로 단일 이미지보다는 관제 시스템의 실시간 영상에 적용되는 경우가 많다. 따라서 실시간 영상 처리가 중요한 분야이며, 일반적으로 Full HD (1920×1080) 해상도의 30 FPS 영상을 처리해야 한다. 고성능의 컴퓨팅 환경에서는 안개 제거 모델을 영상의 모든 프레임에 적용하는 것이 가능하지만, edge 디바이스나 미니 PC, AI 박스와 같이 성능 제약이 있는 환경에서는 모델을 매 프레임마다 적용할 경우, 처리 속도가 영상의 FPS를 따라가기 어려워 실시간성을 확보하기 힘들다. 따라서 본 논문에서는 저성능 컴퓨팅 환경에서도 실시간 안개 제거가 가능하도록 간헐적 추론 방식을 제안한다.

본 논문에서 제안하는 안개 제거 모델은 end-to-end 방식이 아닌 airlight A와 transmission t를 모델을 통해 각각 예측한 뒤, 안개 생성 원리의 역과정을 이용하여 안개가 제거된 이미지를 얻게 된다. CCTV와 같이 고정된 카메라의 경우, depth map은 거의 변하지 않으며 안개의 색상과 안개의 질기는 느린 속도로 변화하기 때문에 airlight와 transmission을 매 프레임마다 추론할 필요가 없다. 따라서 별도의 스레드 혹은 프로세스에서 모델의 추론이 진행되는 동안에는 최근 프레임에 대하여 모델에서 예측했던 airlight와 transmission을 이용하여 모든 프레임에 대한 안개 제거가 가능하다.

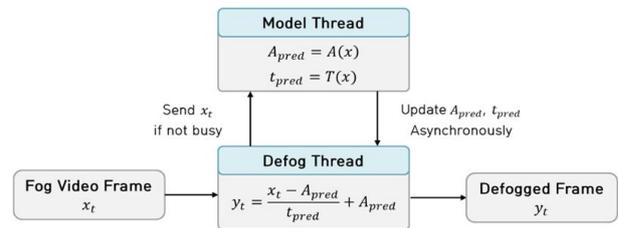


Fig. 3. Intermittent inference pipeline using multiple threads for real-time video inference on low-performance devices.

IV. Experiments

1. Fog Image Synthesis Pipeline

동일한 시각에 같은 장소에서 안개가 낀 영상과 안개가 없는 영상을 동시에 취득하는 것은 어려운 일이다. 따라서 실제 안개 이미지와 안개가 없는 정답 이미지 대신 깨끗한 이미지로부터 특정 색상과 질기의 안개를 합성하여 사용하였다. 안개 이미지 합성 모델링을 위해서는 airlight와 transmission을 설정해야 한다. Airlight는 매 합성 시마다 R, G, B 값을 각각 $[0.35, 1]$ 의 범위에서 무작위로 선택하여 다양한 색상의 안개가 합성되도록 설정하였다. 안개의 밀도를 나타내는 안개 계수 β 를 $[0.5, 3.5]$ 의 범위에서 무작위로 선택하고, 이미 학습되어 있는 depth estimation 모델을 이용하여 depth map을 추론한 뒤 transmission을 계산하였다.



Fig. 4. Intermediate images of our fog image synthesis pipeline.

이러한 방식을 이용하면 안개 유무 이미지쌍 뿐만 아니라 정확한 airlight의 색상과 transmission도 얻을 수 있어 end-to-end 방식의 모델뿐만 아니라 airlight 예측 모델, transmission 예측 모델을 모두 학습시킬 수 있다.

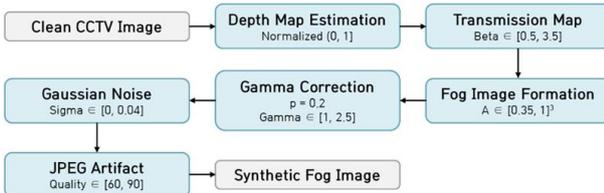


Fig. 5. Our fog image synthesis pipeline to build CCTV-like defog dataset.

Airlight와 transmission를 이용하여 fog image formation modeling 방식으로 합성된 안개 이미지를 얻을 수 있으며 다양한 augmentation을 적용하였다. 먼저

조도가 다른 상황을 위해 gamma correction을 적용하였고, 현실의 카메라 노이즈를 표현하기 위해 적절한 가우시안 노이즈를 더하였다. 마지막으로 CCTV 영상의 낮은 비트레이트로 인해 생기는 인코딩 노이즈 효과를 내기 위해 JPEG 압축 효과를 추가하였다. 구체적인 합성 파이프라인은 Fig. 5와 같다.

2. Dataset

우리는 공공 CCTV로부터 안개가 없는 다양한 환경에서의 영상을 캡처하여 1279장의 깨끗한 이미지들을 취득한 뒤, 해당 프레임들에 안개 합성 파이프라인을 적용하여 모델 학습 시에 on-line 합성을 통해 1279장의 깨끗한 이미지로부터 매번 새로운 안개 이미지쌍을 생성하여 모델을 학습했다. 생성된 안개 이미지쌍은 다양한 scene을 표현하기 위해 512×512 크기로 random crop 혹은 resize되어 학습에 사용되었다.

3. Experimental Settings

본 실험은 Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz와 NVIDIA RTX 3090 GPU, 256GB RAM의 환경에서 진행되었다. Airlight 모델과 transmission 모델은 총 1,000 에포크 동안 학습되었으며, learning rate는 0.0001로 설정하였고 Adam optimizer를 이용하여 학습하였다. 또한 배치 크기는 1로 설정하였다.

V. Results

1. Visible Distance Increment



Fig. 6. Visible distance comparison between fog image (left) and defogged image (right).

안개 제거의 주 목적은 가시거리 증가를 통해 물체 인식 능력과 영상의 가시성을 향상시키는 것이다. 그러므로 안개 제거를 통해 어느 정도의 가시거리 증가량을 확보할 수 있는지를 확인해야 한다. 적절한 가시거리의 안개 영상을 얻기 위하여 안개가 없는 맑은 환경에서 5m부터 40m까지 거리 표지를 세워둔 후, 정확한 거리 측정을 위해 사람이

거리 표지들을 교차로 걸어가는 영상을 촬영하였다. 해당 영상의 depth map을 계산하고 airlight와 안개 계수를 적절히 조절하여 안개 영상을 합성하였다. 이를 통해 가시거리가 각각 5m, 10m, 15m인 안개 영상 샘플을 얻을 수 있었다. 또한 각각의 영상 샘플에 대해 제안 모델을 적용한 뒤 가시거리를 측정하여 가시거리 증가율을 계산하였다.

가시거리의 측정은 COCO 2017 데이터셋 [11] 에 대해 학습된 object detection 모델인 resnet-50 backbone의 DETR [12] 을 이용하였다. DETR이 영상에서 10프레임 연속으로 person 클래스의 물체를 0.5 이하의 신뢰도로 탐지하거나 전혀 탐지하지 못했을 때를 영상 속 사람이 가시거리를 벗어난 상황으로 설정하고, 해당 시점의 프레임을 추출하여 최대 가시거리를 측정하였다.



Fig. 7. The following images are from synthetic fog video (left) and defogged video (right) with visible distances of 5m, 10m, and 15m from top to bottom. These frames represent the moments when a pedestrian moves out of the visible range.

Table 1. Visible Distance Increment

	Before defog	After defog	Increment (%)
Visible Distance	5m	10m	100%
	10m	19m	90%
	15m	27m	80%
	Average	-	90%

가시거리가 5m였던 안개 영상은 안개 제거 모델로 처리한 뒤, 가시거리가 10m로 증가하였으며, 10m 안개 영상은 19m로, 15m 안개 영상은 27m로 증가하였다. 평균적으로 90%의 가시거리가 증가하였음을 확인하였다.

2. Comparisons with Baselines

2.1 Inference speed comparison

대표적인 단일 이미지 안개 제거 데이터셋인 RESIDE [13] 의 subset인 SOTS Outdoor에 대하여 높은 성능을 나타내는 모델들인 Dehamer [4], DehazeFormer [5], ChaIR [6] 을 baseline 모델들로 선정하여 본 논문에서 제안하는 모델과 비교하려 한다. ChaIR 모델은 해당 데이터셋의 테스트셋에 대해서 40.73dB의 높은 PSNR 수치를 나타내는 state-of-the-art 모델이다.

추론 속도를 비교하기 위하여 각각의 FP32 정밀도의 모델들을 FHD (1920×1080) 해상도의 이미지에 대하여 NVIDIA RTX 3090 GPU에서 100회 추론한 뒤, 평균 FPS를 계산하였다.

Table 2. Inference speed comparison with baseline models.

Models	FPS (frames/sec.) ↑
Dehamer [4]	3.026
DehazeFormer [5]	<u>10.58</u>
ChaIR [6]	3.213
Ours	32.60

제안하는 모델은 airlight 모델과 transmission 모델의 두 가지 모델을 모두 추론해야 함에도 모델의 경량화로 인해 다른 모델보다 3배에서 10배 이상의 속도로 추론했다.

Table 3. Quantative comparisons on synthetic fog image testset.

Models	PSNR ↑	SSIM ↑	FID ↓
Fog image	20.65	0.7459	83.01
Dehamer	21.97	0.7905	82.64
DehazeFormer	<u>22.12</u>	<u>0.7935</u>	<u>80.97</u>
CharIR	22.10	0.7933	81.58
Ours	28.94	0.8777	48.43

또한 SOTS Outdoor 데이터셋에 대해 사전 학습된 baseline 모델들과 제안하는 방식으로 학습된 모델을 fog image formation modeling 방식으로 합성된 안개 이미지 100장으로 구성된 테스트 데이터셋에 대하여 추론하였다. 모델의 결과가 합성 이전의 원본 이미지와 얼마나 유사한 지 비교하기 위하여 full-reference quality metric인 YCbCr 공간에서의 PSNR과 RGB 공간에서의 SSIM, FID를 측정하였으며 그 결과는 Table 3과 같다. 모든 metric에서 제안하는 모델이 합성된 안개 테스트 데이터셋에 대하여 안개를 합성하기 전의 깨끗한 이미지와 가장

비슷한 결과를 나타냈다.

또한, 저사양 기기에서의 간헐적 추론 방식의 유효성을 알아보기 위해 13th Gen Intel(R) Core(TM) i7-1360P @ 2.20 GHz의 CPU와 Intel(R) Iris(R) Xe Graphics의 내장 GPU가 장착된 미니 PC 환경에서 안개 제거 모델을 일반 추론과 간헐적 추론의 두 가지 방식을 적용하여 처리 속도를 비교하였다. 일반 추론 방식으로 모든 프레임에 대해 모델을 적용했을 시 초당 2.025 프레임을 처리할 수 있었으나 간헐적 추론 방법을 적용한 결과, 초당 28.62 프레임을 처리했다. 따라서 간헐적 추론으로 저사양 디바이스에서도 실시간 안개 제거가 가능함을 확인할 수 있었다.

하지만 간헐적 추론 방식은 일반 CCTV와 같이 고정된 카메라 영상에 대해서는 일반 추론 방식과 유사한 결과를 나타내지만, PTZ 카메라와 같이 카메라가 크게 움직이는 상황에서는 급격히 변화하는 transmission을 빠르게 추론하지 못해 품질이 떨어지게 된다.

2.2 Visual comparison on real-world images

제안하는 모델과 사전 학습된 기존 모델들이 현실의 안개 이미지에서 어느 정도의 안개 제거 성능을 갖고 있는지 시각적으로 비교하기 위하여, 실제 CCTV로부터 안개 상황의 영상을 취득하여 모델들로 추론하였다.

실제 환경에서는 안개가 있는 영상과 없는 영상을 동시에 취득하는 것이 불가능하기 때문에 PSNR이나 SSIM, FID와 같은 full-reference metric들을 사용하여 정량적인 비교를 할 수 없다. 따라서 본 연구에서는 안개 이미지와 모델 추론 결과를 정성적으로 비교하였다.

Fig. 5의 추론 결과를 비교해보면 일반적인 무채색 계열의 안개 이미지의 경우, baseline 모델들이 안개를 제거하는 하나 결과 이미지가 어둡게 되거나 제거되는 안개량이 매우 적다. 하지만 제안하는 모델은 자연스러운 색감의 안개 제거 이미지를 도출한다.

또한 새벽과 같이 안개의 색상이 푸른 환경에서 취득된 안개 이미지의 경우, baseline 모델들은 안개를 인식하지 못해 Fig. 5에서 볼 수 있듯이 안개를 전혀 제거하지 못했다. 하지만 제안하는 모델은 다양한 airlight에 대해 합성된 데이터셋으로 학습되어 자연스러운 안개 제거 이미지를 추론할 수 있었다.

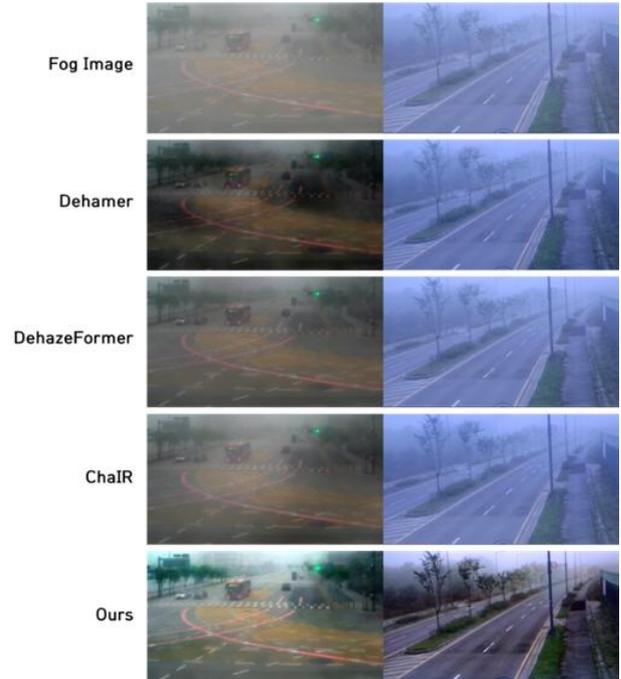


Fig. 8. Visual comparison of baselines and our model on real-world images.

VI. Conclusions

본 논문에서는 고해상도 안개 영상에 대하여 실시간 안개 제거 처리가 가능한 모델과 협동적 네트워크를 이용한 모델 학습 및 추론 방법, 데이터셋 합성 방식을 제안하였다. 기존의 딥러닝 기반 안개 제거 모델들이 갖는 문제점인 느린 추론 속도 문제와 현실의 영상에 대한 성능 저하 문제를 각각 모델의 기능별 분리 및 경량화, depth map과 degradation을 이용한 합성 데이터셋 생성 방법을 통해 개선하였으며 저성능의 디바이스에서도 실시간 안개 처리가 가능하도록 하는 간헐적 추론 방식을 제안하였다. 실제 CCTV의 고해상도 안개 이미지에 대한 처리 속도와 결과를 기존의 모델들과 비교하여, 제안한 방법의 효과를 확인하였다. 하지만 협동적 네트워크를 이용한 방식은 fog image formation modeling을 기반으로 하기에 이미지 내에서 균일한 색상과 밀도의 안개만 존재하는 상황을 가정하므로 homogeneous defogging에는 효과적이지만, 안개가 균일하지 않은 non-homogeneous defogging에서는 성능이 저하된다는 한계점이 존재한다.

ACKNOWLEDGEMENT

This work was supported by the Technology Innovation Program (20019466, Development of Main computing system dedicated to integrate video codec device, autonomous flight, object detect/recognition, communication system, FC) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).

REFERENCES

- [1] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, Vol. 33, No. 12, pp. 2341-2353, June 2009. DOI: 10.1109/cvpr.2009.5206515
- [2] Y. Xu, J. Wen, L. Fei, and Z. Zhang, "Review of video and image defogging algorithms and related studies on image restoration and enhancement," *Ieee Access*, Vol. 4, pp. 165-188, Dec. 2016. DOI: 10.1109/ACCESS.2015.2511558
- [3] H. Israël, and F. Kasten, "Koschmieders theorie der horizontalen sichtweite," *Die Sichtweite im Nebel und die Möglichkeiten ihrer künstlichen Beeinflussung*, pp. 7-10, 1959. DOI: 10.1007/978-3-663-04661-5_2
- [4] C. L. Gui, Q. Yan, S. Anwar, R. Cong, W. Ren, and C. Li, "Image dehazing transformer with transmission-aware 3d position embedding," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5812-5820, Jun. 2022. DOI: 10.1109/cvpr52688.2022.00572
- [5] Y. Song, Z. He, H. Qian, and X. Du, "Vision transformers for single image dehazing," *IEEE Transactions on Image Processing*, Vol. 32, pp. 1927-1941, Mar. 2023. DOI: 10.1109/tip.2023.3256763
- [6] Y. Cui, and A. Knoll, "Exploring the potential of channel interactions for image restoration," *Knowledge-Based Systems*, Vol. 282, pp. 111156, Dec. 2023. DOI: 10.1016/j.knosys.2023.111156
- [7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, Vol. 63, No. 11, pp. 139-144, Oct. 2020. DOI: 10.1145/3422622
- [8] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, G. Sylvain, U. Jakob, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, Oct. 2020. DOI: 10.48550/arxiv.2010.11929
- [9] L. Chen, X. Chu, X. Zhang, and J. Sun, "Simple baselines for image restoration," *European conference on computer vision*, pp. 17-33, 2022. DOI: 10.1007/978-3-031-20071-7_2
- [10] O. Ronneberger, Olaf, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," In *Proceedings of Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference, part III 18*, pp. 234-241, Nov. 2015. DOI: 10.1007/978-3-319-24574-4_28
- [11] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, "Microsoft coco: Common objects in context," In *Proceedings of the Computer Vision-ECCV 2014: 13th European Conference, Part V 13*, pp. 740-755, Sep. 2014. DOI: 10.1007/978-3-319-10602-1_48
- [12] N. Carion, F. Massa, G. Synnaeve, N. Usunier, and A. Kirillov, "End-to-end object detection with transformers," In *European conference on computer vision*, pp. 213-229, Nov. 2020. DOI: 10.1007/978-3-030-58452-8_13
- [13] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking Single-Image Dehazing and Beyond," *IEEE Transactions on Image Processing*, Vol. 28, No. 1, pp. 492-505, Aug. 2018. DOI: 10.1109/tip.2018.2867951

Authors



Sanghyun Jung received the B.S. and M.S. degrees in Computer Science from Korea Advanced Institute of Science and Technology in 2021 and 2023, respectively.

Mr. Jung is currently a researcher in ENERZAI, Korea. He is interested in image restoration, image generation and video restoration.