

Dark channel prior와 고주파 정보 보정을 결합한 Denoising Diffusion Bridge Model 기반 안개 제거

정현정, *이민식

한양대학교 전자공학과, 한양대학교 ERICA 전자공학부

hounja10@hanyang.ac.kr, mleepaper@hanyang.ac.kr

Dehazing Based on a Denoising Diffusion Bridge Model Integrating Dark Channel Prior and High-Frequency Information Compensation

Jeong Hyeon Jeong, *Lee Min Sik

Hanyang Univ., Hanyang ERICA Univ.

요약

안개 환경에서 촬영된 이미지는 빛의 산란으로 인해 화질이 저하되어 영상 내 사물과 사람을 식별하기 어려운 경우가 많다. 이러한 문제를 해결하기 위해 본 논문에서는 고전적인 안개 제거 기법을 통해 전송 맵을 생성하고, 고주파 성분을 강조하는 블록을 활용하여 안개 영상 내 객체의 세부 정보를 강화한 후 이를 Diffusion Bridge Model에 적용하는 접근 방식을 제안한다. 제안한 방법은 영상 내 안개를 효과적으로 제거함으로써 객체와 사람의 시각적 식별 성능을 향상하는 것을 목표로 한다.

I. 서 론

안개 영상은 공기 중 입자에 의한 빛의 산란으로 인해 일반적인 영상에서 제공되는 거리감이 저하될 뿐만 아니라, 영상 내 객체와 사람을 명확하게 인식하기 어려운 문제를 가진다. 이러한 문제를 해결하기 위해 고전적인 방법으로는 안개 영상이 가지는 RGB 채널의 특성을 활용하여 안개를 제거하는 기법(Dark channel prior, DCP)[1]과 같은 통계적 기법들이 제안됐으며, 최근에는 딥러닝 모델을 학습하여 안개를 제거하는 방식[2]이 활발히 연구되고 있다. 본 논문에서는 이 두 가지 접근 방식을 결합하여, 고전적인 방법에서 사용되는 전송 맵과 주파수 보상 블록을 Diffusion Bridge 모델에 통합함으로써 영상 내 안개를 효과적으로 제거하는 새로운 방법을 제안한다.

II. 본론

가. Dark Channel Prior 기반 전송 맵 추정

안개 없는 영상에서 안개 영상이 구성될 때의 기본적인 물리 수식은 다음과 같다.

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

여기서, I 는 안개 영상, J 는 안개가 없는 깨끗한 영상, t 는 전송맵, A 는 대기광, x 는 영상의 화소 위치를 나타낸다. 전송맵은 다음과 같이 구성된다.

$$t(x) = e^{-\beta d(x)} \quad (2)$$

여기서, β 는 산란 계수이며, $d(x)$ 는 깊이 맵을 의미한다.

본 논문에서 사용할 고전적인 방법 중 하나인 DCP로 전송맵을 얻어 해당 전송맵을 Denoising Diffusion Bridge Model(DDBM)의 UNet에 조건으로 넣는 방법을 제안한다. 먼저 DCP는 영상에서 안개가 존재하지 않는 화소의 위치의 RGB 채널 중 적어도 하나의 값은 거의 0에 가깝다라는 통계적인 관찰을 통해 전송맵을 얻게 된다. 전송맵을 구하는 식은 다음과 같다.

$$\tilde{t}(x) = 1 - w \min_c (\min_{y \in \Omega(x)} (\frac{I^c(y)}{A^c})) \quad (3)$$

여기서, c 는 채널, \tilde{t} 는 DCP로 구해진 전송맵, w 는 안개를 제거하는 정도를 낮추는 값으로 안개가 과도하게 제거되어 영상 자체의 자연스러움을 해치는 것을 방지하기 위해 넣으며 w 는 0.95로 사용한다. Ω 는 지역적 패치로 DCP로 전송맵을 만들었을 때 디테일 보존을 위해 사용한다.

나. 주파수 보정 기반 Feature 전달을 위한 UNet 구조

자연 영상에서 고주파 성분은 영상의 세부적인 디테일과 경계 정보를 나타내며, 안개와 같은 노이즈가 포함될 경우 이러한 고주파 성분은 쉽게 손상될 수 있다. 또한 일반적인 딥러닝 모델은 학습 과정에서 저주파 성분을 우선적으로 학습하는 경향이 있어, 고주파 정보의 복원이 제한될 수 있다. 이를 보완하기 위해 본 논문에서는 고주파 성분을 강조하고 복원하기 위한 주파수 보상 블록(Frequency Compensation Block, FCB)을 사용한다. FCB는 입력 특징에 대해 다중 가우시안 지역 필터를 적용한 후, 이를 간의 차이를 계산함으로써 고주파 성분을 효과적으로 추출한다.

$$s_{k,\sigma} = G_{k \times k}^\sigma * s \quad (4)$$

$$s' = [s, s - s_3, s_3 - s_5, s_5 - s_7] \quad (5)$$

여기서, s 는 입력 feature, k 는 가우시안 커널의 크기이며, σ 는 가우시안의 표준편차이다.

이후 위에서 뽑은 주파수 성분들을 학습 가능한 가중치로 결합하는 블록이다. 식은 다음과 같다.

$$\bar{s} = w_1 s + w_2 (s - s_3) + w_3 (s_3 - s_5) + w_4 (s_5 - s_7) \quad (6)$$

여기서, w_i 는 학습 가능한 파라미터이다.

FCB는 기존 UNet 구조에서 사용되는 skip connection을 대체하여 적용된다. 기존 UNet에서는 인코더의 특징이 디코더로 단순히 복사되어 전달되는 반면, FCB를 적용한 구조에서는 주파수 보정이 수행된 특징이 디코더로 전달된다. 이를 통해 디코더는 고주파 성분이 강화된 특징을 입력으로 받아, 영상의 세부 구조와 경계 정보를 보다 효과적으로 학습할 수 있다.

다. 영상 안개 제거를 위한 Denoising Diffusion Bridge Model

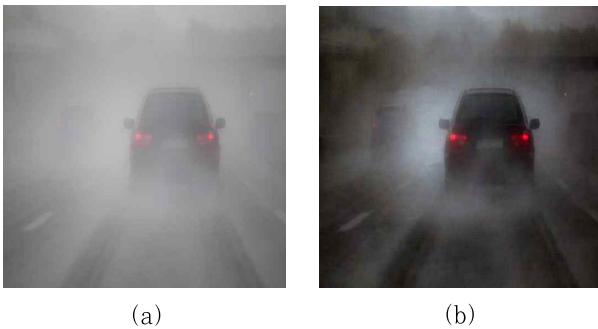
안개 제거 문제에서는 입력 영상과 목표 영상 간의 전역적인 분포 차이를 안정적으로 모델링하는 생성 구조가 함께 요구된다. 이에 본 연구에서는 고주파 정보가 강화된 특징을 효과적으로 활용하기 위해, 입력 영상 분포에서 목표 영상 분포로의 확률적 변환을 직접 학습하는 Denoising Diffusion Bridge Model(DDBM)을 기반으로 전체 복원 과정을 도입하는 것을 제안한다.

diffusion bridge의 역 시간 확률 미분 방정식은 조건부 분포를 따르는 확률 과정으로 표현되며, 다음과 같이 주어진다.

$$dx_t = [f(x_t, t) - g^2(t)(s(x_t, t, y, T) - h(x_t, t, y, T))]dt + g(t)d\hat{w}_t \quad (7)$$

x_t 는 중간 상태 영상, x_T 는 입력 안개 영상, $s(x_t, t, y, T)$ 는 조건부 score를 근사하는 신경망, $h(x_t, t, y, T) = \nabla_{x_t} \log p(x_T | x_t)|_{x_t=x, x_T=y}$ 는 Doob's h-transform[4]에 의해 도입된 drift 보정 항으로 특정 종단점 x_T 를 갖도록 중간 상태의 분포를 정의하는 역할을 하며, 이는 역과정에서 목표 영상으로의 복원을 가능하게 한다. 이러한 diffusion bridge의 역과정을 실제로 구현하기 위해서는, 조건부 분포에 대한 score 함수를 학습하는 과정이 필요하며 다음과 같은 수식으로 나타낸다.

$$\begin{aligned} L(\theta) &= E_{x_0, x_T, x_t} [\|s_\theta(x_t, x_T, t) - \nabla_{x_t} \log q(x_t | x_0, x_T)\|_2^2] \quad (8) \\ s_\theta(x_t, x_T, T) &\text{는 조건부 분포 } q(x_t | x_T) \text{의 score를 근사하는 신경망이다.} \\ \text{이를 통해 모델은 입력 영상 } x_T \text{로부터 목표 영상 } x_0 \text{으로의 역 시간 과정} \\ \text{을 복원할 수 있다.} \end{aligned}$$



	MUSIQ[6]	NIQE[7]
RESIDE- β [5]	20.256	4.4

라. 실험

실험에는 안개가 존재하는 영상과 그렇지 않은 영상 쌍으로 구성된 RESIDE 데이터셋[5]을 사용하여 학습과 테스트를 진행하였다. 해당 데이터셋은 총 69,130장의 안개 영상과 2,062장의 안개가 없는 영상으로 구성되어 있으며, 하나의 깨끗한 영상으로부터 다양한 안개 밀도를 적용하여 다수의 안개 영상을 생성한 것이 특징이다. 학습 과정에서는 모든 영상을 512×512 크기로 고정하여 사용하였으며, 총 400K iteration 동안 모델을 학습시켰다. 정성적 실험 결과는 그림 1에 제시되어 있으며, 그림 1(a)의 안개 영상과 비교했을 때 제안한 방법으로 복원된 영상은 전반적인 가시성과 객체의 경계가 향상된 것을 확인할 수 있다. 또한 정량적 평가는 표 1과 같이 MUSIQ와 NIQE 지표를 사용하여 수행하였다.

III. 결론

본 논문에서는 영상 내 안개 제거를 위해 고전적인 안개 제거 기법에서

사용되는 전송 맵과 주파수 보상 블록을 Diffusion Bridge Model(DDBM)에 통합하는 새로운 접근 방법을 제안하였다. 제안한 방법은 주파수 정보를 활용하여 영상의 세부 구조를 효과적으로 복원함으로써, 실험 결과 안개가 포함된 영상에서 전반적인 가시성이 향상됨을 확인하였다. 그러나 안개 제거 과정에서 영상의 밝기와 색감이 일부 감소하여 객체 식별이 어려워지는 한계 또한 관찰되었다. 향후 연구에서는 이러한 한계를 극복하기 위해 영상의 색상 정보를 보다 효과적으로 보존하면서 안개를 제거하는 방법을 연구할 예정이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임
(IITP-2026-RS-2020-II201741)

참 고 문 헌

- [1] He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." IEEE transactions on pattern analysis and machine intelligence 33.12 (2010): 2341-2353.
- [2] Wang, Jing, et al. "Frequency compensated diffusion model for real-scene dehazing." Neural Networks 175 (2024): 106281.
- [3] Zhou, Linqi, et al. "Denoising diffusion bridge models." arXiv preprint arXiv:2309.16948 (2023).
- [4] Doob, Joseph L., and J. I. Doob. Classical potential theory and its probabilistic counterpart. Vol. 262. New York: Springer, 1984.
- [5] Li, Boyi, et al. "Benchmarking single-image dehazing and beyond." IEEE transactions on image processing 28.1 (2018): 492-505.
- [6] Mittal, Anish, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain." IEEE Transactions on image processing 21.12 (2012): 4695-4708.
- [7] Mittal, Anish, Rajiv Soundararajan, and Alan C. Bovik. "Making a "completely blind" image quality analyzer." IEEE Signal processing letters 20.3 (2012): 209-212.