

깊이 기반 3D 안개 합성과 시간적 일관성

Depth-Aware 3D Fog Synthesis with Temporal Consistency

Abstract

We propose a novel method for generating three-dimensional (3D) fog effect that overcomes the limitations of existing approaches. Previous fog synthesis techniques mainly focus on 2D images and typically produce either homogeneous fog or non-homogeneous fog based on random noise, resulting in limited depth perception and difficulty in applying fog consistently to video sequences. In contrast, the proposed method generates non-homogeneous fog while incorporating depth-dependent fog density, enabling realistic volumetric appearance. Additionally, a temporal filtering mechanism is introduced to ensure smooth temporal transitions, allowing the synthesized fog to vary naturally over time and be seamlessly integrated into videos. Experimental results demonstrate that our approach produces visually convincing 3D fog with improved realism and temporal coherence compared to conventional methods.

요약

본 논문에서는 기존 기법의 한계를 극복하는 새로운 3차원(3D) 안개 합성 방법을 제안한다. 기존 안개 합성 기법들은 주로 2차원 영상에 초점을 맞추며, 균일한(homogeneous) 안개 또는 랜덤 노이즈 기반의 비균질(non-homogeneous) 안개를 생성하는 방식이 일반적이다. 이러한 방식은 깊이감을 충분히 제공하지 못하며, 동영상에 자연스럽게 안개를 적용하기 어렵다는 한계가 있다. 이에 반해, 제안하는 방법은 랜덤 노이즈 기반 비균질 안개를 생성함과 동시에 깊이에 따라 안개 밀도가 달라지는 특성을 부여하여 보다 사실적인 부피감을 표현한다. 또한 temporal filter를 도입하여 시간적으로 자연스럽게 변화하는 안개 효과를 실현함으로써 동영상에도 seamlessly 적용할 수 있다. 실험 결과, 제안 기법은 기존 방법 대비 향상된 사실감과 시간적 일관성을 갖춘 3D 안개 합성 성능을 보여준다.

Key words : 3D Fog Synthesis, Non-Homogeneous Fog, Depth-Aware Fog Density, Temporal Filtering, Video Fog Synthesis

*

★ Corresponding author

E-mail : , Tel. :

+ These authors contributed equally to this work.

※ Acknowledgment

This work was supported by

Manuscript received Nov. 28, 2024; revised Dec. 15, 2024; accepted Dec. 18, 2024

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서론

최근 컴퓨터 비전 및 자율주행 분야에서는 악천후 환경에서의 성능 저항 문제를 완화하기 위해 합성 기반 데이터 생성 기법이 활발히 연구되고 있으며, 특히 안개(fog) 환경은 명암 대비 저하와 구조적 정보 손실을 유발하기 때문에, 현실적인 합성 기법의 필요성이 지속적으로 제기되고 있다. 기존 안개 합성 연구는 주로 2D 이미지 기반 접근법에 집중됐으며, 대표적으로 대기 산란 모델(atmospheric scattering model)에 기반한 물리적 영상 복원 연구와 Perlin noise를 오버레이하여 비균질 패턴을 묘사하는 절차적 합성 방식이 널리 사용됐다[1, 2]. 그러나 이러한 2D 기반 접근은 이미지 평면 단위에서 독립적으로 처리되기 때문에 깊이에 따른 누적 산란 효과, 실제 환경에서 나타나는 3D 공간적 비균질성, 그리고 동영상 기반 합성 시 필연적으로 발생하는 temporal flickering 문제를 해결하기 어렵다. 이로 인해 생성된 안개는 볼륨 기반의 물리적 현실성과 시간적 연속성을 갖추지 못하여, 실제 환경과의 차이가 큰 제한점을 가진다.

이러한 한계를 극복하기 위해 일부 연구들은 volumetric rendering 기반의 3D 안개 합성 또는 신경망을 활용한 implicit density modeling을 시도하고 있다. 예를 들어 volumetric path tracing을 이용한 안개 렌더링 기법[3]이나, NeRF 기반 환경 효과 모델링 연구[4]는 더욱 물리적으로 정합적인 재현을 시도하지만, 과도한 계산 비용, 복잡한 장면 표현 요구, 그리고 특정 데이터 조건에 대한 의존성 등으로 인해 범용적 안개 합성 파이프라인으로 활용하기 어렵다.

본 연구는 이러한 기존 2D 중심 기법과 복잡한 3D 렌더링 기반 기법 사이의 간극을 메우기 위해, 깊이 기반 밀도 조절, random noise를 활용한 non-homogeneous volumetric field 생성, 그리고 temporal filtering을 통한 시간적 안정성 확보를 통합한 새로운 3D 안개 합성 방법을 제안한다. 제안된 기법은 3D 공간에서 비균질 밀도를 갖는 fog volume을 구성한 뒤, 이를 깊이 맵과 결합하여 퍽셀별 누적 투과율을 계산함으로써 기존 2D 방식의 한계였던 깊이에 따른 자연스러운 안개 누적 효과 및 공간적 텍스처 변화를 구현한다. 또

한 temporal filter를 활용하여 동영상에서 프레임 간 fog field의 미세한 변화를 안정적으로 제어함으로써 flickering이 없는 자연스러운 시간적 일관성을 제공한다.

안개 합성 방식의 차이를 시각적으로 보여주기 위해, 본 연구는 homogeneous 방식, 2D noise 기반 비균질 방식, 그리고 제안한 3D volumetric 접근을 비교한 그림을 제시한다. 그림 1은 단일 프레임만을 보여주므로 flickering과 같은 시간적 차이는 supplementary 동영상에서 확인할 수 있다. 그림에서 보이듯 기존 방식은 깊이 변화와 공간적 구조 표현에 한계가 있는 반면, 제안한 방법은 volumetric 비균질성과 깊이 기반 감쇠를 통합하여 정교한 시각적 결과를 제공한다.



Fig. 1. Comparison of three fog synthesis methods
그림 1. 세 가지 안개 합성 방법의 비교

II. 본론

1. Homogeneous Fog Synthesis

Homogeneous 안개 합성은 장면 전체에 동일한 매질 밀도(scattering coefficient)를 가정하는 고전적 접근으로, 물리적 이론[1]에 기반해 발전했다. 이 방식은 입력 이미지의 깊이 맵을 이용해 퍽셀 단위 transmittance를 지수 함수 형태로 계산하고, atmospheric light와 선형 결합하여 합성 이미지를 생성한다. 이러한 모델은 계산 효율이 높고 구현이 단순하다는 장점이 있으나, 공간적 비균질성이나 난류 기반 구조 변화 등 실제 환경의 volumetric 특성을 반영할 수 없다. 또한 프레임 단위로 독립적으로 합성이 이루어지기 때문에, 동영상 처리 시 temporal flickering이 쉽게 발생하며, 3D 구조적 깊이 감과 시간적 일관성이 요구되는 현대적 합성 시나리오에서는 한계가 명확하다. 이러한 이유로 homogeneous 모델은 여전히 기초적 비교 기준으로 활용되지만, 현실적이고 입체적 안개 표현을 위해서는 비균질적이고 volumetric 접근이 필요하다.

2. Non-Homogeneous Fog Synthesis

Non-homogeneous 안개 합성은 공간적 밀도 변화를 모델링하여 현실적 안개 분포를 재현하려는 접근으로, homogeneous 모델의 단조로운 구조를 보완한다. 일반적으로 Perlin noise, Fractal noise, Worley noise 등 절차적 패턴[2]이 깊이 기반 산란 모델 위에 중첩되며, 장면 내 국소적 높도 차이를 표현하기 위해 다중 주파수 노이즈를 혼합하는 방식이 널리 사용된다. 이러한 방법은 비교적 낮은 계산 비용으로 비균질 패턴을 생성할 수 있으나, 실제 대기 중의 volumetric 변화를 정확히 반영하지 못하고, noise field가 이미지 평면에서 정의되기 때문에 3D 공간적 구조가 손실된다는 한계가 존재한다. 또한 절차적 노이즈 기반 합성은 프레임 간 독립적으로 적용될 경우 시각적 일관성을 보장하기 어려워, 동영상에서 flickering과 패턴 도약(jumping artifact)이 쉽게 발생한다. 그럼에도 불구하고 non-homogeneous 방식은 homogeneous 모델 대비 현실감이 높고 구현이 간단해 다양한 연구에서 baseline으로 활용된다.

3. Temporal Consistency in Fog Synthesis

안개 합성에서 temporal consistency는 동영상 기반 합성에서 필수적인 요소로, 프레임 간 안개 밀도와 구조가 불연속으로 변화하면 인지적 위화감과 시각적 artifact가 발생한다. 기존 방식을 살펴보면, 2D 기반 합성은 각 프레임을 독립적으로 처리하기 때문에 depth noise, noise field 문제 등으로 인해 flickering이 빈번하게 나타난다. 이를 개선하기 위해 optical flow 기반 warping, recurrent filtering, temporally correlated noise field 생성 등 다양한 접근이 제안됐으나[5], 대부분 이미지 평면에서의 후처리에 의존하기 때문에 3D 공간을 통한 볼륨 일관성(volumetric coherence)을 근본적으로 확보하기 어렵다. 특히 non-homogeneous 안개에서는 국소적 밀도 변화가 프레임마다 불규칙하게 변동하기 때문에 temporal smoothing만으로는 artifact 제거에 한계가 존재한다. 그 결과, 실제 환경에서 관찰되는 연속적 밀도 변화, 자연스러운 움직임, 3D 공간에 의해 제약되는 구조적 일관성을 동시에 충족하기 위한 더욱 체계적인 접근이 요구된다.

4. Proposed Method

본 연구에서는 공간적·시간적으로 비균질한 안개 분포를 3D 볼륨 형태로 직접 생성하고, 이를 깊이 기반 전송 모델과 결합하여 동영상 전체에서 일관된 안개 효과를 합성하는 새로운 3D fog synthesis 프레임워크를 제안한다. 제안된 방법은 기존 2D 기반 fog synthesis의 구조적 한계를 해결하기 위해 (A) 3D FFT 기반 volumetric noise field 생성, (B) space-time Gaussian filtering 기반 비균질 분포 생성, (C) depth-aware transmission modulation, (D) Koschmieder 대기 전파 모델을 이용한 최종 합성의 네 단계를 포함한다. 이 접근은 전통적 대기 산란 모델[1], 절차적 노이즈 기반 방법[2], 그리고 동영상 temporal filtering 연구[5]를 확장하여, 3D 볼륨 정보와 시간적 안정성을 동시에 반영한다는 점에서 차별적이다.

A. 3D Volumetric Fog Field Generation

우선 임의의 시공간 볼륨 $N(t, y, x)$ 에 대해 정규분포 기반 3D 랜덤 노이즈를 생성한다.

$$N(t, y, x) \sim \delta(0, 1) \quad (1)$$

이 노이즈는 3D FFT를 통해 주파수 공간으로 변환되며,

$$F(N) = \text{FFT}(N) \quad (2)$$

식 (3)에 나와 있는 공간 주파수 (f_x, f_y) 와 시간 주파수 (f_t) 를 포함한 3D Gaussian 필터를 적용하여 고주파 노이즈를 억제한다. 이는 Perlin noise 기반 절차적 합성 방식[2]과 달리 temporal axis까지 포함한 3D filtering을 수행하므로 동영상에서 발생하는 flickering 문제를 근본적으로 줄이는 특징을 갖는다.

$$G(f_x, f_y, f_t) = e^{-\left(\frac{f_x^2}{2\sigma_x^2} + \frac{f_y^2}{2\sigma_y^2} + \frac{f_t^2}{2\sigma_t^2}\right)} \quad (3)$$

Gaussian filtering이 적용된 주파수 표현은 식 (4)에 나와 있으며, 역 FFT를 통해 volumetric fog density field $V(t, y, x)$ 를 복원한다.

$$F_{\text{filtered}} = F(N) \cdot G \quad (4)$$

$$T_{\text{final}}(t, y, x) = T_{\text{base}}(t, y, x) \cdot T_{\text{depth}}(y, x) \quad (9)$$

B. Contrast/Gamma Normalization

정규화된 볼륨 V 는 contrast 계수 c 와 gamma 보정 γ 를 통해 비균질성의 강도를 제어한다. 이 단계는 식 (5)에 나와 있으며, participating media의 absorption coefficient를 조절하는 과정과 유사한 역할을 수행하며[3], 실제 환경에서 관찰되는 농도 변화를 다양하게 재현할 수 있다. V_{\min} 과 V_{\max} 은 V 의 최솟값과 최댓값이다.

$$T_{\text{base}}(t, y, x) = \left(\frac{V - V_{\min}}{V_{\max} - V_{\min}} \right)^{\gamma} \cdot c \quad (5)$$

C. Depth-Aware Transmission Modeling

제안한 방법의 두 번째 핵심은 3D 비균질 안개에 depth-dependent transmission을 추가하여 실제 안개가 거리 증가에 따라 누적되는 효과를 모델링하는 것이다.

영상의 수직 방향을 따라 near-to-far 깊이 비율 $p \in [0,1]$ 을 정의하고, 식 (6)과 식 (7)에 나와 있는 exponential과 sigmoid 모델을 통해 장면 구조에 적합한 깊이 곡선을 선택할 수 있다.

$$d(y) = \frac{e^{\alpha p(y)} - 1}{e^{\alpha} - 1} \quad (6)$$

$$d(y) = \frac{s(y) - s_{\min}}{s_{\max} - s_{\min}}, s(y) = \frac{1}{1 + e^{-k(p(y) - p_0)}} \quad (7)$$

최종 깊이 기반 transmission은 다음과 같이 계산되며, 이는 전통적 대기 산란 이론[1]의 핵심 요소인 $e^{-\beta d}$ 형태의 빛 감쇠 모델을 반영한 것이다.

$$T_{\text{depth}}(y, x) = e^{-\beta d(y)} \quad (8)$$

본 연구에서는 비균질 안개와 깊이 기반 transmission을 곱셈적으로 결합하여 최종 전달 맵을 생성한다. 이 결합은 식 (9)에 나와 있으며, 비균질 volumetric 텍스처, 깊이에 따른 누적 감쇠, 시간적 일관성을 동시에 반영한다는 점에서 기본 방식들과 뚜렷한 차별점을 가진다.

D. Rendering via Koschmieder's Model

마지막으로 Koschmieder 모델[1]을 기반으로 안개가 도입된 프레임을 합성한다.

$$I_{\text{hazy}} = I \cdot T_{\text{final}} + A(1 - T_{\text{final}}) \quad (10)$$

여기서 A 는 atmospheric light이며, 영상의 전역적인 밝기 및 색조 영향을 결정한다. 이 과정을 모든 프레임에 적용하면 시간적으로 부드럽고 공간적으로 비균질하며 깊이감이 보존된 3D 안개 영상이 생성된다.

그림 2는 본 연구에서 제안한 3D 비균질 안개 합성 알고리즘의 전체 흐름을 개략적으로 나타낸 것이다.

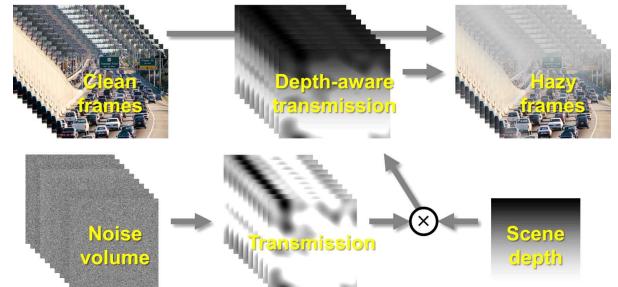


Fig. 2. Overview of the proposed fog synthesis method
그림 2. 제안한 안개 합성 알고리즘 개요

5. Evaluation

본 연구에서는 제안한 3D volumetric 안개 합성 기법의 성능을 가시적(qualitative) 및 정량적(quantitative) 측면에서 평가한다. 비교 대상은 homogeneous 기반 합성 방식과 2D noise 기반 비균질 합성 방식의 두 가지이다.

A. Qualitative Evaluation

그림 3은 세 가지 안개 합성 방식에 대한 대표적인 프레임 비교를 보여주며, 제안한 방법의 depth-aware 특성을 확인하기 위해 exponential 모델과 sigmoid 모델의 결과를 모두 함께 제시한다. 비교 결과, homogeneous 방식은 깊이 변화가 충분히 반영되지 않고, 2D noise 기반 비균질 방식은 공간적 패턴이 인위적으로 반복되는 경향을

보인다. 반면 제안한 방법은 깊이 기반 감쇠와 volumetric 비균질성을 동시에 반영하여 더욱 자연스럽고 현실적인 안개를 생성한다. 특히 exponential 모델은 먼 거리의 점진적 감쇠를 매끄럽게 표현하고, sigmoid 모델은 수평선이나 급격한 깊이 변화가 포함된 장면에서 자연스러운 밀도 변화를 구현한다. 이를 통해 제안한 방법의 시각적 우수성을 직관적으로 검증할 수 있다.

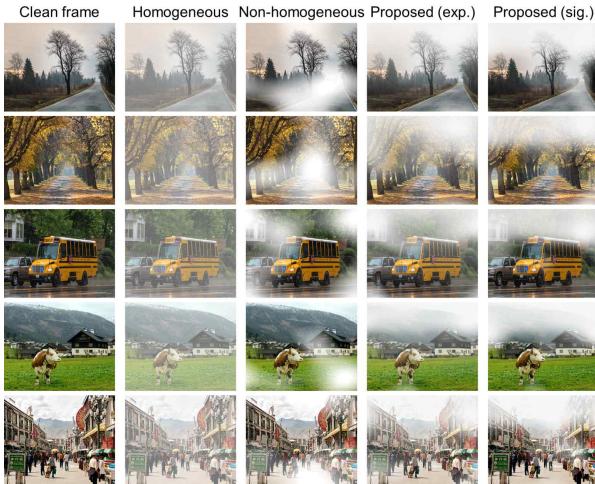


Fig. 3. Qualitative comparison results
그림 3. 가시적 평가 결과

A. Qualitative Evaluation

정량적 평가에서는 각 안개 합성 방식으로 생성된 안개 영상에 대해 기존 안개 제거 알고리즘 [6]을 적용하고, 그 결과를 원본 영상과 비교하여 PSNR 및 SSIM[7]을 측정한다. 공정한 비교를 위해, 세 가지 안개 합성 방식 모두 합성 단계에서 비슷한 수준의 초기 PSNR/SSIM(약 11.62/0.741)을 갖도록 파라미터를 조정하여 동일한 난이도의 안개 강도를 설정하였다. 그림 4는 세 방식으로 생성된 안개 영상에 대해 안개 제거를 수행한 예시를 보여준다.

본 평가는 더 낮은 PSNR/SSIM이 더 제거하기 어려운, 즉 더 현실적인 안개가 합성됨이라는 기존 연구의 관점을 따른 것이다. 즉, 안개 제거 알고리즘이 높은 성능을 내지 못할수록 합성된 안개가 더 복잡하고 실제 환경에 가깝다는 것을 의미한다.

표 1은 그림 4의 영상에 대한 PSNR 및 SSIM 값을 요약한 결과이다. 두 지표 모두에서 제안한

방식이 가장 낮은 수치를 기록했으며, 이는 제안한 3D volumetric 방식이 두 비교 기법보다 현실적이고 제거하기 어려운 안개 패턴을 생성함을 시사한다. 이러한 결과는 제안한 기법이 깊이 기반 감쇠, 비균질 volumetric 구조, temporal consistency를 모두 반영함으로써 기존 2D 기반 방법을 능가하는 합성 품질을 제공한다는 점을 정량적으로 확인한다.



Fig. 4. Dehazing results for the synthesized fog images
그림 4. 합성된 안개 영상에 대한 안개 제거 결과

Table 1. PSNR/SSIM results for Fig. 4
표 1. 그림 4의 PSNR/SSIM 결과

	No.	Homo.	Non-homo.	Proposed (exp.)	Proposed (sig.)
P S N R	1	32.26	11.26	16.96	15.76
	2	24.06	11.83	9.91	10.00
	3	28.65	9.91	9.05	8.27
	4	29.95	13.19	13.17	11.78
	5	32.27	15.61	14.38	13.49
		Avg.	29.44	12.36	11.86
S S I M	1	0.99	0.77	0.90	0.90
	2	0.90	0.72	0.68	0.68
	3	0.96	0.69	0.65	0.65
	4	0.97	0.83	0.76	0.80
	5	0.99	0.88	0.81	0.86
		Avg.	0.96	0.78	0.76

III. 결론

본 연구에서는 3D volumetric noise 생성, depth-aware transmission 모델링, 그리고 temporal filtering을 결합한 새로운 안개 합성 기법을 제안하였다. 제안된 방법은 기존의

homogeneous 및 2D noise 기반 비균질 합성 방식이 가지는 깊이 감쇠 부족, 공간적 구조 단순성, 동영상 적용 시 flickering 발생 등의 근본적 한계를 효과적으로 해결한다. 가시적 평가에서는 제안한 방식이 깊이에 따른 자연스러운 감쇠와 공간적 비균질성을 구현하며, 기존 방법보다 시각적으로 더 현실적인 안개 패턴을 생성함을 확인하였다. 또한 정량적 평가를 통해 제안한 방법이 PSNR 및 SSIM이 가장 낮은 값을 기록함으로써, 실제 환경에 가까운 복잡한 안개를 합성한다는 점을 정량적으로 검증하였다.

종합하면, 본 연구의 3D volumetric 안개 합성 기법은 깊이·공간·시간의 세 축을 모두 고려한 새로운 합성 프레임워크로, 기존 2D 기반 접근을 체계적으로 확장한다. 향후 연구에서는 더 다양한 실제 환경 모델링, 동적 기상 변화 시뮬레이션, 그리고 학습 기반 메커니즘과의 결합을 통해 더욱 일반화된 안개 합성 프레임워크로 발전시킬 수 있을 것이다.

References

- [1] S. G. Narasimhan and S. K. Nayar, “Chromatic framework for vision in bad weather,” *2000 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 598–605, 2000.
DOI: 10.1109/CVPR.2000.855874
- [2] K. Perlin, “An image synthesizer,” *ACM SIGGRAPH Computer Graphics*, vol. 19, no. 3, pp. 287–296, 1985.
DOI: 10.1145/325165.325247
- [3] H. W. Jensen and P. H. Christensen, “Efficient simulation of light transport in scenes with participating media,” *1998 Annual Conference on Computer Graphics and Interactive Techniques*, pp. 311–320, 1998.
DOI: 10.1145/280814.280925
- [4] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi and R. Ng, “NeRF: Representing scenes as neural radiance fields for view synthesis,” *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

DOI: 10.1145/3503250

- [5] J. Han, G. Sim, G. Kim, H. S. Lee, K. Choi, Y. Han and S. Cho, “DC-VSR: Spatially and Temporally Consistent Video Super-Resolution with Video Diffusion Prior,” *2025 SIGGRAPH Conference Papers*, pp. 1–11, 2025.

DOI: 10.1145/3721238.3730719

- [6] Y. Song, Z. He, H. Qian and X. Du, “Vision Transformer for Single Image Dehazing,” *IEEE Transactions on Image Processing*, vol. 32, pp. 1927–1941, 2023.

DOI: 10.1109/TIP.2023.3256763

- [7] W. Zhou, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

DOI: 10.1109/TIP.2003.819861

BIOGRAPHY