



일반논문 (Regular Paper)

방송공학회논문지 제31권 제3호, 2026년 5월 (JBE Vol.31, No.3, May 2026)

<https://doi.org/10.5909/JBE.2026.31.3.461>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

Positional Encoding, Sine 활성화 함수, Fourier Reparameterization 기반 뉴럴라이트필드 표현력 분석

장 영 진^{a)}, 정 현 민^{a)†}

Analysis of Neural Light Field Representation via Positional Encoding, Sine Activation Functions, and Fourier Reparameterization

Young Jin Jang^{a)} and Hyunmin Jung^{a)†}

요 약

본 연구는 Neural Light Field(NeLF)의 저주파 편향으로 인한 고주파 표현 한계를 극복하기 위해, Positional Encoding(PE), Sine 활성화 함수, 그리고 Fourier Reparameterization(FR) 기반 표현 기법을 NeLF에 적용하고 효과를 분석한다. Stanford Light Field 데이터셋을 이용해 Dense, Sparse 조건을 구성하고 기존 모델과 세 가지 개선 모델(PE, PE+SIN, SIN+FR)의 재구성 성능을 비교한다. Dense 조건에서는 PE+SIN이 가장 높은 PSNR을 기록하였으며, SIN+FR이 근소한 차이로 우수한 성능을 보였다. Sparse 조건에서는 PE+SIN이 큰 성능 저하를 보인 반면, SIN+FR은 2.56dB의 적은 PSNR 저하로 가장 안정적인 재구성 품질을 보였다. 결과적으로 SIN+FR의 표현 기법이 NeLF의 재현력 향상과 시점 밀도 감소 상황에 강인한 성능을 보임을 확인하였다.

Abstract

This study aims to address the high-frequency representation limitations of Neural Light Fields (NeLF) caused by inherent low-frequency bias by applying three frequency-based representation techniques-Positional Encoding (PE), sine activation functions, and Fourier Reparameterization (FR)-to the NeLF framework and analyzing their effectiveness. Using the Stanford Light Field dataset, we construct dense and sparse experimental conditions and compare reconstruction performance among the original NeLF and three enhanced models (PE, PE+SIN, and SIN+FR). In the dense setting, the PE+SIN model achieves the highest PSNR, while SIN+FR exhibits similarly strong performance with only a marginal difference. In the sparse setting, PE+SIN suffers a significant performance drop, whereas SIN+FR shows the most stable reconstruction quality with only a 2.56dB PSNR decrease. These results confirm that the SIN+FR representation method effectively enhances NeLF's reconstruction capability and provides robust performance under reduced viewpoint density.

Keyword : Implicit Neural Representation, Neural Light Field, Multi-Layer Perceptron

Copyright © 2026 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

애플 비전 프로와 삼성 갤럭시 XR의 연이은 출시, 스마트폰 업계를 대표하는 두 기업이 차세대 모바일 플랫폼으로서 확장현실(eXtended Reality, XR)에 주목하고 있음을 보여준다. 이전의 가상현실(Virtual Reality, VR) 기술과 달리, 최근에는 보다 상용화된 장비가 등장함에 따라 시장의 기대감도 높은 수준이다. 다만, 여전히 콘텐츠 부족에 대한 우려가 존재하며, 이에 따라 XR 장비에서 활용할 다양한 콘텐츠를 확보하는 것이 중요한 이슈로 대두되고 있다. 특히, 포인트 클라우드나 메쉬(Mesh)^[1] 기반의 기존 3차원 미디어가 가진 근본적인 한계를 극복하기 위한 대안으로, 가우시안 스피래딩^[2], Neural Radiance Field(NeRF)^[3] 등 차세대 3차원 미디어 기술에 대한 관심이 집중되고 있다. 그러나 이러한 기술들 역시 해결해야 할 다양한 과제를 내포하고 있다.

본 연구는 이 중 NeRF의 Implicit Neural Representation (INR)^[4] 개념과 라이트필드 기술이 결합된 Neural Light Field(NeLF)^[5,6]에 초점을 둔다. 전통적인 라이트필드는 자유 공간을 통과하는 광선(light ray)을 정의하고, 이를 재조합하여 원하는 시점의 자유로운 뷰를 생성하는 기술이다^[7]. 이는 3차원 지오메트리를 직접 추정할 필요 없이 자유시점 뷰를 생성할 수 있다는 점에서 장점을 가진다. 하지만 전통적인 라이트필드는 모든 광선 데이터를 다시점 이미지 형태로 취득, 저장하며, 이는 라이트필드 데이터의 획득 난이도를 높이고, 데이터의 보관, 처리 비용을 극도로 높인다. 이러한 점들은 라이트필드 상용화의 걸림돌로 작용해 왔다.

최근 제안된 NeLF는 이러한 라이트필드의 한계를 상당 부분 해소한다는 점에서 주목도가 높다. INR은 좌표계를

입력 받아 해당 좌표의 신호를 뉴럴 네트워크를 통해 예측하는 접근 방식으로, 예를 들어 이미지의 경우 픽셀 좌표(x, y)를 입력으로 RGB 신호를 예측한다. INR의 뉴럴 네트워크로는 대체로 Multi-Layer Perceptron(MLP)가 활용된다. NeRF^[3]는 이를 3차원 공간 좌표(x, y, z)와 시선 방향(θ, ϕ)을 입력으로 밀도와 색을 예측하는 형태로 확장하여, 높은 수준의 3차원 재구성 성능을 보여준다. NeLF^[5,6] 역시 광선의 좌표(u, v, s, t)를 입력으로 대응되는 광선의 색을 예측하는 방식으로 적용된다. 이러한 구조는 전통적인 라이트필드에서 방대한 형태로 저장되었던 다시점 이미지를 하나의 MLP로 대체함으로써 저장 효율을 크게 향상시키며, 획득 과정에서도 정교한 카메라 배열 대신 자유롭게 촬영된 일부 광선만으로 전체 라이트필드를 복원할 수 있어 실용성이 높다. 반면, NeLF는 속도가 느린 단점을 가진다. 렌더링 측면에서 하나의 광선을 예측하기 위해 한 번의 MLP 연산이 필요한데, 이는 하나의 뷰를 생성하기 위해서는 픽셀 수만큼의 MLP 연산이 요구됨을 의미한다. 이러한 문제는 INR 기반 접근 방법들이 공통적으로 가지는 문제이며, 이를 해결하기 위한 연구들은 주로 MLP 규모를 줄이거나, feature의 일부를 외부 구조에서 학습 및 저장하는 방식으로 연산 부담을 줄인다^[8,9,10].

본 논문은 NeLF를 대상으로 표현 효율을 개선하기 위한 방안을 연구한다. 저주파 편향은 INR 기반 접근의 공통적인 또 다른 문제로, 고주파 세부 정보를 충분히 재현하지 못하는 한계가 있다. 이미지 기반 INR 분야에서는 좌표와 신호 간 표현 능력을 향상시키기 위한 다양한 기법들이 제시되어 왔다. 대표적으로 Positional Encoding(PE)^[11]은 입력 좌표를 다양한 주파수의 sine, cosine 신호로 확장하여 고주파 성분 학습을 돕는 방식으로, NeRF에서도 학습 성능 개선 효과가 확인된 바 있다. 본 연구에서는 PE뿐 아니라, Sine 활성화 함수^[4], Fourier Reparameterization(FR)^[12] 기법 등을 NeLF에 적용하여 라이트필드의 표현 및 학습 효율의 개선 여부를 실험적으로 검증 및 분석하는 것을 목표로 한다. 본 연구는 2차원 이미지 위주로 다루어지던 기존의 INR 기법들을 4차원 뷰포인트 공간인 NeLF 환경에 적용하여 라이트필드에 최적화된 아키텍처를 탐색한다는 점에서 의의를 가진다. 특히, 현실적인 데이터 취득의 한계인 Sparse 조건에서 기법 간의 성능 변화를 비교 분석함으로써

a) 서울과학기술대학교 스마트ICT융합공학과(Dept. of Smart ICT Convergence Engineering, Seoul National University of Science and Technology)

‡ Corresponding Author : 정현민(Hyunmin Jung)
E-mail: hmjung@seoultech.ac.kr
Tel: +82-2-970-6457
ORCID: <https://orcid.org/0000-0001-8216-5842>

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2026-25491604).

· Manuscript March 3, 2026; Revised April 13, 2026; Accepted April 13, 2026.

써, 향후 NeLF 모델 최적화 연구를 위한 실증적 기반을 제공하고자 한다. NeLF의 표현 능력 향상은 궁극적으로 NeLF 기반 3차원 미디어 콘텐츠의 활용도를 높이는 데 기여할 것으로 기대한다.

II. Neural Light Field 구조 및 동작 소개

그림 1은 라이트필드에서 사용되는 광선의 좌표계와 NeLF의 구조를 보여준다. 라이트필드는 자유 공간 상에 두 개의 평행한 평면을 가정하며, 이를 통과하는 광선을 두 평면에서의 두 점 (u, v) 와 (s, t) 로 정의한다. 해당 좌표는 NeLF의 MLP 네트워크에 입력으로 들어가며, MLP는 이의 출력으로 해당 광선의 RGB 값을 예측한다. NeRF가 3차원 위치의 밀도와 색을 동시에 예측하는 구조를 가지는 것과 달리, NeLF는 광선에 대한 색 정보만을 직접적으로 예측한다. 이로 인해 NeLF는 볼륨 렌더링 과정이 필요하지 않으며, 상대적으로 단순한 네트워크 구조를 갖는다. 그러나 단순 좌표계와 색 정보 사이의 관계만을 학습하는 NeLF 구조는 깊이 또는 기하 구조에 대한 부가적인 제약 정보가 부족하다. 이러한 특성으로 인해 입력 시점 밀도가 낮은 Sparse한 데이터셋 환경에서는, 3차원 공간 밀도와 색상을 동시에 모델링하는 NeRF에 비해 예측 정확도가 낮은 경향을 보인다.

NeLF의 표현력에 영향을 미치는 요인 중 하나는 MLP의 크기이다. MLP의 크기는 각 은닉층의 길이(Width)와 은닉층의 개수(Depth)로 정의된다. 일반적으로 MLP의 규모가 클수록 높은 capacity를 바탕으로 정확한 표현이 가능해진다^[13]. 그러나 하나의 광선 색상을 예측하는 데 필요한 MLP

연산량이 증가함에 따라 전체 렌더링의 속도는 저하된다. 반대로, MLP의 크기를 축소할 경우 연산량이 감소하여 처리 속도는 향상되지만, 네트워크의 표현력이 제한되어 표현 정확도가 다소 낮아질 수 있다. 이와 같이 MLP의 크기는 표현 정확도와 연산 효율 간의 상충 관계(Trade-off)를 가진다. MLP의 각 은닉층은 활성화 함수를 포함하며, 일반적으로 ReLU 함수가 사용된다. 그러나 은닉층의 개수가 증가함에 따라 깊은 신경망에서 정보 소실이나 학습 불안정 문제가 발생할 수 있다. 이러한 문제를 보완하기 위해, 본 구조에서는 skip connection을 포함하여 학습 안정성을 향상시킨다.

III. Neural Light Field 고주파 성분 개선 방안

MLP 네트워크를 이용한 INR 기반 모델은 일반적으로 저주파 성분을 우선적으로 학습하는 경향을 가지며, 이로 인해 고주파 성분 학습에 어려움을 겪는 것으로 알려져 있다^[14]. 이러한 특성은 세밀한 텍스처나 급격한 색 변화가 포함된 장면에서 표현 성능 저하로 이어질 수 있다. INR 분야에서는 이러한 문제를 완화하기 위해 다양한 주파수 기반 개선 기법들이 제안되어 왔으나, NeLF 구조를 대상으로 고주파 표현력 향상을 체계적으로 분석한 연구는 상대적으로 제한적이다. 본 논문에서는 NeLF에서의 고주파 성분 표현 능력을 향상시키기 위해 기존 INR 연구에서 제안된 주파수 기반 표현 기법들을 NeLF 구조에 적용하고, 그 효과를 시점 밀도 변화에 따른 실험을 통해 분석하고자 한다. 이하에서는 본 연구에서 적용한 세 가지 기법을 순차적으로 설명한다.

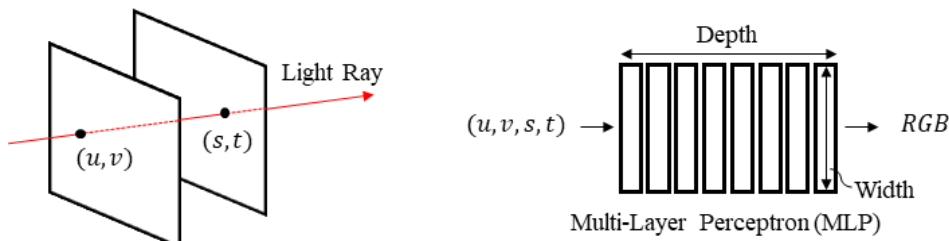


그림 1. 광선 좌표 (u, v, s, t) 를 입력으로 RGB 값을 예측하는 NeLF MLP 구조
 Fig. 1. MLP-based NeLF model that predicts RGB values from light ray coordinates (u, v, s, t)

1. 입력 좌표계의 Positional Encoding

PE는 입력 좌표 p 를 다중 주파수의 sine 및 cosine 함수로 확장하여 고차원 특징 벡터로 변환하는 기법이다. 이를 통해 좌표 변화에 대한 주기적 변동성을 입력 단계에서 명시적으로 제공하며, 복잡한 색상 변화나 세밀한 구조를 보다 효과적으로 표현할 수 있는 입력 표현을 생성한다.

$$y(p) = (\sin(2^0\pi p), \cos(2^0\pi p), \sin(2^1\pi p), \cos(2^1\pi p), \dots, \sin(2^{L-1}\pi p), \cos(2^{L-1}\pi p)) \tag{1}$$

수식 (1)에서 L 은 PE를 통해 확장되는 주파수의 개수를 의미하며, 각 입력 차원에 대해 총 $2L$ 개의 주기 신호가 생성된다. 본 연구에서는 $L=4$ 의 주파수 개수를 사용하여 (u, v, s, t) 의 네 좌표에 각각 적용함으로써 입력 차원을 총 $4 \times 2L = 32$ 차원으로 확장하였다. 이러한 입력 확장은 MLP가 좌표 변화에 따른 비선형적인 색상 변화를 보다 풍부하게 모델링할 수 있도록 돕는다. 다만, PE에서 사용되는 주파수는 사전에 고정된 값으로 설정된다는 특징을 가진다.

2. Sine Activation

기존 INR 연구에서 주로 사용되는 ReLU 활성화 함수와 달리, SIREN(Sinusoidal Representation Network)^[4]은 sine 함수를 활성화 함수로 사용하는 신경망 구조를 제안하였다. Sine 함수는 본질적으로 주기적인 특성을 가지며, 입력 값의 변화에 따라 반복적인 출력 패턴을 형성한다. 이러한 특성은 고주파 신호를 표현하는 데 적합함이 실험적으로 입증된 바 있다. SIREN은 입력 좌표에 별도의 주파수 기반 인코딩을 적용하지 않더라도, 활성화 함수의 주기적 특성만으로 네트워크 내부에서 고주파 성분을 안정적으로 학습할 수 있음을 보였다. 이는 입력 좌표 공간을 확장하는 PE 방식과 달리, 네트워크 구조 자체의 설계를 통해 고주파 표현을 유도한다는 점에서 차별성을 가진다. 다만 Sine 활성화 함수를 사용하는 구조는 가중치 초기화 방식이나 주파수 스케일 설정에 따라 학습 안정성에 민감하게 반응하며, 데이터 특성에 따라 성능 편차가 발생할 수 있다는 한계를 포함한다.

3. Fourier Reparameterization

FR^[12]은 좌표 기반 신경망에서 사용되는 가중치 행렬을 주파수 관점에서 재구성하여, 다양한 주파수 성분을 효율적으로 표현할 수 있도록 설계된 구조이다. 기존 MLP에서는 가중치 행렬이 직접 학습되지만, FR은 이를 학습 가능한 행렬과 사전에 정의된 Fourier basis 행렬의 곱으로 재구성한다.

FR 기반 구조의 중요한 특징은 입력을 단순히 주파수 공간으로 확장하는 방식이 아니라, 주파수 성분의 기여도 자체를 학습 대상으로 포함시킨다는 점이다. 이를 통해 특정 주파수 대역의 진폭(amplitude) 및 위상(phase) 변화가 네트워크 파라미터에 의해 조절될 수 있으며, 이는 고정된 주파수 집합을 사용하는 PE나 단일 주파수 스케일에 의존하는 Sine 활성화 기반 표현보다 높은 표현 유연성을 제공한다. 기존 연구에서는 이러한 재파라미터화 구조가 학습 초기 단계에서 특정 주파수 대역에 과도하게 치우치는 현상을 완화하고, 수렴 속도 및 최종 성능을 향상시킬 수 있음을 보였다. 본 연구에서는 FR 기반 구조를 Sine 활성화 함수와 결합한 SIN+FR layer를 NeLF에 적용하여, 시점 밀도 변화에 따른 성능 특성을 실험적으로 분석한다. SIN+FR 구조

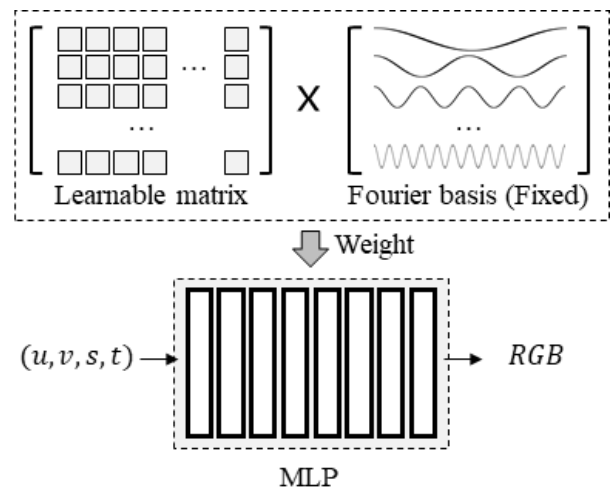


그림 2. 학습 가능한 행렬과 고정된 Fourier basis의 곱으로 가중치를 구성하는 FR 구조
Fig. 2. FR structure that constructs weights as the product of a learnable matrix and a fixed Fourier basis

는 넓은 주파수 스펙트럼을 포괄하면서도 각 주파수 성분의 기여도를 데이터 특성에 맞게 조절할 수 있어, 제한된 시점 정보 환경에서도 안정적인 표현 성능을 제공할 것으로 기대된다.

IV. 실험 결과

1. 실험 환경

본 연구에서는 NeLF와 세 가지 고주파 표현력 개선 방법의 성능 비교를 위해 Stanford Light Field 데이터셋^[15]을 사용한다. 해당 데이터셋은 다양한 물체 및 장면을 대상으로, 규칙적인 격자 형태의 다중 시점 이미지로 구성된 대표적인 라이트필드 데이터셋이다. 실험에서는 학습 데이터의 조밀도 변화에 따른 성능 분석을 위해, 전체 라이트필드 데이터셋의 17×17 시점을 모두 고려한 Dense 조건, 전체 라이트필드 중 4×4의 일부만을 사용하는 데이터셋을 Sparse 조건으로 설정하여 실험한다. Sparse 조건에서는 학습에 사용되는 이미지 개수 감소와 더불어 인접 이미지 사이의 간격이 증가하므로, 모델 입장에서는 학습 난이도가 높아지는 조건에 해당한다.

모든 실험에서는 동일한 MLP 구조를 사용하되, 입력 표

현 방식과 주파수 처리 방법에 따라 세 가지 비교 모델을 구성하여 성능을 비교한다. 첫 번째 비교 모델은 NeLF에 PE를 적용한 모델로, 입력 좌표를 다중 주파수의 sine, cosine 함수로 확장하여 MLP에 입력한다. 이 경우 기존 MLP 구조 대비 입력 차원만 증가하며, 그 외의 네트워크 구조 및 학습 조건은 동일하게 유지하였다. 두 번째 비교 모델인 PE+SIN 모델은 입력 좌표에 PE를 적용하고, MLP의 활성화 함수로 Sine 활성화 함수를 사용한 구조로, 입력 단계와 네트워크 내부 모두에서 주기적 표현을 활용한다. 세 번째 비교 모델은 Sine 활성화 함수를 기반으로, FR을 적용한 SIN+FR 기반 NeLF 구조이다. 본 모델은 입력 좌표 확장 방식 대신, 네트워크 가중치 공간 자체를 주파수 관점에서 재구성함으로써 고주파 성분 학습을 보완한다.

2. Dense 조건에서의 고주파 성분 개선 방안 적용 결과

표 1은 Dense 조건을 대상으로 학습한 NeLF 모델들의 정량적 성능을 샘플별로 비교한다. 각 행의 샘플별 비교에서 굵은 글씨와 밑줄로 표시된 값이 PSNR, SSIM 값이 가장 높은 경우, 굵은 글씨는 두 번째로 높은 경우를 나타낸다. PSNR 결과를 보면, 고주파 성분 반영을 위해 도입한 세 가지 모델 모두 기존 모델 대비 PSNR이 상승하는 결과

표 1. Dense 조건 대상 NeLF의 재구성 화질 성능 비교
 Table 1. Comparison of NeLF reconstruction quality in the dense condition

	PSNR (dB)				SSIM			
	Original	PE	PE+SIN	SIN+FR	Original	PE	PE+SIN	SIN+FR
beans	43.75	43.32	42.60	42.92	0.985	0.984	0.984	0.985
bracelet	37.73	38.93	40.15	40.02	0.985	0.987	0.988	0.989
bulldozer	37.53	38.72	40.75	40.29	0.943	0.946	0.957	0.964
bunny	43.77	43.82	44.24	44.11	0.970	0.970	0.971	0.975
chess	41.49	42.24	42.65	42.23	0.975	0.977	0.980	0.979
flowers	39.50	40.83	41.15	41.08	0.957	0.961	0.962	0.964
gem	40.94	42.08	43.14	42.73	0.969	0.971	0.974	0.975
knights	36.23	36.85	37.73	37.21	0.968	0.972	0.977	0.971
tarot	26.72	26.81	26.72	26.84	0.903	0.906	0.917	0.923
tarot_small	33.70	35.97	38.57	37.97	0.966	0.977	0.987	0.986
treasure	36.02	38.56	38.39	40.36	0.961	0.968	0.968	0.972
truck	41.16	42.50	42.93	42.70	0.957	0.963	0.964	0.964
average	38.21	39.22	39.92	39.87	0.962	0.965	0.969	0.971

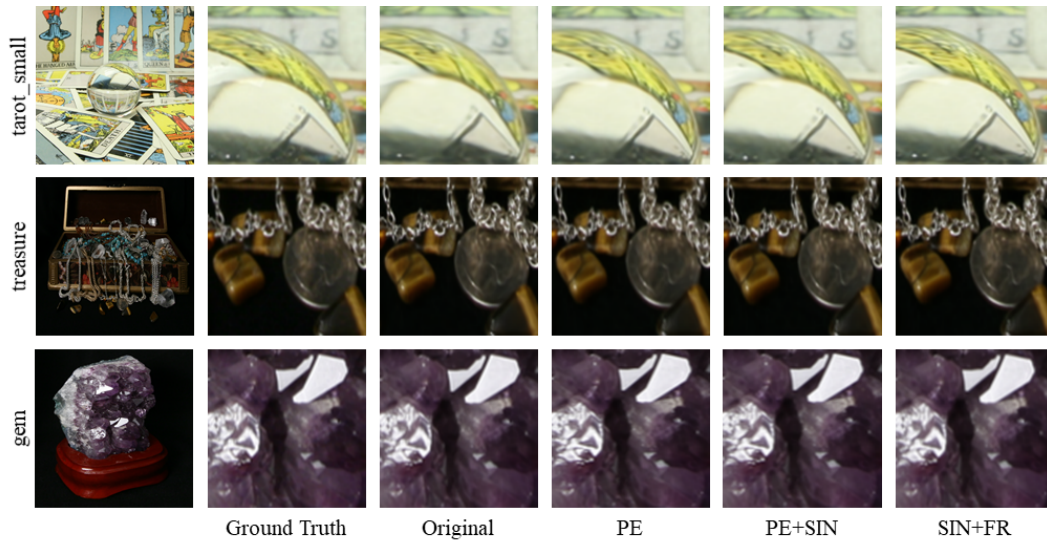


그림 3. Dense 조건 대상 NeLF의 재구성 결과의 질적 비교
 Fig. 3. Qualitative comparison of NeLF reconstruction results in the dense case

를 보인다. 특히, PE+SIN 방법이 대부분의 샘플에서 가장 높은 PSNR을 기록하였으며, 평균 성능 또한 가장 우수하다. SIN+FR은 PE+SIN에 비해 PSNR이 다소 낮게 측정되었으나 두 방법의 차이는 0.1dB 이하로 매우 적었다. SSIM의 비교에서도 유사한 경향을 보이며, 다만 SSIM의 경우 SIN+FR이 PE+SIN 대비 우수한 성능을 보인다.

그림 3은 동일한 Dense 조건에 대한 각 모델의 재구성 결과를 질적으로 비교한다. 첫 번째 행의 tarot_small 샘플에 대한 결과를 살펴보면, PE+SIN 모델은 유리 구체 경계와 카드 패턴 영역에서 Original 및 PE 모델 대비 선명한 경계와 안정적인 텍스처 표현을 보인다. 하지만 이의 차이는 매우 미미하며, 표 1의 PSNR 결과에서 볼 수 있듯 전반적으로 33dB 이상의 높은 재구성 성능에 따른 결과이다. 두 번째, 세 번째 행의 treasure, gem의 경우 PSNR 차이는 더욱 적으며, 질적 비교에서도 차이가 거의 없음을 확인할 수 있다.

3. Sparse 시점 조건에서의 성능 및 일반화 분석

표 2는 Sparse 조건에서 고주파 성분 개선 기법을 적용한 NeLF 모델의 정량적 성능을 비교한다. Sparse 조건은 인접 시점 간의 간격이 멀어져 학습 난이도가 증가하며, 표 1,

2의 Original 결과에서 볼 수 있듯이 평균 PSNR과 SSIM이 크게 떨어지는 결과를 확인할 수 있다. PE와 PE+SIN 역시 성능 저하가 크게 발생하며, 각각 8.74dB, 7.05dB의 PSNR 저하와 0.107, 0.095의 SSIM 저하를 보인다. SIN+FR 역시 Dense 조건에 비해 화질 저하가 발생하지만 낙폭이 2.56dB와 0.016으로 상대적으로 낮으며, 그 결과 모든 샘플에서 가장 우수한 성능을 보인다. PE 단독으로 사용된 경우 평균적으로 가장 낮은 정량적 성능을 보인다. 샘플에 따라 beans, flowers, knights, tarot 샘플에서는 Original이 PE, PE+SIN보다 더 우수한 성능을 보인다.

그림 4는 Sparse 조건에서의 질적 비교를 보여준다. 첫 번째 행의 tarot_small 샘플을 보면, 이전 Dense 조건과 달리 방법에 따른 화질 저하가 두드러진다. Original, PE의 경우 유리구슬 위의 문자가 구분이 되지 않을 정도로 재구성 성능이 낮아짐을 확인할 수 있으며, PE+SIN, SIN+FR은 유사한 정도의 화질을 보인다. 이러한 경향은 두 번째 행의 gem에서도 동일하게 나타난다. 반면, 세 번째, 네 번째 행의 knights와 flowers 샘플의 경우 PE+SIN의 성능 저하가 두드러지는 결과를 보이며, SIN+FR은 이 두 샘플에서도 비교적 우수한 화질을 보인다.

종합적으로, Dense, Sparse 조건에서 각 기법의 주파수 처리 구조에 따른 재구성 성능의 차이는 다음과 같이 분석

표 2. Sparse 조건 대상 NeLF의 재구성 화질 성능 비교
 Table 2. Comparison of NeLF reconstruction quality in the sparse condition

	PSNR (dB)				SSIM			
	Original	PE	PE+SIN	SIN+FR	Original	PE	PE+SIN	SIN+FR
beans	40.95	36.40	37.09	41.85	0.978	0.968	0.971	0.982
bracelet	28.61	22.76	29.53	36.50	0.938	0.869	0.959	0.984
bulldozer	31.75	26.59	29.57	37.47	0.887	0.839	0.871	0.943
bunny	40.28	40.82	42.69	43.20	0.953	0.961	0.966	0.968
chess	35.72	35.74	40.93	41.45	0.952	0.960	0.974	0.974
flowers	36.86	36.50	29.22	39.49	0.943	0.946	0.870	0.955
gem	32.12	36.41	38.75	39.45	0.913	0.954	0.964	0.966
knights	29.03	25.75	25.33	31.78	0.880	0.804	0.769	0.941
tarot	21.41	13.57	14.47	22.87	0.722	0.288	0.273	0.847
tarot_small	25.78	24.98	33.49	35.43	0.852	0.861	0.968	0.975
treasure	29.24	27.01	33.08	37.82	0.910	0.893	0.945	0.964
truck	35.92	39.22	40.35	40.46	0.925	0.952	0.956	0.957
average	32.30	30.48	32.87	37.31	0.904	0.858	0.874	0.955

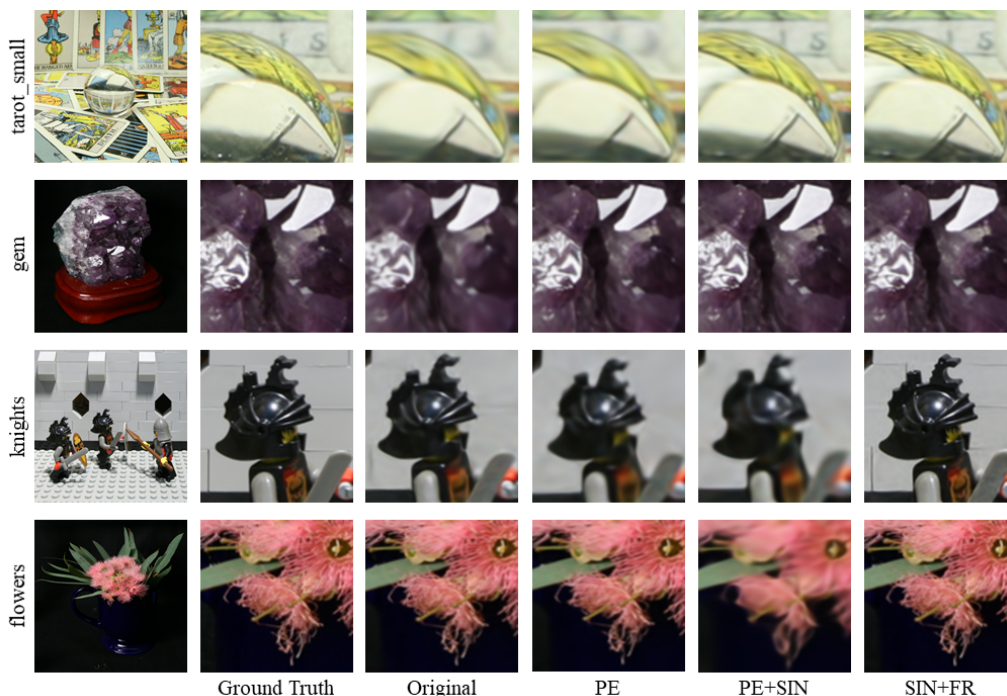


그림 4. Sparse 조건 대상 NeLF의 재구성 결과의 질적 비교
 Fig. 4. Qualitative comparison of NeLF reconstruction results in the sparse case

된다. 뷰포인트 정보가 충분한 Dense 조건에서는 주파수 확장과 주기적 활성화가 시너지를 내는 PE+SIN은 높은 고주파 표현력을 보인다. 그러나 정보량이 부족한 Sparse 조건에서 PE와 Sine 활성화 함수는 각각 고정된 주파수나 단

일 주파수 스케일의 한계를 가지며, 이는 과적합으로 인한 화질 저하를 야기한다. 반면, SIN+FR은 가중치를 학습 가능한 행렬과 다양한 고정 주파수 성분 행렬의 곱으로 구성하는데, 이는 충분하지 않은 신호에서도 대상에 유효한 주

파수 성분의 가중치를 유연하게 조절 및 학습함으로써, 불필요한 과적합을 방지하고 성능 저하를 줄인다. 결과적으로, 아주 조밀한 데이터 확보에 제약이 따르는 실제 응용 환경을 고려할 때, 입력 데이터의 밀도 변화에 대한 민감도를 낮추고 일관된 복원 품질을 유지하는 SIN+FR의 강건성은 매우 중요한 실용적 이점을 제공한다.

V. 결론 및 향후 계획

본 논문은 NeLF가 가지는 저주파 편향으로 인한 고주파 표현 한계를 개선하기 위해, 주파수 기반 표현 기법인 PE, Sine 활성화 함수, 그리고 FR 구조를 NeLF에 적용하고 학습 이미지 밀도 변화에 따른 성능을 비교 분석한다. 실험 결과, Dense 조건에서는 입력 좌표의 주파수 확장과 주기적 활성화가 결합된 PE+SIN이 가장 높은 재구성 성능을 보였으며, SIN+FR이 근소한 차이로 뒤를 이어 주파수 기반 표현의 효과성을 확인할 수 있었다. 반면 Sparse 조건에서는 PE+SIN이 7.05dB의 큰 성능 저하를 보인 데 비해, SIN+FR은 2.56dB 수준의 제한적인 저하만을 나타내며 가장 높은 PSNR을 기록하였다. 이는 SIN+FR이 입력 밀도 변화에 더 높은 강인성을 지님을 보여주는 중요한 결과이다. 본 연구는 주파수 기반 표현 기법이 NeLF의 고주파 표현 한계를 실질적으로 보완할 수 있음을 실험적으로 입증하였으며, 제한된 뷰포인트 환경에서도 안정적인 복원 성능을 유지할 수 있음을 확인하였다. 이는 LF 획득 시 필요한 촬영 시점 수를 줄여 데이터 수집 부담과 학습 비용을 완화하는 데 기여할 수 있다. 향후 연구에서는 FR의 주파수 구성 최적화, 다중 해상도 기반 주파수 구조의 확장, 그리고 실시간 렌더링을 고려한 경량화 모델 설계를 통해 NeLF의 확장성 및 실용성을 더욱 높일 수 있을 것으로 기대한다.

참고 문헌 (References)

- [1] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," In Proceedings of the IEEE International Conference on Robotics and Automation, pp.1-4, 2011.
doi: <https://doi.org/10.1109/ICRA.2011.5980567>
- [2] Y. Bao, T. Ding, J. Huo, Y. Liu, Y. Li, and W. Li, "3D Gaussian Splatting: Survey, Technologies, Challenges, and Opportunities," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 35, No. 7, pp. 6832-6852, 2025.
doi: <https://doi.org/10.1109/TCSVT.2025.3538684>
- [3] 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, 2022.
doi: <https://doi.org/10.1145/3503250>
- [4] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit Neural Representations with Periodic Activation Functions," Advances in neural information processing systems, Vol.33, pp.7462-7473, 2020.
doi: <https://doi.org/10.48550/arXiv.2006.09661>
- [5] Z. Li, L. Song, C. Liu, J. Yuan, and Y. Xu, "NeuLF: Efficient Novel View Synthesis with Neural 4D Light Field," In Proceedings of the Eurographics Symposium on Rendering, pp.1-11, 2022.
doi: <https://doi.org/10.2312/sr.20221156>
- [6] Y.-D. Kim and H. Jung, "Effective Image Complexity Measurement for Predicting View Synthesis Performance of Neural Light Field," JBE Vol.29, No.5, September 2024.
doi: <https://doi.org/10.5909/JBE.2024.29.5.691>
- [7] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," In Seminal Graphics Papers: Pushing the Boundaries, pp.453-464, 2023.
doi: <https://doi.org/10.1145/237170.237200>
- [8] T.Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," ACM Transactions on Graphics, vol. 41, no. 4, pp. 1-15, 2022.
doi: <https://doi.org/10.1145/3528223.3530127>
- [9] C. Sun, M. Sun, and H.-T. Chen, "Direct Voxel Grid Optimization: Super-Fast Convergence for Radiance Fields Reconstruction," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5459-5469, 2022.
doi: <https://doi.org/10.1109/CVPR52688.2022.00538>
- [10] I.-G. Jeong and H. Jung, "Neural light fields with N-dimensional voxel grids: a performance evaluation across voxel grid dimension," IEICE Electronics Express, vol. 22, no. 9, pp. 1-6, 2025.
doi: <https://doi.org/10.1587/ele.22.20250141>
- [11] M. Tancik et al., "Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains", Advances in Neural Information Processing Systems, Vol.33, pp.7537-7547, 2020.
doi: <https://doi.org/10.48550/arXiv.2006.10739>
- [12] K. Shi, X. Zhou, and S. Gu, "Improved Implicit Neural Representation with Fourier Reparameterized Training", In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.25985-25994, 2024.
doi: <https://doi.org/10.1109/CVPR52733.2024.02455>
- [13] I.-G. Jeong and H. Jung, "Scalable Neural Light Field With Layer Add-ons of Multilayer Perceptron," IEEE MultiMedia, vol. 32, no. 3, pp. 60-71, 2025.

doi: <https://doi.org/10.1109/MMUL.2025.3581588>

- [14] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, "On the Spectral Bias of Neural Networks", Proceedings of the 36th International Conference on Machine Learning, Vol.97, pp.5301-5310, 2019.

doi: <https://doi.org/10.48550/arXiv.1806.08734>

- [15] B. Wilburn et al., "High performance imaging using large camera arrays," In Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference, pp.765-776, 2005. doi: <https://doi.org/10.1145/1073204.1073259>

저 자 소 개



장 영 진

- 2026년 2월 : 서울과학기술대학교 전자IT미디어공학과 학사
- ORCID : <https://orcid.org/0009-0008-6476-0072>
- 주관심분야 : 실감미디어(VR, AR, XR, 메타버스), 영상처리, 컴퓨터 비전



정 현 민

- 2014년 2월 : 경희대학교 전자전파공학과 학사
- 2016년 2월 : 서울대학교 전기정보공학부 석사
- 2020년 8월 : 서울대학교 전기정보공학부 박사
- 2023년 3월 ~ 현재 : 서울과학기술대학교 스마트ICT융합공학과 조교수
- ORCID : <https://orcid.org/0000-0001-8216-5842>
- 주관심분야 : 실감미디어(VR, AR, XR, 메타버스), 영상처리, 컴퓨터 비전