일반논문 (Regular Paper)
방송공학회논문지 제31권 제3호, 2026년 5월 (JBE Vol.31, No.3, May 2026)
https://doi.org/10.5909/JBE.2026.31.3.501
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

멀티뷰 3D 재구성을 위한 조명 불일치 보정 방법 조사

Priyanka Muthukumaran[a)], 류 은 석[a)‡]

A Survey of Lighting Inconsistency Correction Methods for Multi-View 3D Reconstruction

Priyanka Muthukumaran[a)] and Ryu Eun-Seok[a)‡]

요 약

조명 불일치는 다시점 영상으로부터의 3D 재구성에 근본적인 도전을 제기하며, 기하학 추정을 왜곡하는 광도측정 불일치, 재구성 가정을 위반하는 시점 의존적 효과, 정확한 새로운 시점 합성을 방해하는 조명 변화로 나타난다. 신경 렌더링 접근법, 특히 Neural Radiance Field (NeRF)와 3D Gaussian Splatting (3DGS)은 이러한 도전들을 처리하기 위한 강력한 프레임워크로 부상하였다. 그러나 조명 보정을 위한 재구성 파이프라인에서의 이들의 관계와 효과성은 여전히 충분히 규명되지 않은 상태이다. 본 논문은 전처리, 재구성 중, 사후처리 문제로서 조명 불일치를 다루는 신경 렌더링 방법들에 대한 체계적인 서베이를 제공한다. 본 논문은 조명 불일치 보정을 위한 방법들을 재구성 파이프라인의 단계에 따라 검토하며, 3DGS 기반 분해 방법, NeRF 기반 외관 모델링 방법, 확산 기반 relighting 기법, 트랜스포머 기반 영상 향상 방법 등 대표적인 방법군을 중심으로 분석한다. 이러한 방법군은 동일한 개념적 범주의 병렬적 분류가 아니라, 전처리, 재구성 중 처리, 사후처리 단계에서 서로 다른 기술적 역할과 실용적 요구사항을 담당한다. NeRF 기반 방법은 재구성 과정에서 외관 변화를 모델링하는 데 효과적이고, 3DGS 기반 방법은 명시적 분해와 relighting을 가능하게 하며, 확산 기반 기법은 강한 영상 수준의 relighting prior를 제공하고, 트랜스포머 기반 방법은 저조도 및 노출 보정을 위한 효율적인 전처리를 제공한다. 우리는 재구성 파이프라인에 조명 보정을 통합하기 위한 기술적 지침과 유망한 연구 방향을 제시한다.

Abstract

Lighting inconsistencies fundamentally challenge 3D reconstruction from multi-view imagery, manifesting as photometric inconsistencies that corrupt geometry estimation, view-dependent effects that violate reconstruction assumptions, and illumination variations that prevent accurate novel view synthesis. Neural rendering approaches, particularly Neural Radiance Fields (NeRF) and 3D Gaussian Splatting (3DGS), have emerged as powerful frameworks for handling these challenges. However, their relationship and effectiveness for lighting correction in reconstruction pipelines remain poorly characterized. This paper provides a systematic survey of neural rendering methods addressing lighting inconsistencies such as preprocessing, during reconstruction, and postprocessing problems. We examine representative method families used for lighting inconsistency correction across different stages of the reconstruction pipeline, including 3DGS-based decomposition methods, NeRF-based appearance modeling methods, diffusion-based relighting techniques, and transformer-based image enhancement methods. Our analysis shows that these method families address different technical roles and practical requirements within the pipeline. NeRF-based methods are effective for joint geometry-appearance modeling during reconstruction, 3DGS-based methods enable explicit decomposition and relighting, diffusion-based techniques provide strong image-level relighting priors, and transformer-based enhancement methods offer efficient preprocessing for low-light or exposure correction.

Keyword : Neural Rendering, 3D Reconstruction, Lighting Correction, NeRF, 3D Gaussian Splatting

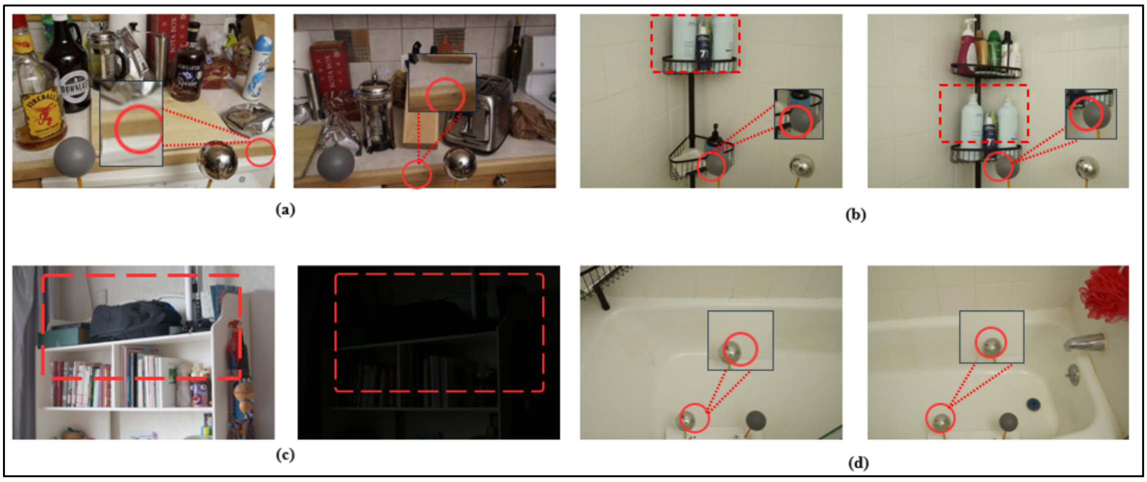

그림 1. 조명 불일치 예시로 노출 변화, 시점 의존 효과, 저조도 조건, 다중 시점 광도 불일치를 보여준다. 강조 표시된 영역과 확대 주석은 각 불일치 유형과 관련된 주요 시각적 단서를 나타낸다. (a) 노출 변화: 시점 간 밝기 변화로 인해 특징 매칭의 신뢰성이 저하된다 [21]. (b) 시점 의존 효과: 정반사 하이라이트와 그 겉보기 위치가 시점에 따라 변화한다 [21]. (c) 저조도 조건: 어두운 영상은 노이즈와 시각 정보 손실의 영향을 받는다 [22]. (d) 다중 시점 광도 불일치: 동일한 장면 지점이 시점별로 서로 다른 조명 상태로 나타난다 [21].

Fig. 1. Lighting inconsistency examples showing exposure variations, view-dependent effects, low-light conditions, and multi-view photometric inconsistency. The highlighted regions and zoomed annotations indicate the key visual cues associated with each inconsistency type. (a) Exposure variations: feature matching becomes unreliable due to brightness changes across views [21]. (b) View-dependent effects: specular highlights and their apparent positions change with viewpoint [21]. (c) Low-light conditions: dark images suffer from noise and missing visual information [22]. (d) Multi-view photometric inconsistency: the same scene point appears under inconsistent illumination across views [21].

I. Introduction

Multi-view 3D reconstruction fundamentally relies on the assumption of photometric consistency whereby the same surface point should exhibit consistent appearance across different viewpoints. Lighting inconsistencies violate this core assumption through exposure variations, view-dependent specular effects, cast shadows, and low-light conditions as illustrated in Figure 1. These inconsistencies corrupt reconstruction at multiple stages: feature matching fails when brightness varies across views, stereo correspondence becomes unreliable, structure-from-motion produces inaccurate poses, and multi-view stereo generates incorrect depth estimates.

Traditional graphics pipelines separate geometry, materials, and lighting through inverse rendering. However, this requires controlled capture conditions or sophisticated decomposition algorithms. Neural rendering approaches learn scene representations directly but face the lighting-geometry entanglement problem without explicit constraints. Consequently, learned representations bake lighting effects into geometry and material estimates, preventing relighting and producing

---

a) 성균관대학교 실감미디어공학과(Department of Immersive Media Engineering, Sungkyunkwan University)
‡ Corresponding Author : 류은석(Eun-Seok Ryu)
E-mail: esryu@skku.edu
Tel: +82-2-760-0677
ORCID: https://orcid.org/0000-0003-4894-6105

※This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Graduate School of Metaverse Convergence support program(IITP-2026-RS-2023-00254129) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)
※This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00457605)
· Manuscript May 7, 2026; Revised May 15, 2026; Accepted May 15, 2026.

incorrect appearance under novel illumination.

Neural Radiance Fields (NeRF) [1] revolutionized novel view synthesis through implicit volumetric representations but introduced implicit encoding of lighting into learned radiance fields. 3D Gaussian Splatting (3DGS) [2] provided alternative through explicit representations, facilitating physically based decomposition. Despite rapid progress, systematic understanding of how these paradigms address lighting challenges remains lacking. Are NeRF and 3DGS competing approaches or complementary techniques? When should correction be preprocessing versus postprocessing? This survey addresses these questions through systematic analysis of state-of-the-art methods (2023- 2026). Given the 8-page constraint, we focus on methods representing distinct technical paradigms, selected based on publication in top venues, technical novelty, demonstrated effectiveness, and distinct contributions. Our key contributions are: (1) a systematic taxonomy that organizes lighting inconsistency correction methods by pipeline integration stage, (2) a literature-based comparative analysis that identifies quality-efficiency and integration trade-offs across representative method families, and (3) practical discussion of method selection and future hybrid research directions for illumination-robust multi-view reconstruction workflows. In this survey, we use pipeline stage as the primary organizing principle and discuss representative method families within that structure, including NeRF-based, 3DGS-based, diffusion-based, and transformer-based approaches.

Ⅱ. Related Work

The lighting decomposition problem can be formulated as separating observed image

$$\mathbf{I}(\mathbf{x}) = \mathbf{L}(\mathbf{x}) \odot \mathbf{R}(\mathbf{x}) + \mathbf{E}(\mathbf{x}) \tag{1}$$

where L represents illumination, R represents surface reflectance, and E captures residual effects including shadows and inter-reflections. In 3D reconstruction, this extends to multi-view consistency: the same surface point p should satisfy

$$\mathbf{I\_}\boldsymbol{i}(\boldsymbol{\pi}\mathbf{\_}\boldsymbol{i}(\mathbf{p})) \approx \mathbf{I\_}\boldsymbol{j}(\boldsymbol{\pi}\mathbf{\_}\boldsymbol{j}(\mathbf{p})) \tag{2}$$

across views *i* and *j* under consistent lighting, where π represents perspective projection. Lighting variations violate this assumption, requiring either preprocessing (normalizing images before reconstruction), during-reconstruction (joint optimization of geometry and appearance), or postprocessing (decomposing learned representations) correction approaches with distinct trade-offs. Evaluation employs photometric accuracy metrics (PSNR measuring pixel-level correctness, SSIM assessing structural similarity, LPIPS capturing perceptual quality), reconstruction quality metrics (depth error, normal consistency, multi-view reprojection error), and computational efficiency metrics (frames per second, memory consumption, training time). Standard benchmarks include MIT Multi-Illumination [21] for indoor relighting evaluation, LOL-V1/V2 [22] for low-light enhancement assessment, and DeepBlending [23] for multi-view 3D consistency testing.

1. 3D Gaussian Splatting Methods

3DGS methods leverage explicit Gaussian primitives for postprocessing decomposition after initial reconstruction. SSD-GS [3] achieves state-of-the-art physically based decomposition by separating appearance into four components: diffuse (Lambertian reflection with learned albedo), specular (anisotropic microfacet BRDFs with Fresnel terms), shadow (occlusion-aware computation with learned refinement), and subsurface scattering (dipole-based models for translucent materials). Training employs progressive integration to avoid local minima, with per-scene optimization requiring hours but enabling real-time rendering (100+ fps) after decomposition. GI-GS [4] addresses global illumination through deferred shading architecture where Gaussian rasterization produces geometry buffers followed by Monte Carlo path tracing. This enables accurate soft shadows, color bleeding, and ambient occlusion through recursive ray evaluation. Luminance-GS [5] handles exposure inconsistencies via per-view affine color transformations and adaptive tone curves, crucial for multi-view captures with automatic exposure. Generalizable approaches eliminate per-scene optimization. RelightAnyone [6] achieves zero-shot relighting without OLAT capture through two-stage processing: constructing flat-lit 3DGS from casual multi-view captures, then predicting physically based parameters using priors learned from synthetic datasets. GRGS [7] focuses on human relighting with lighting-aware geometry refinement producing lighting-invariant geometry and physically grounded neural rendering combining physics-based shading with neural networks. Specialized variants include GTAvatar [8] for UV-mapped editable avatars and GaRe [9] for outdoor scenes with time-of-day dependent lighting decomposition.

2. Neural Radiance Field Methods

NeRF method addresses lighting through during-reconstruction appearance modeling, jointly optimizing geometry and illumination representations. The foundational NeRF [1] learns continuous volumetric functions

$$\mathbf{F}: (x, y, z, \theta, \varphi) \rightarrow (\sigma, c) \tag{3}$$

mapping positions and viewing directions to density and radiance, but bakes lighting into learned radiance values preventing disentanglement. NeRF-W [10] introduces per-image appearance embeddings ℓ_i enabling reconstruction from unconstrained photo collections where lighting varies across captures. Each training image receives learned latent code modulating radiance function, absorbing illumination differences that would otherwise corrupt geometry estimation. This enables processing tourist photos, historical archives, and internet imagery with arbitrary lighting but lacks physical interpretability for accurate relighting. IllumiNeRF [11] incorporates diffusion priors from 2D relighting models, using pretrained diffusion to guide 3D training without explicit OLAT capture or inverse rendering. BLiRF [12] separates dynamic lighting from geometry using band-limited constraints, decomposing radiance as

$$\mathbf{L}(x, y, z, \theta, \varphi, t) = \mathbf{L}_s(x, y, z, \theta, \varphi) + \mathbf{L}_t(t) \tag{4}$$

where static radiance captures object appearance, and temporal modulation represents illumination changes. Aleth-NeRF [13] models illumination-dependent information visibility through concealing fields $C(x, y, z, \ell)$, preventing dark regions from being misinterpreted as empty space in low-light reconstruction.

3. Diffusion and Transformer Based Methods

Diffusion models provide robust preprocessing through learned generative priors from large-scale training. GenLit [14] reformulates relighting as video generation using Stable Video Diffusion, enabling temporal consistency through video generation mechanisms. Given input image and target lighting (environment map or text), the model generates video transitioning from original to target illumination, with frames providing consistently relit views. LumiNet [15] achieves state-of-the-art indoor relighting (28.4 dB in PSNR on MIT Multi-Illumination [21]) through latent intrinsics decomposition combined with ControlNet-based diffusion. The method decomposes inputs into latent albedo, normals, and lighting, then operates on these intrinsics for fine-grained illumination control. Neural Gaffer [16] provides an end-to-end diffusion model with environment map conditioning, learning implicit relationships between the inputs, target lighting, and output appearance without explicit decomposition. Transformer methods enable efficient preprocessing for real-time scenarios. IAT [17] achieves remarkable efficiency (90K parameters, 0.004s inference, 250 fps) by generating ISP parameters rather than directly predicting enhanced images. The parametric output enables temporal interpolation across frames and invertibility for raw measurement recovery. Retinexformer [18] combines Retinex theory ($I = R \odot L$ decomposition) with dual transformer branches processing low-frequency illumination and high-frequency reflectance separately. IllumFlow [19] combines Rectified Flow with Retinex for low-light enhancement, while modern CNN architectures (ConvNeXt-based approaches, hybrid CNN-Transformer designs) and HVI color space [20] address specific efficiency and representation challenges.

Ⅲ. Comparative Analysis

1. Taxonomy Framework

In this paper, methods are organized by pipeline integration stage as shown in Figure 2, revealing fundamental trade-offs more clearly than architectural categorization. Preprocessing methods (diffusion models, transformers) operate independently on images without geometric awareness, enabling integration with traditional reconstruction tools (COLMAP, OpenMVS) at cost of potential multi-view inconsistencies. During-reconstruction methods (NeRF-W, appearance-conditioned 3DGS) jointly optimize geometry and appearance leveraging multi-view constraints, achieving better consistency but requiring specialized implementations. Postprocessing methods (physically based 3DGS) decompose trained representations after reconstruction, enabling flexible relighting but assuming successful initial reconstruction despite lighting inconsistencies. Within this framework, we identify four complementary technical categories: Physically based 3DGS methods perform postprocessing decomposition into diffuse, specular, shadow, and scattering components. Per-scene optimization requires hours but enables real-time rendering (100+ fps) and explicit component control for relighting applications. Representative methods include SSD-GS (comprehensive four-component separation), GI-GS (global illumination via path tracing), and

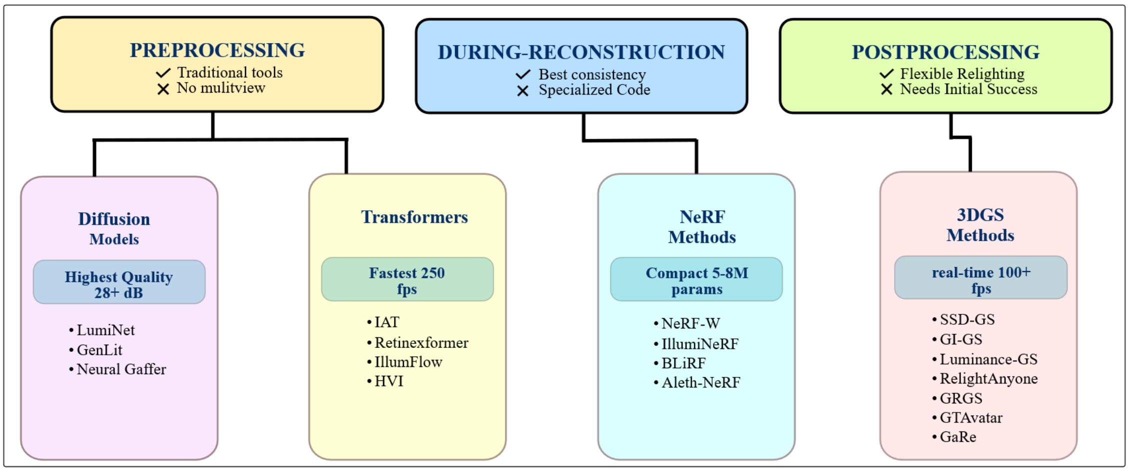

그림 2. 조명 불일치 보정 방법의 파이프라인 중심 분류를 보여준다. 검토된 접근법은 먼저 통합 단계에 따라 전처리, 재구성 중, 후처리로 구성되며, 이후 확산 기반 relighting, Transformer 기반 enhancement, NeRF 기반 appearance modeling, 3DGS 기반 decomposition 방법을 포함한 대표적인 방법군을 통해 설명된다.

Fig. 2. Pipeline-oriented taxonomy of lighting inconsistency correction methods. The reviewed approaches are first organized by integration stage: preprocessing, during reconstruction, and postprocessing and then illustrated using representative method families, including diffusion-based relighting, transformer-based enhancement, NeRF-based appearance modeling, and 3DGS-based decomposition methods.

Luminance-GS (exposure normalization). The explicit Gaussian representation facilitates direct association of lighting properties with geometric primitives. Appearance-conditioned NeRF methods operate during reconstruction through learned embeddings or lighting representations, providing compact representations (5-8M parameters) but slow inference (0.1 fps). Joint optimization of geometry and appearance leverages multi-view constraints preventing lighting variations from corrupting geometric estimates, essential for unconstrained photo collections. Representative methods include NeRF-W (per-image appearance embeddings), IllumiNeRF (diffusion prior integration), BLiRF (band-limited temporal separation), and Aleth-NeRF (concealing fields for low-light). Efficient transformer preprocessing enables real-time normalization (250 fps) with lightweight architectures (90K parameters) accepting quality reduction (21.3 dB in PSNR on LOL-V1 [22]). Parametric ISP generation or Retinex decomposition maintains temporal consistency for video sequences. Representative methods include IAT (extreme efficiency), Retinexformer (one-stage Retinex), and IllumFlow (Rectified Flow integration). Suitable for SLAM and online reconstruction where computational budget precludes expensive preprocessing.

2. Comparative Analysis Framework

Our comparative framework examines representative methods across three dimensions relevant to reconstruction applicability. Quality assessment considers photometric metrics such as PSNR, SSIM, and LPIPS on benchmarks where reported, together with qualitative analysis of lighting-specific effects such as shadow handling, specular behavior, and robustness under low-light conditions. Efficiency-related discussion focuses primarily on reported computational behavior in the literature, with particular attention to inference speed where available, while other aspects such as training cost, memory usage, and hardware dependency are discussed qualitatively only when explicitly reported in the original studies. Integration analysis examines pipeline compatibility (preprocessing, during-reconstruction, and postprocessing), multi-view consistency mechanisms, validation requirements, and common failure modes. Rather than establishing absolute rankings, this framework is intended to identify broad practical trade-offs and integration considerations across different method families. In this context, real-time scenarios such as SLAM and online reconstruction generally favor computationally lightweight preprocessing methods, whereas quality-oriented applications may tolerate higher computational cost for stronger appearance modeling or relighting capability. Likewise, resource-constrained deployment often benefits from lightweight architectures, although such considerations are discussed in this survey only at the level reported in the original literature, not through a unified experimental comparison.

Ⅳ. Experimental Results

1. Quantitative Performance Comparison

Table 1 presents quantitative comparisons across different benchmarks under different experimental settings, compiled from original publications. Results demonstrate fundamental quality-speed trade-offs determining applicability boundaries. (Note: the table includes only the top 5 state-of-the-art methods).

표 1. 문헌 기반 정량 요약
Table 1. Literature-Based Quantitative Summary

Method	Dataset	PSNR (dB)	SSIM	Speed (FPS)
SSD-GS	NR Hints [24]	26.8	0.85	100
GI-GS	Synthetic GI	26.5	0.84	100
NeRF-W	Phototourism	27.1	0.87	0.1
LumiNet	MIT Multi-Illumination	28.4	0.91	0.2
IAT	LOL-V1	21.3	0.79	250

The quantitative values in Table 1 are compiled from the original publications and are presented here as representative results under each method's own experimental setting. They were not obtained by retraining or re-evaluating all methods under a unified benchmark in this survey. Therefore, differences in benchmark dataset, input resolution, hardware configuration, and evaluation protocol should be considered when interpreting these values, particularly for computational indicators such as FPS. Accordingly, Table 1 should be understood as a literature-based summary of reported results rather than a strictly controlled head-to-head comparison. In particular, FPS values quoted from the original papers are not directly comparable across methods because they depend strongly on hardware, input resolution, implementation details, and inference configuration. Recent work such as FlashGS [25] further highlights that rendering efficiency in 3DGS can vary substantially with implementation design and target resolution, reinforcing the limitation of direct cross-paper speed comparison under heterogeneous settings.

Under this literature-based interpretation, the reported values still suggest broad trade-offs between image quality and computational cost across different method families. Diffusion-based methods generally emphasize strong relighting quality, although often with expensive inference settings. Lightweight enhancement models such as IAT prioritize computational efficiency for image preprocessing. 3DGS-based methods often target fast rendering after scene optimization, while NeRF-based methods typically involve higher reconstruction or inference cost. These observations provide useful practical context for understanding design trade-offs and possible application scenarios, but they should not be interpreted as a direct quantitative ranking across methods.

2. Pipeline Stage Analysis

Preprocessing approaches enable traditional pipeline integration without specialized implementations. Diffusion models and transformers process images independently before reconstruction, enabling use of established tools (COLMAP for SfM, OpenMVS/ACMM for dense stereo, standard meshing algorithms). This preserves well-tested geometric reconstruction while addressing illumination through enhancement learned. However, single-image processing lacks multi-view geometric awareness, corrections applied independently per-image may not maintain photometric consistency across views essential for stereo matching and bundle adjustment. Validation requirements include checking multi-view photometric consistency after preprocessing, detecting out-of-distribution inputs requiring fallback strategies, and verifying no detail loss or hallucination artifacts affecting reconstruction accuracy. Use cases include scenarios where severe initial lighting variations prevent feature matching (requiring preprocessing to enable traditional feature detection), integration with established reconstruction pipelines without modification (accepting preprocessing limitations for compatibility), and situations where computational budget permits preprocessing latency but not per-scene optimization (preprocessing once, then using standard MVS). Diffusion excels when quality is paramount, and per-image latency (1-5 seconds) is acceptable; Transformer-based methods are suitable when real-time preprocessing is needed, although their reported quality depends on the specific low-light or exposure-correction benchmark. During-reconstruction methods achieve best multi-view consistency through joint optimization but increase solution space complexity. Appearance parameters (per-image embeddings in NeRF-W, lighting codes in appearance-aware 3DGS) expand optimization, potentially causing convergence difficulties if initialization is poor, or training captures have insufficient coverage. However, joint optimization prevents lighting variations from corrupting geometry by explicitly modeling that same spatial location should have consistent geometry but variable appearance across images, bright spots from specular highlights don't produce spurious geometry bumps, shadows don't create incorrect concavities. These methods prove essential when preprocessing alone cannot ensure photometric consistency, and traditional reconstruction would fail from severe lighting variations. Postprocessing decomposition assumes successful initial reconstruction, trading this assumption for flexible relighting capabilities. Physically based 3DGS methods operate after reconstruction completion, separating learned Gaussians into lighting components. This enables relighting under novel illumination and material editing without re-reconstruction. Physical validation opportunities exist, extracted shadows should align with reconstructed geometry occlusions; specular highlights must respect surface normal orientations, enabling detection of reconstruction artifacts requiring refinement.

Since lighting correction is inherently perceptual, qualitative behavior is also important when interpreting the reviewed methods. Diffusion-based relighting methods can produce strong global illumination changes but may introduce view-inconsistent artifacts when applied independently to multi-view images. Transformer-based enhancement methods are efficient for low-light or exposure correction but may over smooth textures in difficult regions. NeRF-based methods better model appearance variation during reconstruction, while 3DGS-based methods allow more explicit control over relighting and decomposed appearance components. Therefore, the qualitative behavior of each method family depends strongly on the correction stage and application scenario.

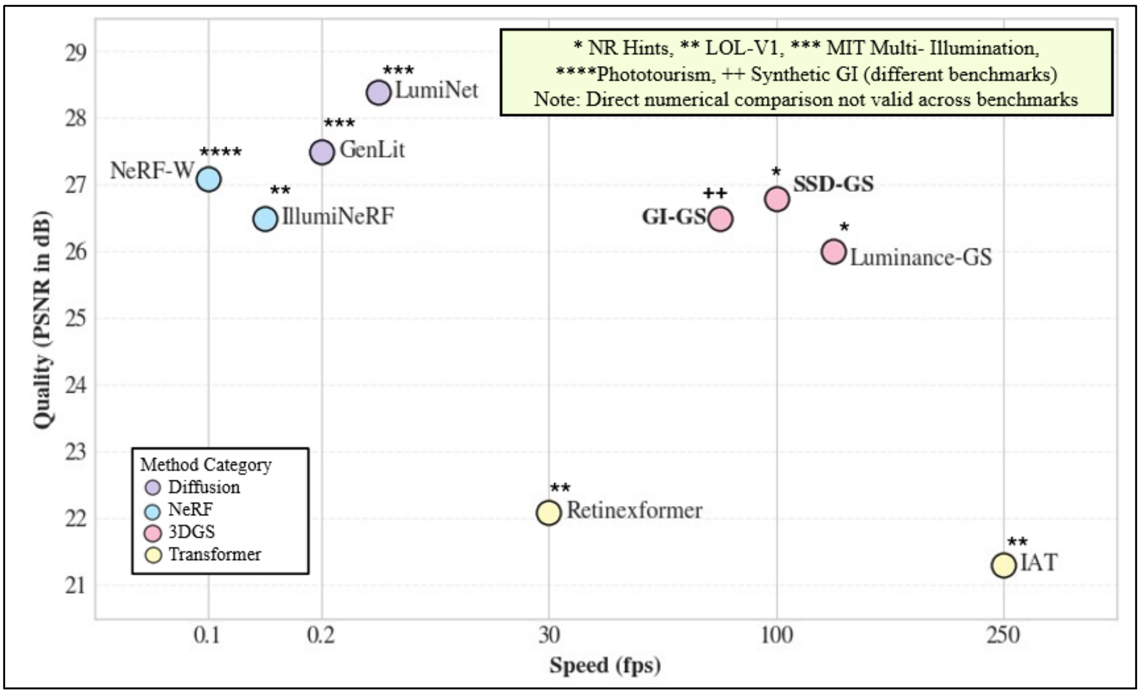

그림 3. 서로 다른 벤치마크에서 보고된 PSNR 및 FPS 값을 사용한 문헌 기반 성능 trade-off 시각화를 보여준다. 마커 주석은 각 방법과 관련된 벤치마크를 나타낸다. 해당 값들은 서로 다른 데이터셋, 하드웨어 설정, 평가 프로토콜에서 수집된 것이므로, 이 그래프는 직접적인 정량적 순위가 아니라 참고용 시각화로 해석되어야 한다.

Fig. 3. Literature-based performance trade-off visualization using reported PSNR and FPS values from different benchmarks. Marker annotations indicate the benchmark associated with each method. Since the values are compiled from different datasets, hardware settings, and evaluation protocols, the plot should be interpreted as an indicative visualization rather than a direct quantitative ranking.

Ⅴ. Conclusion

This paper surveyed recent methods for addressing lighting inconsistencies in multi-view 3D reconstruction. Instead of treating NeRF, 3DGS, diffusion, and transformer models as directly comparable categories, this survey organized the literature according to the stage at which lighting correction is applied: preprocessing, during reconstruction, and postprocessing. This pipeline-oriented view clarifies the different roles of representative method families. In addition, this survey provides practical guidance for selecting and combining complementary methods according to reconstruction stage, illumination condition, and application requirements. NeRF-based methods are useful for modeling appearance variation during reconstruction. 3DGS-based methods support explicit decomposition and relighting after reconstruction. Diffusion-based methods provide strong image-level relighting priors. Transformer-based enhancement methods offer efficient preprocessing for low-light and exposure correction. The fundamental insight is that implicit volumetric functions (NeRF) and explicit primitives (3DGS) address different reconstruction requirements—neither dominates universally, and optimal strategy matches method characteristics to scenario constraints including computational budget, capture conditions, and output requirements. The survey shows that lighting inconsistency correction is not a single-method problem. Each method family has different strengths and limitations depending on capture conditions, computational budget, and the desired output. Preprocess ing methods are easy to integrate with existing reconstruction pipelines, but they may not preserve multi-view consistency. During-reconstruction methods better exploit geometric constraints, but they require specialized optimization. Postprocessing methods enable flexible relighting and appearance editing, but they depend on the quality of the initial reconstruction. Several future research directions remain important. First, multi-view-aware preprocessing is needed so that image enhancement or relighting does not introduce view-inconsistent artifacts. Second, joint optimization of geometry, material, and illumination should be further explored to reduce lighting-geometry entanglement. Third, physics-informed neural representations could improve interpretability by incorporating rendering constraints while retaining learning-based flexibility. Fourth, efficient relighting and decomposition methods are needed for real-time or resource-constrained applications such as SLAM, mobile scanning, and immersive media streaming. Finally, standardized benchmarks and evaluation protocols are necessary for fair comparison across datasets, hardware settings, input resolutions, and reconstruction tasks. These directions indicate that future progress will likely come from hybrid methods that combine the multi-view consistency of reconstruction-based models with the visual robustness of modern image enhancement and relighting techniques.

References

[1] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R., NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, European Conference on Computer Vision (ECCV), pp. 405-421, Aug., 2020.
doi: https://doi.org/10.1145/3503250

[2] Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G., 3D Gaussian Splatting for Real-Time Radiance Field Rendering, ACM Transactions on Graphics, Vol. 42, No. 4, pp. 139-1, 2023.
doi: https://doi.org/10.1145/3592433

[3] Zheng, I., Tang, G., Doronin, A., Teal, P. D., and Zhang, F. L., SSD-GS: Scattering and Shadow Decomposition for Relightable 3D Gaussian Splatting, The Fourteenth International Conference on Learning Representations (ICLR), 2026.
doi: https://doi.org/10.48550/arXiv.2604.13333

[4] Chen, H., Lin, Z., and Zhang, J., GI-GS: Global Illumination Decomposition on Gaussian Splatting for Inverse Rendering, arXiv preprint arXiv:2410.02619, 2024.
doi: https://doi.org/10.48550/arXiv.2410.02619

[5] Cui, Z., Chu, X., and Harada, T., Luminance-GS: Adapting 3D Gaussian Splatting to Challenging Lighting Conditions with View-Adaptive Curve Adjustment, Proceedings of the Computer Vision and Pattern Recognition Conference, pp. 26472-26482, 2025.
doi: https://doi.org/10.1109/CVPR52734.2025.02465

[6] Xu, Y., Rao, P., Weiss, S., Zoss, G., Gross, M., Theobalt, C., Habermann, M., and Bradley, D., RelightAnyone: A Generalized Relightable 3D Gaussian Head Model, arXiv preprint arXiv: 2601.03357, 2026.

doi: https://doi.org/10.48550/arXiv.2601.03357

[7] Sun, Y., Zhang, S., Wang, C., Zheng, S., Li, Z., and Ji, X., GRGS: Generalizable and Relightable Gaussian Splatting for Human Novel View Synthesis, arXiv preprint arXiv:2505.21502, 2025.
doi: https://doi.org/10.48550/arXiv.2505.21502

[8] Baert, K., Younes, M., Bourel, F., Christie, M., and Boukhayma, A., GTAvatar: Bridging Gaussian Splatting and Texture Mapping for Relightable and Editable Gaussian Avatars, arXiv preprint arXiv:2512.09162, 2025.
doi: https://doi.org/10.1111/cfg.70351

[9] Bai, H., Zhu, J., Jiang, S., Huang, W., Lu, T., Li, Y., Guo, J., Fu, R., Guo, Y., and Chen, L., GARE: Relightable 3D Gaussian Splatting for Outdoor Scenes from Unconstrained Photo Collections, arXiv preprint arXiv:2507.20512, 2025.
doi: https://doi.org/10.48550/arXiv.2507.20512

[10] Martin-Brualla, R., Radwan, N., Sajjadi, M. S., Barron, J. T., Dosovitskiy, A., and Duckworth, D., NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7210-7219, 2021.
doi: https://doi.org/10.1109/CVPR46437.2021.00713

[11] Zhao, X., Srinivasan, P. P., Verbin, D., Park, K., Martin-Brualla, R., and Henzler, P., IllumiNeRF: 3D Relighting without Inverse Rendering, Advances in Neural Information Processing Systems, Vol. 37, pp. 42593-42617, 2024.
doi: https://doi.org/10.48550/arXiv.2406.06527

[12] Ramasinghe, S., Shevchenko, V., Avraham, G., and Van Den Hengel, A., BLiRF: Bandlimited Radiance Fields for Dynamic Scene Modeling, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 38, No. 5, pp. 4641-4649, Mar., 2024.
doi: https://doi.org/10.1609/aaai.v38i5.28264

[13] Cui, Z., Gu, L., Sun, X., Ma, X., Qiao, Y., and Harada, T., Aleth-NeRF: Illumination Adaptive NeRF with Concealing Field Assumption, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 38, No. 2, pp. 1435-1444, Mar., 2024.
doi: https://doi.org/10.1609/aaai.v38i2.27908

[14] Bharadwaj, S., Feng, H., Becherini, G., Fernandez Abrevaya, V., and Black, M. J., GenLit: Reformulating Single-Image Relighting as Video Generation, Proceedings of the SIGGRAPH Asia 2025 Conference Papers, pp. 1-12, Dec., 2025.
doi: https://doi.org/10.1145/3757377.3763970

[15] Xing, X., Groh, K., Karaoglu, S., Gevers, T., and Bhattad, A., LumiNet: Latent Intrinsics Meets Diffusion Models for Indoor Scene Relighting, Proceedings of the Computer Vision and Pattern Recognition Conference, pp. 442-452, 2025.
doi: https://doi.org/10.48550/arXiv.2412.00177

[16] Jin, H., Li, Y., Luan, F., Xiangli, Y., Bi, S., Zhang, K., Xu, Z., Sun, J., and Snavely, N., Neural Gaffer: Relighting Any Object via Diffusion, Advances in Neural Information Processing Systems, Vol. 37, pp. 141129-141152, 2024.
doi: https://doi.org/10.52202/079017-4481

[17] Cui, Z., Li, K., Gu, L., Su, S., Gao, P., Jiang, Z., Qiao, Y., and Harada, T., You Only Need 90K Parameters to Adapt Light: A Light Weight Transformer for Image Enhancement and Exposure Correction, arXiv preprint arXiv:2205.14871, 2022.
doi: https://doi.org/10.48550/arXiv.2205.14871

[18] Cai, Y., Bian, H., Lin, J., Wang, H., Timofte, R., and Zhang, Y., Retinexformer: One-Stage Retinex-Based Transformer for Low-Light Image Enhancement, Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 12504-12513, 2023.
doi: https://doi.org/10.1109/ICASSP49357.2023.10096765

[19] Wei, W., Jia, X., Feng, X., Wang, W., and Wang, R., IllumFlow: Illumination-Adaptive Low-Light Enhancement via Conditional Rectified Flow and Retinex Decomposition, arXiv preprint arXiv:2511.02411, 2025.
doi: https://doi.org/10.48550/arXiv.2511.02411

[20] Yan, Q., Feng, Y., Zhang, C., Pang, G., Shi, K., Wu, P., Dong, W., Sun, J., and Zhang, Y., HVI: A New Color Space for Low-Light Image Enhancement, Proceedings of the Computer Vision and Pattern Recognition Conference, pp. 5678-5687, 2025.
doi: https://doi.org/10.48550/arXiv.2502.20272

[21] Murmann, L., Gharbi, M., Aittala, M., and Durand, F., A Dataset of Multi-Illumination Images in the Wild, Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4080-4089, 2019.
doi: https://doi.org/10.1109/ICCV.2019.00418

[22] Wei, C., Wang, W., Yang, W., and Liu, J., Deep Retinex De- composition for Low-Light Enhancement, arXiv preprint arXiv: 1808.04560, 2018.
doi: https://doi.org/10.48550/arXiv.1808.04560

[23] Hedman, P., Ritschel, T., Drettakis, G., and Brostow, G., Scalable Inside-Out Image-Based Rendering, ACM Transactions on Graphics, Vol. 35, No. 6, pp. 1-11, 2016.
doi: https://doi.org/10.1145/2980179.2982420

[24] Zeng, C., Chen, G., Dong, Y., Peers, P., Wu, H. and Tong, X. Relighting neural radiance fields with shadow and highlight hints. In ACM SIGGRAPH 2023 Conference Proceedings pp. 1-11. 2023, July.
doi: https://doi.org/10.1145/3588432.3591482

[25] Feng, Guofeng, et al. “Flashgs: Efficient 3d gaussian splatting for large-scale and high-resolution rendering.” Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.
doi: https://doi.org/10.1109/CVPR52734.2025.02482

저자소개

Priyanka Muthukumaran

- 2019년 8월 ~ 2023년 6월 : Computer Science and Engineering, Anna University 학사
- 2024년 9월 ~ 현재 : 성균관대학교 실감미디어공학과 석사과정
- ORCID : https://orcid.org/0009-0006-9007-3909
- 주관심분야 : 3D Multiview Capturing, Illumination Issues, Sensor Fusion based light and depth acquisition

류은석

- 1999년 8월 : 고려대학교 컴퓨터학과 학사
- 2001년 8월 : 고려대학교 컴퓨터학과 석사
- 2008년 2월 : 고려대학교 컴퓨터학과 박사
- 2008년 3월 ~ 2008년 8월 : 고려대학교 연구교수
- 2008년 9월 ~ 2010년 12월 : 조지아공대 박사후과정
- 2011년 1월 ~ 2014년 2월 : InterDigital Labs Staff Engineer
- 2014년 3월 ~ 2015년 2월 : 삼성전자 수석연구원/파트장
- 2015년 3월 ~ 2019년 8월 : 가천대학교 컴퓨터공학과 조교수
- 2019년 9월 ~ 현재 : 성균관대학교 실감미디어공학과 학과장, 지능형 멀티미디어 연구센터 센터장
- ORCID : https://orcid.org/0000-0003-4894-6105
- 주관심분야 : 멀티미디어 시스템, 비디오 코딩 및 국제 표준, HMD/VR 응용분야