

Review

Not peer-reviewed version

3D Reconstruction Techniques and the Impact of Lighting Conditions on Reconstruction Quality: A Comprehensive Review

[Dimitar Rangelov](#)*, [Sierd Waanders](#), [Kars Waanders](#), [Maurice van Keulen](#), [Radoslav Miltchev](#)

Posted Date: 1 April 2025

doi: 10.20944/preprints202504.0129.v1

Keywords: 3D reconstruction; Lighting Conditions; Reconstruction Quality; Structure from Motion (SfM); Multi-View Stereo (MVS); Neural Radiance Fields (NeRF); 3D Gaussian Splatting (3DGS); Photometric Consistency; Computer Vision; Illumination Variability.



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

3D Reconstruction Techniques and the Impact of Lighting Conditions on Reconstruction Quality: A Comprehensive Review

Dimitar Rangelov ^{1,2,*}, Sierd Waanders ^{1,3}, Kars Waanders ^{1,3}, Maurice van Keulen ²
and Radoslav Miltchev ⁴

¹ Technologies for Criminal Investigations, Saxion University of Applied Sciences, 7513 AB Enschede, The Netherlands

² Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 7522 NB Enschede, The Netherlands

³ Police Academy of The Netherlands, 7334 AC Apeldoorn, The Netherlands

⁴ Faculty of Industrial Technology, Technical University of Sofia, 1756 Sofia, Bulgaria

* Correspondence: d.g.rangelov@utwente.nl

Abstract: Three-dimensional (3D) reconstruction has become a fundamental technology in applications ranging from cultural heritage preservation and robotics to forensics and virtual reality. As these applications grow in complexity and realism, the quality of the reconstructed models becomes increasingly critical. Among the many factors that influence reconstruction accuracy, lighting conditions at capture time remain one of the most influential yet widely neglected variables. This review provides a comprehensive survey of classical and modern 3D reconstruction techniques, including Structure from Motion (SfM), Multi-View Stereo (MVS), Photometric Stereo, and recent neural rendering approaches such as Neural Radiance Fields (NeRF) and 3D Gaussian Splatting (3DGS), while critically evaluating their performance under varying illumination conditions. We describe how lighting-induced artifacts such as shadows, reflections, and exposure imbalances compromise reconstruction quality, and how different approaches attempt to mitigate these effects. Furthermore, we uncover fundamental gaps in current research, including the lack of standardized lighting-aware benchmarks and the limited robustness of state-of-the-art algorithms in uncontrolled environments. By synthesizing knowledge across fields, this review aims to gain a deeper understanding of the interplay between lighting and reconstruction and provides research directions for the future that emphasize the need for adaptive, lighting-robust solutions in 3D vision systems.

Keywords: 3D reconstruction; lighting conditions; reconstruction quality; Structure from Motion (SfM); Multi-View Stereo (MVS); Neural Radiance Fields (NeRF); 3D Gaussian Splatting (3DGS); photometric consistency; computer vision; illumination variability

1. Introduction

Three-dimensional (3D) reconstruction is a foundational task of computer vision, enabling the creation of precise digital replicas of real scenes from visual data, most often from 2D images or image sequences. It is a core component of a wide variety of application fields, ranging from recording of historical artifacts for the preservation of heritage, via the interpretation of the world for autonomous robotics, the realistic digital environments needed for virtual and augmented reality, and the precise documentation of crime scenes for investigation and legal proceedings [1–4]. The diversity of application fields has given rise to the development of numerous reconstruction pipelines, ranging from the classical photogrammetric approaches up to advanced neural rendering models. Among the most widely used ones are Structure from Motion (SfM) [5–7], reconstructing camera poses and sparse geometry from a sequence of images, and Multi-View Stereo (MVS) [8,9], densifying these

representations with pixel correspondences. In the recent years, learning-based approaches such as Neural Radiance Fields (NeRF) [10] and 3D Gaussian Splatting (3DGS) [11,12] attracted attention due to their ability for generating novel views and reconstructing complex geometries with high photorealism.

Despite these advancements, a critical yet often underexplored variable continues to impact the fidelity and reliability of 3D reconstructions is lighting conditions. Illumination directly influences the appearance of surfaces in images, affecting colour consistency, feature visibility, and shading cues that many reconstruction algorithms depend upon. Variations in lighting, such as cast shadows, reflections, specular highlights, or exposure differences, can lead to significant degradation in point cloud density, surface smoothness, and overall geometric accuracy. For example, areas under poor or uneven lighting may result in missing geometry due to lack of discernible features, while overexposed regions might introduce false depth information. Although some techniques assume stable lighting or rely on preprocessing for normalization, these assumptions rarely hold in uncontrolled or outdoor environments. As a result, reconstruction models often struggle with robustness when deployed in the wild, and the impact of lighting on reconstruction performance remains an active area of investigation with limited systematic analysis in existing literature.

Theoretically, this review brings together understanding from classical and state-of-the-art 3D reconstruction pipelines by examining assumptions and constraints with respect to illumination. We contrast the means by which methods represent or circumvent illumination, their photometric robustness, and the underlying geometric or learning-based theoretical frameworks that govern their sensitivity to light variation.

From a practical point of view, this work is a practitioner's handbook for data acquisition specialists and researchers carrying out 3D reconstruction. By summarizing empirical findings from the literature, we present practical information on the effect of illumination on reconstruction outcomes under real conditions. This includes the detection of common photometric artifacts, data acquisition best practices, and guidance on the selection of appropriate reconstruction methods for varying illumination conditions.

The remainder of this paper is structured as follows. Section 2 introduces relevant 3D reconstruction approaches, comparing classical and state-of-the-art ones. Section 3 is reserved for the contribution of the light conditions at capture and reconstruction of the images, e.g., photometric challenges and modelling approaches. Section 4 shows empirical results and comparative studies measuring the reconstruction quality for different light conditions. In Section 5, we present a critical discussion of the current limitations and methodological gaps. Section 6 introduces future research directions, and Section 7 concludes the review with the most significant insights and proposals.

2. Relevant Techniques

Three-dimensional reconstruction has evolved on a spectrum of methodologies, from the classical geometric approaches to the most recent neural rendering approaches. All these classes operate on different assumptions on scene geometry, image generation, and environmental factors, particularly lighting. In this chapter, we give a formal description of the most common reconstruction methods and a concise description of their principles, their strengths, and their known illumination sensitivities.

2.1. *Traditional Geometry-Based Methods*

2.1.1. Structure from Motion (SfM)

One of the earliest and most influential approaches for 3D reconstruction from unorganized collections of images is Structure from Motion. Structure from Motion estimates sparse 3D structure and camera poses by detecting and matching salient features across views. SURF [13–15] or SIFT [16,17] feature detectors would usually be employed, and bundle adjustment would be used for camera parameter and point location optimization. Despite being incredibly powerful for textured

scenes with uniform illumination, SfM is sensitive to photometric inconsistencies. Feature matching can be caused by shadows, low contrast, and reflections, leading to sparse or false reconstructions.

2.1.2. Multi-View Stereo (MVS)

Once sparse reconstruction is obtained with the assistance of SfM, Multi-View Stereo can densify the point cloud by estimating depth maps from overlapping parts of the images [18]. MVS relies on photometric consistency, i.e., the fact that corresponding pixels from views must be of the same look. Photometric consistency can be violated with the variation of light, and MVS is particularly sensitive, therefore, to illumination variation, specularities, and shadows. Despite some advanced versions incorporating surface regularization or shading models, the technique remains fundamentally light-sensitive.

2.1.3. Photometric Stereo

Photometric Stereo reconstructs high-resolution surface orientation data by viewing the same stationary scene with a number of images under differently directed lights, with the light source being known and the surface reflectance being Lambertian [19,20]. Even though effective for the recovery of high-resolution fine geometry, the technique is heavily constrained and not suitable for unstructured environments. In addition, it depends on a known light configuration, which is not practical for most real cases outside the laboratory.

2.2. *Learning-Based and Neural Rendering Methods*

2.2.1. Volumetric and Depth-Fusion Approaches

Learning-based volumetric methods make use of CNNs for learning voxelized scene representations or depth maps from images [21]. Some illumination invariances can be learned by these models, but these will be data-driven heavily and may not generalize across new illumination conditions unless trained specifically for the purpose.

2.2.2. Neural Radiance Fields (NeRF)

NeRF renders a scene as a continuous volumetric function parameterized by a neural network describing colour and density at any point in 3D, conditioned on position and view direction. NeRF has performed extremely well on view synthesis and scene reconstruction, particularly under constrained conditions. NeRF, however, relies on a stationary illumination condition for input images and is not invariant to dynamic light or shadows unless specifically adapted (e.g., with relighting-aware extensions). As it integrates lighting into the learned radiance field, generalization across new light conditions or robustness against changing light remains a core challenge.

2.2.3. 3D Gaussian Splatting (3DGS)

3D Gaussian Splatting is a recent rendering method that is based on the scene being represented as a group of 3D Gaussian anisotropic primitives with colour and transparency information. 3D Gaussian Splatting is capable of real-time rendering and high-visual-fidelity rendering. Despite the fact that it lacks a native dynamic or varying light mechanism, its explicit management of spatial data leaves room for the inclusion of light-aware rendering pipelines. However, current implementations make the assumption of uniform lighting on the training images and may be suboptimal for strong illumination changes.

3. Lighting Conditions in 3D Reconstruction

Lighting is one of the most critical factors influencing the performance of image-based 3D reconstruction techniques. Most reconstruction pipelines, whether geometric or learning-based, rely heavily on visual cues derived from pixel intensity, texture, shading, and colour consistency. Variations in lighting during image capture can significantly alter these visual features, leading to mismatches, depth estimation errors, and loss of geometric fidelity. In this section, we analyse how lighting conditions interfere with reconstruction quality and explore the typical assumptions, challenges, and emerging solutions associated with this problem.

3.1. Photometric Challenges in Reconstruction

Photometric consistency is a foundational assumption in many 3D reconstruction pipelines, particularly those based on geometry-driven approaches such as Structure from Motion (SfM) and Multi-View Stereo (MVS). This assumption implies that the appearance of a point in the 3D scene remains similar across multiple images taken from different viewpoints. In mathematical terms, it assumes Lambertian reflectance properties, where a surface reflects light uniformly in all directions, and image intensities vary only due to geometric transformations, not lighting changes. However, in real-world conditions, this assumption is frequently violated.

One of the most significant disruptions to photometric consistency arises from shadowing. Shadows can obscure important features, drastically altering their local contrast or causing them to disappear entirely from certain viewpoints. This makes feature detection and matching highly unreliable in those regions. Furthermore, self-shadowing, where parts of an object cast shadows onto themselves, creates apparent changes in surface structure that confuse depth estimation algorithms.

Another challenge is introduced by specular highlights. Unlike diffuse reflections, specular reflections depend on both the viewing angle and the light source direction, causing bright spots that move across the surface depending on the camera pose. These view-dependent effects violate the assumption of photometric invariance and lead to false correspondences in stereo matching, introducing ghosting artifacts or distorted geometry [22–24]

Exposure variation also significantly affects photometric stability, especially in uncontrolled environments or when using automatic camera settings. Changes in exposure between frames can result in varying brightness and contrast levels, even when the scene remains static. This leads to inconsistent pixel values across views, degrading the effectiveness of both feature-based and dense reconstruction methods [25]. In learning-based approaches, inconsistent lighting between training views can confuse the neural model, causing it to encode lighting as part of the scene geometry or material, which reduces generalization and reconstruction accuracy.

Colour temperature shifts further complicate the reconstruction process. Scenes illuminated by different light sources (e.g., daylight, incandescent, LED) exhibit colour casts that alter RGB values in non-linear ways. For example, an object may appear warmer under tungsten lighting and cooler under daylight, even though its geometry is unchanged. Without proper white balance or colour correction, these variations interfere with feature descriptor matching and learning-based texture synthesis.

Additionally, dynamic lighting environments, where the light source moves or changes intensity over time, introduce further variability. This is particularly problematic for outdoor scenes captured over extended periods, such as drone footage, time-lapse scanning, or forensic scene reconstruction under mixed natural and artificial lighting.

In essence, any violation of the photometric consistency assumption introduces noise, ambiguity, or missing data into the reconstruction pipeline. While robust algorithms may handle minor lighting differences, significant photometric inconsistencies often require specialized preprocessing or architectural adaptations to preserve reconstruction quality.

3.2. Modeling and Mitigating Lighting Effects

As lighting inconsistencies represent a major source of error in 3D reconstruction, various strategies have been developed to model, mitigate, or adapt to these photometric variations. These approaches can broadly be classified into preprocessing techniques, photometric-invariant descriptors, shading-aware reconstruction models, and neural methods with explicit lighting modeling. Each strategy offers different trade-offs between robustness, computational complexity, and generalizability.

3.2.1. Preprocessing and Normalization

One of the most accessible methods to deal with lighting inconsistencies is image preprocessing before feeding data into a reconstruction pipeline. This often includes techniques such as:

- Histogram equalization, which standardizes the intensity distribution across images to reduce contrast disparities.
- White balance correction, which adjusts the image to a standard neutral gray, reducing colour temperature shifts due to different light sources.
- Retinex theory-based methods, which attempt to decompose an image into illumination and reflectance components, preserving structural detail while suppressing lighting effects [26]

While these methods are relatively easy to apply and computationally efficient, they may also degrade valuable photometric cues necessary for certain types of depth estimation, such as stereo matching that relies on subtle shading gradients. Furthermore, they can introduce artifacts or overcorrect in scenes with mixed lighting or high dynamic range.

3.2.2. Photometric-Invariant Feature Descriptors

To address the problem at the algorithmic level, researchers have developed illumination-invariant descriptors for feature detection and matching. Examples include:

- Gradient-based descriptors (e.g., normalized gradient orientation), which are less sensitive to absolute intensity.
- Colour-invariant representations, which transform RGB values into chromaticity coordinates, isolating hue from illumination.
- Moment-based features like Zernike moments and local phase information, which capture structural patterns rather than raw intensity [27,28]

These descriptors enhance robustness against brightness and colour variation, particularly in SfM and MVS pipelines that rely on consistent keypoint detection. However, they may lose precision in textured regions and often trade fine detail for robustness, which can result in smoother but less accurate reconstructions.

3.2.3. Shading-Aware Reconstruction Techniques

Another line of defence involves explicitly modelling how lighting interacts with surface geometry, which is particularly relevant in shading-aware reconstruction. For instance:

- Intrinsic image decomposition techniques aim to separate an image into reflectance and shading layers, isolating object colour from illumination effects [29]
- Photometric bundle adjustment, which jointly optimizes camera parameters and surface appearance under varying lighting, has been applied in some hybrid SfM systems [30]

These methods allow reconstruction pipelines to reason about scene lighting rather than suppressing it, enabling more accurate modeling of shape and texture. However, they often require prior knowledge about scene materials or light sources, and are computationally expensive. Many of these techniques also struggle with non-Lambertian surfaces (e.g., glossy or transparent materials).

3.2.4. Learning-Based Approaches with Lighting Awareness

Recent advances in deep learning have enabled data-driven approaches to model and disentangle lighting from geometry. Several notable techniques include:

- NeRF-W (Neural Radiance Fields in the Wild), which augments NeRF with appearance embeddings that account for lighting variations between views, enabling more robust reconstructions in unconstrained environments [31]
- Neural reflectance decomposition models, such as NeRD [32] which aims to learn separate latent representations for shape, material, and illumination.
- Relighting-aware networks, trained with synthetic datasets under variable lighting, to infer canonical scene representations that can generalize across different illumination setups.

These models represent a significant leap toward real-world robustness but come with their own limitations. They often require large and diverse training datasets, multi-view supervision, or calibrated lighting information. Additionally, while many of these methods are highly effective at novel view synthesis, their utility in generating metrically accurate 3D geometry is still under active research.

3.2.5. Hybrid and Adaptive Systems

Some recent pipelines aim to combine a number of strategies, unifying preprocessing, shading models, and data learning for the purpose of building adaptive systems that can function under a wide range of light. Real-time SLAM for robotics [33], for example, has begun incorporating learned light correction modules, and some of the uses of the technology for forensics and heritage purposes include the application of light rigs with software-level compensation. These hybrid methods show great potential but are still rare and application specific.

3.3. *Controlled vs Uncontrolled Environments*

The effect of the illumination conditions on the performance of 3D reconstruction is intimately connected with the imaging scenario control level. In photogrammetry studios, laboratories, and dedicated scanning facilities, the light can be adjusted for minimum variation and artifacts. This enables reconstruction techniques to function close to their ideal operating conditions. In the other situation, uncontrolled ones such as outdoor scenes, catastrophe zones, public areas, and crime scenes for forensics, light is typically unpredictable and non-uniform, introducing significant reconstruction challenges.

3.3.1. Controlled Environments

Controlled environments are characterized by deliberate and often standardized illumination conditions. These typically include:

- Uniform diffuse lighting to minimize harsh shadows and highlights.
- Fixed camera exposure and white balance settings to ensure photometric consistency.
- Stationary or synchronized lighting setups, such as ring lights, softboxes, or dome lighting systems that eliminate directional shadows.

Such settings exist for the digitization of heritage, factory inspection, and laboratory-based 3D scanning, for which geometric accuracy and repeatability are critical. In such cases, the illumination is incorporated into the design of the capture pipeline, and the calibration is undertaken with utmost care such that photometric artifacts are eliminated or minimized [34].

Reconstruction under constrained conditions will yield high-fidelity results with dense point clouds, clean meshes, and accurate surface textures. In fact, data captured under such conditions form the test basis for most reconstruction algorithms. However, models trained or tuned on these datasets may fail to generalize when deployed in the field.

3.3.2. Uncontrolled Environments

In contrast, uncontrolled environments are inherently variable, dynamic, and complex in terms of lighting. These conditions are common in:

- Outdoor environments, where illumination is affected by the time of day, weather, and shadows cast by surrounding structures.
- Indoor scenes with mixed lighting, such as a combination of natural light from windows and artificial bulbs of varying colour temperatures.
- Forensic or emergency situations, where documentation must be conducted quickly without the ability to manipulate lighting.

Under these conditions, photometric inconsistencies become particularly prominent. The direction, intensity, and spectrum of light can vary from one picture to the next, even for a series of pictures. Furthermore, the presence of moving shadows, reflective, and translucent surfaces, and automatic camera exposure control adds complexity.

These unstructured environments directly undermine the efficiency of methods like SfM and MVS, which are reliant on regular key point correspondences and brightness patterns across views. Even learning-based methods trained under stationary illumination can confuse shadow boundaries with object boundaries or confuse specular highlights with geometric features [35]

Additionally, the majority of the neural methods (i.e., NeRF, 3D Gaussian Splatting) assume fixed light for input views, and their performance is compromised when this is not the case. Some of the extensions utilize appearance embeddings or light disentanglement procedures, but these can be data-intensive and require careful capture settings, which are not typically available for real-world fieldwork.

4. Empirical Evidence from the Literature

While theoretical analyses highlight the sensitivity of 3D reconstruction techniques to lighting variations, empirical validation is essential for understanding the practical impact of illumination in real-world and synthetic conditions. This section reviews notable experimental studies, benchmark datasets, and comparative evaluations that explore how lighting affects reconstruction fidelity. By synthesizing these findings, we aim to identify patterns in algorithmic robustness and outline where current methods fall short under photometric variability.

4.1. Benchmark Datasets with Lighting Variability

Several benchmark datasets have been created to support research on 3D reconstruction. However, relatively few of them deliberately introduce lighting variation as a core variable. Among the most relevant is the DTU dataset [36,37], which captures over 120 scenes with multiple camera positions and different lighting setups. The dataset includes scans with strong shadows, side lighting, and ambient-only configurations. These variations have proven useful for testing robustness in Multi-View Stereo (MVS) pipelines, revealing substantial performance drops under directional lighting.

Similarly, the Middlebury Multi-Illumination Stereo dataset [38–43] provides stereo pairs of static objects photographed under up to 64 different lighting conditions using a dome rig. This setup enables isolated analysis of how lighting direction impacts correspondence matching and disparity map quality. Studies using this dataset have demonstrated that traditional stereo algorithms often fail under directional lighting due to strong cast shadows and specular reflections.

However, mainstream benchmarks like Tanks and Temples [44] or ETH3D [45,46] focus primarily on geometric diversity and scene scale rather than photometric diversity. These datasets are invaluable for testing algorithmic scalability but provide limited insights into lighting robustness. As a result, evaluations based solely on such benchmarks may overestimate the real-world generalizability of reconstruction methods.

4.2. Comparative Evaluations under Varying Light

Controlled experimental studies have shown consistent evidence that illumination variability significantly affects reconstruction accuracy across algorithmic categories.

For example, Yoon and Kweon (2006) [47] tested local and global stereo matching algorithms under simulated lighting changes. They found that global optimization methods with adaptive cost aggregation, such as graph cuts and belief propagation, showed better resilience compared to local approaches, though none were entirely immune to strong photometric shifts.

In Goesele et al. (2007) [48], the authors explored how uncalibrated MVS performs under lighting-induced noise. They observed that reconstructions degraded visibly in shadowed regions, with increased surface roughness and geometric holes in areas lacking texture or uniform illumination. Their study emphasized the need for illumination-aware confidence maps or visibility weighting.

Recent neural approaches have attempted to mitigate lighting effects through learned compensation. NeRF in the Wild (NeRF-W) [31] introduced per-image embeddings that allowed the model to adjust to appearance changes, improving reconstructions from unconstrained photo collections. Although it offered more visually coherent results than vanilla NeRF under variable lighting, some structural accuracy was sacrificed in the process. Similarly, 3D Gaussian Splatting [11] has been used in high-fidelity rendering scenarios but has not yet been systematically tested for lighting invariance, and its robustness under photometric variation remains an open question.

4.3. Summary of Observed Impacts

A cross-study synthesis highlights several consistent trends:

- Traditional methods (SfM, MVS) perform best under stable, diffuse lighting and are particularly vulnerable to shadows, reflections, and low-contrast regions.
- Dense stereo algorithms degrade under directional lighting due to ambiguity in correspondence matching, especially in textureless or reflective areas.
- Neural rendering techniques like NeRF and its variants demonstrate improved visual coherence under photometric inconsistency, but this often comes at the expense of structural precision and requires large, well-sampled datasets.
- Benchmark limitations continue to constrain our ability to comprehensively evaluate lighting robustness, as most public datasets lack intentional photometric diversity.
- Few studies provide quantitative metrics for lighting robustness, and visual inspection remains a dominant (and subjective) evaluation method in many papers.

Overall, the existing body of empirical work confirms the substantial impact of lighting on 3D reconstruction quality. However, there remains a clear need for standardized lighting-aware benchmarks, large-scale comparative studies, and robust evaluation metrics to enable fair and reproducible assessments of algorithmic performance under real-world illumination conditions.

5. Discussion

The analysis presented in the previous sections reveals that lighting conditions play a central yet under-addressed role in determining the effectiveness of 3D reconstruction techniques. Although both classical and modern approaches have evolved significantly in terms of geometric modelling and visual fidelity, photometric robustness remains a persistent limitation. In this section, we critically discuss the core challenges, trade-offs, and methodological gaps in current research, with a focus on practical implications and future research trajectories.

5.1. Sensitivity of Classical vs Neural Methods

Classical geometry-based pipelines such as Structure from Motion (SfM) and Multi-View Stereo (MVS) are inherently sensitive to lighting inconsistencies due to their reliance on photometric

consistency and feature-based matching. While these methods can produce metrically accurate reconstructions under ideal conditions, their performance deteriorates sharply in the presence of shadows, exposure variations, and specularities. Preprocessing and feature engineering can help to some extent, but their effect is limited when faced with complex lighting interactions or mixed-illumination scenes.

In contrast, learning-based approaches, particularly neural rendering techniques like NeRF and its variants, show greater resilience to minor photometric inconsistencies. By learning scene-dependent appearance models, these methods can infer structure and texture even when lighting varies across views. However, this flexibility comes with notable drawbacks. First, the radiance fields learned by models like NeRF inherently encode lighting into the scene representation, which makes them fragile when generalizing to new lighting conditions. Second, these methods are computationally expensive, require dense image sampling, and often depend on scene-specific training. Third, they tend to prioritize view synthesis realism over metric accuracy, which may not be acceptable in applications like forensics, engineering, or metrology.

Overall, while neural methods offer improved robustness in terms of visual appearance, they are not inherently immune to lighting-induced reconstruction errors and often lack interpretability or error diagnostics found in traditional geometric pipelines.

5.2. Lack of Lighting-Aware Benchmarks and Metrics

One of the most significant bottlenecks in advancing photometric robustness in 3D reconstruction is the lack of standardized benchmarks that include diverse and controlled lighting conditions. Most existing datasets are captured in ideal or stable environments, meaning that algorithmic performance under lighting variation is rarely tested or reported. This gap leads to overfitting of methods to photometric consistency, reducing their applicability in real-world scenarios.

Furthermore, current evaluation metrics, such as point cloud accuracy, completeness, or re-projection error, do not capture lighting-specific degradations like shadow-induced sparsity, reflectance-related misalignments, or photometric inconsistency across views. There is a pressing need to develop lighting-aware evaluation metrics that can assess not only geometric precision but also the consistency and realism of surface appearance under different lighting conditions.

5.3. Practical Implications and Application Constraints

In practical applications such as crime scene documentation, cultural heritage digitization, or drone-based 3D mapping, reconstruction often takes place in uncontrolled lighting environments. The operator may not have the luxury of adjusting lighting, setting fixed exposures, or capturing redundant views. Under such constraints, robustness to photometric conditions becomes essential, not optional.

However, few existing reconstruction pipelines are optimized for such scenarios. Field users are often left to rely on manual trial-and-error adjustments or empirical best practices (e.g., photographing at dawn or dusk to avoid harsh shadows), which adds operational complexity and uncertainty. Moreover, many neural techniques currently assume computational infrastructure and data volume not readily available in mobile or real-time settings.

These limitations underscore a broader issue: the disconnect between research-oriented reconstruction techniques and real-world deployment needs, particularly in environments where lighting control is infeasible.

5.4. Toward Lighting-Resilient Reconstruction Pipelines

To bridge the gap between theoretical robustness and practical usability, future reconstruction systems should integrate multiple strategies, including:

- Photometric calibration tools during data acquisition.

- Self-supervised learning frameworks that can adapt to lighting variation without explicit relighting supervision.
- Confidence-based fusion techniques, which down-weight or exclude views with inconsistent lighting.
- Cross-modal data integration, combining visual, depth, and reflectance data to mitigate reliance on intensity-based matching.

Moreover, interdisciplinary collaboration between computer vision, optics, and machine learning communities will be essential for developing principled models of illumination, capable of generalizing across environments and use cases.

6. Future Directions

The problem of light variation for 3D reconstruction is a multidimensional research problem. Arguably, the most critical necessity is the formulation of reconstruction models that can learn light invariant representations. Ideally, these representations should factorize geometry, material, and light in a structured manner, making the representations stronger and enabling applications such as scene relighting or adaptive view synthesis. Recent work on neural rendering has made some early attempts toward this goal, but most of the current models still leverage supervised data such as light direction or reflectance ground truth. A key opportunity lies in exploring unsupervised and weakly supervised learning approaches, particularly those informed by physical models of light interaction and scene geometry.

Another important direction is the creation of benchmark datasets that explicitly incorporate photometric diversity. Most current benchmarks favor geometric or textural variation but don't take into consideration the complex dynamics of light. Datasets for the future would be improved if scenes were captured under a wide range of light conditions, both indoors, under control, and outdoors, under natural, dynamic conditions. High-quality synthetic datasets with realistic light and material interactions can also be utilized for training improved generalization models, particularly if coupled with domain adaptation techniques.

Real-time reconstruction systems that work effectively in the real world, with its unpredictable and varying light conditions, would also become increasingly desired. Embedded photometric calibration features, such as ambient light property detection from sensors or camera metadata that adjusts for exposure and white balance, would be beneficial for such systems. Beyond calibration, dynamic frame selection strategies that prefer consistent input images could render reconstruction less sensitive with little post-processing. Real-time adaptability will be important for forensics, public safety, and mobile robotics, where the user is not controlling the surrounding conditions.

Adding other sensing modalities can further reduce the reliance on ambient light. Mixing vision with infrared, LiDAR, or even polarisation sensor data is especially promising. These modalities are less sensitive to the fluctuation of light and can be utilized to complement photometric data, especially for challenging surface properties like specularly or translucency.

Finally, domain-specific toolkits and procedures should be designed. In crime scene documentation, for instance, preservation of heritage, or remote drone-based inspections, controlling the light or re-capturing data is typically not feasible. Lightweight relighting models, optimized capture procedures, or portable light gear should be tailored for such environments and incorporated into everyday workflows. In the meantime, exchange with the community of computer vision researchers and the community of field workers will be critical for ensuring that innovations in the technical domain take into consideration the practical limitations of the real world. It will involve a blend of algorithmic design, data collection, and application-driven insight. By not seeing light variation as a problem, but as a fundamental consideration for 3D reconstruction, future systems will be constructed stronger, truer, and better prepared for the subtleties of real-world deployment.

7. Conclusions

Lighting conditions constitute a key but yet often underappreciated variable in the field of 3D reconstruction. Even with the significant advances made with classical geometric approaches and newer neural rendering approaches, this review has shown that photometric variation is still a serious issue for reconstruction quality and consistency. Across all technique classes, be it structure-from-motion, multi-view stereo, or even the newer ones like NeRF and 3D Gaussian Splatting, lighting is the most important determinant of the fidelity of the reconstruction, particularly for uncontrolled environments.

Through the analysis of theoretical principles and empirical research, we identified shared problems such as sparsity caused by shadows, specular reflection artifacts, and general degradation in correspondence matching under non-uniform lighting. Learning-based approaches are more robust than their classical counterparts but sacrifice structural correctness or require additional training data and supervision. In addition, the lack of photometrically varied benchmarks and illumination-sensitive evaluation metrics has prevented the community from systematically comparing and pushing the state of the art.

This review has also indicated promising directions for future research, including learning representations invariant to light, photometric dataset construction, multimodal sensor fusion, and domain-specific reconstruction toolkits. Ultimately, the ability to produce high-fidelity 3D reconstructions under disparate light conditions is not only a technological objective. It is a step toward real-world applicability across forensics, archaeology, robotics, and beyond.

By advancing the boundaries of 3D reconstruction with consideration for lighting, the field can reach toward systems that not only appear realistic but also geometrically correct and deployable in the unpredictable environments where they will be of most utility.

Author Contributions: Conceptualization, K.W., S.W., R.M., M.v.K. and D.R.; methodology, D.R.; software, K.W., S.W. and D.R.; validation, K.W., S.W. and D.R.; formal analysis, D.R.; investigation, K.W., S.W. and D.R.; resources, D.R. and R.M.; data curation, D.R.; writing—original draft preparation, D.R.; writing—review and editing D.R., M.v.K. and R.M.; visualization, K.W. and S.W.; supervision, M.v.K. and R.M.; project administration, D.R.; funding acquisition, D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received financial support from the University of Twente and research group Technologies for Criminal Investigations, part of Saxion University for Applied Sciences and the Police Academy of the Netherlands. Also, this research is part of a project funded by the Police and Science grant of the Police Academy of the Netherlands and Stichting Saxion - Zwaartepunt Veiligheid & Digitalisering from Saxion University of Applied Sciences.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: Special gratitude to the University of Twente in the Netherlands, Technology for Criminal Investigations research group part of Saxion University of Applied Sciences and the Police Academy in the Netherlands, the Technical University of Sofia in Bulgaria and all researchers in the CrimeBots research line part of research group Technologies for Criminal Investigations. Also, thanks for the dedicated opportunity and funding from the Police Academy with the Police and Science grant.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SfM	Structure from Motion
MVS	Multi-View Stereo
NeRF	Neural Radiance Fields
GS	Gaussian Splatting
CNN	Convolutional Neural Network
NeRF-W	Neural Radiance Fields in the Wild
NeRD	Neural Reflectance Decomposition
SLAM	Simultaneous Localization and Mapping
RGB	Red-Green-Blue (colour channels)
DTU	Technical University of Denmark (dataset)
ETH3D	ETH Zurich 3D Reconstruction Benchmark
SIFT	Scale-Invariant Feature Transform
SURF	Speeded Up Robust Features

References

1. N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: Exploring photo collections in 3D," *ACM Trans Graph*, vol. 25, no. 3, pp. 835–846, Jul. 2006, doi: 10.1145/1141911.1141964/SUPPL_FILE/P835-SNAVELY-LOW.MOV.
2. E. Grilli, F. Menna, and F. Remondino, "A REVIEW OF POINT CLOUDS SEGMENTATION AND CLASSIFICATION ALGORITHMS," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2-W3, no. 2W3, pp. 339–344, Feb. 2017, doi: 10.5194/ISPRS-ARCHIVES-XLII-2-W3-339-2017.
3. F. Remondino and S. El-hakim, "Image-based 3D Modelling: A Review," *The Photogrammetric Record*, vol. 21, no. 115, pp. 269–291, Sep. 2006, doi: 10.1111/J.1477-9730.2006.00383.X.
4. D. Rangelov, J. Knotter, and R. Miltchev, "3D Reconstruction in Crime Scenes Investigation: Impacts, Benefits, and Limitations," *Lecture Notes in Networks and Systems*, vol. 1065 LNNS, pp. 46–64, 2024, doi: 10.1007/978-3-031-66329-1_4.
5. A. Shalaby, M. Elmogy, and A. Abo El-Fetouh, "Algorithms and Applications of Structure from Motion (SfM): A Survey," *International Journal of Computer and Information Technology*, p. 358, 2017, Accessed: Mar. 31, 2025. [Online]. Available: www.ijcit.com
6. J. L. Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited", Accessed: Mar. 31, 2025. [Online]. Available: <https://github.com/colmap/colmap>.
7. A. Eltner and G. Sofia, "Structure from motion photogrammetric technique," *Developments in Earth Surface Processes*, vol. 23, pp. 1–24, Jan. 2020, doi: 10.1016/B978-0-444-64177-9.00001-1.
8. F. Wang et al., "Learning-based Multi-View Stereo: A Survey," Aug. 2024, Accessed: Mar. 31, 2025. [Online]. Available: <https://arxiv.org/abs/2408.15235v2>
9. Y. Zhang, J. Zhu, and L. Lin, "Multi-View Stereo Representation Revist: Region-Aware MVSNet".
10. 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," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12346 LNCS, pp. 405–421, 2020, doi: 10.1007/978-3-030-58452-8_24.
11. B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," *ACM Trans Graph*, vol. 42, no. 4, p. 14, Aug. 2023, doi: 10.1145/3592433.
12. G. Chen and W. Wang, "A Survey on 3D Gaussian Splatting," Jan. 2024, Accessed: Mar. 31, 2025. [Online]. Available: <https://arxiv.org/abs/2401.03890v6>

13. S. Wu and B. Feng, "Parallel SURF Algorithm for 3D Reconstruction," pp. 153–157, Apr. 2019, doi: 10.2991/SMONT-19.2019.34.
14. X. Wang, W. Cao, C. Yao, and H. Yin, "Feature Matching Algorithm Based on SURF and Lowes Algorithm," Chinese Control Conference, CCC, vol. 2020-July, pp. 5996–6000, Jul. 2020, doi: 10.23919/CCC50068.2020.9189468.
15. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features".
16. K. Peng, X. Chen, D. Zhou, and Y. Liu, "3D reconstruction based on SIFT and Harris feature points," 2009 IEEE International Conference on Robotics and Biomimetics, ROBIO 2009, pp. 960–964, Dec. 2009, doi: 10.1109/ROBIO.2009.5420735.
17. P. Du, Y. Zhou, Q. Xing, and X. Hu, "Improved SIFT matching algorithm for 3D reconstruction from endoscopic images," Proceedings of VRCAI 2011: ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications to Industry, pp. 561–564, 2011, doi: 10.1145/2087756.2087864.
18. Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," IEEE Trans Pattern Anal Mach Intell, vol. 32, no. 8, pp. 1362–1376, 2010, doi: 10.1109/TPAMI.2009.161.
19. C. Yu, Y. Seo, and S. W. Lee, "Photometric Stereo from Maximum Feasible Lambertian Reflections," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6314 LNCS, no. PART 4, pp. 115–126, 2010, doi: 10.1007/978-3-642-15561-1_9.
20. H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," vol. 11, 1994.
21. A. Caglayan and A. B. Can, "Volumetric Object Recognition Using 3-D CNNs on Depth Data," IEEE Access, vol. 6, pp. 20058–20066, Mar. 2018, doi: 10.1109/ACCESS.2018.2820840.
22. R. Nair, A. Fitzgibbon, D. Kondermann, and C. Rother, "Reflection Modeling for Passive Stereo".
23. R. Yang, M. Pollefeys, and G. Welch, "Dealing with Textureless Regions and Specular Highlights-A Progressive Space Carving Scheme Using a Novel Photo-consistency Measure".
24. D. N. Bhat and S. K. Nayar, "Stereo and Specular Reflection," Int J Comput Vis, vol. 26, no. 2, pp. 91–106, 1998, doi: 10.1023/A:1007940725322/METRICS.
25. I. Konovalenko et al., "Influence of Uneven Lighting on Quantitative Indicators of Surface Defects," Machines 2022, Vol. 10, Page 194, vol. 10, no. 3, p. 194, Mar. 2022, doi: 10.3390/MACHINES10030194.
26. E. H. Land, "The retinex theory of color vision," Sci Am, vol. 237, no. 6, pp. 108–128, 1977, doi: 10.1038/SCIENTIFICAMERICAN1277-108.
27. H. D. Vankayalapati, S. Kuchibhotla, M. S. K. Chadalavada, S. K. Dargar, K. R. Anne, and K. Kyandoghere, "A Novel Zernike Moment-Based Real-Time Head Pose and Gaze Estimation Framework for Accuracy-Sensitive Applications," Sensors 2022, Vol. 22, Page 8449, vol. 22, no. 21, p. 8449, Nov. 2022, doi: 10.3390/S22218449.
28. K. N. Oo and A. K. Gopalkrishnan, "Zernike Moment Based Feature Extraction for Classification of Myanmar Paper Currencies," ISCIT 2018 - 18th International Symposium on Communication and Information Technology, pp. 208–213, Dec. 2018, doi: 10.1109/ISCIT.2018.8587967.
29. J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," IEEE Trans Pattern Anal Mach Intell, vol. 37, no. 8, pp. 1670–1687, Aug. 2015, doi: 10.1109/TPAMI.2014.2377712.
30. C. Wu, S. Agarwal, B. Curless, and S. M. Seitz, "Multicore Bundle Adjustment", Accessed: Mar. 31, 2025. [Online]. Available: <http://grail.cs>.
31. R. Martin-Brualla, N. Radwan, M. S. M. Sajjadi, J. T. Barron, A. Dosovitskiy, and D. Duckworth, "NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 7206–7215, 2021, doi: 10.1109/CVPR46437.2021.00713.

32. M. Boss, R. Braun, V. Jampani, J. T. Barron, C. Liu, and H. P. A. Lensch, "NeRD: Neural Reflectance Decomposition from Image Collections".
33. Z. Zhao et al., "Light-SLAM: A Robust Deep-Learning Visual SLAM System Based on LightGlue under Challenging Lighting Conditions," May 2024, Accessed: Mar. 31, 2025. [Online]. Available: <http://arxiv.org/abs/2407.02382>
34. F. Remondino and A. Rizzi, "Reality-based 3D documentation of natural and cultural heritage sites-techniques, problems, and examples," *Applied Geomatics*, vol. 2, no. 3, pp. 85–100, Jul. 2010, doi: 10.1007/S12518-010-0025-X/FIGURES/11.
35. M. Tancik et al., "Block-NeRF: Scalable Large Scene Neural View Synthesis," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022-June, pp. 8238–8248, 2022, doi: 10.1109/CVPR52688.2022.00807.
36. "DTU Dataset | Papers With Code." Accessed: Mar. 31, 2025. [Online]. Available: <https://paperswithcode.com/dataset/dtu>
37. R. Jensen, A. Dahl, G. Vogiatzis, E. Tola, and H. Aanaes, "Large Scale Multi-view Stereopsis Evaluation", Accessed: Mar. 31, 2025. [Online]. Available: <http://roboimagedata.imm.dtu.dk/>
38. D. Scharstein et al., "High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth", Accessed: Mar. 31, 2025. [Online]. Available: <http://vision.middlebury.edu/stereo/data/2014/>.
39. H. Hirschmüller and D. Scharstein, "Evaluation of Cost Functions for Stereo Matching".
40. D. Scharstein and C. Pal, "Learning Conditional Random Fields for Stereo", Accessed: Mar. 31, 2025. [Online]. Available: <http://vision.middlebury.edu/stereo/data/>.
41. D. Scharstein and R. Szeliski, "High-Accuracy Stereo Depth Maps Using Structured Light," vol. 1, pp. 195–202, 2003, Accessed: Mar. 31, 2025. [Online]. Available: <http://www.middlebury.edu/stereo/>.
42. D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms", Accessed: Mar. 31, 2025. [Online]. Available: www.middlebury.edu/stereo.
43. "vision.middlebury.edu/stereo/data." Accessed: Mar. 31, 2025. [Online]. Available: <https://vision.middlebury.edu/stereo/data/>
44. A. Knapitsch, J. Park, Q. Y. Zhou, and V. Koltun, "Tanks and temples: Benchmarking large-scale scene reconstruction," *ACM Trans Graph*, vol. 36, no. 4, Jul. 2017, doi: 10.1145/3072959.3073599/SUPPL_FILE/PAPERS-0083.MP4.
45. T. Schöps, T. Sattler, and M. Pollefeys, "BAD SLAM: Bundle Adjusted Direct RGB-D SLAM", Accessed: Mar. 31, 2025. [Online]. Available: www.eth3d.net.
46. T. Schöps et al., "A Multi-View Stereo Benchmark with High-Resolution Images and Multi-Camera Videos", Accessed: Mar. 31, 2025. [Online]. Available: www.eth3d.net.
47. K. J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Trans Pattern Anal Mach Intell*, vol. 28, no. 4, pp. 650–656, Apr. 2006, doi: 10.1109/TPAMI.2006.70.
48. M. Goesele, B. Curless, and S. M. Seitz, "Multi-View Stereo Revisited".

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.