Intelligent Evolution of Global Illumination Models and Their Applications in Virtual Power Scenarios

Wei Guo and Luping Wang*

Laboratory of 3D Scene Understanding and Visual Navigation, School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China

Abstract: Illumination models are fundamental techniques in computer graphics for simulating realistic lighting effects and are widely applied in virtual reality, game development, and cinematic visual effects. With the continuous advancement of hardware capabilities and rendering algorithms, illumination models have become increasingly important in both real-time and offline rendering pipelines. This paper presents a comprehensive review of recent advances in illumination modeling and systematically analyzes representative studies in this field. Particular attention is given to neural-network-based light baking, dynamic global illumination, environment light mapping, neural rendering with Transformer architectures, illumination decomposition and editing, low-light image enhancement, real-time global illumination for VR/AR environments, as well as emerging high-order illumination representations and performance optimization strategies. Furthermore, this paper provides an in-depth discussion of intelligent illumination reconstruction for virtual power scenarios. Specifically, we introduce a semantic–pixel coupled probabilistic multi-cue illumination estimation model (SPC-PMC) designed for power operation simulation systems, together with its corresponding three-dimensional visualization framework. Through comparative analysis of these techniques, this study examines their underlying principles, characteristics, advantages, and limitations. Combined with illustrative figures and mathematical formulations, the paper highlights recent progress in improving rendering quality and computational efficiency, while also outlining future research directions in intelligent illumination modeling.

Keywords: Illumination model, Virtual power scenario, Power operation, Intelligent operation and maintenance, Digital twin.

1. INTRODUCTION

Illumination models constitute a fundamental component of computer graphics for simulating realistic lighting effects. Their primary objective is to approximate physical lighting phenomena in the real world through mathematical formulations and computational algorithms. Traditional illumination models, such as the Phong and Blinn–Phong models, can efficiently reproduce basic shading effects; however, they exhibit inherent limitations when handling complex global illumination (GI) phenomena. As a central research topic in computer graphics, computer vision, and virtual reality, global illumination aims to simulate physical light transport processes, including propagation, reflection, and refraction, in order to generate highly realistic visual results. By accounting for multiple light interactions within a scene, GI techniques can reproduce sophisticated visual effects such as soft shadows, interreflections, and refractions. Nevertheless, accurate GI computation is typically computationally intensive, making real-time rendering challenging. In recent years, advances in graphics hardware and rendering algorithms have significantly improved both rendering quality and computational efficiency. In particular, the introduction of deep neural networks has opened new directions for illumination modeling, optimization, and acceleration.

In the power systems domain, the rapid development of next-generation power systems has substantially increased both the operational complexity and safety risks associated with power-related tasks, thereby imposing higher requirements on personnel training and emergency response capabilities. As a key approach for improving operational safety, intelligent power-operation training has evolved from conventional theoretical instruction and two-dimensional multimedia demonstrations toward highly immersive virtual power environments based on virtual reality (VR), augmented reality (AR), and digital twin technologies [1-3]. By constructing three-dimensional digital twins of substations, transmission towers, and other operational environments, high-fidelity power-operation simulation systems can provide trainees with risk-free, repeatable, and highly interactive training platforms. Such systems offer significant practical value in reducing training costs and mitigating operational safety risks. However, existing virtual power environments primarily emphasize accurate geometric reconstruction and static texture refinement, while largely overlooking the dynamic and time-varying evolution of real-world illumination conditions. As illumination plays a critical role in human visual perception of spatial geometry, material appearance, and environmental atmosphere [4-5], the quality of illumination reconstruction directly affects the realism and immersion of simulation

*Address correspondence to this author at the Laboratory of 3D Scene Understanding and Visual Navigation, School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China;
E-mail: 15110240007@fudan.edu.cn

Received on 24-06-2026 Accepted on 17-07-2026 Published on 27-07-2026

E-ISSN: 2978-6436/26

systems. Most existing power-operation simulation platforms still rely on predefined light sources or idealized static skybox configurations embedded in commercial rendering engines. Such simplified virtual illumination settings fail to reproduce realistic environmental conditions corresponding to actual operational scenarios, including varying weather conditions and temporal lighting changes. More importantly, they cannot accurately capture complex visual phenomena under extreme operating conditions, such as intense backlighting glare, geometric shadow occlusions caused by large steel frameworks, and intricate shadow interactions among transmission lines. Consequently, leveraging ubiquitous historical monitoring images to inversely estimate real-world illumination conditions and reconstruct physically consistent lighting environments for virtual power scenarios has become a critical challenge in enhancing the immersion and practical effectiveness of power-operation simulation training systems [1-4].

Beyond immersive training, illumination-consistent scene reconstruction is increasingly important for emerging digital-twin power systems. In modern intelligent energy infrastructures, digital twins are expected to provide not only accurate geometric representations of substations and transmission corridors but also realistic environmental awareness that reflects dynamic operating conditions. Accurate illumination reconstruction can improve visual consistency between physical assets and their digital counterparts, thereby supporting intelligent inspection, autonomous robotic navigation, equipment status interpretation, and human–machine interaction in virtual environments. Consequently, illumination estimation has evolved from a rendering-oriented task into an enabling technology for perception-driven digital-twin applications in smart power systems.

This paper presents a comprehensive review of recent advances in illumination modeling and reconstruction, with particular emphasis on their relevance to virtual power environments, intelligent energy systems, and digital-twin applications. The review covers the evolution of global illumination techniques, deep-learning-based illumination estimation, lightweight rendering frameworks for VR/AR platforms, and illumination reconstruction methods for complex outdoor power scenes. Furthermore, the paper discusses the proposed Semantic–Pixel Coupled Probabilistic Multi-Cue (SPC-PMC) framework and analyzes its potential for enhancing illumination-aware digital-twin construction, intelligent inspection, and immersive operation-training systems in future smart power networks.

2. TRADITIONAL FUNDAMENTAL ILLUMINATION MODELS

2.1. Global Illumination and Neural Rendering

Global illumination (GI) simulates complex light interactions within a scene, including multiple reflections, refractions, and indirect light transport. Among various GI approaches, path tracing is one of the most representative physically based rendering algorithms. It estimates illumination by stochastically sampling light transport paths and can generate highly realistic visual results. However, its computational cost remains substantial, particularly when handling complex light transport phenomena such as caustics. Path tracing computes the final pixel radiance by simulating the propagation of light rays throughout a scene. Its theoretical foundation is based on the rendering equation proposed by Kajiya:

$$L_o\left(x,\omega_o\right) = L_e\left(x,\omega_o\right) + \int_{\Omega} L_i\left(x,\omega_i\right)\rho\left(x,\omega_o,\omega_i\right)\cos\theta_i d\omega_i \quad (1)$$

where $L_o(x,\omega_o)$ denotes the outgoing radiance, $L_e(x,\omega_o)$ represents the emitted radiance, and $L_i(x,\omega_o)$ denotes the incident radiance. $\rho(x,\omega_o,\omega_i)$ is the bidirectional reflectance distribution function (BRDF), while θ_i denotes the angle between the surface normal and the incident direction ω_i.

Path tracing estimates the integral through Monte Carlo sampling of a large number of light transport paths to obtain physically accurate rendering results. Nevertheless, the computational burden is extremely high. As reported in [6], even high-end GPUs equipped with dedicated ray-tracing hardware can only afford a limited number of samples per pixel per frame in real-time applications. To alleviate rendering noise and improve efficiency, various denoising techniques have been proposed. For example, the method discussed in [6] reduces noise by reusing samples from neighboring pixels and adjacent frames. Although such approaches improve rendering quality to some extent, they still remain computationally expensive for mid-range and low-end hardware platforms. Figure **1** illustrates a global illumination result generated using path tracing.

Neural rendering has emerged in recent years as a promising rendering paradigm that leverages neural networks to learn illumination distributions and light transport behaviors. Unlike traditional physically based rendering approaches, neural rendering methods can generate high-quality images with significantly improved inference efficiency after training, making them particularly suitable for real-time rendering in dynamic environments. The neural-network-based global illumination framework proposed in [7] performs rendering by learning illumination characteristics from a

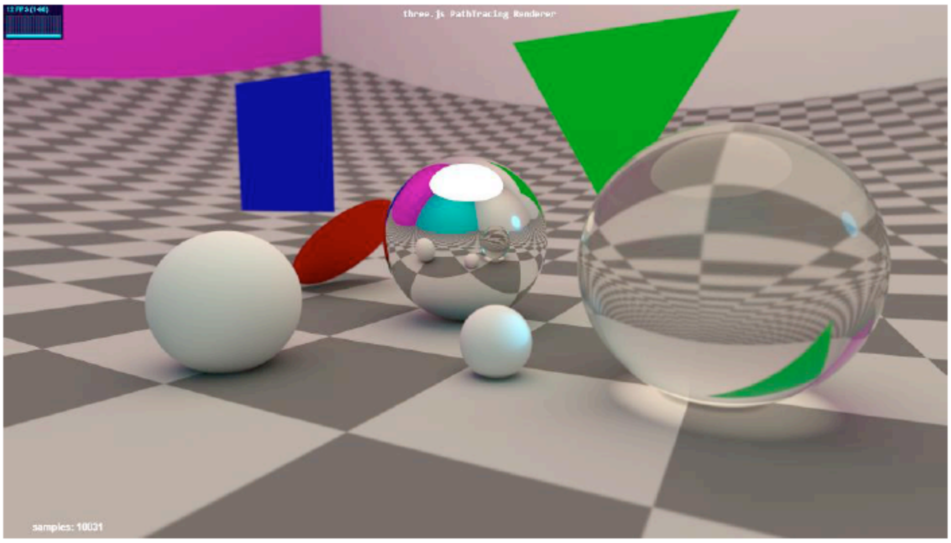

Figure 1: Global Illumination Effect Generated via Path Tracing.

large corpus of real-world images. To improve training efficiency, the method introduces an active exploration strategy based on Markov Chain Monte Carlo (MCMC) sampling, which guides sample generation within a high-dimensional sampling space toward regions that contribute most effectively to network optimization.

Specifically, the sampling process is guided by an objective function designed to prioritize informative training samples. The objective function is defined as the product of the reconstruction loss and the gradient norm:

$$\text{Objective Function} = \text{Loss} \times \text{Gradient Norm} \quad (2)$$

During training, the sampling strategy is continuously updated to generate samples that better satisfy the optimization requirements of the neural network, thereby accelerating convergence and improving rendering performance. In addition, the framework incorporates adaptive resolution scheduling and a self-adjusting sample reuse strategy, enabling interactive rendering of dynamic scenes after approximately 5–18 hours of training. The method can effectively handle complex hard-light transport paths while supporting dynamic scene navigation and user interaction. The network architecture is built upon a modified PixelGenerator framework, which learns a mapping from scene inputs to final pixel radiance values. The model is optimized using a combination of (L_1) reconstruction loss and structural similarity (SSIM) loss. Table **1** presents a comparative analysis of the effectiveness of active exploration and sample reuse strategies.

Similarly, the study in [11] proposes a neural-network-based illumination model that learns scene representations to predict lighting effects. During training, the model leverages large-scale scene datasets to capture illumination characteristics and underlying physical regularities, enabling accurate prediction of spatial illumination distributions under varying scene conditions. Compared with traditional approaches, neural-network-based illumination models demonstrate superior capability in handling complex lighting variations in challenging scenes, thereby improving both rendering realism and computational efficiency.

In traditional illumination modeling, Monte Carlo methods and importance sampling are two widely used numerical techniques for efficiently estimating complex light transport integrals. Monte Carlo methods form the foundation of path tracing and estimate illumination by stochastic sampling. As a cornerstone of physically based rendering, this approach has been extensively adopted in industry-level rendering systems, including modern GPU-accelerated path tracers such as Mitsuba 2 [7]. Importance sampling, as a variance reduction

Table 1: Comparison of the Effects of Active Exploration and Sample Reuse

Method	Training Time (Hours)	Rendering Quality(DSSIM)
Uniform Sampling	18	0.0241
Active Exploration	5	0.0141

Table 2: Comparison between Monte Carlo Methods and Importance Sampling

Method	Advantages	Disadvantages
Monte Carlo Method	Simple and easy to implement; suitable for complex scenarios	low convergence speed; significant noise
Importance Sampling	Improves sampling efficiency; reduces noise	Requires pre-computation; difficult to handle complex paths

strategy within the Monte Carlo framework, improves computational efficiency by preferentially sampling light transport paths that contribute more significantly to the final rendering result. In path tracing, Monte Carlo sampling is used to estimate global illumination by randomly tracing light paths within a scene. However, naive uniform sampling is computationally expensive and often exhibits high variance. By introducing importance sampling, sampling is biased toward directions and paths that have higher contributions to the final radiance, thereby reducing both computational cost and rendering noise. Table **2** presents a comparison between Monte Carlo methods and importance sampling.

With the continuous advancement of deep learning, Transformer architectures have been increasingly introduced into illumination modeling due to their strong capability in capturing long-range dependencies and complex relational structures. The method proposed in [14] presents a Transformer-based illumination model aimed at further improving rendering realism and computational efficiency. By leveraging the self-attention mechanism, the model effectively captures long-range dependencies between different spatial regions in a scene, which is crucial for accurately modeling global illumination under complex lighting conditions. In contrast to traditional convolutional neural networks, whose receptive fields are inherently limited, convolutional operations struggle to capture global illumination interactions in a single forward pass. The self-attention mechanism in Transformers enables each spatial location to attend to all other regions in the scene when estimating illumination, thereby providing a more holistic modeling of light transport, including propagation and reflection effects.

Specifically, the proposed model encodes scene illumination information into a set of feature tokens, which are then processed by a Transformer network. Through a Multi-Head Self-Attention mechanism, the model extracts illumination features from multiple representation subspaces, thereby enhancing its ability to model complex lighting patterns. For instance, in scenes containing multiple light sources and reflective surfaces, different attention heads can specialize in capturing distinct physical factors, such as light source influence regions, reflection paths, and occlusion relationships between objects. During training, a joint optimization strategy combining contrastive learning and reconstruction loss is employed. Contrastive learning enhances the model's discriminative capability under varying illumination conditions, enabling it to capture intrinsic illumination characteristics. Meanwhile, the reconstruction loss ensures consistency between the predicted illumination and the ground-truth rendering results. This joint training strategy leads to improved performance across multiple illumination-related tasks.

Experimental results demonstrate that, in complex indoor lighting scenarios, the proposed Transformer-based illumination model outperforms traditional convolutional neural network-based approaches in accurately modeling light propagation and reflection. It significantly reduces lighting artifacts and shadow inconsistencies. Moreover, for scenes with complex geometries and material properties, the proposed model achieves superior global and local consistency. In particular, it better reproduces physically plausible effects such as sunlight passing through windows and multiple interreflections between objects, exhibiting higher realism and accuracy. Figure **2** illustrates a qualitative comparison of the Transformer-based illumination model.

The RenderFormer model proposed in [15] extends Transformer architectures to neural rendering on triangle meshes, leveraging self-attention mechanisms to capture illumination interactions across mesh primitives. By explicitly modeling long-range dependencies among geometric elements, the method significantly improves view synthesis quality in complex scenes involving specular reflection and refraction. Experimental results demonstrate that RenderFormer achieves a PSNR improvement of approximately 1.5–2 dB over CNN-based approaches in novel view synthesis tasks. In particular, it exhibits superior performance in long-range light transport scenarios involving multiple bounces. Moreover, RenderFormer preserves fine-grained geometric details and produces

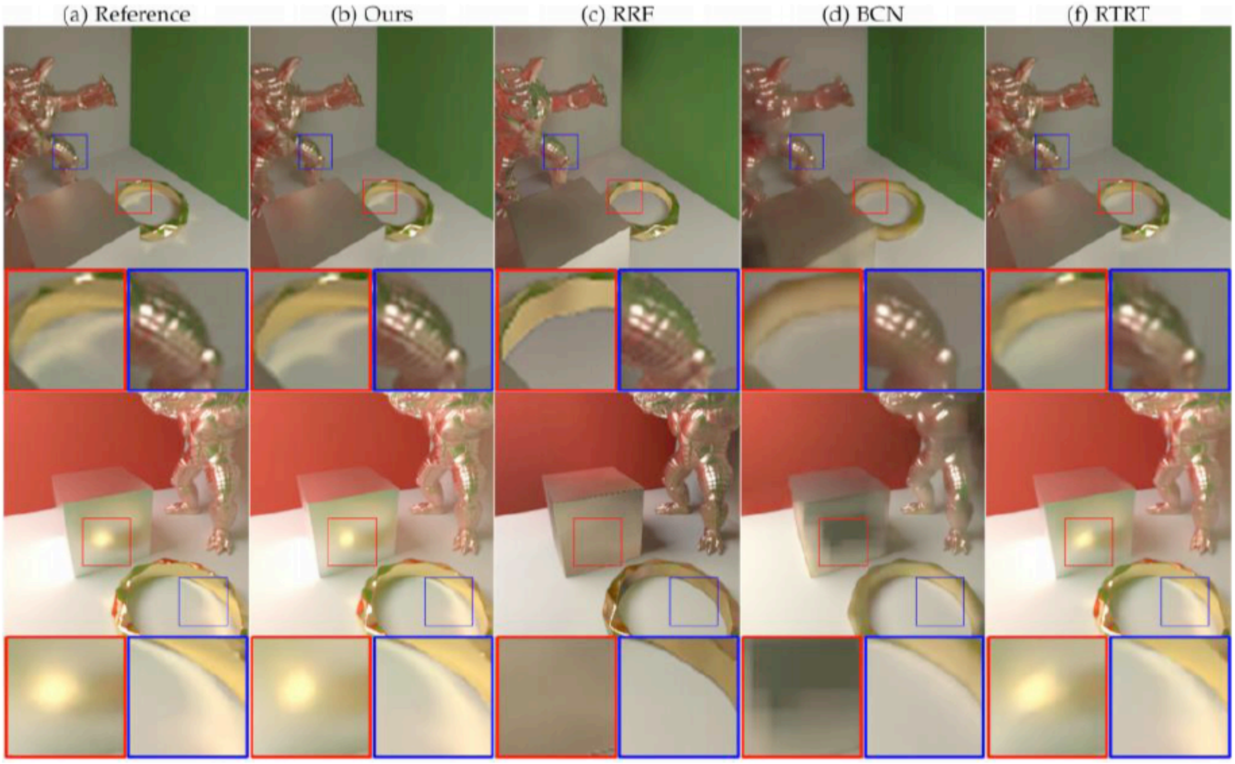

Figure 2: Comparison of Results for Transformer-based Illumination Models.

smoother shading transitions, effectively reducing artifacts and rendering distortions, thereby improving both visual fidelity and perceptual coherence.

2.3. Light Probe–Based Illumination Models

To enable real-time rendering, a class of global illumination approximation techniques based on light probes has been widely studied. Light probes provide a compact representation of scene illumination by encoding lighting information into sparse spatial caches, making them highly suitable for real-time applications. As described in [8], light probes can be represented in various forms, including cube maps, octahedral maps, spherical Gaussians, and spherical harmonics (SH) expansions. Different probe representations encode different levels of illumination detail. For instance, irradiance probes typically use SH basis functions to represent low-frequency lighting efficiently and support dynamic scene updates; however, they are primarily limited to diffuse interreflection. To enable glossy indirect reflection, Křivánek *et al*. introduced cached and interpolated directional incident radiance representations. Building upon this idea, light field probes further extend irradiance probes by incorporating geometric information, thereby reducing light leaking artifacts. Despite their efficiency, traditional light probe methods remain limited in handling complex light transport phenomena. High-frequency effects such as long glossy paths are often ignored or poorly approximated. In addition, their accuracy is highly dependent on probe density, representation capacity, and reconstruction strategies, and they are typically restricted to specific transport types such as diffuse reflection.

Taking spherical harmonics (SH) as an example, illumination is represented using a compact set of coefficients corresponding to orthogonal basis functions on the sphere. These coefficients can be linearly combined to approximate incident light fields. However, SH representations struggle to model high-frequency lighting variations; insufficient coefficient orders lead to ringing artifacts, restricting their applicability mainly to highly rough or diffuse materials. For spherical Gaussians, illumination is modeled as a weighted sum of Gaussian lobes defined on the unit sphere, as shown in Eq. (3):

$$P_{lc} = \sum_{j}^{N} G\left(v_i; u_j, \lambda_j, \mu_{jc}\right) = \sum_{j}^{N} \mu_{jc} e^{\lambda_j (v_i \cdot u_j - 1)} \tag{3}$$

where G $(v_i; u_j, \lambda_j, \mu_{jc})$ denotes an individual spherical Gaussian component. Here, u represents the lobe axis direction, μ denotes amplitude, and λ controls sharpness. When using fixed-direction spherical Gaussians or SH representations, several limitations arise, including the inability to reposition lobes to more optimal directions. Consequently, illumination reconstruction may exhibit aliasing artifacts when light sources move dynamically.

Rodriguez *et al*. [9] proposed a Glossy Probe Reprojection method that precomputes all light transport paths and performs reprojection at runtime, enabling interactive global illumination for opaque objects in static scenes. The method introduces an adaptive probe parameterization strategy to reduce memory consumption, and employs an efficient reflection estimation and online search scheme to gather view-dependent texture samples. Furthermore, a two-stage convolution process is applied to mitigate sharp boundary artifacts in glossy reflection occlusions.

Experimental results demonstrate that the method performs effectively in handling complex glossy light transport paths, achieving high-quality approximations of ground-truth illumination at interactive frame rates. Compared with competing approaches, it provides superior image quality under equivalent runtime constraints. Figure **3** illustrates qualitative results of the Glossy Probe Reprojection method, while Table **3** summarizes its performance comparison.

In [6], the authors proposed a real-time global illumination framework that represents precomputed local light fields using spherical Gaussians. A convolutional neural network (CNN) is trained to predict the parameters of the spherical Gaussian representation, enabling efficient high-quality rendering with relatively low memory and computational overhead in static scenes. In addition, the method introduces an illumination-weighted environment visibility algorithm, which enables high-fidelity distant environment reflections while accounting for both material properties and visibility constraints.

The work in [8] further combines precomputed light probes with deep learning for real-time global illumination rendering. The proposed framework reconstructs high-quality global illumination from low-quality probe representations through a gradient-based reflection search algorithm and a neural image reconstruction network. During probe querying, the gradient-based reflection search explicitly considers parallax variations to avoid incorrect

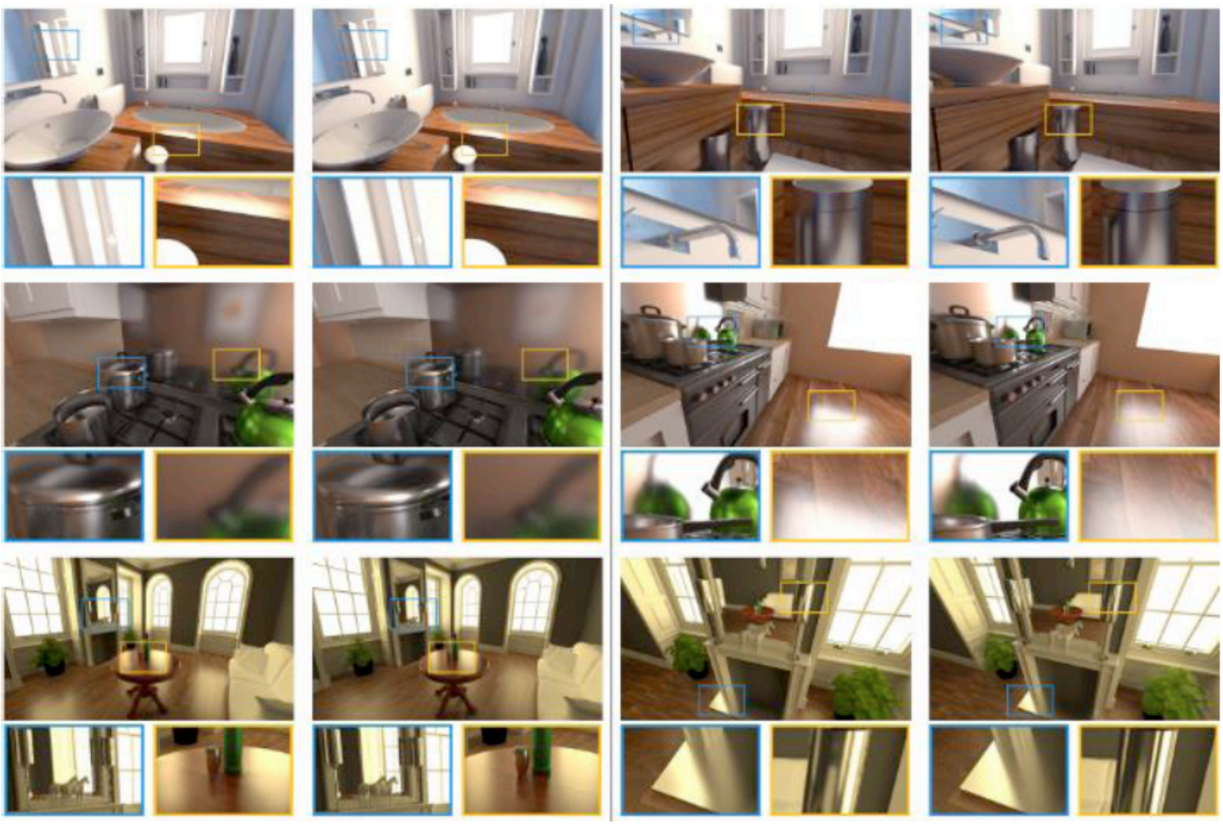

Figure 3: Visual Results of the Glossy Probe Reprojection Method.

Table 3: Performance Comparison of the Glossy Probe Reprojection Method

Method	Frame Rate (fps)	Memory Usage (GB)	Rendering Quality (DSSIM)
Traditional Light Probe	30	2.5	0.093
Glossy Probe	60	5.6	0.027

reflections in glossy regions. The optimal reflection position is iteratively estimated, and its validity is defined by Eq. (4):

$$C(x_k) = \frac{px_k + zx_k}{\|px_k + zx_k\|} \cdot N(x_k) \tag{4}$$

where $N(x_k)$ denotes the surface normal at position x_k, and the value of $C(x_k)$represents the confidence score of each candidate point x_k. The iterative search terminates once either the predefined threshold Tmax or the maximum iteration number Nmax is reached.

The neural image reconstruction module is built upon a dedicated convolutional neural network architecture together with multiple G-buffers to recover high-quality images from low-quality inputs. A G-buffer modulation module is introduced to better fuse multimodal features. The overall architecture adopts a typical encoder–decoder structure consisting of multiple convolutional and deconvolutional layers for hierarchical feature extraction and fusion. During training, the network is optimized using a hybrid loss function composed of pixel-wise (L_1) loss, perceptual loss (LPIPS), and temporal consistency loss:

$$\xi = \lambda_{L1}\xi_{L1} + \lambda_{LPIPS}\xi_{LPIPS} + \lambda_T\xi_T \tag{5}$$

where ξ_{L1} denotes the pixel-wise L1distance, ξ_{LPIPS}represents the perceptual loss, and $\xi_T$denotes the temporal loss. The weighting coefficients λ_{L1}, λ_{LPIPS}, and $\lambda_T$are used to balance the contributions of different loss terms. Through these strategies, the method effectively handles complex glossy reflection paths in real-time rendering scenarios and achieves interactive performance exceeding 30 frames per second at 1080P resolution.

The study in [12] further explores more efficient processing strategies for light probes using deep learning. By introducing an improved neural network architecture, the method can extract illumination features more effectively from probe data, thereby improving rendering quality. Experimental results demonstrate superior capability in preserving illumination details across different scenes, particularly in environments containing complex geometries and material properties. For instance, in indoor scenes with diverse materials and intricate light interactions, the method produces more accurate reflections and shadow effects compared with conventional approaches, resulting in higher visual realism.

Recent advances in inverse rendering have demonstrated that physically based material reconstruction requires accurate modeling of indirect illumination effects. The NeGIL framework proposed in [13] separates direct environmental lighting and indirect illumination into surface light fields, enabling global illumination simulation without expensive path tracing operations. Its key innovation lies in a Monte Carlo differentiable rendering framework that models indirect illumination at occluded points through a shared multilayer perceptron (MLP). This design enables end-to-end physically based material reconstruction in static scenes, as illustrated in Figure **4**. Experimental results show that NeGIL improves novel-view synthesis PSNR by approximately 2–4 dB over conventional approaches in scenes containing complex glossy reflections.

2.4. Point-Light Clustering and Neural Residual Modeling

In the field of point-light-based global illumination, the method proposed in [10] based on Well-Separated Pair Decomposition (WSPD) represents one of the state-of-the-art clustering approaches. By constructing spatial clusters that satisfy the geometric separation condition r(C)<ϵ g d(p,C) the method achieves an

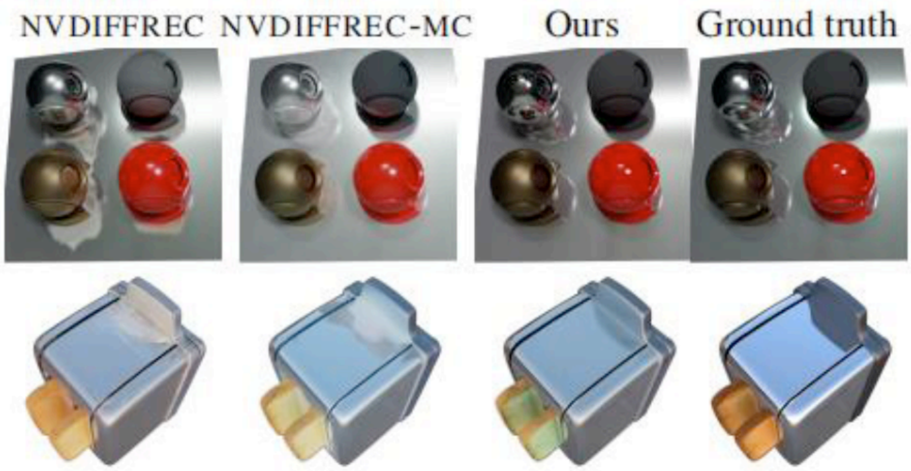

Figure 4: Visualization of the NeGIL-estimated plausible material reflection regions.

effective balance between rendering accuracy and computational efficiency. Its core innovation lies in reformulating complex global illumination computations as a geometric separability problem and efficiently organizing virtual point lights (VPLs) using a compressed octree structure.

As illustrated in Figure **5**, this adaptive spatial partitioning strategy dynamically adjusts clustering precision according to the distance between the shading point and the corresponding light cluster. Consequently, the light sources within each cluster maintain geometric coherence from the perspective of the shading point, significantly reducing the overall computational complexity of global illumination evaluation.

The effectiveness of the WSPD-based approach has been validated through multiple benchmark experiments. As summarized in Table **4**, when rendering complex scenes containing up to 500,000 VPLs, the method achieves approximately 3–5× speedup compared with the conventional Lightcuts algorithm while maintaining comparable visual quality. The superior performance primarily originates from its hierarchical dual-stage optimization framework. During preprocessing, a well-separated pair decomposition structure is constructed with a computational complexity of $O(n\log n+n\epsilon^{-3})$while the rendering stage performs near-constant-time illumination evaluation through efficient nearest-neighbor queries.

This design makes the WSPD framework particularly suitable for large-scale dynamic scenes. Furthermore, several auxiliary techniques, including normal-direction subclustering and boundary visibility caching, effectively address challenges associated with anisotropic materials and complex shadow boundaries.

The theoretical foundation of the WSPD framework is equally noteworthy, as it provides rigorous guarantees for error control. For Lambertian materials, the relative error upper bound of the WSPD approximation is theoretically given by $O(\epsilon)\sum Is/d2$ where the bound remains independent of overall scene complexity, thereby ensuring stable rendering behavior across scenes of different scales. These properties make the WSPD framework highly promising for applications in game engines, cinematic visual effects, and architectural visualization. At the same time, they also highlight several important directions for future research, particularly in extending the framework to highly specular materials and dynamically changing environments.

From the perspective of inverse rendering, Hadadan *et al*. [16] proposed a neural radiance prior framework that replaces the expensive multi-bounce

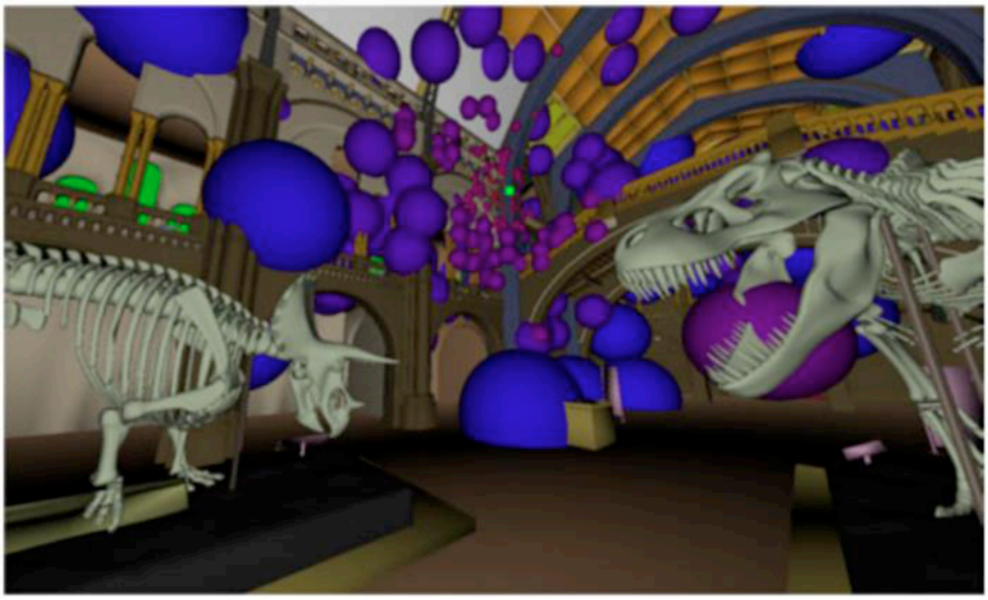

Figure 5: Spatial Cluster Structure Visualization.

Table 4: Performance Comparison of Point Light Clustering Methods (500K VPLs)

Method	Preprocessing Time	Rendering Time	Memory Usage	Applicable Scenarios
Lightcuts	0.4s	515s	0.6GB	Medium-scale Static Scenes
WSPD (ε=0.25)	52s	63s	1.1GB	Large-scale Dynamic Scenes

integration process in conventional differentiable path tracing with a neural radiance residual term. The network is trained in a self-supervised manner by minimizing the residual norm of the rendering equation. During inference, only one- or two-bounce sampling is required to satisfy energy conservation constraints, while scene parameters such as BRDFs and illumination are optimized through standard automatic differentiation. This formulation completely avoids the excessive memory overhead caused by long-path gradient propagation in traditional differentiable rendering pipelines. Experimental results demonstrate that the proposed method can recover physically plausible material and illumination parameters comparable to path replay backpropagation (PRB) within approximately 260 minutes, while maintaining memory consumption independent of path length. This work significantly improves the practicality of inverse rendering under complex global illumination conditions.

Meanwhile, Mo *et al*. [17] addressed the loss of cross-frequency specular details in real-time forward rendering and proposed a dual-frequency feature fusion strategy. The framework maintains two parallel feature streams corresponding to low-frequency diffuse components and high-frequency specular components, respectively. After obtaining specular intersection points through a single ray-tracing pass, a kernel-prediction CNN performs multi-resolution low-pass filtering on the high-frequency branch, followed by zero-enhanced linear interpolation with the low-frequency features. This design enables continuous reflection reconstruction across the full roughness range from 0 to 1. In addition, the method adopts a two-stage training strategy that first stabilizes low-frequency convergence before progressively introducing high-frequency detail learning. This effectively suppresses artifacts commonly observed in traditional approaches, such as overly sharp specular contours and light leaking effects. Experimental evaluations on scenes such as Hall and Coffee demonstrate that the method achieves 36–41 dB PSNR within 20–26 ms rendering time, consistently outperforming OIDN and AE baselines. More importantly, it represents one of the first neural rendering frameworks capable of faithfully reconstructing dynamic multi-frequency reflections in real time.

Taken together, these two research directions—one focusing on inverse rendering and the other on forward rendering—collectively reveal the significant potential of replacing explicit long-path light transport simulation with neural residual modeling or learnable feature fusion mechanisms. The former implicitly encodes physical constraints into neural radiance priors, thereby alleviating the substantial computational burden of path integration in inverse rendering. The latter maintains both low-frequency energy consistency and high-frequency reflection sharpness through learnable frequency decomposition and fusion within forward rendering pipelines. Complementary to each other, these approaches establish a unified and efficient neural paradigm for global illumination computation in both real-time and offline rendering scenarios.

2.5. Illumination Reconstruction for Power and Industrial Environments

Recent advances in illumination estimation and neural rendering have significantly improved the ability to reconstruct complex lighting conditions in outdoor environments. However, the majority of existing studies focus primarily on natural landscapes, urban streets, autonomous driving datasets, and synthetic benchmark scenes. Comparatively, illumination reconstruction in power and industrial environments has received far less attention despite its increasing importance in intelligent infrastructure management, digital-twin construction, and industrial visual perception. Power-system environments present several unique challenges for illumination reconstruction. Unlike conventional outdoor scenes, substations, transmission corridors, and distribution facilities contain large numbers of metallic structures, conductors, transformers, insulators, and support frames. These components generate highly complex geometric configurations characterized by dense occlusions, severe self-shadowing, and spatially varying reflectance properties. In addition, metallic equipment often exhibits strong specular reflections that are difficult to model using assumptions commonly adopted in natural-scene illumination estimation frameworks. As a result, accurately recovering environmental illumination in power scenarios requires not only robust physical lighting models but also effective integration of scene semantics and structural information[58].

One of the most important application areas is intelligent power inspection. Modern power utilities increasingly rely on unmanned aerial vehicles (UAVs), mobile inspection robots, and vision-based monitoring systems to perform routine inspections of transmission lines and substation equipment. The accuracy of object detection, defect recognition, and condition assessment is often influenced by environmental illumination variations. Reliable illumination reconstruction can provide physically consistent lighting information for image enhancement, scene understanding, and visual perception algorithms, thereby improving the robustness of intelligent inspection systems under diverse weather and lighting conditions.

Another emerging application is digital-twin power systems. Digital twins aim to establish synchronized virtual representations of physical power infrastructures by integrating sensing, simulation, and visualization technologies. To maintain consistency between physical and virtual environments, realistic lighting reconstruction plays a critical role in scene rendering, visual perception, and operator interaction. Illumination-consistent digital twins enable more accurate visualization of equipment status, facilitate remote inspection and maintenance planning, and improve the realism of virtual operation environments. As power grids evolve toward intelligent and autonomous operation, illumination-aware digital twins are expected to become an important component of next-generation energy management systems.

Illumination reconstruction is also closely related to virtual power training and safety education. Immersive training platforms based on virtual reality (VR), augmented reality (AR), and mixed reality (MR) technologies are increasingly adopted for personnel training in substations, transmission systems, and emergency-response scenarios. The realism of these virtual environments strongly depends on the accuracy of environmental lighting simulation. Physically consistent illumination reconstruction can improve depth perception, material appearance, shadow realism, and overall scene fidelity, thereby enhancing training effectiveness and user immersion.

Although recent advances in neural rendering, inverse rendering, and data-driven illumination estimation have demonstrated remarkable performance in generic environments, their applicability to power-system scenarios remains limited. Existing methods typically lack mechanisms for incorporating power-specific semantic priors, equipment geometry constraints, and industrial scene characteristics into the illumination inference process. Consequently, there remains a significant gap between current illumination estimation technologies and the practical requirements of intelligent power applications. Addressing this gap represents an important research direction for future illumination reconstruction systems and provides the primary motivation for the power-oriented SPC-PMC framework discussed in the following section.

3. CORE CHALLENGES AND TECHNICAL REQUIREMENTS FOR ILLUMINATION RECONSTRUCTION IN VIRTUAL POWER SCENARIOS

With the rapid development of next-generation power systems, the operational complexity and safety risks associated with power-related tasks have continued to increase, placing higher demands on practical skills training and emergency response capabilities for power industry personnel. As a key approach to improving operational safety, intelligent power-operation training has evolved from traditional theoretical instruction and two-dimensional multimedia demonstrations toward highly immersive virtual power environments based on virtual reality (VR), augmented reality (AR), and digital twin technologies [1–5]. By constructing three-dimensional digital twins of substations, transmission towers, and other operational infrastructures, high-fidelity simulation systems provide trainees with risk-free, repeatable, and highly interactive training environments, offering substantial practical value in reducing training costs and mitigating operational hazards.

Within this evolution, the realism and immersion of digital twin environments directly determine the effectiveness of practical training. However, existing virtual power scene construction methods primarily focus on high-precision geometric reconstruction and static texture refinement, while largely overlooking the dynamic and time-varying evolution of real-world illumination conditions. As illumination fundamentally governs human perception of spatial geometry, material appearance, and environmental atmosphere [5], the quality of illumination reconstruction directly affects the realism and situational fidelity of simulation systems. Most existing power-operation simulators still rely on predefined light sources or idealized static skybox configurations embedded in commercial rendering engines. Such homogeneous virtual lighting setups cannot accurately reproduce real environmental conditions corresponding to actual operational scenarios, including weather variations and temporal illumination changes. More critically, they fail to reconstruct fine-grained visual phenomena under extreme working conditions, such as intense backlighting glare, geometric shadow occlusions caused by large steel truss structures, and complex shadow interactions among transmission lines.

In practical power-operation scenarios, strong backlighting may severely impair the visual judgment of field operators, while shadows cast by specific equipment structures may interfere with infrared thermography and inspection viewpoint selection. Consequently, the overly simplified and static treatment of physical illumination in existing simulation systems creates a substantial visual gap between virtual training environments and real operational conditions, significantly limiting the practical effectiveness of intelligent power-operation training. From a technical perspective, applying illumination estimation techniques to virtual power scenarios involves three major challenges.

First, fine-grained semantic segmentation of high-spatial-frequency structures remains a major bottleneck for accurate illumination prediction. Power environments such as substations and transmission towers contain dense metallic trusses, highly intertwined transmission lines, and small-scale insulator structures, all of which exhibit extremely high spatial frequencies and severe self-occlusion characteristics. Traditional shallow machine learning–based segmentation methods lack sufficient high-level semantic abstraction capability and therefore frequently suffer from boundary fragmentation, pixel adhesion, and large-area segmentation holes when processing such highly textured and multi-scale industrial structures. As a result, they fail to generate accurate semantic masks for illumination-aware geometric reasoning [5, 18].

Second, the failure of conventional multi-cue fusion strategies often leads to severe distortions in physical illumination inversion. Due to the highly cluttered background structures in power environments, classical Bayesian fusion frameworks tend to amplify estimation errors under conditions involving dense industrial infrastructures. On the one hand, blurred sky boundaries and object occlusions distort the luminance distribution of sky pixels, resulting in directional bias during physical sky model fitting [57]. On the other hand, segmentation errors in vertical surface masks directly introduce geometric inconsistencies into surface-normal voting mechanisms, thereby degrading illumination estimation accuracy.

Third, mapping illumination estimation results from two-dimensional image space into three-dimensional scene space remains a critical challenge for intuitive visualization and digital twin reconstruction. Illumination cues extracted from historical monitoring images inherently exist in two-dimensional image space. Accurately transforming estimated solar positions and multi-region illumination directions from 2D sky dome representations into physically consistent 3D hemispherical illumination models, while simultaneously performing quantitative projection and shadow analysis in the absence of detailed geometric models, remains a major obstacle in establishing a complete “image-to-scene” illumination reconstruction pipeline.

3.1. Deep Semantic Segmentation and Illumination Estimation for Complex Industrial Environments

In industrial-scale virtual simulation systems, the accuracy of illumination estimation heavily depends on precise extraction of scene geometry and semantic information. In particular, for power-operation environments such as substations and transmission towers, which exhibit extremely high spatial frequencies and severe self-occlusion, dense steel trusses and intertwined transmission lines frequently cause severe segmentation errors and region discontinuities in conventional shallow-learning-based methods. These inaccuracies further introduce directional bias and geometric distortion into physically based illumination inversion pipelines [4, 18]. To address this limitation, recent deep neural network architectures such as YOLOv11 employ multi-scale feature extraction and high-level semantic abstraction through deep convolutional structures, enabling end-to-end fine-grained pixel-level segmentation for complex outdoor industrial scenes. These models can accurately extract key geometric-semantic masks, including sky regions, ground regions, and vertical surfaces with multiple orientations [4]. This deep-learning-driven spatial semantic mechanism provides a robust front-end foundation for extending classical probabilistic multi-cue (PMC) illumination estimation frameworks to highly complex industrial digital twin environments. More importantly, it substantially alleviates the self-occlusion and boundary adhesion problems caused by metallic truss structures and densely intersecting transmission lines.

Zhao *et al*. [23] proposed a hybrid global illumination and deep-learning rendering pipeline for immersive VR environments. Building upon the conventional ray-tracing formulation:

$$L_o\left(p,w\right)=L_e\left(p,w\right)+\int_{\Omega}f_r\left(p,w,w'\right)\left(L_i\left(p,w'\right)\right)-L_a\left(p\right)dw \tag{6}$$

The framework introduces a hybrid octree–KD-tree acceleration structure for efficient visibility queries. When dynamic object or light-source displacement occurs, only the affected leaf nodes are recomputed, while an incremental update strategy is employed to avoid full-frame re-rendering.

The neural rendering module adopts a ResNeXt-UNet architecture as the generator network and is optimized using the adversarial objective:

$$\min_{G}\max_{D}\left(D,G\right)=E_{x\sim p_{data}(x)}\left[\left(\log D\left(x\right)\right)\right]+E_{z\sim p_g(z)}\left[\left(\log\left(1-D\left(G(z)\right)\right)\right)\right] \tag{7}$$

Integrated within Unity and Unreal Engine pipelines, the framework performs a single forward inference pass with approximately 4 ms latency. Experimental results demonstrate that the method increases rendering performance from 40 fps to 55 fps in indoor daytime scenes and from 28 fps to 40 fps in outdoor nighttime environments. In addition, subjective user evaluation scores exceeded 9.0 out of 10, indicating strong perceptual realism and immersion quality.

At the application level within the power industry, recent advances in VR/AR and AI-integrated intelligent training technologies are fundamentally reshaping skill development paradigms for power-system operators. Existing studies have introduced large-language-model-driven interactive virtual instructors into three-dimensionally reconstructed substation environments, enabling multimodal training interactions through speech and gesture guidance. Experimental results indicate that such approaches significantly improve trainee engagement and knowledge acquisition while maintaining relatively low cognitive load [33]. To address the limited interactivity of conventional grid communication and maintenance training systems, Zhao *et al*. [34] combined artificial intelligence and augmented reality to develop the AA-COMT framework, which integrates three-dimensional modeling, dynamic scene understanding, and adaptive human–computer interaction into a unified intelligent training environment. From the perspective of personalized education, Lin *et al*. [35] integrated low-fidelity digital-twin VR experiential learning, GPT-4-driven zero-shot affective analysis, and retrieval-augmented personalized instructional content generation into a unified training framework. In addition, finite-state automata were employed to dynamically adjust task difficulty according to trainee performance accuracy. Experimental evaluations demonstrated that the proposed method simultaneously improved learning accuracy while reducing overall training duration.

Yoo *et al*. [36, 37] systematically investigated AI-integrated AR systems as intelligent industrial companions for workers. They proposed a general conceptual framework incorporating activity understanding, user modeling, and adaptive reasoning mechanisms. Furthermore, by synchronously collecting multimodal data through AR head-mounted displays and wearable sensors, the studies revealed strong correlations among visual scanning efficiency, physiological stress indicators, and professional skill proficiency. Phadke *et al*. [38] further explored the influence of sequential AR/VR exposure on industrial skill acquisition. Their results showed that VR-first training significantly reduces cognitive load during subsequent AR-assisted operational guidance while improving task confidence and execution efficiency. Consequently, a VR-to-AR training sequence was recommended to jointly optimize spatial reasoning capability and procedural execution performance. At the organizational management level, Witkowski [39] constructed a personalized employee training framework integrating VR, AR, and AI technologies based on a questionnaire survey involving 106 industrial enterprises together with an extensive literature review. The study identified simulation of safety-critical tasks, enhanced immersive engagement, and AI-driven personalized content adaptation as the primary benefit dimensions of next-generation industrial training systems.

3.2 Dynamic Global Illumination Techniques

Dynamic Global Illumination (DGI) has become a major research direction in real-time rendering. Conventional global illumination techniques suffer from substantial efficiency limitations in dynamic scenes, since illumination must typically be recomputed whenever scene geometry or lighting conditions change. To address this issue, numerous dynamic global illumination frameworks have been proposed.

Majercik *et al*. [12] introduced Dynamic Diffuse Global Illumination (DDGI), which approximates indirect illumination through dynamically updated irradiance fields. DDGI employs hardware-accelerated ray tracing to update irradiance probes in real time, enabling support for dynamic lighting and scene changes. However, the computational overhead remains significant for large-scale environments. To further improve efficiency, Majercik *et al*. proposed an extended DDGI framework by combining sparse world-space probes with screen-space resampling based on ReSTIR. Through spatiotemporal sample reuse, ReSTIR enables efficient reuse of a large number of illumination samples within constant-time budgets, thereby substantially improving sampling efficiency in dynamic GI computation. The framework reformulates global illumination as a direct-lighting sampling problem by treating DDGI probes as secondary light sources, resulting in highly efficient indirect illumination evaluation. Experimental results demonstrate that the method significantly improves computational efficiency while preserving high rendering quality.

Several years later, Majercik *et al*. [22] further unified ReSTIR direct illumination sampling and DDGI-based indirect illumination caching into a single resampling framework for dynamic diffuse global illumination. The method maintains a per-pixel resampling reservoir in screen space, where candidate samples are generated from both emissive geometry sampling and DDGI secondary radiance contributions. Variance is controlled through bounded geometric terms, while multiple importance sampling (MIS) is used to combine different illumination sources. This unified formulation enables low-noise full global illumination rendering within approximately 31 ms for 1080p scenes.

Compared with conventional serial pipelines that separately process direct ReSTIR sampling and indirect DDGI evaluation, the proposed approach significantly reduces visibility artifacts caused by insufficient probe resolution while preserving approximate first-bounce glossy reflections. As a result, the framework provides a practical “high-quality global illumination with minimal configuration” solution for modern real-time game engines.

Dynamic illumination simulation in natural environments introduces even greater challenges for real-time rendering systems. The work in [23] employed the lattice Boltzmann method (LBM) to simulate wind-field interactions with grass blades and solved multiple-scattering transport using a plane-parallel radiative transfer model. In this framework, vegetation is modeled as a continuous distribution of microscopic scattering elements. Radiative transport solutions corresponding to different bending angles ((¥theta_b)) are precomputed offline, while runtime rendering is achieved through interpolation over a five-dimensional shading table: （S(z,α,θ_b,θ_n,φ_n)） Although preprocessing for a single frame requires approximately 15 minutes, the method maintains interactive rendering performance under (256 ¥times 256) directional discretization, achieving approximately 10.75 seconds per frame. This work provides a physically plausible solution for dynamic global illumination in large-scale natural environments.

Lü *et al*. [24] proposed a real-time approximate global illumination framework for dynamic scenes based on a hybrid “accurate direct illumination + simplified indirect illumination” strategy. Direct illumination is first computed using convolutional soft shadow maps, producing color, normal, and position textures. Subsequently, several hundred “pixel lights” are sampled from these textures using a Halton sequence. Hard shadow maps are then rendered on simplified geometry to approximate single-bounce indirect illumination, which is finally combined with direct lighting contributions through weighted blending. Benefiting from the ability of modern GPUs to perform thousands of low-resolution rasterization operations per second, the method achieves real-time rendering performance of 25–35 fps on NVIDIA GTX 280 hardware for dynamic scenes containing indirect shadows. This lightweight framework provided an efficient global illumination solution for early-generation graphics hardware and established the foundation for subsequent “simplification–resampling” rendering paradigms.

Environment mapping is another widely adopted technique for representing and evaluating distant illumination in both offline and real-time rendering applications. Traditional environment mapping approaches generally rely on high-resolution HDR images to encode scene illumination; however, such representations often suffer from high computational and memory costs during real-time rendering. To improve efficiency, numerous compression and sampling techniques for environment maps have been proposed. For example, the method in [6] introduced a Rotation-Equivariant Conditional Spherical Neural Field framework for environment lighting representation. By learning priors of natural illumination distributions, the method enables efficient environment-map representation and sampling.

More recently, Rodriguez-Pardo *et al*. [20] proposed NEnv, a neural environment-mapping framework that combines normalizing flows with implicit neural representations. The normalizing-flow module maps uniformly distributed samples onto the illumination distribution of the environment map while simultaneously providing probability density estimation for Monte Carlo importance sampling. Meanwhile, the implicit neural representation compresses environment lighting into a compact differentiable function, significantly reducing memory consumption. Experimental results demonstrate that NEnv outperforms conventional methods in both sampling efficiency and environment-map reconstruction quality, enabling high-fidelity global illumination rendering.

The field-of-view limitation of screen-space global illumination (SSGI) can be alleviated through multi-view input strategies. The SAwareSSGI framework proposed in [25] employs a six-camera capture setup to obtain 360° feature buffers and utilizes a GAN-based architecture to generate surround-aware indirect illumination maps, as illustrated in Figure **6**. The generator adopts a U-Net architecture augmented with CoordConv layers to preserve spatial awareness during illumination synthesis. On an NVIDIA GTX 1660 Ti GPU, the method generates 1080p illumination maps within only 3.32 ms. Compared with conventional SSGI approaches, SAwareSSGI achieves substantial improvements in both SSIM (0.908) and PSNR (30.31 dB), while maintaining illumination continuity even when light sources move outside the visible screen region.

3.4. Research Gap and Motivation

The literature reviewed in the previous sections demonstrates that significant progress has been achieved in illumination modeling, global illumination rendering, inverse rendering, and data-driven illumination estimation. Classical physically based

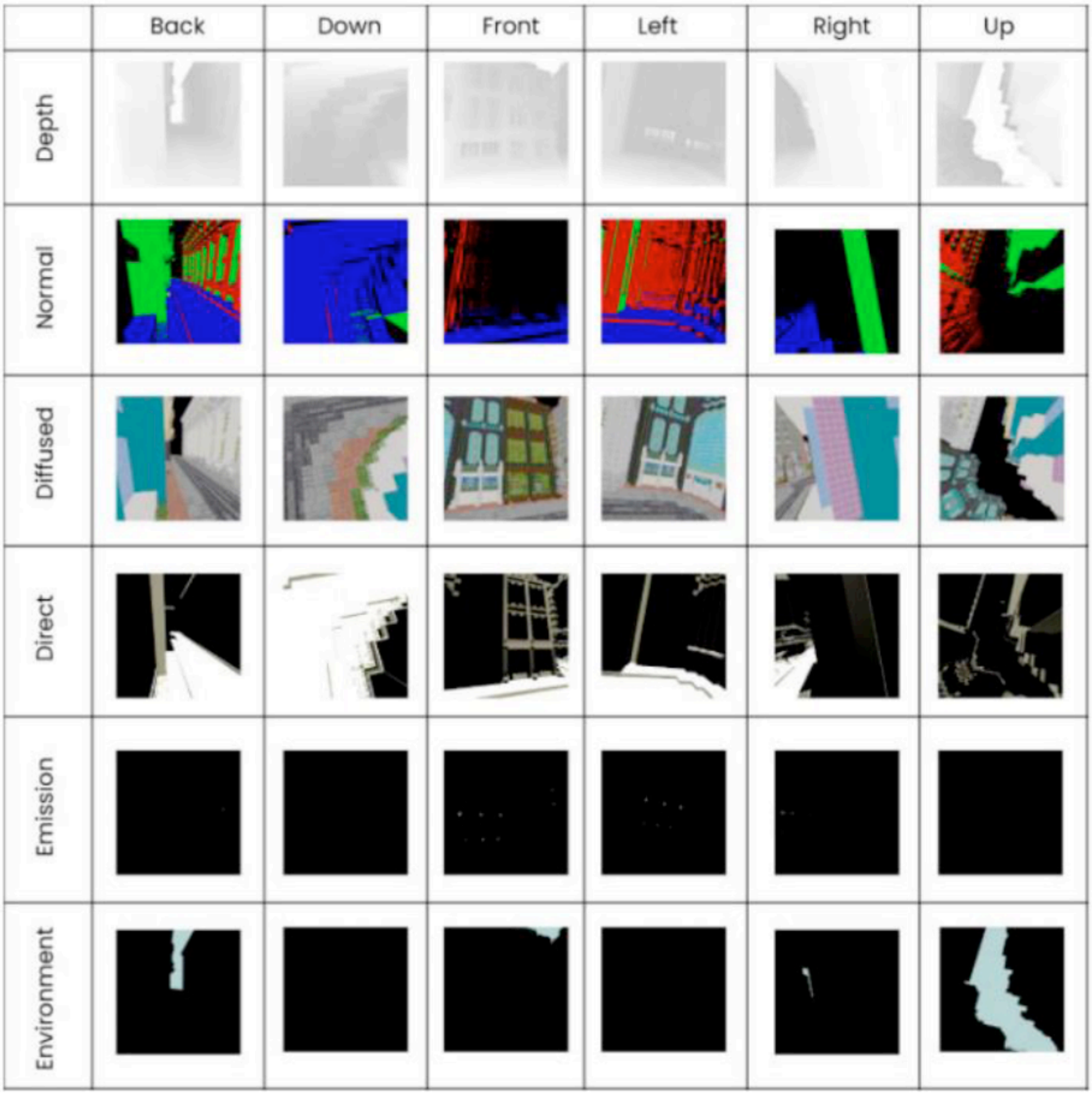

Figure 6: 360° Feature Buffer Map Captured by a Six-Camera Cluster.

approaches, such as the Perez sky model, the Hošek–Wilkie model, and the Lalonde–Matthews framework, provide interpretable representations of outdoor illumination and have shown reliable performance in natural environments. More recently, deep learning and neural rendering techniques have further improved the ability to infer complex lighting distributions from images, enabling efficient illumination reconstruction for a wide range of computer vision and graphics applications. However, most existing methods are developed and evaluated primarily on generic outdoor scenes, urban environments, or synthetic benchmarks, leaving their applicability to power-system scenarios largely unexplored.

Power-system environments exhibit several distinctive characteristics that fundamentally challenge conventional illumination estimation frameworks. Transmission towers, substation equipment, conductors, and metallic support structures form highly cluttered geometric configurations with dense occlusions and complex shadow interactions. In addition, extensive metallic surfaces introduce strong specular reflections and spatially varying radiometric responses that violate the assumptions commonly adopted in natural-scene illumination models. These factors often lead to inaccurate semantic segmentation, unreliable sky-region extraction, and unstable cue aggregation, particularly for probabilistic multi-cue approaches that rely on the accurate identification of sky, ground, and vertical surface regions. Consequently, the performance of existing methods may degrade significantly when applied to power-system imagery, limiting their effectiveness in digital-twin construction, intelligent inspection, immersive training, and other perception-driven applications.

These observations reveal a critical research gap between general-purpose illumination estimation methods and the practical requirements of virtual power scenes. Existing approaches typically lack explicit mechanisms for incorporating power-system structural priors, equipment-specific semantic information, and industrial scene constraints into the illumination inference process. To address this limitation, this paper proposes the Semantic–Pixel Coupled Probabilistic Multi-Cue (SPC-PMC) framework. Building upon the physical interpretability of traditional probabilistic multi-cue illumination estimation, SPC-PMC introduces deep semantic priors and pixel-level structural constraints to improve the robustness of illumination inference in highly occluded and geometrically complex power-system environments. By coupling semantic understanding, scene geometry, and physical lighting models, the framework aims to provide more reliable illumination reconstruction for intelligent energy systems, digital-twin power grids, and next-generation virtual power applications.

4. SEMANTIC–PIXEL COUPLED PROBABILISTIC MULTI-CUE ILLUMINATION ESTIMATION MODEL

This study further investigates the feasibility of constructing a Semantic–Pixel Coupled Probabilistic Multi-Cue (SPC-PMC) illumination estimation framework for virtual power-system environments. The theoretical foundation of the proposed model originates from the Probabilistic Multi-Cue (PMC) outdoor illumination estimation framework introduced by Lalonde *et al*. [4], which has become a cornerstone paradigm for single-image illumination inference owing to its strong physical interpretability and probabilistic representation capability. The central principle of PMC lies in its region-cue-based formulation, which decomposes the high-dimensional and highly ill-posed panoramic illumination estimation problem into multiple independent statistical inference tasks associated with specific semantic regions. This formulation substantially reduces the complexity of illumination inference from a single-view image.

For sky-region cues, the framework relies on a physically grounded sky illumination model [57], where the statistical relationship between sky luminance distribution and solar position is exploited to estimate solar zenith and azimuth angles through nonlinear fitting in a discretized parameter space. The framework further quantifies solar visibility and occlusion conditions based on the estimated sky model. For vertical-surface cues, the underlying mechanism is based on a statistical voting strategy. Specifically, highly illuminated vertical surfaces are assumed to indicate approximate solar incidence directions aligned with their surface normals. By aggregating luminance observations from multiple wall regions with different orientations, the framework constructs a probabilistic prior distribution for solar azimuth estimation. Unlike deterministic illumination estimation approaches, this probabilistic illumination representation models lighting conditions as continuous probability distributions associated with solar position and visibility rather than single deterministic solutions. Consequently, the framework naturally captures estimation uncertainty arising from incomplete information in monocular imagery.

However, the robustness of PMC fundamentally depends on the accuracy of region-cue extraction. In power-system environments such as substations and transmission towers, the presence of densely distributed metallic trusses, intertwined transmission lines, and severe self-occlusion introduces extremely high spatial-frequency structures. Under such conditions, conventional shallow-learning-based segmentation methods [18] frequently suffer from severe region fragmentation, boundary ambiguity, and semantic void artifacts, thereby introducing directional bias into physical sky-model fitting and geometric distortion into vertical-surface normal voting. To address these limitations, the proposed SPC-PMC framework preserves the physically interpretable probabilistic cue-fusion mechanism of classical PMC while incorporating the state-of-the-art YOLOv11 deep paired neural architecture to establish a fine-grained spatial-semantic-driven segmentation mechanism. By replacing conventional shallow segmentation pipelines with robust semantic masks generated through end-to-end deep learning, the proposed framework provides a reliable technical foundation for transferring classical region-cue illumination estimation theory into highly complex virtual power-system digital-twin environments.

4.1. Fast Global Illumination Based on Microfacet Models

In global illumination rendering, accurate simulation of specular reflection through microfacet models plays a critical role in enhancing material realism. The Filterable Discrete Stochastic Microfacets framework proposed in [46] focuses on simplifying the computation of specular reflection by precomputing scattering probability distributions over stochastic microfacet structures, thereby significantly improving the efficiency of specular global illumination rendering. The core innovation of this method lies in its reformulation of the bidirectional reflectance distribution function (BRDF). Conventional BRDF evaluation requires continuous simulation of highly complex light transport interactions, whereas the proposed framework decomposes the BRDF into a filterable discrete representation:

$$BRDF(\omega_i,\omega_o)=\sum_{N}^{k=1} w_k\cdot f_k(\omega_i,\omega_o) \tag{8}$$

Here, ω_i and ω_o denote the incident and outgoing light directions, respectively; w_k represents the weighting coefficient associated with basis function f_k; and the basis functions are predefined scattering primitives designed to conform to microfacet light-transport behavior. Through this discrete reformulation, specular reflection evaluation is transformed from continuous high-dimensional light-transport integration into a weighted combination of discrete basis functions.

Within the real-time rendering pipeline, spatial filtering techniques are applied to adaptively process the basis functions according to the pixel footprint size. Specifically, for small pixel footprints, the contribution of individual basis functions is evaluated with high precision, whereas for large footprints, multiple basis functions are filtered and blended to efficiently generate physically consistent specular responses. In automotive material rendering scenarios, such as metallic paint and pearlescent coatings, the proposed framework demonstrates substantial improvements over conventional approaches. While preserving fine-grained specular characteristics—including anisotropic reflection behavior and gloss variation—the method achieves a PSNR improvement of approximately 2.3 dB and significantly reduces rendering noise. Consequently, highly reflective materials such as vehicle paint and chrome-plated components exhibit more realistic specular appearance in real-time environments, providing an efficient solution for high-quality global illumination rendering in applications including games and virtual simulation systems.

The two-level radiance caching framework proposed by Tatzgern *et al*. [53], which is based on surface-space caching, represents an important advancement in real-time global illumination. By directly storing radiance information within texture-space surface caches, the framework enables efficient processing of highly complex illumination environments. Compared with conventional approaches, this design avoids the inherent limitations of screen-space and world-space caching methods, particularly mitigating reprojection artifacts and light-leakage problems commonly encountered in dynamic scenes. The core innovation of the framework lies in its hierarchical dual-cache architecture, where the primary cache stores incident radiance while the secondary cache maintains outgoing reflected radiance. The primary cache captures radiative energy arriving directly from illumination sources onto surfaces, whereas the secondary cache records radiance after surface reflection events. This decoupled design enables efficient handling of multi-bounce illumination and glossy reflection effects.

Mathematically, the framework computes the final pixel irradiance through a carefully designed radiative transport formulation incorporating accurate visibility evaluation and radiance accumulation mechanisms. Experimental results demonstrated that the method achieves approximately 5–10% performance improvement at (512 ¥times 512) resolution while preserving high visual fidelity. The Stochastic Substitute Tree (SST) framework proposed by the Tatzgern research group [53] further introduces an innovative probabilistic sampling strategy for real-time global illumination. The method constructs a hierarchical probabilistic representation describing clustered virtual point lights (VPLs), thereby enabling highly efficient illumination sampling and evaluation. The key insight of SST is to reformulate complex global illumination computation as an optimization problem over substitute-light distributions. By hierarchically organizing probabilistic illumination representatives, the framework significantly reduces the computational complexity associated with large-scale VPL evaluation while maintaining accurate approximation of indirect illumination contributions.

The rendering results of the SST method under different scenarios are illustrated in Figure **7**. From left to right, the primary cache, secondary cache, and final

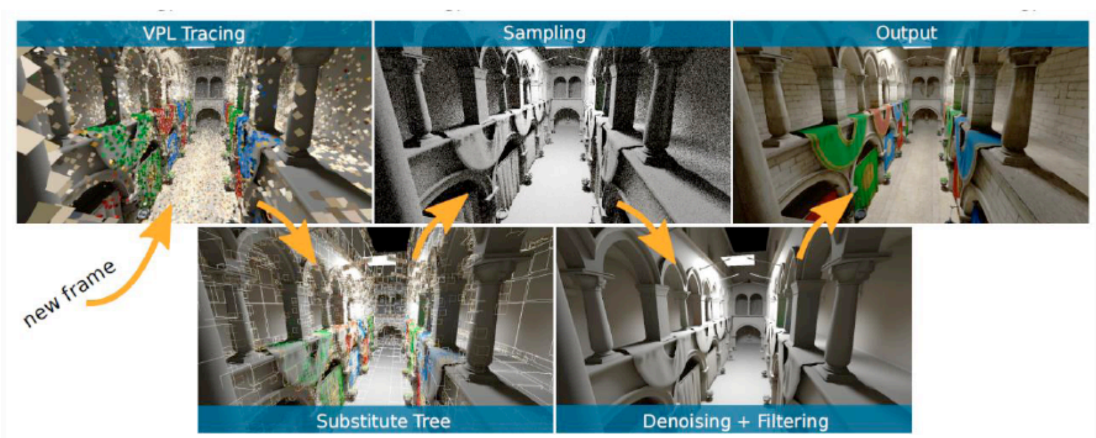

Figure 7: Rendering Results of the SST Method.

rendered image are shown. The method effectively handles complex illumination effects, including specular reflections, intricate shadowing, and multi-bounce light transport.

In its mathematical formulation, SST models illumination clustering using a Gaussian distribution and computes lighting contributions via an explicit probability density function. The system efficiently supports dynamic light sources and animated scenes, achieving real-time performance at 36 FPS on an NVIDIA RTX 3090 Ti GPU. Compared with conventional light-cut-based methods, SST achieves up to an 8× performance improvement. Furthermore, the framework integrates autoencoder-based denoising and temporal filtering techniques, effectively mitigating Monte Carlo noise and enabling high-quality real-time global illumination.

4.2. Point-Based and Progressive Global Illumination

In recent years, extended variants of point-based global illumination (PBGI) have provided new paradigms for realistic rendering of complex materials. Wang *et al*. [48] proposed Wavelet Point-Based Global Illumination (WPBGI), which introduces Haar wavelets into radiance field representations, significantly improving rendering efficiency for non-diffuse effects. In this formulation, outgoing radiance is represented as a sparse combination of directional wavelet coefficients:

$$L_{oj}(\omega) = \sum_{l} c_l \psi_l(\omega), \quad \tilde{L}_{oj}(\omega) = C_j \Psi(\omega) \tag{9}$$

where C_j denotes the wavelet coefficient vector and Ψ represents the Haar basis functions. Through hierarchical wavelet coding, the radiance at internal nodes is expressed as a weighted average of child coefficients, *i.e.*, $C_j = (C_j^- + C_j^+)/2$, reducing memory consumption by approximately 50% compared with traditional spherical harmonic-based PBGI representations. Combined with adaptive micro-buffer (AMB) techniques, which dynamically adjust sampling resolution according to BRDF glossiness and incident light intensity, the method preserves high-frequency details while achieving up to a 10× speedup over bidirectional path tracing (BPT). Experimental results show that WPBGI completes rendering of the Kitchen scene in 3.77 hours with an MSE of 7.948×10^{-5}, compared with 11.2 hours for BPT, as shown in Figure **8**.

To address real-time requirements in dynamic scenes, Laurent *et al*. [47] proposed Forward Light Cuts (FLC), which enables on-the-fly generation of virtual point lights (VPLs) through the coordinated use of geometry shaders and tessellation units. The method partitions scene triangles into subsets $L^0,\ldots,L^N$ according to area-based stochastic subdivision, governed by:

$$P(t_i \in L^k) = A(t_i) / S_k, S_k = S_0 \mu^k \tag{10}$$

where S_0 denotes a base area threshold and μ is a geometric growth factor. By constructing a multi-scale radiance cache directly on the GPU pipeline, FLC achieves a rendering rate of 35 ms per frame on the PowerPlant scene, representing a 56% speedup over Deep Screen Space (DSS) methods. However, the approach does not explicitly handle indirect shadowing and is primarily limited to unoccluded diffuse illumination, as illustrated in Figure **9**.

Progressive global illumination has recently become a key breakthrough in real-time rendering. Dammertz *et al*. [49] proposed a progressive point-light global illumination system that integrates three complementary techniques: point-based diffuse interreflection, specular gathering, and a caustics histogram approach for residual light transport. The system adopts a constant-memory progressive approximation scheme, requiring no precomputation or parameter tuning, making it particularly suitable for artist-driven workflows.

The framework introduces a decomposition of path space into four disjoint subsets: (1) EDL and EDD(S|D)L paths handled by point-light evaluation, (2)

(a) BPT, 11.2h (b) PPM, 11.55h (c) WPBGI, 3.77h (d) DBPT, 5.56h

Figure 8: Comparison of kitchen scene rendering results.

Figure 9: Comparison between the proposed method and indirect illumination-aware approaches.

EDS(S|D)L paths processed via caustics histograms, (3) ES+D+(S|D)L paths addressed using specular gathering, and (4) ESL paths directly evaluated. This decomposition enables complementary interaction between sub-methods while preserving their individual strengths. A key advantage of progressive rendering is the decoupling between memory consumption and image quality, allowing convergence to a stable solution even under complex lighting conditions, including caustics, interreflections, and dispersion effects.

4.3. 360° Real-Time Environmental Illumination Estimation for Augmented Reality

Alhakamy and Tuceryan [40] proposed the CubeMap360 system, which captures real-world illumination in real time using 360° fisheye video streams, enabling millisecond-level synchronization of direct and indirect lighting components. The captured illumination is encoded into dynamically updateable cube maps. The method first applies histogram-based adaptive thresholding to detect high-intensity regions in panoramic images, followed by spherical inverse projection to estimate the light direction vector:

$$x = \frac{\sin\phi\cos\theta}{\pi \times dh}, y = \frac{\cos\phi}{\pi \times dw}, z = \frac{\sin\phi\cos\theta}{\pi \times 90} \tag{11}$$

The real-time panoramic frames are periodically (every 10 frames) converted into six-face cube maps, which are then directly used in OpenGL-based image-based lighting (IBL) pipelines driven by HDR exposure parameters. User studies demonstrate that the method achieves illumination direction estimation errors within 1.4° under typical indoor dual-light, window illumination, and outdoor sunlight conditions. Moreover, it enables visually indistinguishable fusion between virtual objects and real shadows at 1080p@30 FPS on mobile platforms.

For mobile augmented reality applications, Zhao *et al*. [41] proposed the LitAR dual-field illumination reconstruction framework, which separately models near-field spatially varying illumination and far-field directional illumination. By leveraging multi-resolution projection and anchor extrapolation strategies, the point cloud projection latency is reduced from approximately 3 seconds to 14.6 ms, achieving superior visual consistency compared with ARKit-based pipelines. In outdoor relighting scenarios, Gardner *et al*. [42] proposed NeuSky, which treats sky pixels as direct observations of distant illumination. By integrating neural directional distance functions, the method enables efficient differentiable sky visibility estimation and jointly optimizes geometry, material, illumination, and visibility in an end-to-end framework, achieving state-of-the-art relighting performance on the NeRF-OSR benchmark.

Performance optimization remains a critical challenge in real-time global illumination. Cosin-Ayerbe *et al*. [43] proposed a voxel-based dynamic global illumination method that voxelizes the scene and applies ray tracing for illumination computation. By separating static and dynamic geometry, the method reduces ray tracing overhead and improves efficiency via spatiotemporal resampling. Experimental results demonstrate high-quality GI rendering in complex scenes while maintaining interactive frame rates. Similarly, Deferred Voxel Shading [44] leverages a voxelized representation within a deferred rendering pipeline. Combined with hardware-accelerated voxel traversal, the method achieves real-time global illumination for scenes exceeding 100K triangles. On an NVIDIA GTX 1660 Ti GPU, it achieves up to 40 FPS, representing a 25% improvement over conventional deferred shading approaches. The system separates voxels into static and dynamic categories, where static voxels are precomputed and dynamic voxels are updated in real time, achieving a balance between accuracy and efficiency.

4.4. Experimental Analysis and Preliminary Validation of SPC-PMC

The SPC-PMC model proposed in this paper offers significant advantages in both theoretical innovation and engineering application. Compared to the original PMC method, its core strengths lie in two areas: first, by introducing a highly robust deep semantic segmentation stream, it completely overcomes the interference caused by intricate trusses and interlaced conductors—common in power utility environments—during region mask extraction; second, it discards traditional Bayesian probability fusion strategies (which perform poorly on complex industrial images) in favor of an innovative, context-aware conditional decision mechanism. This mechanism utilizes prediction results for the sky and vertical surface regions to determine solar visibility and final position output, respectively, thereby significantly enhancing the algorithm's robustness in estimating light source direction from a single image. To comprehensively validate the SPC-PMC model's effectiveness, a rigorous end-to-end experimental workflow—encompassing dataset construction, model training, batch inversion, and 3D visualization—was conducted. A foundational aspect of this study involved the meticulous, fully manual semantic segmentation and annotation of over 300 images depicting typical power utility operational scenarios; this created a specialized high-level semantic dataset for power utility environments, which was then used for end-to-end deep segmentation training of the YOLOv11 network. Subsequently, the trained network was employed to generate masks and batch predictions for 110 historical surveillance images of substations and other power utility sites; these outputs were then fed into the improved SPC-PMC algorithm to quantitatively calculate solar angles (zenith and azimuth). Finally, a mapping mechanism converted the 2D calculation results into absolute illumination directions within a 3D sky hemisphere model, enabling shadow projection simulation analysis based on simple geometric structures. Validation against manually annotated "ground truth" solar coordinates—derived from analyzing shadow directions, object illumination, and specular highlights in the original images—demonstrated that the SPC-PMC model achieved significantly better performance than traditional comparative models in terms of solar angle deviation and radiometric color difference metrics. The illumination directions mapped onto the 3D sky hemisphere model align closely with physical reality, enabling a highly realistic 3D reconstruction of phenomena such as elongated equipment shadows and strong backlighting effects at specific operational moments. Experimental results fully demonstrate that the proposed method successfully establishes an intelligent "image-to-scene" lighting evolution workflow, providing technical support—characterized by both physical rigor and intuitiveness—for intelligent training in power utility operations.

The core value of the SPC-PMC model and its associated 3D lighting prediction and visualization method lies in bridging the gap between "single 2D images" and "3D volumetric lighting and shadows," thereby offering broad cross-domain application prospects within the diverse digital scenarios of new power systems and smart grids. In the future, this technology can serve as a core algorithm embedded in intelligent training systems, enhancing practical drill effectiveness by realistically reproducing specific backlighting and shadowing conditions. Furthermore, it can deeply empower intelligent power system O&M (operations and maintenance) by helping drones and inspection robots mitigate visual interference caused by light and shadow, while providing precise digital twin-based compensation for errors in infrared temperature measurements caused by direct sunlight. Beyond this, the method's application can extend to the coordinated management of source, grid, load, and storage in microgrids; by extracting cues regarding highly dynamic environmental radiance from distributed or substation-based surveillance imagery, it supports visual-level physical and meteorological sensing for the short-term output analysis of regional micro-photovoltaic plants. Additionally, during the foundational phase of grid digitalization, the method can be used to evaluate dynamic sunlight exposure and architectural shading effects for high-voltage transmission corridors and substation sites, offering spatial lighting-based decision support for the full-lifecycle digital planning of grid infrastructure. In summary, the SPC-PMC model holds promise as a

Table 5: Accuracy Analysis of 3D Illumination Estimation

Evaluation Dimension	Statistical Indicators	Sample Size	Proportion
Interval distribution statistics	High precision (< 22.5°)	32	29.1%
	Low precision (22.5°- 45°)	36	32.7%
	There is an error (＞45°)	42	38.2%
Global statistical error	**Average angle error (°)**	**Median angular error (°)**	**Standard deviation (°)**
—	39.586	42.539	29.375

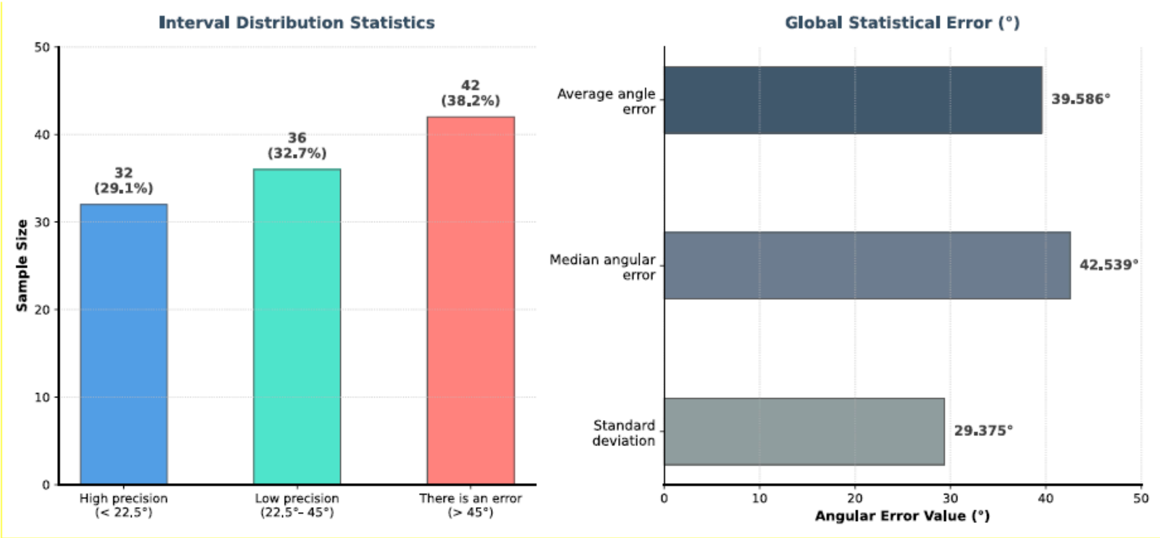

Figure 10: Visualization of 3D Lighting Estimation Accuracy Analysis.

ubiquitous visual computing tool, injecting physically rigorous digital momentum into the multidimensional spatial sensing and digital twin evolution of future smart grids.

Table **5** and Figure **10** present the quantitative results of 3D illumination direction estimation using the proposed SPC-PMC model across 110 test samples, revealing its performance characteristics under various operating conditions. Regarding the distribution of results, the model exhibits a distinct polarization: under typical conditions—characterized by clear weather and large wall surfaces directly facing the camera—the model performs exceptionally well, with 29.1% of samples falling into the high-precision range (error < 22.5°), thereby demonstrating its peak accuracy in standard specialized industrial scenarios. In scenarios involving side-facing walls or perspective distortion (22.5° < error < 45°), which account for 32.7% of the samples, the model effectively manages physical errors through its maximum likelihood optimization mechanism based on regional probability matrices. Conversely, in complex electrical environments (error > 45°), accounting for 38.2% of samples, mask segmentation errors caused by dense metal substation structures and complex occlusions trigger false peaks during optimization, representing the algorithm's primary bottleneck.

From a global statistical perspective, the test set yields a mean angular error of 39.586°, a median of 42.539°, and a standard deviation of 29.375°. The proximity of the mean to the median indicates overall algorithmic robustness—free from catastrophic estimation failures—thereby validating the effectiveness of the conditional decision-making mechanism. Meanwhile, the significant standard deviation intuitively reflects the physical nature of industrial lighting cues, which are heavily dependent on geometric structure; the model excels in geometrically ideal scenarios but faces constrained effectiveness when dealing with complex, irregular equipment and conditions involving strong scattering or occlusion. In summary, while these metrics confirm the physical consistency of the SPC-PMC architecture, they also objectively highlight the rigorous challenges posed by complex scenarios, pointing the way toward future optimizations involving strengthened geometric constraints and multi-region joint modeling.

5. FROM 2D ILLUMINATION ESTIMATION TO 3D MAPPING AND SHADOW ANALYSIS

To enable end-to-end “image-to-scene” illumination reconstruction, the SPC-PMC framework further constructs a pipeline that maps 2D illumination cues into a 3D representation space for visualization and shadow analysis. Based on high-quality semantic masks produced by a YOLOv11 segmentation network, sky regions are fitted using a physical sky model to estimate solar zenith and azimuth angles, while vertical surfaces employ a multi-directional voting scheme to independently estimate solar azimuth priors. The estimated solar position from the 2D sky dome is then projected onto a 3D hemispherical sky model, enabling intuitive visualization of illumination direction in 3D space. Given the reconstructed illumination direction, simplified geometric primitives (*e.g.*, transmission towers, beams, and equipment enclosures) are used for shadow projection and quantitative illumination analysis under operational conditions. This enables realistic simulation of long-shadow occlusion and strong backlighting effects in electrical operation training environments, significantly enhancing visual realism and interpretability in digital twin systems.

For validation, a dedicated high-level semantic dataset of power grid scenes is constructed via full manual annotation of representative electrical operation images. The YOLOv11 network is then trained in an end-to-end manner for fine-grained segmentation. The trained model is applied to unseen substation surveillance images for mask prediction, which are subsequently fed into the improved SPC-PMC pipeline for solar parameter estimation. The estimated solar zenith and azimuth angles are compared against ground-truth annotations derived from manual interpretation of shadow orientation, illumination distribution, and specular highlights. Experimental results show that SPC-PMC achieves significantly lower angular deviation and radiometric error compared with baseline methods. The reconstructed illumination direction in the 3D hemispherical model exhibits strong physical consistency, accurately reproducing long-shadow occlusion and high-intensity backlighting effects in real-world scenarios. These results validate the effectiveness of the proposed framework in bridging the gap between 2D observations and 3D physically consistent illumination reconstruction for intelligent power system training.

5.1. Shadow Layer Decomposition in Global Illumination

Witten *et al*. [45] proposed a Global Illumination Shadow Layer decomposition framework aimed at balancing rendering efficiency and visual fidelity in complex scenes. The method decomposes shadows into direct and indirect components, enabling separate optimization of each contribution. Direct shadows arise from occlusion of primary light paths (*e.g.*, sharp solar occlusion), resulting in hard shadow boundaries with high contrast. Indirect shadows originate from multi-bounce light transport, producing soft and smooth shading effects.

To model the fusion of both components, the radiance formulation is defined as:

$$L_{\text{shadow}}(x) = L_{\text{direct}}(x) \cdot \alpha(x) + L_{\text{indirect}}(x) \cdot (1 - \alpha(x)) \tag{12}$$

where $L_{direct}(x)$ and $L_{indirect}(x)$ denote direct and indirect radiance contributions, respectively, and $\alpha(x)$is a shadow blending coefficient that depends on surface roughness. Higher roughness leads to stronger indirect illumination contribution, consistent with physical scattering behavior.

Through a *layered computation and dynamic blending* strategy, this method achieves a favorable trade-off between rendering quality and computational cost in complex scenarios (*e.g.*, scenes with dynamic objects and intricate geometry in film production or game engines). It preserves fine-grained detail in direct shadows while producing naturally smooth indirect shadows, thereby providing a physically more plausible solution for global illumination shadow rendering.

5.2. Dynamic Height-Field Global Illumination

Nowrouzezahrai and Snyder [56] proposed a dynamic height-field global illumination method that achieves efficient lighting computation through a multi-resolution pyramid representation and a normalized Legendre polynomial basis. A key advantage of this method lies in its strong support for dynamic geometry and its efficient algorithmic design.

Figure **11** illustrates the rendering results in complex scenes, including an indoor room, maple tree shadows, and a power plant model, demonstrating the method's robustness across varying scene complexities.

The system constructs two key pyramid structures: a height variation pyramid for direct shadow computation and a radiance pyramid for indirect illumination. Mathematically, illumination is represented using compact NLP coefficient vectors, significantly reducing memory bandwidth requirements. Experimental results show that at 512×512 resolution, the method achieves 36 fps, representing a 3–6×

Figure 11: Application Results in Complex Scenes.

performance improvement over conventional approaches. This efficiency enables real-time global illumination for complex height-field scenes and provides a new solution for terrain rendering and special effects.

Zhu *et al.* [54] proposed WishGI, a system optimized for low-end and mobile platforms. The method introduces an efficient spherical harmonic fitting scheme combined with an inverse probe distribution strategy for static global illumination computation. WishGI encodes radiance information into a compact data structure using a lightweight spherical harmonic coefficient representation. The system achieves comparable visual quality to traditional 42 KB directional lightmaps while requiring only 5 KB of memory. Tests on a Samsung Galaxy S21 demonstrate a rendering performance of 55.67 fps with a 95% reduction in memory usage, enabling high-quality global illumination on resource-constrained mobile devices.

Luksch *et al*. [55] proposed an incremental baking global illumination method that extends traditional one-shot baking into a frame-wise progressive update scheme. The core idea is to construct an editable octree-clustering structure for each virtual point light (VPL) while recording full light transport paths, allowing localized recomputation when geometry, materials, or lighting conditions change. The algorithm uses invalid radiance flux as a priority metric, updating light sources with the highest perceptual error first and gradually converging toward global consistency. Combined with a GPU rasterization accumulation pipeline, it achieves a smooth transition from direct lighting to multi-bounce illumination within 1–2 seconds, enabling flicker-free interactive lighting design workflows. This method bridges the gap between traditional lightmap baking, which lacks adaptability to dynamic edits, and real-time GI approaches with limited accuracy, providing a scalable framework for editable precomputed global illumination.

6. CONCLUSION AND FUTURE PERSPECTIVES

6.1. Cross-Domain Application Prospects for Digital Twin Power Grids

The primary contribution of the proposed SPC-PMC framework and its 3D illumination prediction and visualization pipeline lies in establishing a complete inversion pathway from a single 2D image to a physically interpretable 3D lighting representation. This capability demonstrates significant potential across a wide range of digital twin applications in next-generation power systems and smart grids. In intelligent power-operation training systems, the proposed framework can be integrated into virtual simulation platforms to faithfully reconstruct task-specific illumination conditions, including backlighting effects and complex shadow distributions. By reproducing realistic visual environments corresponding to specific times and weather conditions, the method addresses the visual discontinuity introduced by the static lighting configurations commonly adopted in existing training simulators, thereby improving training realism and operational preparedness. In intelligent power-grid inspection and maintenance, the proposed approach can enhance the performance of unmanned aerial vehicles (UAVs) and autonomous inspection robots by mitigating visual interference caused by challenging illumination conditions. More accurate modeling of environmental lighting can improve defect detection reliability while providing digital-twin-based compensation mechanisms for solar-radiation-induced errors in infrared thermography, thereby increasing the accuracy and consistency of temperature measurements.

Beyond these applications, the proposed framework may also benefit microgrid energy-management scenarios involving the coordinated operation of generation, grid, load, and energy-storage resources. By extracting environmental radiometric information from distributed monitoring imagery, the method can provide visual-physics-based meteorological indicators for short-term photovoltaic power forecasting. Furthermore, during grid planning and infrastructure development, the framework can be employed to evaluate dynamic solar exposure and shading effects associated with transmission corridors and substation locations, supporting lifecycle-oriented digital planning and decision-making. Collectively, these capabilities suggest that SPC-PMC can serve as a general-purpose visual computing framework for future smart-grid systems, providing physically grounded illumination awareness and enhancing the evolution of digital twin environments through improved spatial perception and environmental understanding.

Accurate illumination reconstruction further supports the synchronization between physical power infrastructures and their virtual counterparts in digital-twin environments. By maintaining illumination consistency across real and virtual spaces, the reconstructed lighting information enhances scene realism, improves visual perception accuracy, and facilitates intelligent inspection, condition assessment, and operator training in power-system digital twins.

From a broader digital-twin perspective, Cavus *et al*. [58] conducted a comprehensive PRISMA-based review of digital twin technologies for power-grid resilience enhancement, identifying four major challenges: data latency, computational complexity, limited interoperability with conventional SCADA systems, and cybersecurity risks. Their study further outlined a roadmap toward self-healing intelligent power grids. Similarly, Barreto and Aoki [59] employed the ProKnow-C methodology to analyze 2,207 candidate publications and identified seven major application domains, including real-time decision support, cybersecurity, predictive maintenance, energy management, forecasting, fault handling, and grid resilience. Within this broader context, illumination estimation and global illumination technologies can be regarded as essential components of the visual perception layer in digital twin systems. Their importance is expected to increase in applications such as equipment-state visualization, AR-assisted inspection, and immersive emergency-response training, where they serve as a critical bridge between physical visual observations and digital spatial representations.

6.2. Future Research Directions

This review has presented a comprehensive overview of the evolution of illumination modeling, spanning traditional physically based approaches and modern neural rendering paradigms. Representative techniques include path tracing, light probes, virtual point light clustering, neural baking, Transformer-based architectures, illumination decomposition, dynamic global illumination, environment lighting representations, low-light enhancement, VR/AR rendering, and rendering acceleration techniques. Building upon these developments, this work further investigated intelligent illumination reconstruction for virtual power-system environments and introduced the theoretical foundations, implementation methodology, and experimental validation of the Semantic-Pixel Coupled Probabilistic Multi-Cue (SPC-PMC) framework. The results demonstrate the potential of integrating deep learning with physically grounded illumination modeling for industrial digital twin applications. While conventional approaches continue to offer strong physical interpretability and theoretical rigor, deep-learning-based methods provide substantial advantages in computational efficiency, robustness, and scalability, particularly for complex materials, dynamic environments, and multimodal perception tasks.

Despite the significant progress achieved in rendering quality and computational performance, several important challenges remain. First, as scene complexity and environmental dynamics continue to increase, reducing computational overhead while maintaining visual fidelity remains a fundamental research problem. This challenge is especially critical for resource-constrained platforms such as mobile devices, wearable systems, and edge-computing hardware, where lightweight algorithms and efficient data structures are essential. Second, although neural rendering techniques have demonstrated remarkable capabilities, limitations persist in terms of cross-scene generalization, physical consistency, and dependence on large-scale training data. Future research should explore more robust network architectures and training paradigms, including physics-guided regularization strategies and diffusion-based generative frameworks that explicitly incorporate physical constraints. In this direction, Liang *et al*. [60] recently proposed Diffusion Renderer, which unifies inverse rendering and forward rendering within a video diffusion framework. By introducing domain embeddings, a single model can generate multiple G-buffer representations, suggesting a future trend toward unified scene representations with diverse downstream outputs.

Furthermore, dynamic global illumination techniques still exhibit limitations when handling complex material interactions, indirect shadowing, and caustic effects. Future systems may benefit from hybrid rendering pipelines that combine hardware-accelerated ray tracing with neural radiance field representations, enabling physically informed illumination modeling while preserving real-time performance. Environment-lighting representations also require further improvement in terms of sampling efficiency and compression capability. Neural-field-based and implicit scene representations offer promising directions for high-dimensional environment-light modeling, potentially enabling low-latency, high-fidelity global illumination estimation. For industrial applications such as power systems, several domain-specific challenges deserve particular attention. Future studies should investigate semantic segmentation architectures specifically designed for high-frequency industrial structures, illumination inversion strategies incorporating multi-view geometric constraints, and lightweight deployment frameworks suitable for edge devices.

Finally, with the rapid adoption of XR technologies and digital twin systems, future illumination models are expected to move beyond purely photometric objectives and incorporate semantic scene understanding together with human visual perception mechanisms. Such developments may lead to adaptive intelligent lighting systems capable of dynamically optimizing visual realism, immersion, and user experience. For virtual power-system environments in particular, future illumination reconstruction techniques should be more tightly integrated with physics-based equipment models. In this setting, illumination simulation would not only serve visual rendering purposes but also provide physically consistent optical inputs for advanced applications, including thermal analysis of electrical equipment, photovoltaic power forecasting, and vision-based contamination assessment of insulators. Such integration would further strengthen the role of illumination modeling as a foundational component of next-generation digital twin power systems.

DECLARATIONS

Funding

This work was supported by the NSFC Project (Project Nos. 62003212).

Data Availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

Not applicable. No additional ethical approval was required for this research.

Consent for Publication

Written consent to participate and written consent to publish were obtained from all participants.

Competing Interests

The authors declare no competing interests.

REFERENCES

[1] Y. Gao, Y. Wang, N. Yang, Q. Wang, B. Javadi, Q. Ai, and J. Zhu, "Multi-timescale distributed control for multi-energy virtual power plant clusters via cloud-edge collaboration," Appl. Energy, vol. 414, Art. no. 127835, Jul. 2026. https://doi.org/10.1016/j.apenergy.2026.127835

[2] Y. Gao, X. Hu, T. Lü, Q. Ai, and X. He, "Digital twin control strategy for multi-energy virtual power plant considering vehicle-grid interaction and dynamic carbon trading," Power Syst. Technol., vol. 50, no. 2, pp. 683-698, Feb. 2026.

[3] S. López Flórez, G. Hernández González, J. Prieto, and F. de la Prieta, "Hybrid physics-LSTM framework for wind power prediction and control in virtual microgrid simulations," IEEE Access, vol. 13, pp. 122175-122186, Jul. 2025. https://doi.org/10.1109/ACCESS.2025.3586798

[4] J-F. Lalonde, A. A. Efros, and S. G. Narasimhan, "Estimating Natural Illumination from a Single Outdoor Image," Proc. IEEE Int. Conf. Comput. Vis., pp. 183-190, Sep. 2009. https://doi.org/10.1109/ICCV.2009.5459163

[5] A. Singh, U. Demirbaga, G. S. Aujla, A. Jindal, H. Sun, and J. Jiang, "Scalable and reliable data framework for sensor-enabled virtual power plant digital twin," IEEE J. Sel. Areas Sensors, vol. 2, pp. 108-120, 2025. https://doi.org/10.1109/JSAS.2025.3540956

[6] R. R. Rodriguez-Pardo, D. Dolonius, U. Assarsson, and E. Sintorn, "Spherical Gaussian Light-field Textures for Fast Precomputed Global Illumination," Comput. Graph. Forum, vol. 39, no. 2, pp. 133-146, 2020. https://doi.org/10.1111/cgf.13918

[7] Diolatzis, J. Philip, and G. Drettakis, "Active Exploration for Neural Global Illumination of Variable Scenes," ACM Trans. Graph., vol. 41, no. 5, Art. no. 171, 2022. https://doi.org/10.1145/3522735

[8] J. Guo, Z. Zong, Y. Song, X. Fu, C. Tao, Y. Guo, and L. Yan, "Efficient Light Probes for Real-time Global Illumination," ACM Trans. Graph., vol. 41, no. 4, Art. no. 202, 2022. https://doi.org/10.1145/3550454.3555452

[9] S. Rodriguez, T. Leimkuhler, S. Prakash, C. Wyman, P. Shirley, and G. Drettakis, "Glossy Probe Reprojection for Interactive Global Illumination," ACM Trans. Graph., vol. 39, no. 6, Art. no. 237, 2020. https://doi.org/10.1145/3414685.3417823

[10] N. Bus, N. H. Mustafa, and V. Biri, "Global illumination using well-separated pair decomposition," Comput. Graph. Forum, vol. 34, no. 8, pp. 88-103, 2015. https://doi.org/10.1111/cgf.12610

[11] K. Doi, Y. Morimoto, and R. Tsuruno, "Global Illumination-Aware Stylised Shading," Comput. Graph. Forum, vol. 40, no. 7, Art. no. 14397, 2021. https://doi.org/10.1111/cgf.14397

[12] Z. Majercik, T. Müller, A. Keller, D. Nowrouzezahrai, and M. McGuire, "Dynamic Diffuse Global Illumination Resampling," Comput. Graph. Forum, vol. 41, no. 1, pp. 158-171, 2022. https://doi.org/10.1111/cgf.14427

[13] N. Patakin, D. Senushkin, A. Vorontsova, and A. Konushin, "Neural global illumination for inverse rendering," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2023. https://doi.org/10.1109/ICIP49359.2023.10222145

[14] D. Gao, H. Mu, and K. Xu, "Neural Global Illumination: Interactive Indirect Illumination Prediction Under Dynamic Area Lights," IEEE Trans. Vis. Comput. Graph., vol. 29, no. 12, 2023. https://doi.org/10.1109/TVCG.2022.3209963

[15] C. Zeng, Y. Dong, P. Peers, H. Wu, and X. Tong, "RenderFormer: Transformer-based neural rendering of triangle meshes with global illumination," in Proc. SIGGRAPH Conf. Papers, 2025, Art. no. 11. https://doi.org/10.1145/3721238.3730595

[16] S. Hadadan, G. Lin, J. Novák, F. Rousselle, and M. Zwicker, "Inverse global illumination using a neural radiometric prior," in Proc. SIGGRAPH Conf. Papers, 2023, Art. no. 11. https://doi.org/10.1145/3588432.3591553

[17] S. Mo, C. Zheng, Z. Lin, D. Xi, Q. Ye, R. Wang, H. Bao, and Y. Huo, "Dual-band feature fusion for neural global illumination with multi-frequency reflections," in Proc. SIGGRAPH Conf. Papers, 2025, Art. no. 11. https://doi.org/10.1145/3721238.3730733

[18] D. Hoiem, A. A. Efros, and M. Hebert, "Recovering Surface Layout from an Image," Int. J. Comput. Vis., vol. 75, no. 1, pp. 151-172, 2007. https://doi.org/10.1007/s11263-006-0031-y

[19] A. Meka, M. Shafiei, M. Zollhöfer, C. Richardt, and C. Theobalt, "Real-time Global Illumination Decomposition of Videos," ACM Trans. Graph., vol. 40, no. 3, Art. no. 22, 2021. https://doi.org/10.1145/3374753

[20] C. Rodriguez-Pardo, J. Fabre, E. Garces, and J. Lopez-Moreno, "NEnv: Neural Environment Maps for Global Illumination," Comput. Graph. Forum, vol. 42, no. 4, 2023. https://doi.org/10.1111/cgf.14883

[21] Y. Fang, W. Zhu, and Q. Zhu, "UGNet: Underexposed images enhancement network based on global illumination estimation," in Proc. IEEE Int. Conf. Vis. Commun. Image Process. (VCIP), 2020. https://doi.org/10.1109/VCIP49819.2020.9301810

[22] Z. Majercik, T. Müller, A. Keller, D. Nowrouzezahrai, and M. McGuire, "Dynamic diffuse global illumination resampling," in Proc. SIGGRAPH Talks, 2021, Art. no. 2. https://doi.org/10.1145/3450623.3464635

[23] N. Max, S. Saito, K. Watanabe, and M. Nakajima, "Rendering grass blowing in the wind with global illumination," Tsinghua

Sci. Technol., vol. 15, no. 2, pp. 133-137, 2010. https://doi.org/10.1016/S1007-0214(10)70042-0

[24] W. Lü, J. Lu, X. Liu, and E. Wu, "A fast method for real-time computation of approximated global illumination," in Proc. 6th Int. Conf. Comput. Graph., Imaging Vis., 2009, pp. 62-68. https://doi.org/10.1109/CGIV.2009.24

[25] J. Noor, A. Mahmud, M. A. Rahman, A. Sifar, F. Y. Mostafa, L. Tasnova, and S. Chellappan, "SAwareSSGI: Surrounding-aware screen-space global illumination using generative adversarial networks," IEEE Access, vol. 12, pp. 139946-139961, 2024. https://doi.org/10.1109/ACCESS.2024.3467102

[26] Z. Shao, K. Zhang, D. Zhao, T. Wang, and T. Lu, "LLFA: Fusing global illumination and local priors for low-light face image enhancement with adaptor," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2025. https://doi.org/10.1109/ICASSP49660.2025.10890118

[27] C. Wyman, S. Parker, P. Shirley, and C. Hansen, "Interactive display of isosurfaces with global illumination," IEEE Trans. Vis. Comput. Graph., vol. 12, no. 2, pp. 186-196, 2006. https://doi.org/10.1109/TVCG.2006.33

[28] J. Hilliard, A. Hilton, and J.-Y. Guillemaut, "HDR environment map estimation with latent diffusion models," arXiv:2507.21261, Jul. 2025._

[29] R. Liang, K. He, Z. Gojcic, I. Gilitschenski, S. Fidler, N. Vijaykumar, and Z. Wang, "LuxDiT: Lighting estimation with video diffusion transformer," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), San Diego, CA, USA, Dec. 2025._

[30] C. Bolduc, J. Philip, L. Ma, M. He, P. Debevec, and J.-F. Lalonde, "Lighting in motion: Spatiotemporal HDR lighting estimation," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2026.

[31] X. Zhao, P. P. Srinivasan, D. Verbin, K. Park, R. Martin-Brualla, and P. Henzler, "IllumiNeRF: 3D relighting without inverse rendering," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), Vancouver, Canada, pp. 42593-42617, Dec. 2024. https://doi.org/10.52202/079017-1349

[32] Y. Wang, S. Song, L. Zhao, H. Xia, Z. Yuan, and Y. Zhang, "CGLight: An effective indoor illumination estimation method based on improved ConvMixer and GauGAN," Comput. Graph., vol. 125, Art. no. 104122, Dec. 2024. https://doi.org/10.1016/j.cag.2024.104122

[33] A. R. Anonymous *et al.*, "Enhancing XR training with interactive AI-powered virtual instructors," IEEE MultiMedia, early access, 2025._

[34] X. Zhao, S. Zheng, L. Wang, Y. Wu, J. Zhao, and Y. Guo, "AA-COMT: An AI-driven AR system for communication operation and maintenance training," Proc. Int. Conf. Intell. Power Syst. (ICIPS), Yichang, China, pp. 963-967, Dec. 2024. https://doi.org/10.1109/ICIPS64173.2024.10900096

[35] Y.-Z. Lin, K. Petal, A. H. Alhamadah, S. Ghimire, M. W. Redondo, D. R. Vidal Corona, J. Pacheco, S. Salehi, and P. Satam, "Personalized education with generative AI and digital twins: VR, RAG, and zero-shot sentiment analysis for Industry 4.0 workforce development," arXiv:2502.14080, Feb. 2025.

[36] S. Yoo, S. Reza, H. Tarashiyoun, A. Ajikumar, and M. Moghaddam, "AI-integrated AR as an intelligent companion for industrial workers: A systematic review," IEEE Access, vol. 12, pp. 191808-191827, 2024. https://doi.org/10.1109/ACCESS.2024.3516536

[37] S. Yoo, C. Harteveld, N. Wilson, K. Jona, and M. Moghaddam, "Multimodal assessment of expertise in AR-guided psychomotor tasks," IEEE Trans. Syst., Man, Cybern.: Syst., vol. 55, no. 11, pp. 8126-8141, Nov. 2025. https://doi.org/10.1109/TSMC.2025.3604693

[38] V. Phadke, C. Harteveld, K. Jona, and M. Moghaddam, "First things first: Effects of sequential AR/VR exposure on skill acquisition in industrial training," Adv. Eng. Inform., vol. 71, Art. no. 104328, 2026. https://doi.org/10.1016/j.aei.2026.104328

[39] K. Witkowski, "Integrating VR, AR and AI into corporate employee training: A study of mixed methods towards personalized learning design," Management, vol. 2026, no. 1, pp. 70-95, Apr. 2026. https://www.management-poland.com/Integrating-VR-AR-and-AI-into-corporate-employee-training-A-study-of-mixed-methods,218508,0,2.html https://doi.org/10.58691/man/218508

[40] A. Alhakamy and M. Tuceryan, "CubeMap360: Interactive global illumination for augmented reality in dynamic environment," in Proc. IEEE Int. Symp. Mixed Augment. Reality (ISMAR), 2019. https://doi.org/10.1109/SoutheastCon42311.2019.9020588

[41] Y. Zhao, C. Ma, H. Huang, and T. Guo, "LitAR: Visually coherent lighting for mobile augmented reality," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), vol. 6, no. 3, Art. no. 153, Sep. 2022. https://doi.org/10.1145/3550291

[42] J. A. D. Gardner, E. Kashin, B. Egger, and W. A. P. Smith, "The sky's the limit: Relightable outdoor scenes via a sky-pixel constrained illumination prior and outside-in visibility," Proc. Eur. Conf. Comput. Vis. (ECCV), ser. LNCS, vol. 15112, Milan, Italy, pp. 126-143, Oct. 2024. https://doi.org/10.1007/978-3-031-72949-2_8

[43] A. Cosin Ayerbe, P. Poulin, and G. Patow, "Dynamic Voxel-Based Global Illumination," Comput. Graph. Forum, vol. 43, no. 1, 2024. https://doi.org/10.1111/cgf.15262

[44] J. Villegas and E. Ramírez, "Deferred voxel shading for real-time global illumination," in Proc. IEEE Pacific Graph. Conf., 2016. https://doi.org/10.1109/CLEI.2016.7833375

[45] F. Desrichard, D. Vanderhaeghe, and M. Paulin, "Global illumination shadow layers," Comput. Graph. Forum, vol. 38, no. 4, pp. 184-191, 2019. https://doi.org/10.1111/cgf.13781

[46] B. Wang, L. Wang, and N. Holzschuch, "Fast global illumination with discrete stochastic microfacets using a filterable model," Comput. Graph. Forum, vol. 37, no. 7, pp. 56-64, 2018. https://doi.org/10.1111/cgf.13547

[47] G. Laurent, C. Delalandre, G. De La Rivière, and T. Boubekeur, "Forward light cuts: A scalable approach to real-time global illumination," Comput. Graph. Forum, vol. 35, no. 4, pp. 80-88, 2016. https://doi.org/10.1111/cgf.12951

[48] B. Wang, X. Meng, and T. Boubekeur, "Wavelet point-based global illumination," Comput. Graph. Forum, vol. 34, no. 4, pp. 144-153, 2015. https://doi.org/10.1111/cgf.12686

[49] H. Dammertz, A. Keller, and H. P. A. Lensch, "Progressive point-light-based global illumination," Comput. Graph. Forum, vol. 29, no. 8, pp. 2504-2515, 2010. https://doi.org/10.1111/j.1467-8659.2010.01786.x

[50] P. Bauszat, M. Eisemann, S. John, and M. Magnor, "Sample-based manifold filtering for interactive global illumination and depth of field," Comput. Graph. Forum, vol. 34, no. 1, pp. 265-276, 2014. https://doi.org/10.1111/cgf.12511

[51] A. Gruson, M. Ribardière, and R. Cozot, "Eye-centered color adaptation in global illumination," Comput. Graph. Forum, vol. 32, no. 7, pp. 112-120, 2013. https://doi.org/10.1111/cgf.12218

[52] J. Kontkanen, E. Tabellion, and R. S. Overbeck, "Coherent out-of-core point-based global illumination," Comput. Graph. Forum, vol. 30, no. 4, pp. 1354-1360, 2011. https://doi.org/10.1111/j.1467-8659.2011.01995.x

[53] W. Tatzgern, B. Mayr, B. Kerbl, and M. Steinberger, "Stochastic substitute trees for real-time global illumination," in Proc. Symp. Interact. 3D Graph. Games, 2020, Art. no. 9. https://doi.org/10.1145/3384382.3384521

[54] J. Zhu, Z. Wu, Q. Zhang, C. Liao, and Z. Huang, "WishGI: Lightweight static global illumination baking via spherical harmonics fitting," ACM Trans. Graph., vol. 44, no. 4, Art. no.

72, 2025.
https://doi.org/10.1145/3730935

[55] C. Luksch, M. Wimmer, and M. Schwärzler, "Incrementally baked global illumination," in Proc. Symp. Interact. 3D Graph. Games, 2019, Art. no. 2.
https://doi.org/10.1145/3306131.3317015

[56] D. Nowrouzezahrai and J. Snyder, "Fast global illumination on dynamic height fields," Comput. Graph. Forum, vol. 28, no. 4, pp. 1132-1139, 2009.
https://doi.org/10.1111/j.1467-8659.2009.01490.x

[57] R. Perez, R. Seals, and J. Michalsky, "All-Weather Model for Sky Luminance Distribution," Solar Energy, vol. 50, no. 3, pp. 235-245, 1993.
https://doi.org/10.1016/0038-092X(93)90017-I

[58] M. Cavus, J. Jiang, A. Allahham, A. G. Rameshrao, E. Scullion, B. D. Malamud, H. Sun, and W. P. Ng, "Digital twins for hazard-resilient power grids: A systematic review and roadmap," Renew. Sustain. Energy Rev., vol. 235, Art. no. 116947, Mar. 2026.
https://doi.org/10.1016/j.rser.2026.116947

[59] N. E. M. Barreto and A. R. Aoki, "Cyber-physical power system digital twins—A study on the state of the art," Energies, vol. 18, no. 22, Art. no. 5960, Nov. 2025.
https://doi.org/10.3390/en18225960

[60] R. Liang, Z. Gojcic, H. Ling, J. Munkberg, J. Hasselgren, C.-H. Lin, J. Gao, A. Keller, N. Vijaykumar, S. Fidler, and Z. Wang, "Diffusion renderer: Neural inverse and forward rendering with video diffusion models," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 26069-26080, Jun. 2025. https://openaccess.thecvf.com/content/CVPR2025/html/Liang_Diffusion_Renderer_Neural_Inverse_and_Forward_Rendering_with_Video_Diffusion_CVPR_2025_paper.html
https://doi.org/10.1109/CVPR52734.2025.02428

https://doi.org/10.31875/2978-6436.2026.02.03