

## IPCD: Intrinsic Point-Cloud Decomposition

Shogo Sato, Takuhiro Kaneko, Shoichiro Takeda, Tomoyasu Shimada,  
 Kazuhiko Murasaki, Taiga Yoshida, Ryuichi Tanida, Akisato Kimura  
 NTT, Inc.

{shg.sato, takuhiro.kaneko, shoichiro.takeda, tomoyasu.shimada,}  
 {kazuhiko.murasaki, taiga.yoshida, ryuichi.tanida, akisato.kimura}@ntt.com

### Abstract

Point clouds are widely used in various fields, including augmented reality (AR) and robotics, where relighting and texture editing are crucial for realistic visualization. Achieving these tasks requires accurately separating albedo from shade. However, performing this separation on point clouds presents two key challenges: (1) the non-grid structure of point clouds makes conventional image-based decomposition models ineffective, and (2) point-cloud models designed for other tasks do not explicitly consider global-light direction, resulting in inaccurate shade. In this paper, we introduce **Intrinsic Point-Cloud Decomposition (IPCD)**, which extends image decomposition to the direct decomposition of colored point clouds into albedo and shade. To overcome challenge (1), we propose **IPCD-Net** that extends image-based model with point-wise feature aggregation for non-grid data processing. For challenge (2), we introduce **Projection-based Luminance Distribution (PLD)** with a hierarchical feature refinement, capturing global-light cues via multi-view projection. For comprehensive evaluation, we create a synthetic outdoor-scene dataset. Experimental results demonstrate that **IPCD-Net** reduces cast shadows in albedo and enhances color accuracy in shade. Furthermore, we showcase its applications in texture editing, relighting, and point-cloud registration under varying illumination. Finally, we verify the real-world applicability of **IPCD-Net**.

### 1. Introduction

Colored point clouds are widely used in applications such as augmented reality (AR) [27] and robotics [19], due to their ability to preserve geometry and appearance [15]. These applications often require texture editing and relighting to enhance visual realism and simulate diverse illuminations [12, 40]. Additionally, point-cloud registration is required for reconstructing digital spaces even under varying illuminations [29, 39]. However, the entanglement of

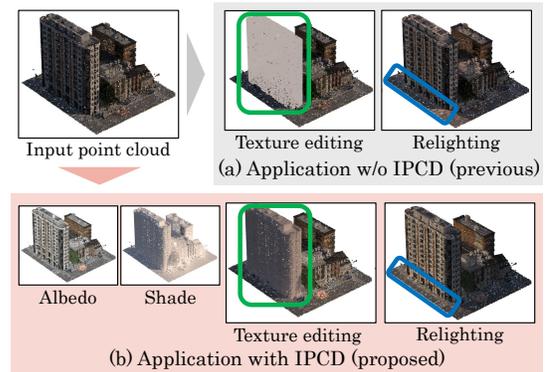


Figure 1. Motivation of Intrinsic Point-Cloud Decomposition (IPCD). (a) Directly using colored point clouds for texture editing and relighting introduces artifacts due to illumination, resulting in unnatural colors (green) and residual shade (blue). To address this issue, (b) we extend image decomposition to point clouds, namely **IPCD**, which decomposes point clouds into albedo and shade. Applying **IPCD** enables more natural texture editing and significantly reduces residual shade in relighting.

surface reflectance (albedo) and illumination (shade) in colored point clouds may hinder these tasks as shown in Fig. 1. With the widespread of three dimensional (3D) scanners equipped with synchronized cameras such as NavVis [1], only colored point clouds are often stored, while the original images are discarded. Thus, techniques for directly decomposing colored point clouds are desired.

In two-dimensional (2D) image domain, intrinsic image decomposition (IID) [3, 20] has been extensively explored to decompose an image into albedo and shade [5, 7]. To apply existing IID models to point cloud, we render the points into images for IID and project the results back to points, but this may lead to low decomposition quality due to occlusions and geometric lacking.

Achieving direct point-cloud decomposition introduces two key challenges: (1) the non-grid structure of point clouds makes conventional image-based decomposition models ineffective, and (2) point-cloud models designed for other tasks do not explicitly consider global-light direction,

resulting in inaccurate shade.

In this paper, we extend the conventional IID problem to colored point clouds, which we refer to as **Intrinsic Point-Cloud Decomposition (IPCD)**. IPCD decomposes a colored point cloud into albedo and shade<sup>1</sup>. To tackle IPCD, we propose **IPCD-Net**, a deep learning model designed to address the aforementioned challenges (1) and (2). First, to address challenge (1), we extend IID models with point-wise feature aggregation based on the point-cloud model for other tasks [47], enabling effective processing of non-grid data. Specifically, we propose  $\text{IPCD-Net}_{\text{Base}}$ , inspired by PoInt-Net [49], which performs IID on RGB-D data via point-cloud representations but computes losses in image space. In contrast,  $\text{IPCD-Net}_{\text{Base}}$  performs IPCD entirely in point-cloud space, allowing direct optimization in the 3D geometric domain. Second, to address challenge (2), we introduce **Projection-based Luminance Distribution (PLD)**. PLD captures global-light cues by projecting the input point cloud from multiple viewpoints on a hemispherical surface and adopting its average luminance as a feature representation, as detailed in Fig. 3. To effectively integrate PLD features, we employ a hierarchical feature refinement that progressively estimates albedo and shade. Then, we propose **IPCD-Net**, which combines  $\text{IPCD-Net}_{\text{Base}}$  and PLD with hierarchical feature refinement to achieve robust IPCD performance.

Additionally, we create a synthetic dataset of colored point clouds with albedo and shade ground truths in outdoor scenes to assess IPCD performance. Based on this dataset, we verify the effectiveness of IPCD-Net, practicality in reducing cast shadows and improving shade color accuracy. Additionally, we verify the practicality of IPCD-Net in texture editing, relighting, and point-cloud registration under varying illumination. Finally, we verify the real-world applicability of IPCD-Net by applying it to real-world colored point clouds captured by a 3D scanner.

The main contributions are summarized as follows.

- We extend IID to colored point clouds and formalize this setting as **Intrinsic Point-Cloud Decomposition (IPCD)**, decomposing a colored point cloud into albedo and shade. To assess IPCD performance, we create an synthetic dataset of colored point clouds with albedo and shade ground truths.
- We propose **IPCD-Net**, an IPCD-specific model that incorporates point-wise feature aggregation to handle non-grid data. IPCD-Net also employs **Projection-based Luminance Distribution (PLD)** with a hierarchical feature refinement to capture global-light cues.
- We verify the practicality of IPCD-Net in texture editing, relighting, and registration under varying illumination. Additionally, we verify its real-world applicability.

<sup>1</sup>Note that, global-light direction is not provided as input, aligning with real-world scenarios where such information is typically unavailable.

## 2. Related works

This section summarizes point-cloud models, IID models, and light-direction estimation, which form the foundation of our IPCD approach.

**Point-cloud models.** Point-cloud models are widely explored in classification [16, 44] and semantic segmentation [46, 53] tasks. Early models employed shared-weight multilayer perceptrons (MLPs) to extract point-wise features [33]. Later, multi-scale grouping was introduced [34] and its further refined version by subsequent architectures [30, 35]. To better handle the non-grid structure of point clouds, PointCNN [21] introduced a grid-like convolutional approach. More recently, Point Transformer (PTv1) [54] leverages self-attention mechanisms [45] for enhanced feature interactions. Point Transformer version 2 (PTv2) [47] further optimized this by grouped vector attention. The recent model [48], enhances computational efficiency using serialized neighbor mapping. The aforementioned models assign per-point labels from colored point clouds, making them as baselines for IPCD, which requires estimating per-point albedo and shade.

**Intrinsic image decomposition (IID).** IID decomposes an image into albedo and shade. Early methods estimated only albedo, and then computed shade via image division [11, 57]. Later approaches simultaneously estimated albedo and shade [22], by employing shared feature extractors with respective decoders for both [24, 42, 56]. Recent work, PoInt-Net [49], performs IID using RGB-D data by leveraging geometric information through point-cloud representation. However, PoInt-Net is designed for image-based IID, not for point-cloud decomposition. Beyond single-image IID, multi-view methods, such as Neural Radiance Fields (NeRF) [28] and 3D Gaussian Splatting [18], have been explored for IID [23, 51, 58]. However, their direct application to IPCD is challenging due to occlusions in point-cloud rendering and per-scene training limitations, which hinder generalization across diverse scenes.

**Light-direct estimation.** The global-light feature is critical cue for accurate albedo and shade decomposition in IPCD. Traditional methods estimate light direction from a single image using shade cues [2] or weak signals from environmental elements [52]. Recent works explore the task from multiple images [8, 55], leveraging global-light consistency. However, these approaches are designed for image-based estimation and are not directly applicable to point clouds. To address this, we propose Projection-based Luminance Distribution (PLD), which captures the correlation between projected image brightness and global-light direction.

## 3. Methods

This section begins with the problem setting for IPCD. We then describe  $\text{IPCD-Net}_{\text{Base}}$ , an expanded version of PoInt-

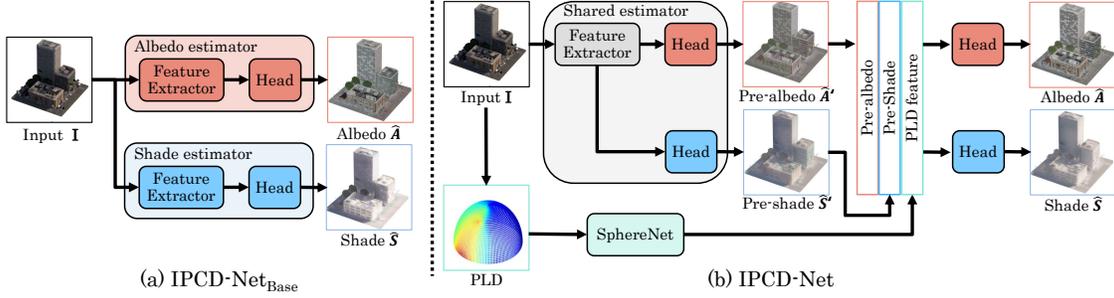


Figure 2. Architectures of (a) IPCD-Net<sub>Base</sub> and (b) IPCD-Net. IPCD-Net<sub>Base</sub> independently estimates albedo  $\hat{A}$  and shade  $\hat{S}$  from the colored point cloud  $I$  using separate albedo and shade estimators. IPCD-Net employs a shared feature extractor to process  $I$ , generating pre-albedo and pre-shade representations ( $\hat{A}'$ ,  $\hat{S}'$ ). The Projection-based Luminance Distribution (PLD) is processed by SphereNet to capture global-light features and enhance IPCD performance. The final albedo  $\hat{A}$  and shade  $\hat{S}$  are obtained by integrating PLD features with the pre-albedo and pre-shade representations through a hierarchical feature refinement.

Net [49] for point cloud decomposition. Finally, we introduce IPCD-Net, which incorporates PLD with a hierarchical feature refinement to enhance decomposition accuracy.

### 3.1. Problem setting

The IPCD task aims to estimate albedo  $\hat{A} \in \mathbb{R}^{N \times 3}$  and shade  $\hat{S} \in \mathbb{R}^{N \times 3}$  from a colored point cloud defined by color  $I \in \mathbb{R}^{N \times 3}$  and position  $P \in \mathbb{R}^{N \times 3}$ , where  $N$  represents the number of the points. Each point  $i$  consists of 3D position  $P_i = (x_i, y_i, z_i)$  and color  $I_i = (I_i^r, I_i^g, I_i^b)$ . The goal is to estimate albedo color  $\hat{A}_i = (\hat{A}_i^r, \hat{A}_i^g, \hat{A}_i^b)$  and shade color  $\hat{S}_i = (\hat{S}_i^r, \hat{S}_i^g, \hat{S}_i^b)$ . During training, the model takes input color  $I$  and position  $P$ , learning to estimate albedo  $\hat{A}$  and shade  $\hat{S}$  using ground-truth albedo  $A \in \mathbb{R}^{N \times 3}$  and shade  $S \in \mathbb{R}^{N \times 3}$ . Unlike NeRF-based approaches that require per-scene optimization, IPCD-Net can perform IPCD on arbitrary point clouds without requiring scene-specific fine-tuning. During inference, the model receives  $I$  and  $P$ , and estimates  $\hat{A}$  and  $\hat{S}$ . To verify the IPCD performance, we prepare IPCD dataset, containing multiple scene of point cloud pairs  $(P, I, A, S)$ , as described in Sec. 4.2. Note that global-light direction and illumination color are not provided as inputs, allowing IPCD models to be applied to real-world scenarios where such information is typically unavailable.

### 3.2. Preliminary: PoInt-Net

PoInt-Net [49] is an intrinsic image decomposition framework trained on RGB-D data  $(I_{\text{img}}, D_{\text{img}})$  paired with ground-truth light direction, shade image  $S_{\text{img}}$ , and albedo image  $A_{\text{img}}$ . This model consists of three modules; DirectionNet, Shader, and AlbedoNet. PoInt-Net first reconstructs colored point clouds from RGB-D input. Then, DirectionNet estimates the light direction, which is used alongside surface normals to compute rough shade. This rough shade is refined by Shader, producing the final shade points. AlbedoNet then estimates albedo points from the

input point cloud. The estimated albedo and shade point clouds are projected back into image space  $(\hat{A}_{\text{img}}, \hat{S}_{\text{img}})$ . We then compute the albedo image loss  $\mathcal{L}_{\text{img}}^{\text{alb}}$  and shade image loss  $\mathcal{L}_{\text{img}}^{\text{shd}}$  by comparing ground-truth images. Given the assumption of a Lambertian surface, the product of albedo image  $\hat{A}_{\text{img}}$  and shade image  $\hat{S}_{\text{img}}$  is expected to match input image color  $I_{\text{img}}$  [4]. Thus, the physical loss  $\mathcal{L}_{\text{img}}^{\text{phy}}$  [25] is also calculated. Additionally to these losses, PoInt-Net incorporates the gradient loss  $\mathcal{L}_{\text{img}}^{\text{gd}}$  and the color cross ratio loss  $\mathcal{L}_{\text{img}}^{\text{ccr}}$  to enforce smoothness and maintain color consistency. PoInt-Net follows a step-by-step learning strategy, where shade estimation modules are trained first, then albedo estimation is trained with shade parameters frozen. DirectionNet, Shader, and AlbedoNet employ MLP-based model [33].

### 3.3. Expansion for IPCD: IPCD-Net<sub>Base</sub>

PoInt-Net [49] addresses intrinsic image decomposition using RGB-D input. This model computes supervision and loss in 2D image space. While this model shows its effectiveness in image-based settings, this model design makes it unsuitable for direct use in the IPCD task, which requires decomposition in unstructured 3D point-cloud space. To bridge this gap, we extend PoInt-Net and implement **IPCD-Net<sub>Base</sub>**, which operates directly on colored point clouds with point-level supervision. The network consists of two primary modules: Shade Estimator and Albedo Estimator as illustrated in Fig. 2. The Shade Estimator predicts shade  $\hat{S}$  directly from the input point cloud  $(P, I)$ , bypassing the explicit light direction estimation in PoInt-Net due to the absence of ground-truth light direction. The Albedo Estimator then estimates albedo  $\hat{A}$ , following PoInt-Net’s AlbedoNet approach. A key difference is that PoInt-Net computes losses in image space, whereas IPCD-Net<sub>Base</sub> computes losses directly in point-cloud space as albedo loss  $\mathcal{L}_{\text{pnt}}^{\text{alb}}$ , shade loss  $\mathcal{L}_{\text{pnt}}^{\text{shd}}$ , and physical loss  $\mathcal{L}_{\text{pnt}}^{\text{phy}}$ .

$$\mathcal{L}_{\text{pnt}}^{\text{alb}} = \|A - \hat{A}\|_F, \quad (1)$$

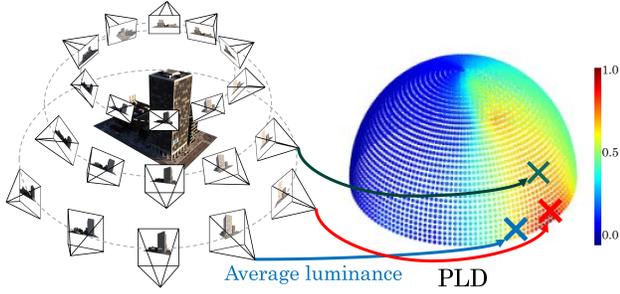


Figure 3. PLD overview. The input point cloud is first projected onto an image plane from a specific viewpoint. By systematically varying the projection angles and computing the average luminance across these views, the PLD is constructed and represented on a hemispherical surface. Projections aligned with the illumination direction result in brighter images with minimal shade, whereas projections from the opposite direction yield darker images dominated by shadowed regions. The color bar represents the luminance intensity from its corresponding view.

$$\mathcal{L}_{\text{pnt}}^{\text{shd}} = \|S - \hat{S}\|_F, \quad (2)$$

$$\mathcal{L}_{\text{pnt}}^{\text{phy}} = \|I - \hat{A} \odot \hat{S}\|_F, \quad (3)$$

where,  $\|\cdot\|_F$  and  $\odot$  represent the Frobenius norm and the Hadamard product, respectively. Due to the non-grid structure of point cloud, IPCD-Net<sub>Base</sub> does not utilize the gradient loss  $\mathcal{L}^{\text{gld}}$  or the color cross ratio loss  $\mathcal{L}^{\text{ccr}}$  employed in PoInt-Net. The training strategy follows PoInt-Net’s step-by-step learning scheme. IPCD-Net<sub>Base</sub> incorporates Ptv2 [47] instead of MLP-based model [33] for further effective feature extraction by grouped vector attention. Ptv2 is selected through preliminary validation in Sec. A.

### 3.4. Improvement for IPCD: IPCD-Net

While IPCD-Net<sub>Base</sub> enables direct decomposition on point clouds, it suffers from two issues: (i) color inconsistency in shade estimation and (ii) leakage between albedo and shade when trained independently. To address these, IPCD-Net introduces a *Projection-based Luminance Distribution (PLD)* to provide implicit global-light cues, and follows a *simultaneous training approach* with a *shared encoder*.

**Projection-based luminance distribution (PLD).** IPCD-Net<sub>Base</sub> often suffers from color inconsistency in shade estimation because it cannot account for global-light cues. In outdoor scenes, illumination color depends on the sun direction, but this information is unavailable in IPCD task setting. To address this, we design PLD to implicitly capture global-light cues by aggregating luminance from multiple viewpoints, as described in Fig. 3. First, we render  $256 \times 256$  images from  $K = 324$  uniformly sampled directions  $(\theta, \phi)$  over the upper hemisphere (e.g.,  $\theta \in 0^\circ, 10^\circ, \dots, 80^\circ, \phi \in 0^\circ, 10^\circ, \dots, 350^\circ$ ), using point clouds with about  $1.0 \times 10^6$  points within the field of view to ensure sufficient geometric fidelity. Each point cloud is

rotated with the matrix  $R(\theta, \phi)$ , while the camera position  $(0, 0, 1)$  and direction  $(0, 0, -1)$  are fixed. Next, we apply orthographic projection onto an image coordinates  $(u, v)$  to generate the projected image  $L(\theta, \phi; u, v)$ . Subsequently, PLD value  $\text{PLD}(\theta, \phi)$  for direction  $(\theta, \phi)$  is calculated as follows.

$$\text{PLD}(\theta, \phi) = \frac{1}{N_P} \sum_{u,v} L(\theta, \phi; u, v), \quad (4)$$

where  $N_P$  is the number of projected image pixels. Finally, we represent  $\text{PLD}(\theta, \phi)$  on a hemispherical surface. Projection in the illumination direction produces a brighter image with fewer shade, while projection in the opposite direction results in a darker image with dominant shade. Importantly, PLD does not assume a one-to-one mapping like “bright to direct sunlight” or “dark to shade”, instead, it aggregates multi-view luminance distributions to derive statistical global-light cues. For PLD processing, standard CNNs are not suitable since they introduce significant distortions when projecting spherical data onto a 2D grid, especially near the poles. Thus, IPCD-Net employs SphereNet [9], directly handling spherical data by adjusting the sampling grid to the sphere’s geometry. SphereNet ensures distortion invariance and preserves the spatial relationships inherent to spherical representations.

**Architecture.** Fig. 2 (b) illustrates the architecture of IPCD-Net, which improves over IPCD-Net<sub>Base</sub>. IPCD-Net<sub>Base</sub> suffers from large model parameters and unstable training. To address this, IPCD-Net adopts a shared encoder that reduces parameters and stabilizes training by avoiding redundant learning of correlated albedo and shade features. Thus, first, the input point cloud  $(P, I)$  is processed by the shared encoder, followed by respective heads to estimate pre-albedo  $\hat{A}'$  and pre-shade  $\hat{S}'$ . For natural embedding of global-light cues into both albedo and shade estimation, we employ a hierarchical feature refinement: intermediate predictions  $\hat{A}'$  and  $\hat{S}'$  are progressively refined with PLD features, allowing PLD to guide decomposition without overwhelming local cues. Second, the PLD feature  $L \in \mathbb{R}^3$  is expanded to match the point cloud size  $\mathbb{R}^{N \times 3}$  and concatenated with  $\hat{A}'$  and  $\hat{S}'$ . Finally, the concatenated features  $\in \mathbb{R}^{N \times 9}$  are processed by the albedo head and shade head to estimate final albedo  $\hat{A}$  and shade  $\hat{S}$ , respectively.

**Training strategy.** IPCD-Net<sub>Base</sub> employs a step-by-step learning strategy, learning shade first, followed by albedo estimation. However, IPCD-Net follows a simultaneous training approach for both albedo and shade to capture their dependencies more effectively.

**Losses.** IPCD-Net estimates pre-albedo  $\hat{A}'$ , albedo  $\hat{A}$ , pre-shade  $\hat{S}'$ , and shade  $\hat{S}$  from input point cloud  $I$ . We additionally define pre-albedo loss  $\mathcal{L}_{\text{pre}}^{\text{alb}}$ , pre-shade loss  $\mathcal{L}_{\text{pre}}^{\text{shd}}$ ,

and pre-physical loss  $\mathcal{L}_{\text{pre}}^{\text{phy}}$ .

$$\mathcal{L}_{\text{pre}}^{\text{alb}} = \|A - \hat{A}'\|_F. \quad (5)$$

$$\mathcal{L}_{\text{pre}}^{\text{shd}} = \|S - \hat{S}'\|_F. \quad (6)$$

$$\mathcal{L}_{\text{pre}}^{\text{phy}} = \|I - \hat{A}' \odot \hat{S}'\|_F, \quad (7)$$

Albedo loss  $\mathcal{L}^{\text{alb}}$ , shade loss  $\mathcal{L}^{\text{shd}}$ , and physical loss  $\mathcal{L}^{\text{phy}}$  are calculated like IPCD-Net<sub>Base</sub>. In summary, the total loss  $\mathcal{L}^{\text{tot}}$  is defined as follows.

$$\mathcal{L}^{\text{tot}} = \mathcal{L}_{\text{pnt}}^{\text{alb}} + \mathcal{L}_{\text{pnt}}^{\text{shd}} + \mathcal{L}_{\text{pnt}}^{\text{phy}} + \lambda (\mathcal{L}_{\text{pre}}^{\text{alb}} + \mathcal{L}_{\text{pre}}^{\text{shd}} + \mathcal{L}_{\text{pre}}^{\text{phy}}), \quad (8)$$

where  $\lambda$  is hyper parameters to balance these losses. Note that, all loss terms could be assigned independent hyper parameters to fine-tune their relative contributions. However, for simplicity and to avoid excessive hyper parameter tuning, we use a single  $\lambda$  to collectively scale the pre-estimation losses.

## 4. Experiments

This section starts with the experimental settings including compared models and implementation details in Sec. 4.1. We then introduce IPCD dataset in Sec. 4.2 and evaluate IPCD performance on the dataset with an ablation study in Sec. 4.3. To demonstrate the practical applications of IPCD models, we showcase texture editing, relighting, and point-cloud registration under varying illumination conditions in Sec. 4.4. Finally, we apply IPCD models trained on synthetic data to real-world scenes for assessing generalization.

### 4.1. Experimental setup

**Compared models.** We compare IPCD-Net against both traditional baselines and recent IID models. First, we set Baseline-A ( $\hat{A} = I, \hat{S} = 1$ ) and Baseline-S ( $\hat{S} = I, \hat{A} = 1$ ) as the simplest baselines, treating the input point cloud  $I$  directly as the estimated albedo or shade, following the typical IID evaluations [4]. Retinex [14] is evaluated by extending it to IPCD. Recent IID models, including NIID-Net [26], CD-IID<sup>2</sup> [5], IID-Anything<sup>2</sup> [7], and GS-IR [23], are applied to the IPCD dataset as references. Specifically, we render the point cloud into images, perform albedo and shade inference in 2D, then project the results back onto the point cloud. This pipeline may cause occlusion or artifacts, hence, it is not a strictly fair comparison to IPCD-Net, which operates directly on point clouds; however, we include these results as informative baselines.

**Implementation details.** All IPCD models are implemented using PyTorch [32] and trained on an NVIDIA H100 GPU. During training, 10,000 points are randomly

<sup>2</sup>Due to its unavailable training code, this model was not fine tuned.

Model	MSE( $10^{-2}$ )↓		MAE( $10^{-1}$ )↓		PSNR↑	
	alb	shd	alb	shd	alb	shd
Baseline-A [4]	18.9	29.1	3.58	4.27	7.57	5.96
Baseline-S [4]	30.4	25.1	5.01	4.23	5.22	6.39
Retinex [14]	19.0	25.3	3.59	4.28	7.55	6.25
NIID-Net [26]	15.2	12.1	2.93	2.46	8.97	9.99
CD-IID <sup>2</sup> [5]	17.3	27.8	3.36	4.45	8.00	5.88
IID-Anything <sup>2</sup> [7]	16.5	27.1	3.23	4.43	8.35	6.04
GS-IR [23]	24.8	13.7	4.41	2.78	6.16	8.91
IPCD-Net <sub>Base</sub>	4.02	5.11	1.58	1.62	14.0	13.5
IPCD-Net	<b>3.03</b>	<b>3.25</b>	<b>1.31</b>	<b>1.37</b>	<b>15.6</b>	<b>15.1</b>

Table 1. Numerical comparison of IPCD models. Recent IID models are limited in their performance due to dimensional reduction through rendering. IPCD models are trained in a supervised manner, leading to better performance than the rule-based and IID models. Especially, IPCD-Net demonstrated superior performance to IPCD-Net<sub>Base</sub> due to its incorporation of PLD, a shared feature extractor, and a hierarchical feature refinement.

sampled from each asset in every iteration to ensure diverse point coverage and address GPU memory limitations. Since the pre-estimation losses primarily serve as auxiliary supervision, we set  $\lambda$  to a lower weight as  $\lambda = 0.1$ . The implementations of IPCD-Net are based on PTv2 [47]<sup>3</sup> and SphereNet [9]<sup>4</sup>. PLD is computed with PyTorch3D renderer [38], by varying both elevation and depression angles in 2.5 degree increments. The point size  $\epsilon$  is set to 0.02.

### 4.2. Dataset preparation

To train and evaluate IPCD models, we introduce the IPCD dataset, a synthetic 3D dataset specifically designed for IPCD. We focused on the most common and practical case: outdoor clear-sky scenes with a few buildings of diffuse materials, illuminated by a sunlit directional light. The dataset construction process consists of the following steps:

1. Prepare surface meshes of synthetic outdoor scenes with albedo colors, and label them as “albedo”.
2. Simulate sunlight based on geographic location and three different time of day to illuminate the albedo meshes, labeling the results as “input”.
3. Remove albedo colors from the meshes and re-illuminate the meshes under identical conditions to obtain “shade”.
4. Export the meshes as point clouds by random sampling with 1,000,000 points on the mesh surface.

Consequently, the dataset includes 90 point-cloud sets derived from 30 unique assets. To prevent data leakage and ensure model generalizability, we divide the dataset into training and test sets, with 23 and 7 assets respectively. Sec. B provides more details on the dataset construction process.

<sup>3</sup>Github page: <https://github.com/Pointcept/PointTransformerV2>

<sup>4</sup>Github page: <https://github.com/ChiWeiHsiao/SphereNet-pytorch>

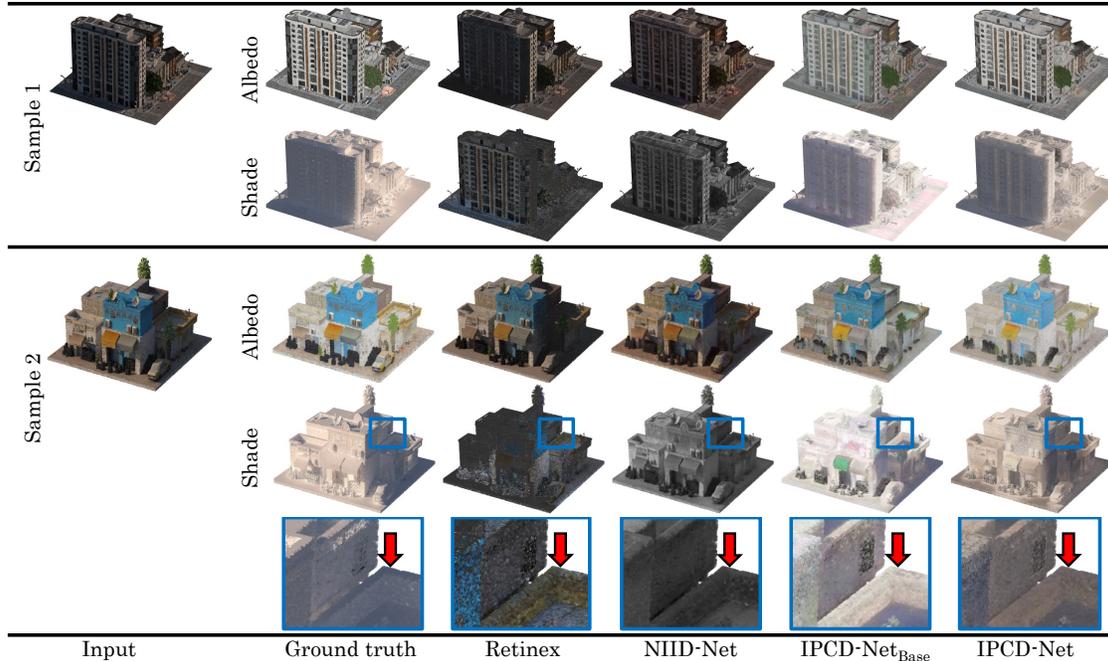


Figure 4. Visual comparison of IPCD performance. Retinex [14] exhibits residual cast shadows, while more recent models mitigate this issue. IID-Anything [7] performs IPCD through point-cloud rendering, but inconsistencies in albedo brightness across rendered images introduce noise, and dimensional reduction during rendering leads to remaining shade. Our IPCD models effectively separate albedo from shade. However, IPCD-Net<sub>Base</sub> lacks explicit light direction, causing some cast shadows to be omitted in shade, as indicated by red arrows. Conversely, IPCD-Net successfully estimates the shade color and cast shadows.

### 4.3. Results and discussion

**Quantitative results.** Tab. 1 summarizes the performance of IPCD-Net and the compared models. Baseline-A and Baseline-S serve as references, directly using the input point cloud as albedo and shade, respectively. Since Retinex primarily aims to reduce attached shadows by smoothing regions with similar neighboring colors, it is less effective on the IPCD dataset, where cast shadows dominate. Recent IID models, including CD-IID, IID-Anything, and GS-IR, apply IPCD through point-cloud rendering. Their performance is limited due to dimensional reduction during rendering, which loses geometric details crucial for intrinsic decomposition. In contrast, IPCD-Net<sub>Base</sub> estimates both albedo and shade directly in point-cloud space, achieves superior quantitative performance. IPCD-Net further enhances results by integrating PLD, a shared feature extractor, and a hierarchical feature refinement. The contributions of each component in IPCD-Net are analyzed in the ablation study.

**Visual results.** Fig. 4 provides qualitative comparisons of albedo and shade estimations by IPCD-Net and the compared models, complementing the quantitative results in Tab. 1. For clarity, we showcase representative examples from each category; additional results are provided in the supplementary materials. Retinex smooths regions of similar colors, producing albedo estimates that closely resemble the input point cloud, failing to remove cast shadows.

Model	MSE( $10^{-2}$ )↓		MAE( $10^{-1}$ )↓		PSNR↑	
	alb	shd	alb	shd	alb	shd
w/o PLD	3.00	3.46	1.30	1.40	15.6	14.9
w/o HFR+PLD	3.24	3.41	1.35	1.38	15.5	14.9
w/o share enc.	3.16	3.32	1.33	1.33	15.4	15.2
IPCD-Net	3.03	3.25	1.31	1.37	15.6	15.1

Table 2. Ablation study to evaluate the contribution of each component in IPCD-Net, including PLD, a shared encoder, and a hierarchical feature refinement (HFR).

Recent IID models exhibit inconsistencies in albedo brightness across rendered images, introducing noise, and dimensional reduction during rendering leads to residual shadows. IPCD-Net<sub>Base</sub> provides visually improved results due to its direct estimation in point-cloud space. However, without explicit global light direction input, IPCD-Net<sub>Base</sub> may omit some cast shadows in shade estimation as indicated by red arrows. Additionally, the lack of light direction input causes the shade color inaccuracy. In contrast, IPCD-Net maintains the original colors while effectively separating albedo from shade, demonstrating its robustness in IPCD. Additional examples are provided in Sec. C.

**Ablation study.** To evaluate the contribution of each component in IPCD-Net, we conducted an ablation study by selectively removing or modifying specific elements. Tab. 2

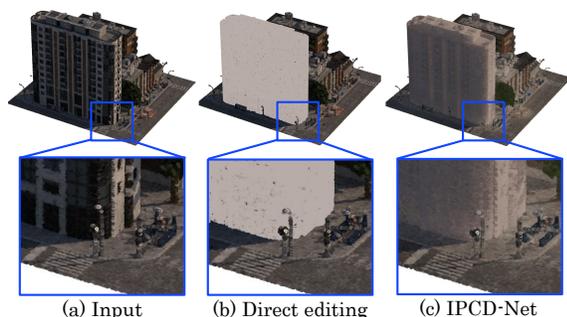


Figure 5. Visual result of texture editing with IPCD. (a) Input point cloud, (b) direct editing, and (c) editing through IPCD-Net are shown. Direct texture editing results in unnatural appearance due to lighting inconsistencies. Editing albedo separated by IPCD-Net, then recombining with shade, achieving natural results.

presents the quantitative impact of PLD, a shared encoder, and a hierarchical feature refinement. PLD contributes to shade estimation performance, especially for shade colors, since the illumination color of outdoor scenes varies with the angle of the sun. The hierarchical feature refinement with pre-albedo and pre-shade concatenation supports accurate albedo estimation from shade priors, hence, "w/o PLD" represents better albedo estimation than "w/o HFR+PLD". The shared-feature extractor captures a common feature representation for both albedo and shade to enhance these estimation performance. Further detailed ablation study is provided in Sec. D.

#### 4.4. Applications

**Texture editing.** Directly modifying point cloud colors often results in unnatural appearances due to the presence of embedded shade effects. To overcome this limitation, IPCD-Net decomposes point clouds into albedo and shade components, enabling albedo editing while preserving the original shade. This decomposition facilitates more natural texture modifications by ensuring that shade remains consistent even after color adjustments. To assess the effectiveness of this approach, we conducted a targeted color adjustment on building structures by altering specific regions of the input illumination  $I$  to a gray tone, referred to as direct editing. Additionally, we modified the underlying albedo  $\hat{A}$  while maintaining the shade  $\hat{S}$ , following IPCD-based editing. Fig. 5 compares direct color editing with IPCD-based editing, demonstrating that direct modifications introduce visual inconsistencies due to unintended shade alterations. In contrast, IPCD-based editing produces realistic textures. Unlike NeRF-based methods [23], which require per-scene optimization, IPCD-Net can perform texture editing across diverse point cloud datasets without requiring scene-specific fine-tuning.

**Relighting.** Altering illumination in colored point clouds without intrinsic decomposition often result in unrealistic appearances due to the entanglement of albedo and shade.

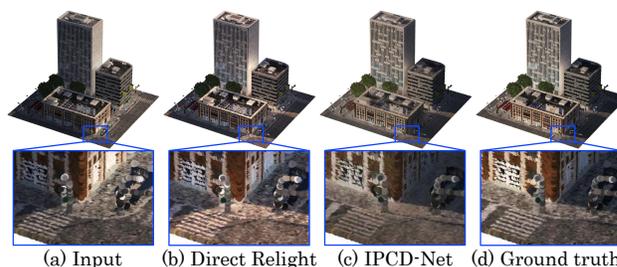


Figure 6. Visual results of relighting application, including (a) input, (b) direct relighting, (c) relighting through IPCD-Net, and (d) ground truth. Direct relighting led to remaining cast shadow of the original points, as indicated by red circle. IPCD models reduce cast shadows and relight points similar to ground truth.

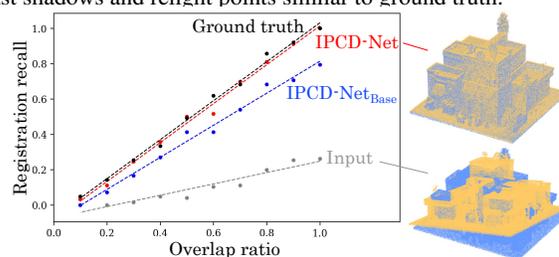


Figure 7. Point-cloud registration under varying illumination. Registration performance is compared using raw input colors and estimated albedos, by registration recall (higher is better). The overlap ratio of input point clouds is varied to assess performance across different registration difficulties. Registration accuracy improves when the estimated albedo closely aligns with the ground truth, mitigating the effects of illumination variations. IPCD-Net achieves registration performance comparable to the ground truth, demonstrating its robustness under varying lighting conditions.

The IPCD framework addresses this issue by decomposing albedo from shade, allowing for the simulation of diverse lighting conditions while preserving the intrinsic color properties of objects. To evaluate the effectiveness of this approach, we utilized the IPCD dataset, which provides albedo under three distinct lighting conditions. Specifically, we selected an input point cloud  $I_1$  consisting of albedo  $A_1$  and shade  $S_1$  and replaced the illumination condition from  $S_1$  to  $S_2$  for relighting. Direct relighting is performed by directly combining  $I_1$  with  $S_2$  due to the inherent challenges in separating albedo and shade. Instead, IPCD-Net first decomposes  $I_1$  into  $\hat{A}_1$  and  $\hat{S}_1$ , after which  $\hat{A}_1$  is recombined with  $S_2$  to generate the relighted point cloud  $\hat{I}_{1 \rightarrow 2}$ . As shown in Fig. 6, direct relighting results in residual cast shadows from the original shade, while IPCD-Net successfully reduces these artifacts and produces colors that closely resemble the ground truth. Notably, relighting through IPCD-Net also simulates accurate sunlit colors due to its precise decomposition. These findings demonstrate the effectiveness of IPCD in enabling realistic and flexible relighting of point clouds by separating albedo from shade.

**Point-cloud registration.** Aligning point clouds captured under varying lighting conditions presents significant chal-

allenges due to illumination-induced discrepancies. To address these inconsistencies, point-cloud registration should be performed using estimated albedo rather than directly relying on the original point clouds, which are affected by illumination variations. To evaluate the effectiveness of this approach, we conducted registration experiments on the IPCD dataset. Specifically, for each test asset, we obtained both the ground-truth albedo and the estimated albedo from IPCD models. Registration was then performed using pairs of albedo maps estimated from different sunlit times of the same asset, considering a total of 7 assets with 3 different time combinations. The alignment process was carried out using ColoredICP [31], varying the overlap ratio, and the registration accuracy was evaluated using registration recall, as illustrated in Fig. 7. When performing registration with the original input point clouds, variations in illumination conditions introduced inconsistencies that made the process more challenging. However, registration performance was significantly improved when using the albedo estimated by IPCD-Net<sub>Base</sub> and IPCD-Net. In particular, IPCD-Net achieved results comparable to the ground truth, demonstrating its robustness in aligning point clouds under varying illumination conditions.

**Real-world scenes.** IPCD-Net can perform IPCD on real-world scenes without requiring scene-specific fine-tuning even when trained exclusively on synthetic data. To demonstrate its robustness, we evaluated IPCD-Net trained on the IPCD dataset using real-world point clouds from the SensatUrban dataset [17], which consists of large-scale urban point clouds. Since the absolute values of ground truth albedo are difficult to obtain due to its inseparability from illumination, we conducted a quantitative evaluation with the relative reflectance between point pairs, following well-known method for real-world scenes [4, 41]. Specifically, we rendered 12 SensatUrban samples and manually annotated 900 point-pairs with relative reflectance labels (darker/lighter/equal). We then evaluated IPCD models using the f1 score [41], showing that IPCD-Net outperforms IPCD-Net<sub>Base</sub> and other models as depicted in Fig. 8. These results confirm IPCD-Net’s robustness and applicability across diverse real-world environments. We further validated our approach on STPLS3D [6], Toronto-3D [43], and KITTI [13] datasets, observing similar trends as detailed in Sec. H.

## 5. IPCD-Net limitation

IPCD-Net leverages PLD instead of explicit light direction, hence this approach has certain limitations. One challenge arises in scenes with highly heterogeneous reflectivity, such as a building with one highly reflective white surface and another with a low-reflective black surface. In such cases, the PLD may inaccurately estimate light direction due to the disproportionate influence of surfaces with strong re-

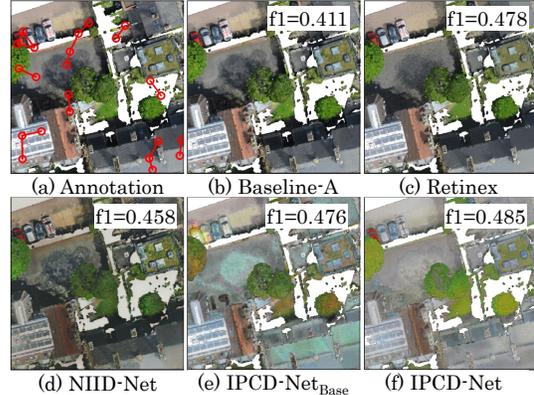


Figure 8. Illustration of annotation and comparison results for real-world scenes of SensatUrban dataset [17]. (a) shows the annotation of relative reflectance between point pairs indicated by red circles. We also showcase (b) Baseline-A, (c) Retinex, (d) NIID-Net, (e) IPCD-Net<sub>Base</sub>, and (f) IPCD-Net. We also show f1 score of each model for all 900 annotated point pairs. IPCD-Net achieves the best performance for visual and quantitative results, reducing cast shadows in real-world scenes.

flectivity differences. However, these extreme cases are uncommon and typically confined to assets with atypical structures, limiting the practical impact of this issue. Another limitation pertains to scenarios involving sparse or occluded point clouds. When background elements are unexpectedly revealed, PLD may struggle to compute accurate luminance values due to missing information. In such cases, leveraging denser point clouds or integrating PLD with inpainting techniques [37, 50] could enhance model performance. Finally, IPCD-Net achieves preferable results in outdoor clear-sky scenes with diffuse materials under directional lighting. However, its performance may degrade in the situations involving complex illumination, such as indoor environments with multiple light sources. Addressing these challenges may require incorporating more sophisticated lighting models or additional data modalities to capture the intricacies of such environments.

## 6. Conclusion

We introduced IPCD, a task for decomposing colored point clouds into albedo and shade. To achieve this, we developed IPCD-Net, employing point-wise feature aggregation and PLD with a hierarchical feature refinement. IPCD-Net significantly outperformed IPCD-Net<sub>Base</sub>, particularly in reducing cast shadows and improving shade color accuracy. It also demonstrated strong performance in texture editing, relighting, and point-cloud registration, with robustness validated on real-world scenes. As a first implementation, we adopted PTV2 as the feature extractor due to its effectiveness in point-cloud processing. However, IPCD-Net is adaptable to various point-cloud models, providing flexibility for further improvements. Future work will explore integrating advanced models [48] to enhance its performance.

## References

- [1] NavVis: Indoor Mapping and Digital Twin Technology. <https://www.navvis.com/>. 1
- [2] Jonathan T Barron and Jitendra Malik. Shape, Illumination, and Reflectance from Shading. In *TPAMI*, pages 1670–1687, 2015. 2
- [3] Harry Barrow, J Tenenbaum, A Hanson, and E Riseman. Recovering Intrinsic Scene Characteristics from Images. *Computer Vision Systems*, 2(2):3–26, 1978. 1
- [4] Sean Bell, Kavita Bala, and Noah Snavely. Intrinsic Images in The Wild. *ACM TOG*, 33(4):1–12, 2014. 3, 5, 8, 17
- [5] Chris Careaga and Yağız Aksoy. Colorful Diffuse Intrinsic Image Decomposition in The Wild. *ACM TOG*, 43(6):1–12, 2024. 1, 5, 13, 17
- [6] Meida Chen, Qingyong Hu, Zifan Yu, Hugues THOMAS, Andrew Feng, Yu Hou, Kyle McCullough, Fengbo Ren, and Lucio Soibelman. STPLS3D: A Large-Scale Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset. In *BMVC*, 2022. 8, 21, 22
- [7] Xi Chen, Sida Peng, Dongchen Yang, Yuan Liu, Bowen Pan, Chengfei Lv, and Xiaowei Zhou. IntrinsicAnything: Learning Diffusion Priors for Inverse Rendering under Unknown Illumination. In *ECCV*, pages 450–467, 2025. 1, 5, 6, 13, 17
- [8] JunYong Choi, SeokYeong Lee, Haesol Park, Seung-Won Jung, Ig-Jae Kim, and Junghyun Cho. MAIR: Multi-View Attention Inverse Rendering with 3D Spatially-Varying Lighting Estimation. In *CVPR*, pages 12900–12909, 2023. 2
- [9] Benjamin Coors, Alexandru Paul Condurache, and Andreas Geiger. SphereNet: Learning Spherical Representations for Detection and Classification in Omnidirectional Images. In *ECCV*, pages 518–533, 2018. 4, 5, 11
- [10] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *CVPR*, pages 5828–5839, 2017. 23
- [11] Qingnan Fan, Jiaolong Yang, Gang Hua, Baoquan Chen, and David Wipf. Revisiting Deep Intrinsic Image Decompositions. In *CVPR*, pages 8944–8952, 2018. 2
- [12] Jian Gao, Chun Gu, Youtian Lin, Zhihao Li, Hao Zhu, Xun Cao, Li Zhang, and Yao Yao. Relightable 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing. In *ECCV*, pages 73–89, 2024. 1
- [13] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In *CVPR*, pages 3354–3361, 2012. 8, 22
- [14] Roger Grosse, Micah K Johnson, Edward H Adelson, and William T Freeman. Ground Truth Dataset and Baseline Evaluations for Intrinsic Image Algorithms. In *ICCV*, pages 2335–2342, 2009. 5, 6, 17
- [15] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennis. Deep Learning for 3D Point Clouds: A Survey. *TPAMI*, 43(12):4338–4364, 2020. 1
- [16] Timo Hackel, N. Savinov, L. Ladicky, Jan D. Wegner, K. Schindler, and M. Pollefeys. SEMANTIC3D.NET: A New Large-Scale Point Cloud Classification Benchmark. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 91–98, 2017. 2
- [17] Qingyong Hu, Bo Yang, Sheikh Khalid, Wen Xiao, Niki Trigoni, and Andrew Markham. SensatUrban: Learning Semantics from Urban-Scale Photogrammetric Point Clouds. *IJCV*, 130(2):316–343, 2022. 8, 20, 22
- [18] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM TOG*, 42(4):139–1, 2023. 2, 13
- [19] Pileun Kim, Jingdao Chen, and Yong K Cho. SLAM-Driven Robotic Mapping and Registration of 3D Point Clouds. *Automation in Construction*, 89:38–48, 2018. 1
- [20] Edwin H. Land and John J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971. 1
- [21] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. PointCNN: Convolution on X-Transformed Points. In *NeurIPS*, 2018. 2
- [22] Zhengqi Li and Noah Snavely. Learning Intrinsic Image Decomposition from Watching The World. In *CVPR*, pages 9039–9048, 2018. 2, 11, 12
- [23] Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. GS-IR: 3D Gaussian Splatting for Inverse Rendering. In *CVPR*, pages 21644–21653, 2024. 2, 5, 7, 13, 17
- [24] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised Image-to-Image Translation Networks. In *NeurIPS*, 2017. 2
- [25] Yunfei Liu, Yu Li, Shaodi You, and Feng Lu. Unsupervised Learning for Intrinsic Image Decomposition from A Single Image. In *CVPR*, pages 3248–3257, 2020. 3
- [26] Jundan Luo, Zhaoyang Huang, Yijin Li, Xiaowei Zhou, Guofeng Zhang, and Hujun Bao. NIID-Net: Adapting Surface Normal Knowledge for Intrinsic Image Decomposition in Indoor Scenes. *IEEE TVCG*, 26(12):3434–3445, 2020. 5, 13, 17
- [27] Bilawal Mahmood, SangUk Han, and Dong-Eun Lee. BIM-based registration and localization of 3D point clouds of indoor scenes using geometric features for augmented reality. *Remote Sensing*, 12(14):2302, 2020. 1
- [28] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *Communications of the ACM*, 65(1):99–106, 2021. 2, 13
- [29] Juncheng Mu, Lin Bie, Shaoyi Du, and Yue Gao. ColorPCR: Color Point Cloud Registration with Multi-Stage Geometric-Color Fusion. In *CVPR*, pages 21061–21070, 2024. 1
- [30] Peiyuan Ni, Wenguang Zhang, Xiaoxiao Zhu, and Qixin Cao. PointNet++ Grasping: Learning An End-to-End Spatial Grasp Generation Algorithm from Sparse Point Clouds. In *ICRA*, pages 3619–3625, 2020. 2
- [31] Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Colored Point Cloud Registration Revisited. In *ICCV*, pages 143–152, 2017. 8, 19
- [32] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming

- Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *NeurIPS*, pages 8026–8037, 2019. 5
- [33] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In *CVPR*, pages 652–660, 2017. 2, 3, 4, 11, 12
- [34] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in A Metric Space. In *NeurIPS*, 2017. 2, 11, 12
- [35] Guocheng Qian, Yuchen Li, Houwen Peng, Jinjie Mai, Hasan Hammoud, Mohamed Elhoseiny, and Bernard Ghanem. PointNeXt: Revisiting PointNet++ with Improved Training and Scaling Strategies. In *NeurIPS*, pages 23192–23204, 2022. 2
- [36] Zheng Qin, Hao Yu, Changjian Wang, Yulan Guo, Yuxing Peng, and Kai Xu. Geometric Transformer for Fast and Robust Point Cloud Registration. In *CVPR*, pages 11143–11152, 2022. 19
- [37] Wentao Qu, Yuantian Shao, Lingwu Meng, Xiaoshui Huang, and Liang Xiao. A Conditional Denoising Diffusion Probabilistic Model for Point Cloud Upsampling. In *CVPR*, pages 20786–20795, 2024. 8
- [38] Nikhila Ravi, Jeremy Reizenstein, David Novotny, Taylor Gordon, Wan-Yen Lo, Justin Johnson, and Georgia Gkioxari. Accelerating 3D Deep Learning with PyTorch3D. *arXiv:2007.08501*, 2020. 5
- [39] Siyu Ren, Xiaodong Chen, Huaiyu Cai, Yi Wang, Haitao Liang, and Haotian Li. Color Point Cloud Registration Algorithm based on Hue. *Applied sciences*, 11(12):5431, 2021. 1
- [40] Manuele Sabbadin, Gianpaolo Palma, Francesco Banterle, Tamy Boubekour, and Paolo Cignoni. High Dynamic Range Point Clouds for Real-Time Relighting. In *Computer Graphics Forum*, pages 513–525, 2019. 1
- [41] Shogo Sato, Yasuhiro Yao, Taiga Yoshida, Takuhiro Kaneko, Shingo Ando, and Jun Shimamura. Unsupervised Intrinsic Image Decomposition with LiDAR Intensity. In *CVPR*, pages 13466–13475, 2023. 8
- [42] Kouki Seo, Yuma Kinoshita, and Hitoshi Kiya. Deep Retinex Network for Estimating Illumination Colors with Self-Supervised Learning. In *LifeTech*, pages 1–5, 2021. 2
- [43] Weikai Tan, Nannan Qin, Lingfei Ma, Ying Li, Jing Du, Guorong Cai, Ke Yang, and Jonathan Li. Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways. In *CVPRW*, pages 202–203, 2020. 8, 21, 22
- [44] Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Thanh Nguyen, and Sai-Kit Yeung. Revisiting Point Cloud Classification: A New Benchmark Dataset and Classification Model on Real-World Data. In *ICCV*, pages 1588–1597, 2019. 2
- [45] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In *NeurIPS*, 2017. 2
- [46] Lei Wang, Yuchun Huang, Yaolin Hou, Shenman Zhang, and Jie Shan. Graph Attention Convolution for Point Cloud Semantic Segmentation. In *CVPR*, pages 10296–10305, 2019. 2
- [47] Xiaoyang Wu, Yixing Lao, Li Jiang, Xihui Liu, and Hengshuang Zhao. Point Transformer V2: Grouped Vector Attention and Partition-Based Pooling. In *NeurIPS*, pages 33330–33342, 2022. 2, 4, 5, 11, 12
- [48] Xiaoyang Wu, Li Jiang, Peng-Shuai Wang, Zhijian Liu, Xihui Liu, Yu Qiao, Wanli Ouyang, Tong He, and Hengshuang Zhao. Point Transformer V3: Simpler, Faster, Stronger. In *CVPR*, 2024. 2, 8
- [49] Xiaoyan Xing, Konrad Groh, Sezer Karaoglu, and Theo Gevers. Intrinsic Appearance Decomposition Using Point Cloud Representation. In *ICCVW*, pages 4232–4236, 2023. 2, 3, 11
- [50] Lequan Yu, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. PU-Net: Point Cloud Upsampling Network. In *CVPR*, pages 2790–2799, 2018. 8
- [51] Kai Zhang, Fujun Luan, Qianqian Wang, Kavita Bala, and Noah Snavely. PhysSG: Inverse Rendering with Spherical Gaussians for Physics-Based Material Editing and Relighting. In *CVPR*, pages 5453–5462, 2021. 2
- [52] Yuxuan Zhang and Jean-Francois Lalonde. Light Source Estimation for Outdoor Images using A Prior on Sun Position. In *CVPR*, pages 9202–9211, 2020. 2
- [53] Yang Zhang, Zixiang Zhou, Philip David, Xiangyu Yue, Zerong Xi, Boqing Gong, and Hassan Foroosh. PolarNet: An Improved Grid Representation for Online LiDAR Point Clouds Semantic Segmentation. In *CVPR*, pages 9601–9610, 2020. 2
- [54] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip HS Torr, and Vladlen Koltun. Point Transformer. In *ICCV*, pages 16259–16268, 2021. 2, 11, 12
- [55] Yiqin Zhao, Sean Fanello, and Tian Guo. Multi-Camera Lighting Estimation for Photorealistic Front-Facing Mobile Augmented Reality. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, pages 68–73, 2023. 2
- [56] Hao Zhou, Xiang Yu, and David W Jacobs. GLoSH: Global-Local Spherical Harmonics for Intrinsic Image Decomposition. In *ICCV*, pages 7820–7829, 2019. 2
- [57] Tinghui Zhou, Philipp Krahenbuhl, and Alexei A Efros. Learning Data-Driven Reflectance Priors for Intrinsic Image Decomposition. In *ICCV*, pages 3469–3477, 2015. 2
- [58] Jingsen Zhu, Yuchi Huo, Qi Ye, Fujun Luan, Jifan Li, Dianbing Xi, Lisha Wang, Rui Tang, Wei Hua, Hujun Bao, et al. I2-SDF: Intrinsic Indoor Scene Reconstruction and Editing via Raytracing in Neural SDFs. In *CVPR*, pages 12489–12498, 2023. 2