



Virtual home staging and relighting from a single panorama under natural illumination

Guanzhou Ji¹ · Azadeh O. Sawyer¹ · Srinivasa G. Narasimhan²

Received: 22 February 2024 / Revised: 19 May 2024 / Accepted: 21 May 2024 / Published online: 11 July 2024
© The Author(s) 2024

Abstract

Virtual staging technique can digitally showcase a variety of real-world scenes. However, relighting indoor scenes from a single image is challenging due to unknown scene geometry, material properties, and outdoor spatially-varying lighting. In this study, we use the High Dynamic Range (HDR) technique to capture an indoor panorama and its paired outdoor hemispherical photograph, and we develop a novel inverse rendering approach for scene relighting and editing. Our method consists of four key components: (1) panoramic furniture detection and removal, (2) automatic floor layout design, (3) global rendering with scene geometry, new furniture objects, and the real-time outdoor photograph, and (4) virtual staging with new camera position, outdoor illumination, scene texture, and electrical light. The results demonstrate that a single indoor panorama can be used to generate high-quality virtual scenes under new environmental conditions. Additionally, we contribute a new calibrated HDR (Cali-HDR) dataset that consists of 137 paired indoor and outdoor photographs. The animation for virtual rendered scenes is available [here](#).

Keywords HDR photography · Photometric calibration · Scene editing · Global illumination · Inverse rendering

1 Introduction

Virtual home staging is crucial in the real estate industry, enabling both the public and real estate agents to assess properties conveniently. Particularly since the COVID-19 pandemic, the use of virtual home tours in the housing market has increased significantly due to work-from-home restrictions [72]. The increasing popularity of omnidirectional cameras has driven increasing research interests in panoramic photography in recent years, and a single panorama provides a complete representation of the surrounding context.

The High Dynamic Range (HDR) technique captures a wide range of pixel information from the real world. The

global rendering method proposed by Debevec [8] offers an image-based rendering model for relighting virtual objects using HDR photographs. Previous studies have focused on estimating 360° HDR environment map directly from Low Dynamic Range (LDR) images for relighting new virtual objects [15, 16, 18]. However, these data-driven approaches often assume linear proportionality between pixel values and scene radiance without considering photometric calibration. Bolduc et al. [2] recently conducted a study that calibrated an existing panoramic HDR dataset with approximate scene luminance levels. In our work, we take this a step further by calibrating the captured HDR panoramas using absolute luminance value (in SI units) measured in each scene. The actual brightness of a scene, measured in luminance (cd/m^2), accurately reflects the light properties in the real world. This calibration ensures that our HDR images accurately represent realistic spatially-varying lighting conditions, distinguishing them from existing indoor panorama datasets [6, 60, 69].

Panoramic images introduce unique challenges for scene understanding due to the distortion caused by equirectangular projection. When dealing with scenes that contain furniture objects, the complexities of 3D scene reconstruction are further amplified. Recent studies on indoor furniture inpainting focus on furniture removal from 2D perspective images [35,

✉ Guanzhou Ji
gji@andrew.cmu.edu

Azadeh O. Sawyer
asawyer@andrew.cmu.edu

Srinivasa G. Narasimhan
srinivas@andrew.cmu.edu

¹ School of Architecture, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

² Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA



Fig. 1 Illustration of Our Rendering Approach: A real-world scene is captured at 11:27 AM (cloudy sky condition). Different furniture objects are removed from the existing room, and the empty scene is refurbished

with new virtual furniture objects [13]. A new panorama is then virtually rendered with edited camera position, outdoor illumination at 01:00 PM (clear sky condition), scene textures, and electrical light

54]. Directly applying these inpainting techniques to furnished panoramas can result in geometric inconsistencies within indoor surfaces. To directly remove furniture objects from a single panorama, our method first segments the panorama into multiple 2D perspective images to identify furniture objects. It then uses 3D indoor planes to guide the inpainting process.

Existing studies on inserting small objects into 2D perspective images have focused on different contexts, such as virtual scenes [59] and real-world scenes [7, 15, 40]. However, there is a lack of automated methods for directly inserting virtual furniture objects into panoramas. We introduce a rule-based layout design that identifies floor boundaries and allows precise placement of new furniture objects through geometric transformations and spatial parameters.

Our inverse rendering approach aims to directly estimate physical attributes, such as 3D geometry, material properties, and illumination, from a single panorama. We focus on realistic virtual staging for indoor scenes and develop a rendering approach that includes a detailed scene model, surface materials, and outdoor spatially-varying light. Our rendering pipeline allows flexible scene editing and generates the new virtual scene with global illumination (Fig. 1).

We demonstrate a complete pipeline that includes data collection, scene design, and virtual home staging to showcase real-world indoor scenes (Fig. 2). This paper is the extended version of our previous work [32]. We further develop the algorithms for editing the outdoor scene and integrate the design applications into our rendering pipeline. Overall, our work makes the following technical contributions:

- (1) A method for calibrating indoor-outdoor HDR photographs, along with a new calibrated HDR (Cali-HDR) dataset that includes 137 scenes.
- (2) An image inpainting method that detects and removes furniture objects from a panorama.
- (3) A rule-based layout design for positioning multiple furniture objects on the floor based on spatial parameters.

- (4) The outdoor sun position is accurately edited during the day, incorporating the full-spectral sky appearance.
- (5) Virtual staging is achieved through the integration of new camera positions, scene texture, and electrical light.

2 Related work

2.1 HDR and photometric calibration

The dynamic range of radiances in a real-world scene spans from 10^{-3} cd/m^2 (starlight) to 10^5 cd/m^2 (sunlight) [48]. In the context of a 2D perspective image, some studies have focused on predicting panoramic HDR environment maps [15], lighting representation [16], and estimating HDR panoramas from LDR images [18]. Considering that HDR images reflect the relative luminance values from the real world, absolute luminance measurement is required for on-site HDR photography to recover scene radiance [10]. To display the absolute luminance value, the captured HDR image requires photometric calibration, which is a means of radiometric self-calibration [42]. Reference planes, such as a matte ColorChecker chart or gray cards, should be positioned within the scene for luminance measurement [43].

2.2 Indoor light estimation

Previous studies on indoor lighting estimation have explored indoor lighting editing [40], material property estimation [63], and the recovery of spatially-varying lighting [17, 39, 51] from a 2D image. Following the global rendering method [9], some studies aim to estimate a 360° indoor HDR environment map from a 2D image and subsequently render the virtual objects [15, 37]. User inputs, such as annotating indoor planes and light sources, have also been utilized to assist scene relighting and object insertion [33]. Zhi et al. focus on the decomposition of light effects in the empty panoramas [70]. While previous studies have extensively focused on

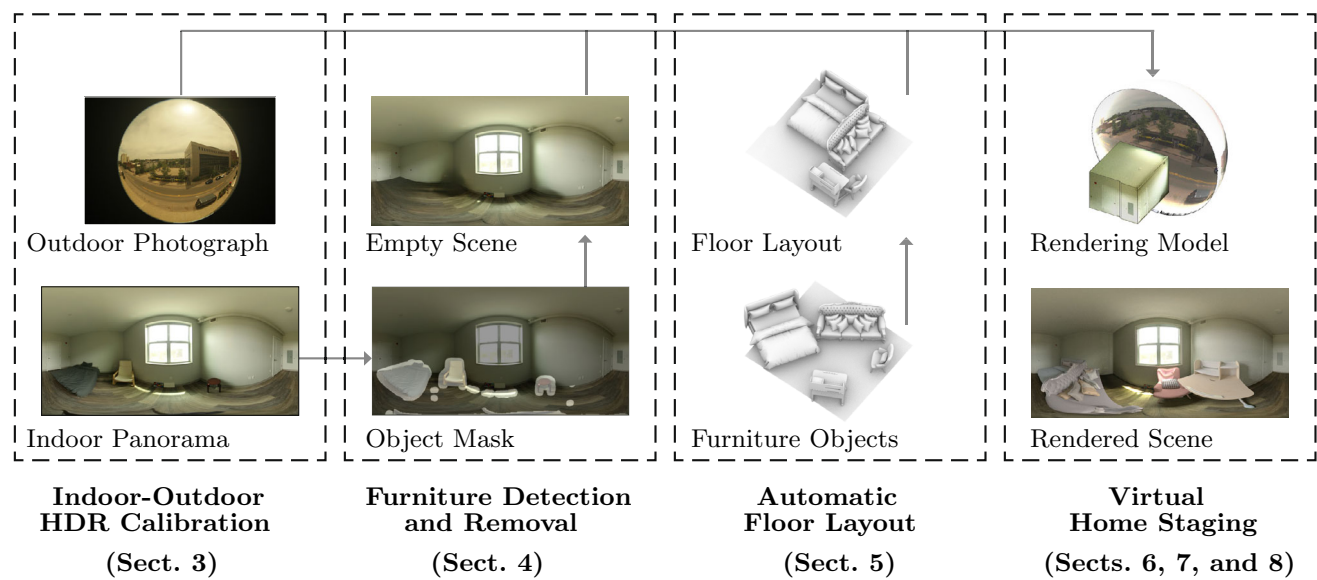


Fig. 2 Our rendering pipeline consists of four modules: Indoor-Outdoor HDR Calibration (Sect. 3) calibrates the captured indoor and outdoor HDR photographs with measured absolute luminance values. Furniture Detection and Removal (Sect. 4) identifies and removes the target fur-

niture objects from the scene. Automatic Floor Layout (Sect. 5) allows the automatic placement of multiple furniture objects. Virtual Home Staging (Sects. 6, 7, and 8) achieves high-quality virtual scenes

global light estimation and 3D object insertion, there is limited research on panoramic global rendering under real-time outdoor illumination.

2.3 Outdoor light estimation

Outdoor illumination includes direct sunlight and sky light. The all-weather sky model proposed by Perez et al. [47] visualizes relative sky luminance at different times. Later, the full spectral sky model achieves color change under different atmospheric conditions [25]. However, the virtual sky model cannot reconstruct the complete outdoor scene content for outdoor illumination, lacking cloud details and surrounding context in the rendered sky image. Given a single LDR image, some studies [24, 68] estimate a 360° HDR environment map as spatially-varying light for inverse rendering and object insertion. Meanwhile, other studies [11, 22, 23, 36, 49] focus on estimating outdoor illumination from multiple images.

2.4 Panoramic furniture removal

The conventional image inpainting method assumes a nearly planar background around the target object, making it unsuitable for indoor scenes with complex 3D room structures. For the case of indoor scenes, the existing state-of-the-art inpainting models, such as LaMa [54], cannot recognize the global structure, including the boundaries of walls, ceilings, and floors. Several approaches have been attempted

to address this challenge: (1) utilizing lighting and geometry constraints [67], (2) using planar surfaces to approximate contextual geometry [26, 34, 35], and (3) estimating an empty 3D room geometry from furnished scenes [66]. These studies have primarily focused on furniture detection and inpainting tasks for 2D perspective images. Panoramic furniture removal needs furniture detection [21], spherical semantic segmentation [65], and image inpainting. Recent studies have focused on furniture removal tasks in both virtual panoramas [19, 20] and real-world panoramas [50].

2.5 3D layout estimation

Estimating a 3D room layout from a single image is a common task for indoor scene understanding. While indoor panorama can be converted into a cubic map [4], the actual 3D layout is oversimplified. Building on this cube map approach, other studies [55, 56] focus on panorama depth estimation using 3D point clouds. Moreover, under the Manhattan world assumption [5], a 360° room layout with separated planar surfaces can be segmented from a single panorama [58, 62, 69, 73]. Moving beyond 3D room layout, detailed scene and furniture geometry can be reconstructed from 2D perspective images [27, 30, 45]. Additionally, when provided with a 2D floor plan, indoor space semantics and topology representations can be generated to create a 3D model [61] and recognize elements in floor layouts [64]. An accurate room geometry allows new furniture objects to be inserted precisely into the existing scene.

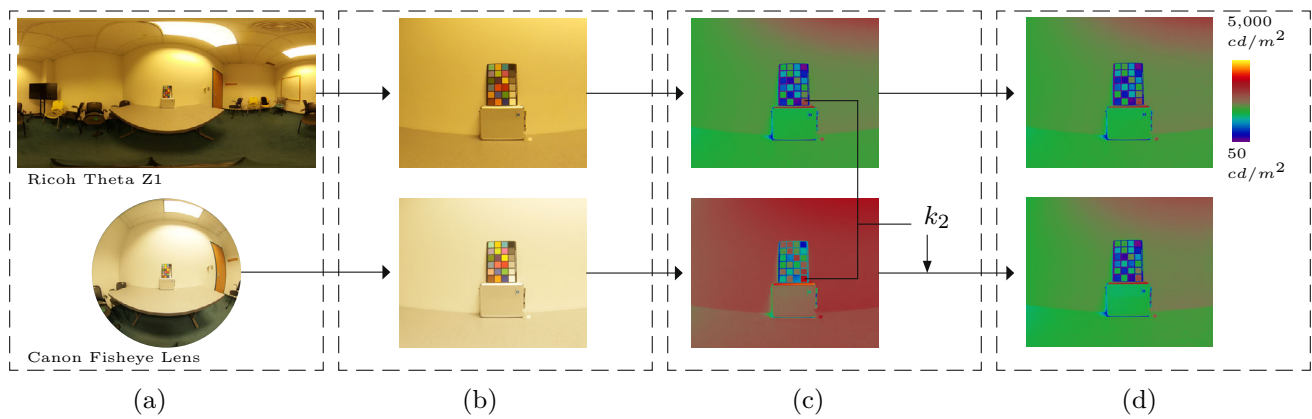


Fig. 3 Photometric Calibration between Two Cameras: (a) Photographs from two cameras. (b) Cropped target regions. (c) Original luminance maps. (d) Luminance maps after scaling the HDR image captured by the fisheye lens with k_2

3 Indoor-outdoor HDR calibration

3.1 Indoor HDR calibration

For indoor scenes, a Ricoh Theta Z1 camera was positioned in the room to capture panoramic HDR photographs. The camera settings were configured as follows: White Balance (Daylight 6500), ISO (100), Aperture (F/5.6), Image Size (6720 x 3360), and Shutter Speed (4, 1, 1/4, 1/15, 1/60, 1/250, 1/1000, 1/4000, 1/8000). To ensure consistency and avoid motion blur during photography, the camera was fixed on a tripod at a height of 1.6m. We placed a Konica Minolta LS-160 luminance meter next to the camera to measure the target luminance on a white matte board. Each HDR photograph needs per-pixel calibration to accurately display luminance values for the scene. The measured absolute luminance value at the selected point is recorded in SI unit (cd/m^2).

As shown in Eq. (1), the luminance value of each pixel is based on CIE XYZ values, based on the standard color space (sRGB) [52] and CIE Standard Illuminants D_{65} . According to the study by Inanici [29], indoor scene luminance (L_i) (cd/m^2) is expressed as:

$$L_i = k_1 \cdot (0.2127 \cdot R + 0.7151 \cdot G + 0.0722 \cdot B) \quad (1)$$

where R , G , and B values are three color channels (RGB) in the captured indoor HDR image. The measured luminance value and the displayed luminance value from the original HDR image are used to calculate the calibration factor (k_1).

3.2 Outdoor HDR calibration

To capture outdoor scenes, a Canon EF 8-15 mm f/4L fisheye lens was installed on a Canon EOS 5D Mark II Full Frame DSLR Camera, and a 3.0 Neutral Density (ND) filter was utilized for capturing direct sunlight with HDR technique [53]. The camera settings were configured as follows: White Bal-

ance (Daylight 6500), ISO (200), Aperture (F/16), Image Size (5616 x 3744), and Shutter Speed (4, 1, 1/4, 1/15, 1/60, 1/250, 1/1000, 1/4000, 1/8000). Due to the diverse outdoor contexts, it is impractical to place a target plane to measure target luminance values.

Each camera has its own fixed camera response curve to merge multiple images with varying exposures into one single HDR image. Rather than performing a separate calibration process for outdoor HDR, our objective is to determine a fixed calibration factor between two distinct cameras and calibrate the outdoor HDR images with indoor luminance measurement. As shown in Fig. 3, we positioned two cameras in an enclosed room under consistent electrical lighting. Following the camera settings of indoor and outdoor HDR photography (Sect. 3), we captured the target Macbeth ColorChecker chart from two cameras, respectively. Then, 2D perspective images displaying the same target regions were cropped from the original images. After merging the two sets of images into HDR photographs, we calculated the difference ratio (k_2) between the target pixel region (white patch) on the HDR photographs obtained from the two cameras. Ultimately, the HDR image captured by the Canon EOS 5D Camera was linearly calibrated with the computed constant value (k_2), and the HDR photographs from the two cameras were calibrated to display the same luminance range. k_2 is a fixed constant when the two camera settings stay the same.

As shown in Eq. (2), outdoor scene luminance (L_o) (cd/m^2) is expressed as:

$$L_o = k_1 \cdot k_2 \cdot (0.2127 \cdot R + 0.7151 \cdot G + 0.0722 \cdot B) \quad (2)$$

where R , G , and B values are three color channels (RGB) in the captured outdoor HDR image. k_1 is the calibration factor determined by the measured luminance target value and the displayed luminance value in the captured indoor HDR image, and k_2 is the computed constant for scaling the outdoor hemispherical image into the indoor panorama.

After linear rescaling, the outdoor HDR photographs are processed through the following steps: (1) vignetting correction that compensates for the light loss in the periphery area caused by the fisheye lens [29], (2) color correction for chromatic changes introduced by ND filter [53], and (3) geometric transformation from equi-distant to hemispherical fisheye image for environment mapping [28].

4 Furniture detection and removal

4.1 Panoramic furniture detection

A single panorama displayed in 2D image coordinates can be transformed into a 3D spherical representation [1, 58], and this process can also be inverted. Building on this concept, our objective is to convert a panorama into a list of 2D images for scene segmentation. Subsequently, we aim to reconstruct the panorama where target furniture objects are highlighted. The selected region on the input panorama I_p is geometrically cropped and transformed into a 2D perspective image, within longitude angle (θ) and latitude angle (ϕ). $\theta \in (-\pi, +\pi)$ and $\phi \in (-0.5\pi, +0.5\pi)$. With the fixed field of view (FOV) and the image dimension of height (h) by width (w), we obtain 2D perspective image set $I = \{I_1, I_2, I_3, \dots, I_i\}$, and the process of equirectangular-to-perspective can be expressed as mapping function S (Eq. (3)):

$$I_i = S(I_p; FOV, \theta, \phi, h, w) \quad (3)$$

After scene segmentation for 2D perspective images, a set of processed images $I' = \{I'_1, I'_2, I'_3, \dots, I'_i\}$ is stitched back to reconstruct a new panorama according to annotated θ and ϕ . The invertible mapping process enables image transformation between equirectangular and 2D perspective representations.

As shown in Fig. 4 (a), one single panorama is segmented into a set of 2D perspective images and segmented per color scheme in semantic segmentation classes [71]. A 3D layout is estimated with separated planer surfaces of the ceiling, wall, and floor textures. The rendering model generates an indoor mask to distinguish the floor and other interior surfaces, and the result highlights the furniture object placed on the floor (Fig. 4 (b)). The target mask includes tripod location, direct sunlight region, and furniture areas (Fig. 4 (c)).

4.2 Furniture removal

For furnished panoramas, we first estimate the 3D room geometry [58] and utilize the indoor planar information in the panoramas to guide the inpainting process. As shown in Fig. 5, our method allows for image inpainting on the original furnished panoramas with surrounding context, while

utilizing the floor boundary as a guiding reference to preserve clear indoor boundaries. One challenge in inpainting the floor texture is when the masked region is distant from nearby pixels, leading to blurring and noise. Unlike walls and ceilings, the floor texture often exhibits a strong pattern with various textures. Thus, we address this issue by treating the floor texture in the indoor scenes as a Near-Periodic Pattern (NPP). Compared to LaMa [54], which is trained on existing 2D image datasets, the NPP model developed by Chen et al. [3] learns the masked region from the provided image. This results in outputs that are optimized based on the content of the input image itself. As demonstrated in Fig. 5, our approach, combined with the LaMa [54] and NPP models [3], effectively recovers the scene context around the detected furniture area. The restored indoor textures, including ceiling, wall and floors, will be incorporated into the 3D rendering model.

5 Automatic floor layout

The indoor furniture layout follows the user's preferences and daily activities. To automate furniture placement, we develop a rule-based method that automatically positions furniture objects on the floor plane using a set of predefined spatial parameters and floor geometry. The location of the window and the corresponding floor edge are known in the estimated floor layout. For the floor geometry, we compute and determine the longest side of the floor as the reference floor edge. The window edge and the reference floor edge are used to precisely guide the placement of furniture on the floor plane.

For each scene, we segment the floor mesh from the captured panorama, and the orientation of each object is determined based on whether it faces the window or indoor walls. For the translation distance, we normalize the distance between the object's dimension and the floor boundary to a range between 0 and 1. This normalization allows the object to be precisely positioned along the wall and window side. Different spatial parameters and orientation combinations can express alternative floor layouts. The rule-based method adapts to various layout rules by recognizing different floor boundaries and placing target objects accordingly within different indoor scenes.

Within the 3D coordinate system, the segmented floor mesh and furniture objects are positioned on the xy plane (Fig. 6). Each furniture object can be represented as a set of point clouds. The task of floor layout design is subject to the constraint of the floor boundary. Each furniture object rotates around the z axis to align with the target floor edge and translates to the designated position, denoted by the valid translation distances t_x and t_y . We transform the 3D point set \mathbf{x}_i to its corresponding transformed point \mathbf{x}'_i in the xy plane,

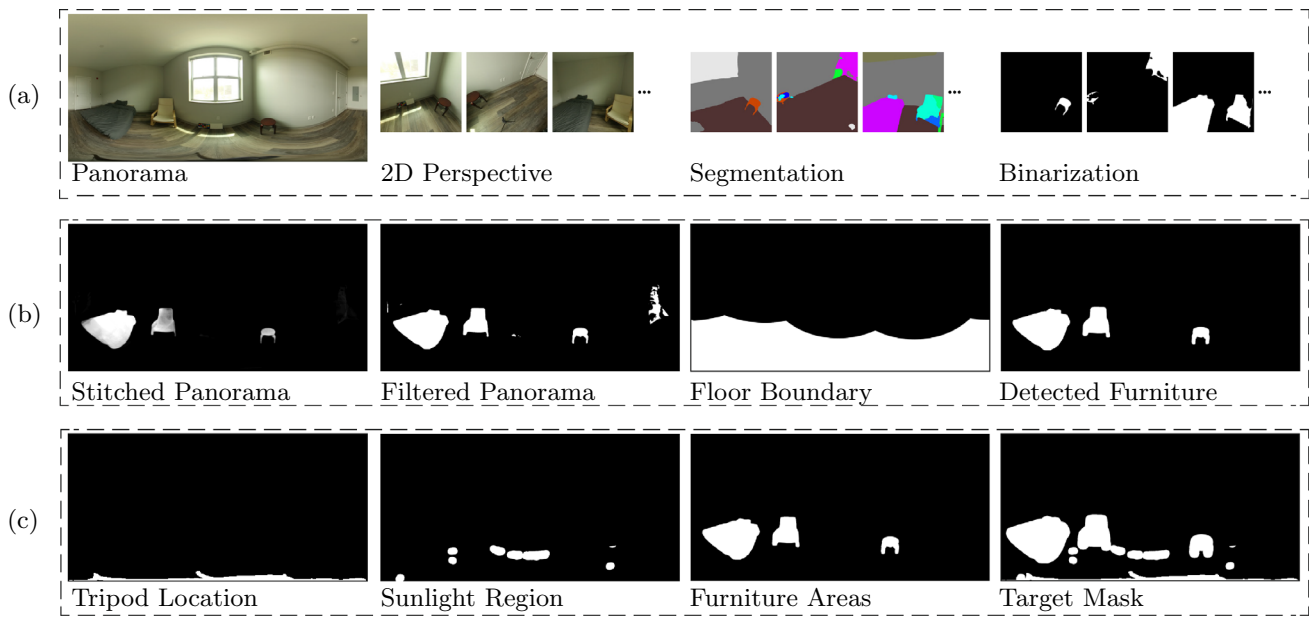


Fig. 4 Panoramic Scene Segmentation: (a) A single panorama is segmented into a set of 2D perspective images, and target furniture objects are detected. (b) The stitched panorama is processed to display furniture contours, and the rendered floor boundary is utilized to filter out solid

contours that are not attached to the floor area. (c) Estimated tripod location [70], direct sunlight region, and the detected furniture areas are combined as the target mask

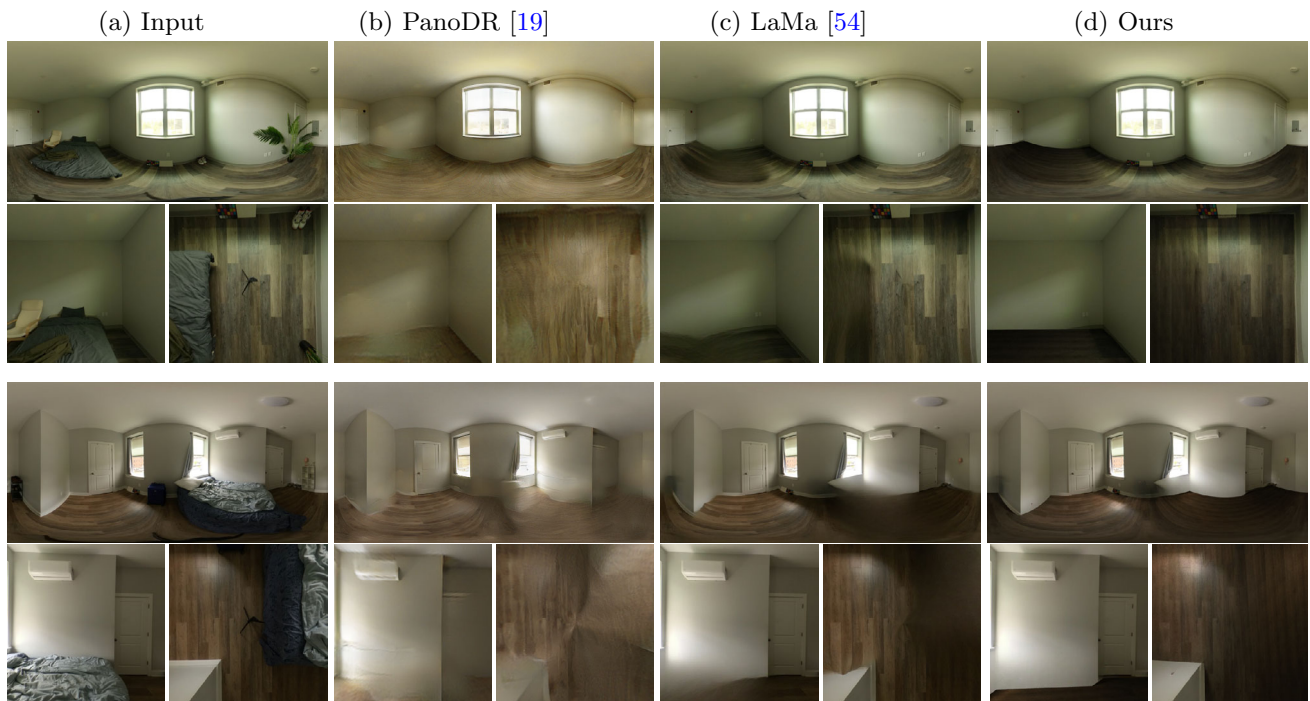


Fig. 5 Comparison of Image Inpainting Methods: The target mask (from Fig. 4 (c)) is paired with input panorama (a) to remove the target region using PanoDR [19] (b), LaMa [54] (c), and our method (d), respectively



Fig. 6 Furniture Layout Alternatives: Given an empty floor mesh, multiple furniture objects are placed on the floor with predefined positions and orientations

by applying the rotation matrix and the translation matrix: $\mathbf{x}_i' = R_z(\theta)\mathbf{x}_i + t$, where $t = [t_x \ t_y \ 0]^T$.

The detailed process of furniture arrangement is expressed in Algorithm 1. The inputs for furniture arrangement include geometry inputs (Floor Mesh (A), Furniture Object (B)) and spatial parameters (Distance to window (d_x), Distance to wall (d_y), Orientation (α)). The algorithm first obtains geometric information from the target floor mesh and 3D furniture object, then rotates and translates a single furniture object to the target location. When there are multiple furniture objects, each object is placed according to its corresponding spatial parameters.

6 Indoor virtual staging

The appearance of an indoor scene depends on various factors, such as indoor and outdoor illumination, room geometry, and materials. Following the Lambertian assumption in previous studies [12, 14, 38, 44], the indoor surfaces' Bidirectional Reflectance Distribution Function (BRDF) is assumed to be constant in all directions.

Algorithm 1 Furniture Arrangement

Require: Floor Mesh (A), Furniture Object (B)

Require: Distance to window (d_x), Distance to wall (d_y), Orientation (α)

Phase 1 - Read Floor Mesh

- 1: **Input:** A
- 2: get the floor vertices in the xy plane
- 3: compute the Euclidean distance between adjacent vertices
- 4: **Output:** the longest floor edge (L), the floor edge with window (W)

Phase 2 - Read Furniture Object

- 5: **Input:** B
- 6: use a bounding box to get the vertices in the xy plane
- 7: compute the Euclidean distance between adjacent vertices
- 8: **Output:** furniture length (l), furniture width (w)

Phase 3 - Geometric Transformation

- 9: **Input:** B, d_x, d_y, α
 - 10: rotate B with α around the z axis to adjust orientation
 - 11: use floor dimension (L, W) and furniture dimension (l, w) to compute valid translation distance
 - 12: use parameter inputs (d_x, d_y) to compute the target translation distance
 - 13: **Output:** translated B'
-

As shown in Eq. (4), a real-world image (I) can be represented as the pixel-wise product of reflectance layer $R(I)$ and shading layer $S(I)$.

$$I = R(I) \odot S(I) \quad (4)$$

In our work, we capture the scene under natural illumination, where the outdoor image serves as the only light source. The indoor panorama is used to estimate the 3D room geometry, paired with segmented surface textures. For each scene, the outdoor image and room geometry are used to render the shading layer $S(I)$ and compute the corresponding reflectance layer $R(I)$ for surface textures.

We build up a physics-based rendering pipeline in Mitsuba platform [31]. Various real-world scenes are refurbished with virtual furniture objects (Fig. 7). Compared to previous scene relighting and object insertion approaches [15, 39, 70], our proposed rendering method integrates complete 3D scene geometry (including both room geometry and furniture objects), outdoor environment map, and material textures, allowing the new furniture objects to be virtually rendered within the scene.

Our rendering approach not only accurately renders the virtual objects but also reconstructs the inter-reflection between the scene and newly inserted objects. It is important to note that as the scene geometry is approximated into individual planar surfaces, certain indoor details, such as curtains or window frames, are simplified as planar surfaces in the rendering model. Overall, our rendering pipeline effectively generates high-quality indoor panoramas while preserving the essential characteristics of real-world scenes.

We compared our rendering method with the conventional global rendering approach for scene relighting (Fig. 8). Meanwhile, a single LDR image is used to generate the corresponding HDR image by Liu et al.'s method [41], and the rendered results are then compared with those obtained from calibrated HDR images.

First, an indoor LDR panorama is used to estimate HDR panorama to relight the virtual furniture objects (Fig. 8(a)), and the objects are rendered with our calibrated indoor HDR panoramas (Fig. 8(b)). While the furniture objects can both be rendered within the indoor context, the conventional method fails to recover the direct illumination for the newly inserted object, such as the bed in Room 2, and the inter-reflection between the new objects and the existing scene is missed. Second, we use one single outdoor image to estimate its HDR image to render the indoor scene (Fig. 8(c)), and our calibrated outdoor HDR fisheye image is used as the light source (Fig. 8(d)). The estimated HDR image (Fig. 8(c)) lacks photometric calibration, necessitating multiple searches for optimal scaling factor, when converting the rendered HDR panorama into an LDR image for visualization.

The camera position and indoor textures can be flexibly customized for virtual staging. As illustrated in Fig. 9, three rooms are rendered at new indoor locations (Position 1 and 2), compared to the captured indoor panoramas. When the camera position is closer to the window, the window pattern is removed and the outdoor scene is visualized (Position 3). The real-time outdoor scene can be observed from various indoor view directions and locations. The 3D room geometry can be customized with new wall and floor textures to visualize the refurbished scenes.

7 Editing sun positions

Under natural illumination, the outdoor scene changes as the sun position shifts at different times of the day. Considering that the paired indoor-outdoor images are captured only at a specific time point, the rendered scene remains fixed at that captured moment. In this section, we focus on editing the direct illumination of the sun in the outdoor scene, allowing the indoor scene to be rendered under varying sun positions. By using GPS location and time of the day, the local sun position can be represented through altitude (ϕ) and azimuth (θ) angles.

$$\begin{aligned} S(\theta, \phi) &= (x, y, z) \\ x &= \cos(\phi) \sin(\theta) \\ y &= \sin(\phi) \\ z &= \cos(\phi) \cos(\theta) \end{aligned} \quad (5)$$

$$\begin{aligned} \theta' &= \arctan \frac{x}{z} \\ \phi' &= \arcsin \frac{y}{\sqrt{x^2 + y^2 + z^2}} \end{aligned} \quad (6)$$

According to the study on spherical coordinate system [58], given a unit sphere (S), the sun position (x, y, z) in the 3D coordinate is computed by ϕ and θ (Eq. (5)). In the equirectangular representation [57], the 3D coordinate can be converted into its corresponding ϕ' and θ' on the 2D image (Eq. (6)), $\theta' \in (-\pi, +\pi)$ and $\phi' \in (-0.5\pi, +0.5\pi)$. Thus, at different time points, the sun position in the real world can be precisely annotated on the captured outdoor image.

The outdoor scene changes not only with sun position but also with color temperature of the sky during the day. To achieve realistic effects of outdoor sky appearance, we integrate the full spectral sky model [25] into our method. The detailed process of editing sun positions is expressed in Algorithm 2.

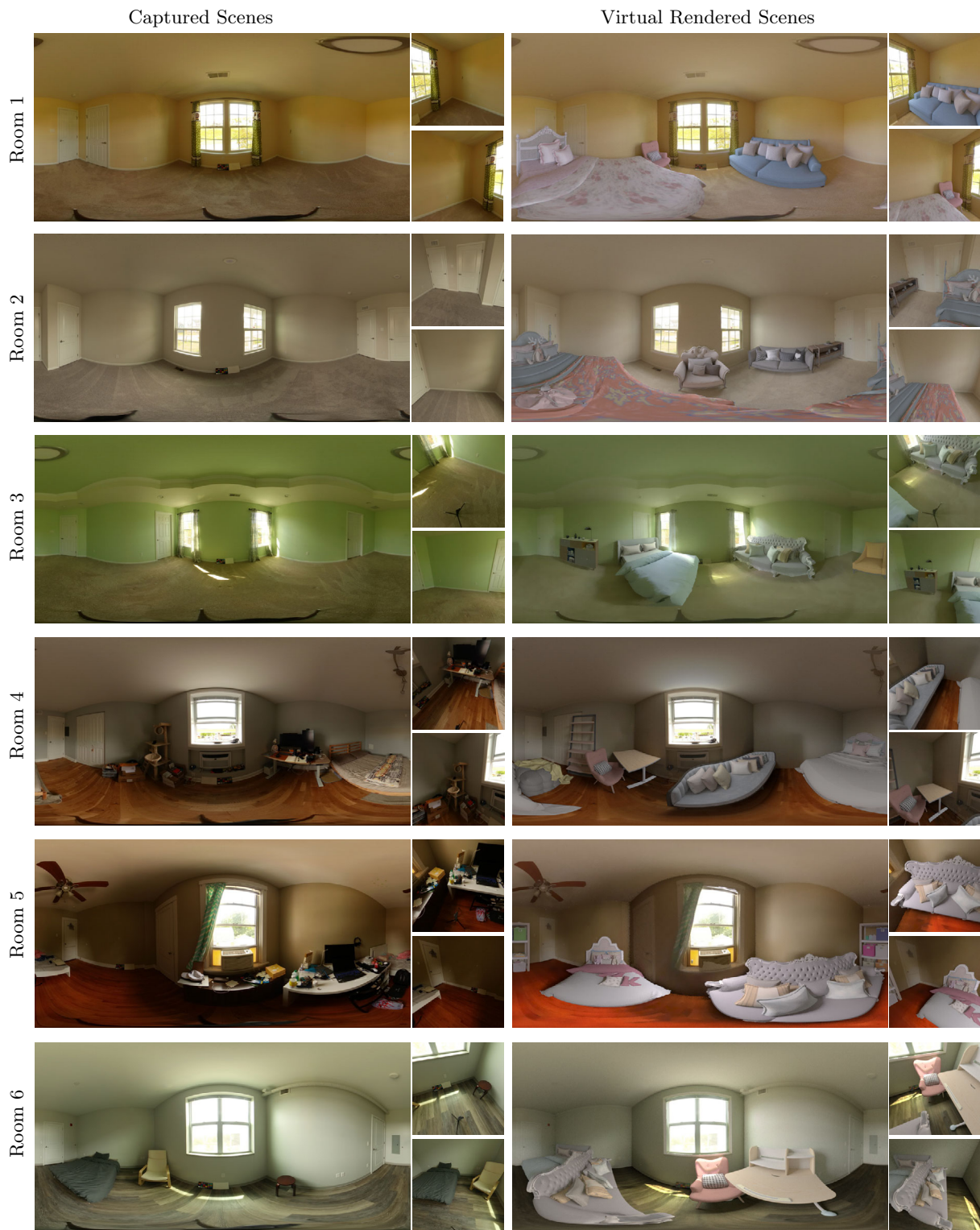


Fig. 7 Photo Gallery of Scene Editing: (left) The captured scenes include empty and furnished rooms. (right) The new scenes are virtually rendered with new furniture objects [13]. In the empty scenes,

furniture objects are arranged according to the 3D layout. For the furnished scenes, existing furniture objects are removed, and then new virtual objects are inserted



Fig. 8 Comparison between our rendering method and the conventional global rendering approaches, with the light source positioned in the lower-left corner for each scenario. Relighting with indoor panoramas ((a) estimated HDR image and (b) calibrated HDR image) fails to

reconstruct the direct illumination and inter-reflection between furniture and the existing scenes, in contrast to the use of outdoor fisheye images ((c) estimated HDR image and (d) our calibrated HDR image)

Algorithm 2 Editing Sun Positions

Require: Captured Outdoor Image (I)

Require: Current Time (t_c), Target Time (t_t), GPS location, Room Orientation (α)

Phase 1 - Editing Sun Position

- 1: **Input:** I
- 2: use GPS location and α to get current sun position (θ_1, ϕ_1) at t_c and target sun position (θ_2, ϕ_2) at t_t in the equirectangular representation
- 3: segment I into sky region (I_s) and non-sky region (I_n)
- 4: translate I_s from (θ_1, ϕ_1) to (θ_2, ϕ_2) and keep I_n consistent
- 5: **Output:** Edited Outdoor Image (I')

Phase 2 - Applying Spectral Change

- 6: **Input:** I'
 - 7: render the Current Full-Spectral Sky (S_c) at t_c and Target Full-Spectral Sky (S_t) at t_t
 - 8: compute chromatic change between S_c and S_t
 - 9: apply chromatic change to I'
 - 10: image inpainting to complete the outdoor image I''
 - 11: **Output:** New Outdoor Image
-

As illustrated in Fig. 10, the captured outdoor image is transformed from the fisheye image into the equirectangular representation. Using GPS location, room orientation, the current time, and the target time as inputs, the corresponding sun positions (θ', ϕ') are annotated on the image. Due to the varying room orientations, adjustment of the azimuth (θ) angle is required for each scene. The outdoor image is segmented into sky and non-sky regions. The sky patch is then translated based on the new sun position. To address the changes in sky color throughout the day, we render full spectral sky images [25] in panoramic perspective to compute the chromatic change in the sky region between the current and target time points. For the non-sky region, we assume the scene context is flat and apply the color change between the rendered ground surfaces to the new image.



Fig. 9 Rendering with New Camera Positions and Indoor Textures: Compared to captured scenes (top row), virtual indoor scenes are rendered at new camera positions (Position 1 and 2), and outdoor scenes are

viewed from indoor locations (Position 3). (bottom row) The existing scenes are refurbished with new indoor textures

In Fig. 11, we demonstrate that using a single input image allows for editing new sun positions in outdoor scenes under different weather conditions and at various time points.

The edited outdoor images serve as new illumination sources for rendering virtual indoor scenes. In clear sky conditions (Fig. 12: Room 1 and Room 2), the virtual rendered scenes show the indoor appearance under dynamic sun positions throughout the day. When the sun is not visible, such as in a cloudy sky condition (Fig. 12: Room 3), a simulated sun mask is added to the cloudy image, providing direct illumination for the indoor space. In this way, even if the captured scene does not have direct illumination from the sun, the rendering pipeline can still visualize the scene under clear sky conditions.

8 Adding electrical light

During the daytime, natural light from outdoor space primarily illuminates indoor spaces, while in the evening electrical lighting fixtures provide illumination. To showcase the indoor appearance during nighttime, we modified the illumination source from the outdoor environment map to electrical light and achieved virtual staging in the evening, when the virtual scene is exclusively illuminated by electrical lighting with accurate spectral data.

In the real world, each electrical lighting fixture corresponds to a distinct spectral power distribution, resulting in unique color temperatures when it's deployed to illuminate the space. For our study, we chose three different electrical

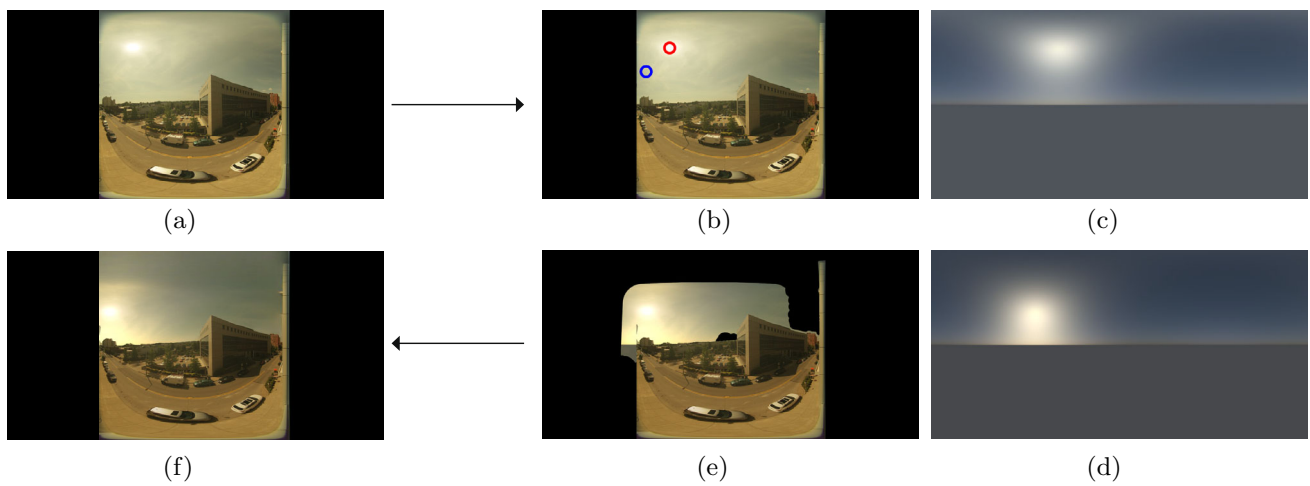


Fig. 10 Process of Editing Outdoor Image: (a) The captured outdoor fisheye image is transformed into the equirectangular representation. (b) We annotate the sun positions at the current time in **Red** and target time in **Blue**, respectively. The full spectral sky images at the current

time (c) and target time (d) are rendered to transfer chromatic change to the new sky appearance. (e) The sky patch is then translated according to the target sun position on the 2D image. (f) Image inpainting [54] is applied to complete the outdoor image



Fig. 11 Results of Editing Outdoor Images: A single image, labeled in **Red**, serves as the input for generating new outdoor images at different times of the day. The sun position is adjusted according to the target time point

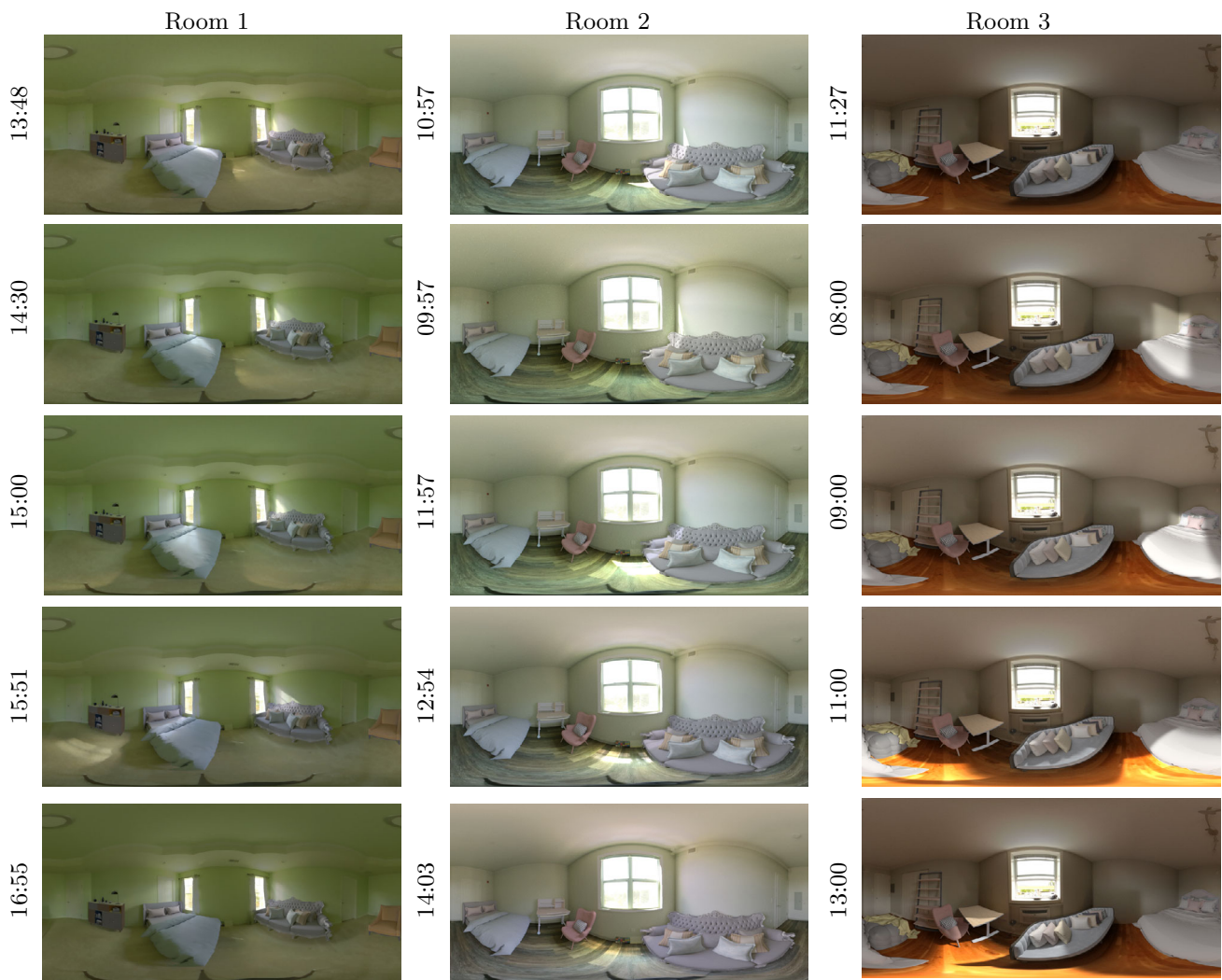


Fig. 12 Virtual Rendered Scenes with New Illumination: (top row) The scenes are rendered using the captured outdoor images. For Room 1 and Room 2, when the sun is visible outside, we edited the outdoor images to

reflect new sun positions. For Room 3, when the outdoor sky is cloudy, we added direct sun illumination onto the outdoor image and edited the sun positions to render the scene

light sources and selected their respective spectral data [46] to illuminate the scene. Our rendering pipeline incorporates the scene geometry and textures obtained from previous steps and then integrates electrical light into the virtual model. As illustrated in Fig. 13, the virtual rendered scenes exhibit different appearances when the electrical lighting changes. When the color temperature of the light source increases from 2700k to 6336k, the rendered night scene transitions from displaying yellowish tones to bluish tones.

9 Discussion

Conventional virtual home staging is labor-intensive and expensive, especially when dealing with an existing furnished scene to create a new refurnished scene. While the

global rendering approach is a commonly used method for scene relighting, it encounters limitations in fully reconstructing global illumination. Typically, new virtual objects are rendered within a 360° environment map and then inserted, along with their surrounding shadows, into the original image. However, since such a rendering process doesn't integrate the complete scene geometry, the resulting image often lacks inter-reflection between the newly inserted objects and the existing scene.

For indoor scenes under natural illumination, light estimation is a challenging task. The outdoor light enters through the window into the indoor space, and the indoor appearance is shaped by the room geometry, materials, and spatially-varying light within the room. To achieve complete global illumination for indoor virtual staging, we focus on a single

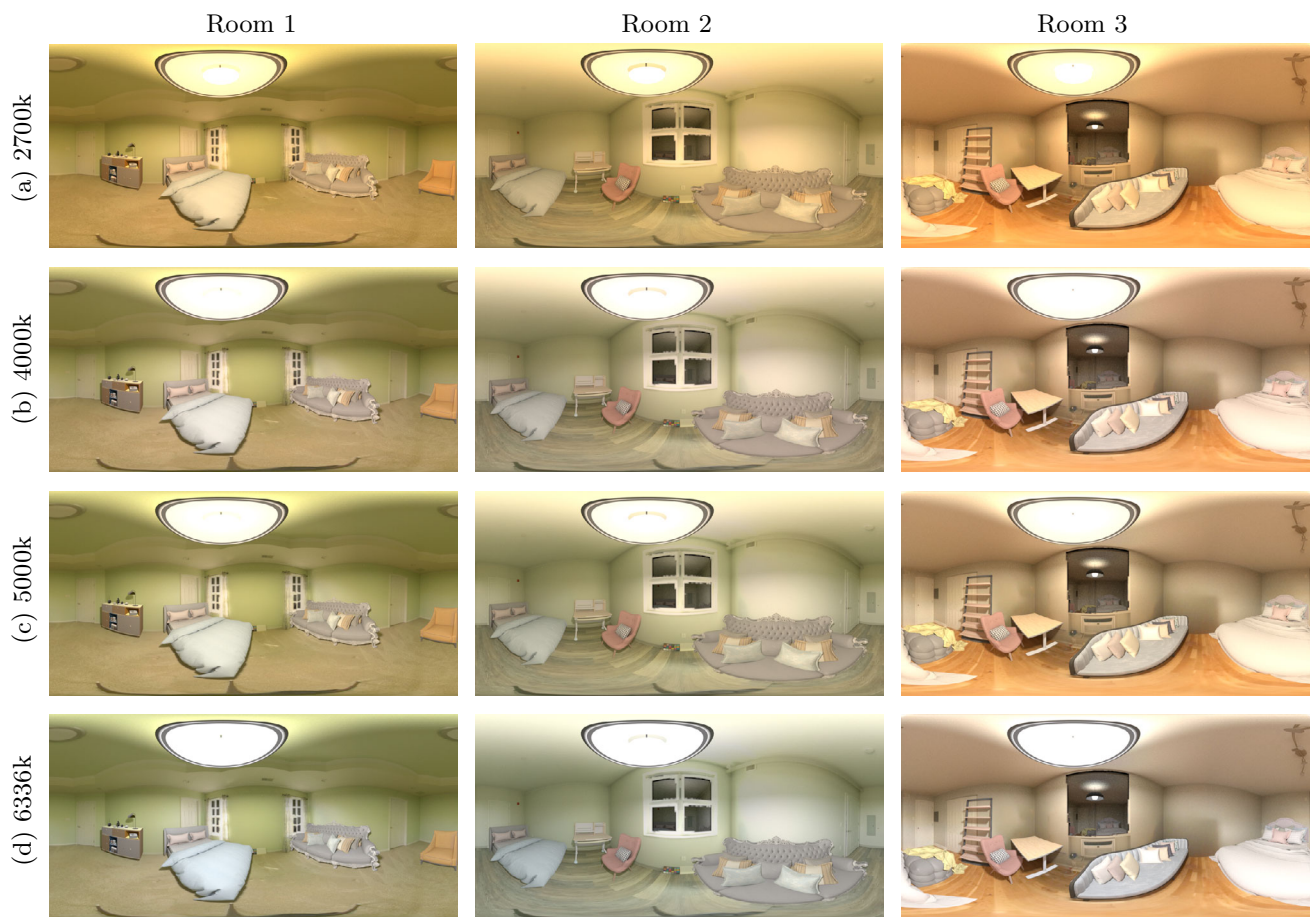


Fig. 13 Rendering Night Scenes with Electrical Light: Four spectral power distribution curves with different color temperatures are selected from the database [46]: (a) LSI LumeLEX 2040-C2M2-6S: LED - 2700K - 96.00 CRI, (b) LSI LumeLEX 2040-C4M2-6S: LED - 4000K

- 95.00 CRI, (c) Philips TLD 36W/95: FL - 5000K - 90.00 CRI, and (d) Endura OT16-3101-WT MR16: LED - 6336K - 76.00 CRI. A single lighting fixture is mounted in the ceiling and paired with different spectral data to illuminate the entire scene

panorama that allows us to understand the indoor scene from all directions.

Following the conventional image-based rendering method, we proposed a rendering approach that takes an indoor panorama along with its corresponding outdoor image as inputs. Our photography technique, supported by photometric calibration, allows the captured HDR image to be linearly scaled into the correct luminance range and ensures the accurate display of scene radiance. The captured indoor panorama is used to estimate 3D layout, materials, and window apertures. Meanwhile, the outdoor image serves as an environment map to render the scene. Leveraging the segmented floor boundary, we systematically arrange multiple furniture objects within the space and render the entire scene geometry, textures, and illumination globally. This approach reconstructs both the direct and indirect illumination between the newly inserted objects and the existing scene. Using the outdoor image as the light source enables direct adjustments

for new sun position and color temperature, providing flexibility in customizing light sources for indoor virtual staging.

9.1 Limitations

This study has several limitations. First, our HDR data collection needs users to set up two different cameras and perform luminance measurement for each scene, which requires labor efforts and a series of photography procedures. Second, when changing the sun position in the outdoor images, we adjust the sky patch to match the new sun position, assuming a flat ground context. In the real world, outdoor scenes include all kinds of geometries and materials. Particularly, in outdoor scenes where trees block the sun and light penetrates the leaves, recovering the subtle light leakage is a complex task. Our method, focusing primarily on editing direct illumination and rendering indoor scenes with new sun positions, cannot fully capture these complexities. Third, the layout estimation specifically is built on the Manhattan-based layout assump-

tion, and 3D room geometry is simplified as multiple planar surfaces. This approach fails to recognize small scene details such as cabinets, countertops, and curtains.

9.2 Future work

Future work aims to improve the virtual staging application and address complex indoor scenarios. Based on our data collection approach, future work will develop a simpler and more cost-effective photography method for capturing the indoor and outdoor scenes. When estimating indoor 3D layouts, the current data-driven approach fails to recognize complete room geometry in certain scenarios, such as the living room and kitchen. Additionally, future research will investigate how to integrate user annotation to improve the accuracy of layout generation. Other work will involve recognizing detailed indoor structures and refurbishing indoor scenes, particularly developing applications to transform kitchen spaces.

10 Conclusion

Virtual staging plays an important role in the housing market and provides an immersive way to showcase houses. Our study demonstrates a complete pipeline for virtual home staging that integrates indoor-outdoor HDR photography, inverse rendering, and scene editing. Key features of our design application include automatic floor layout, camera position adjustment, scene texture replacement, and illumination changes, allowing the existing scene to be rendered with various environmental changes. Throughout this study, we evaluated our virtual staging method across different homes and contributed a new calibrated HDR dataset for future lighting research. Our method offers a robust solution for virtual home staging and contributes to indoor scene relighting, architectural design, and real estate development.

Acknowledgements This work was partially supported by a gift from Zillow Group, USA. The authors thank the Center for Building Performance and Diagnostics (CBPD) and the Illumination and Imaging Laboratory (ILIM) at Carnegie Mellon University.

Author Contributions Conceptualization and Methodology: G.J., A.O.S., and S.G.N.; Data Collection: G.J.; Original Draft Preparation and Writing: G.J., A.O.S., and S.G.N.; Figures and Visualization: G.J., A.O.S., and S.G.N.; Review and Editing: G.J., A.O.S., and S.G.N.; Supervision: A.O.S. and S.G.N. All authors reviewed the manuscript.

Funding Open Access funding provided by Carnegie Mellon University.

Declarations

Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Araújo, A.B.: Drawing equirectangular VR panoramas with ruler, compass, and protractor. *J. Sci. Technol. Arts* **10**(1), 15–27 (2018)
2. Bolduc, C., Giroux, J., Hébert, M. et al.: Beyond the pixel: a photometrically calibrated hdr dataset for luminance and color temperature prediction. arXiv preprint [arXiv:2304.12372](https://arxiv.org/abs/2304.12372) (2023)
3. Chen, B., Zhi, T., Hebert, M., et al.: Learning continuous implicit representation for near-periodic patterns. In: *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, Proceedings, Part XV*, pp. 529–546. Springer, (2022)
4. Cheng, H.T., Chao, C.H., Dong, J.D., et al.: Cube padding for weakly-supervised saliency prediction in 360 videos. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1420–1429 (2018)
5. Coughlan, J.M., Yuille, A.L.: Manhattan world: compass direction from a single image by bayesian inference. In: *Proceedings of the seventh IEEE international conference on computer vision, IEEE*, pp. 941–947 (1999)
6. Cruz, S., Hutchcroft, W., Li, Y., et al.: Zillow indoor dataset: annotated floor plans with 360deg panoramas and 3d room layouts. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2133–2143 (2021)
7. Dastjerdi, M.R.K., Eisenmann, J., Hold-Geoffroy, Y., et al.: Everlight: indoor-outdoor editable HDR lighting estimation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7420–7429, (2023)
8. Debevec, P.: Image-based lighting. In: *ACM SIGGRAPH 2006 Courses*. Association for Computing Machinery, New York, NY, United States, 4–es (2006)
9. Debevec, P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In: *SIGGRAPH 2008 classes*. ACM, pp. 1–10, (2008)
10. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: *SIGGRAPH 2008 classes*, pp. 1–10. ACM, New York, NY, USA (2008)
11. Duchêne S., Riant, C., Chaurasia, G., Moreno, J. L., Laffont, P.-Y., Popov, S., Bousseau, A., Drettakis, G.: Multiview intrinsic images of outdoors scenes with an application to relighting. *ACM Trans. Graph.* **34**(5), 164 (2015). <https://doi.org/10.1145/2756549>
12. Fan, Q., Yang, J., Hua, G., et al.: Revisiting deep intrinsic image decompositions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8944–8952 (2018)
13. Fu, H., Jia, R., Gao, L., et al.: 3d-future: 3d furniture shape with texture. *Int. J. Comput. Vision* **129**, 3313–3337 (2021)
14. Garces, E., Rodriguez-Pardo, C., Casas, D., et al.: A survey on intrinsic images: Delving deep into lambert and beyond. *Int. J. Comput. Vision* **130**(3), 836–868 (2022)

15. Gardner, M.A., Sunkavalli, K., Yumer, E., et al.: Learning to predict indoor illumination from a single image. arXiv preprint [arXiv:1704.00090](https://arxiv.org/abs/1704.00090) (2017)
16. Gardner, M.A., Hold-Geoffroy, Y., Sunkavalli, K., et al.: Deep parametric indoor lighting estimation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 7175–7183 (2019)
17. Garon, M., Sunkavalli, K., Hadap, S., et al.: Fast spatially-varying indoor lighting estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6908–6917, (2019)
18. Gkitsas, V., Zioulis, N., Alvarez, F., et al.: Deep lighting environment map estimation from spherical panoramas. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 640–641 (2020)
19. Gkitsas, V., Sterzentsenko, V., Zioulis, N., et al.: Panodr: spherical panorama diminished reality for indoor scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3716–3726 (2021)
20. Gkitsas, V., Zioulis, N., Sterzentsenko, V., et al.: Towards full-to-empty room generation with structure-aware feature encoding and soft semantic region-adaptive normalization. arXiv preprint [arXiv:2112.05396](https://arxiv.org/abs/2112.05396) (2021)
21. Guerrero-Viu, J., Fernandez-Labrador, C., Demonceaux, C., et al.: What's in my room? object recognition on indoor panoramic images. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 567–573 (2020)
22. Haber, T., Fuchs, C., Bekaer, P., et al.: Relighting objects from image collections. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 627–634 (2009)
23. Hauage, D.C., Wehrwein, S., Upchurch, P., et al.: Reasoning about photo collections using models of outdoor illumination. In: BMVC (2014)
24. Hold-Geoffroy, Y., Sunkavalli, K., Hadap, S., et al.: Deep outdoor illumination estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7312–7321, (2017)
25. Hosek, L., Wilkie, A.: An analytic model for full spectral sky-dome radiance. *ACM Trans. Graph. (TOG)* **31**(4), 1–9 (2012)
26. Huang, J.B., Kang, S.B., Ahuja, N., et al.: Image completion using planar structure guidance. *ACM Trans. Graph. (TOG)* **33**(4), 1–10 (2014)
27. Huang, S., Qi, S., Zhu, Y., et al.: Holistic 3d scene parsing and reconstruction from a single rgb image. In: Proceedings of the European conference on computer vision (ECCV), pp. 187–203, (2018)
28. Inanici, M.: Evaluation of high dynamic range image-based sky models in lighting simulation. *Leukos* **7**(2), 69–84 (2010)
29. Inanici, M.N.: Evaluation of high dynamic range photography as a luminance data acquisition system. *Lighting Res. Technol.* **38**(2), 123–134 (2006)
30. Izadinia, H., Shan, Q., Seitz, S.M.: Im2cad. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5134–5143 (2017)
31. Jakob, W., Speierer, S., Roussel, N., et al.: Mitsuba 3 renderer. <https://mitsuba-renderer.org> (2022)
32. Ji, G., Sawyer, A.O., Narasimhan, S.G.: Virtual home staging: inverse rendering and editing an indoor panorama under natural illumination. In: International Symposium on Visual Computing, Springer, pp. 329–342 (2023)
33. Karsch, K., Hedau, V., Forsyth, D., et al.: Rendering synthetic objects into legacy photographs. *ACM Trans. Graph. (TOG)* **30**(6), 1–12 (2011)
34. Kawai, N., Sato, T., Yokoya, N.: Diminished reality based on image inpainting considering background geometry. *IEEE Trans. Visual Comput. Graph.* **22**(3), 1236–1247 (2015)
35. Kulshreshtha, P., Lianos, N., Pugh, B., et al.: Layout aware inpainting for automated furniture removal in indoor scenes. In: 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), IEEE, pp. 839–844 (2022)
36. Lalonde, J.F., Matthews, I.: Lighting estimation in outdoor image collections. In: 2014 2nd International Conference on 3D Vision, IEEE, pp. 131–138 (2014)
37. LeGendre, C., Ma, W.C., Fyffe, G., et al.: Deeplight: learning illumination for unconstrained mobile mixed reality. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5918–5928 (2019)
38. Li, Y., Brown, M.S.: Single image layer separation using relative smoothness. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2752–2759 (2014)
39. Li, Z., Shafiei, M., Ramamoorthi, R., et al.: Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting and svbrdf from a single image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2475–2484 (2020)
40. Li, Z., Shi, J., Bi, S., et al.: Physically-based editing of indoor scene lighting from a single image. In: European Conference on Computer Vision, Springer, pp. 555–572 (2022)
41. Liu, Y.L., Lai, W.S., Chen, Y.S., et al.: Single-image HDR reconstruction by learning to reverse the camera pipeline. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1651–1660 (2020)
42. Mitsunaga, T., Nayar, S.K.: Radiometric self calibration. In: Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149), IEEE, pp. 374–380, (1999)
43. Moeck, M.: Accuracy of luminance maps obtained from high dynamic range images. *Leukos* **4**(2), 99–112 (2007)
44. Narihira, T., Maire, M., Yu, S.X.: Direct intrinsics: learning albedo-shading decomposition by convolutional regression. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2992–2992 (2015)
45. Nie, Y., Han, X., Guo, S., et al.: Total3dunderstanding: Joint layout, object pose and mesh reconstruction for indoor scenes from a single image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 55–64 (2020)
46. Padfield, J.: Light Sources | SPD Curves | National Gallery, London | Information. <https://research.ng-london.org.uk/scientific/spd/?page=home> (2023)
47. Perez, R., Seals, R., Michalsky, J.: All-weather model for sky luminance distribution-preliminary configuration and validation. *Sol. Energy* **50**(3), 235–245 (1993)
48. Reinhard, E., Heidrich, W., Debevec, P., et al.: High dynamic range imaging: acquisition, display, and image-based lighting. Morgan Kaufmann, San Francisco, CA, USA (2010)
49. Shan, Q., Adams, R., Curless, B., et al.: The visual turing test for scene reconstruction. In: 2013 International Conference on 3D Vision-3DV 2013, IEEE, pp. 25–32 (2013)
50. Slavcheva, M., Gausebeck, D., Chen, K., et al.: An empty room is all we want: Automatic defurnishing of indoor panoramas. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2024)
51. Srinivasan, P.P., Mildenhall, B., Tancik, M., et al.: Lighthouse: Predicting lighting volumes for spatially-coherent illumination. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8080–8089 (2020)
52. Stokes, M., Anderson, M., Chandrasekar, S., et al.: A standard default color space for the internet-srgb. <https://www3.org/Graphics/Color/sRGB> (1996)
53. Stumpfel, J., Jones, A., Wenger, A., et al.: Direct HDR capture of the sun and sky. In: SIGGRAPH 2006 Courses. ACM, 5–es (2006)

54. Suvorov, R., Logacheva, E., Mashikhin, A., et al.: Resolution-robust large mask inpainting with fourier convolutions. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 2149–2159 (2022)
55. Wang, F.E., Hu, H.N., Cheng, H.T., et al.: Self-supervised learning of depth and camera motion from 360 videos. In: Asian Conference on Computer Vision, Springer, pp. 53–68 (2018)
56. Wang, F.E., Yeh, Y.H., Sun, M., et al.: Bifuse: Monocular 360 depth estimation via bi-projection fusion. In: The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
57. Wang, F.E., Yeh, Y.H., Sun, M., et al.: Bifuse: Monocular 360 depth estimation via bi-projection fusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 462–471 (2020)
58. Wang, F.E., Yeh, Y.H., Sun, M., et al.: Led2-net: Monocular 360deg layout estimation via differentiable depth rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12956–12965 (2021)
59. Wu, L., Zhu, R., Yaldiz, M.B., et al.: Factorized inverse path tracing for efficient and accurate material-lighting estimation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3848–3858 (2023)
60. Xiao, J., Ehinger, K.A., Oliva, A., et al.: Recognizing scene viewpoint using panoramic place representation. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 2695–2702 (2012)
61. Yang, B., Jiang, T., Wu, W., et al.: Automated semantics and topology representation of residential-building space using floor-plan raster maps. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* **15**, 7809–7825 (2022)
62. Yang, S.T., Wang, F.E., Peng, C.H., et al.: Dula-net: A dual-projection network for estimating room layouts from a single RGB panorama. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3363–3372 (2019)
63. Yeh, Y.Y., Li, Z., Hold-Geoffroy, Y., et al.: Photoscene: Photorealistic material and lighting transfer for indoor scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18562–18571 (2022)
64. Zeng, Z., Li, X., Yu, Y.K., et al.: Deep floor plan recognition using a multi-task network with room-boundary-guided attention. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9096–9104 (2019)
65. Zhang, C., Liwicki, S., Smith, W, et al.: Orientation-aware semantic segmentation on icosahedron spheres. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3533–3541 (2019)
66. Zhang, E., Cohen, M.F., Curless, B.: Emptying, refurbishing, and relighting indoor spaces. *ACM Trans. Graph. (TOG)* **35**(6), 1–14 (2016)
67. Zhang, E., Martin-Brualla, R., Kontkanen, J., et al.: No shadow left behind: removing objects and their shadows using approximate lighting and geometry. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16397–16406 (2021)
68. Zhang, J., Lalonde, J.F.: Learning high dynamic range from outdoor panoramas. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4519–4528, (2017)
69. Zhang, Y., Song, S., Tan, P., et al.: Panocontext: A whole-room 3d context model for panoramic scene understanding. In: Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part VI 13, Springer, pp. 668–686, (2014)
70. Zhi, T., Chen, B., Boyadzhiev, I., et al.: Semantically supervised appearance decomposition for virtual staging from a single panorama. *ACM Trans. Graph. (TOG)* **41**(4), 1–15 (2022)
71. Zhou, B., Zhao, H., Puig, X., et al.: Semantic understanding of scenes through the ade20k dataset. *International Journal on Computer Vision* (2018)
72. Zillow: Creation of 3d home tours soared in march as stay-at-home orders expanded. <https://www.zillow.com/research/3d-home-tours-coronavirus-26794/> (2020)
73. Zou, C., Colburn, A., Shan, Q., et al.: Layoutnet: reconstructing the 3d room layout from a single rgb image. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2051–2059 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

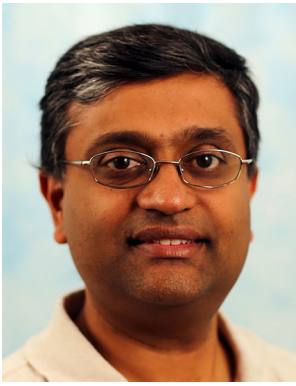


Guanzhou Ji is a PhD candidate at Carnegie Mellon University. He received Master of Architecture from University of Washington in 2019. He focuses on indoor photometry, computational photography, and physics-based rendering, and explores computational methods for indoor virtual staging and scene editing.



Azadeh O. Sawyer is an Assistant Professor and the Track Chair of Building Performance & Diagnostics (BPD) at the Carnegie Mellon University School of Architecture, as well as a Faculty Affiliate at the Wilton E. Scott Institute for Energy Innovation. Azadeh holds a PhD in Architecture Building Technology from the University of Michigan, where she was also a Rackham Predoctoral Fellow and received the Distinguished Dissertation Award. She has a Master of Science in

Architecture from the University of Michigan, a Master of Design Studies in Sustainable Design from the Harvard University Graduate School of Design, and was the 2011 recipient of the Harvard Daniel L. Schodek Award for Technology and Sustainability. Her research focuses on building facades, daylighting, energy performance, and evaluation through simulation and immersive virtual environments. She received a Bachelor of Architecture from the California College of the Arts in San Francisco with distinction in 2008, where she also received the Technology Book Award. Azadeh has been a LEED-accredited professional since 2008.



Srinivasa G. Narasimhan is the U.A. and Helen Whitaker Professor of Robotics at Carnegie Mellon University. He served as Interim Director of the Robotics Institute from Aug 2019 to Dec 2021. He obtained his PhD from Columbia University in Dec 2003. His group focuses on novel techniques for imaging and illumination to enable applications in vision, graphics, robotics, agriculture, intelligent transportation and medical imaging. His works have received a dozen Best Paper or

Best Demo or Honorable mention awards at major conferences.