DOI: 10.1111/cgf.14166
Pacific Graphics 2020
E. Eisemann, A. Jacobson, and F.-L. Zhang
(Guest Editors)

Volume 39 (2020), Number 7

Interactive Design and Preview of Colored Snapshots of Indoor Scenes

Qiang Fu[1], Hai Yan[1†], Hongbo Fu[2], Xueming Li[1‡]

[1]School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications
[2]School of Creative Media, City University of Hong Kong

Abstract

This paper presents an interactive system for quickly designing and previewing colored snapshots of indoor scenes. Different from high-quality 3D indoor scene rendering, which often takes several minutes to render a moderately complicated scene under a specific color theme with high-performance computing devices, our system aims at improving the effectiveness of color theme design of indoor scenes and employs an image colorization approach to efficiently obtain high-resolution snapshots with editable colors. Given several pre-rendered, multi-layer, gray images of the same indoor scene snapshot, our system is designed to colorize and merge them into a single colored snapshot. Our system also assists users in assigning colors to certain objects/components and infers more harmonious colors for the unassigned objects based on pre-collected priors to guide the colorization. The quickly generated snapshots of indoor scenes provide previews of interior design schemes with different color themes, making it easy to determine the personalized design of indoor scenes. To demonstrate the usability and effectiveness of this system, we present a series of experimental results on indoor scenes of different types, and compare our method with a state-of-the-art method for indoor scene material and color suggestion and offline/online rendering software packages.

CCS Concepts

• ***Applied computing*** $\rightarrow$ ***Computer-aided design;*** • ***Computing methodologies*** $\rightarrow$ ***Graphics systems and interfaces; Rendering;***

1. Introduction

In recent years, personalized interior design has become more and more popular. To meet the personalized demands of customers, interior designers have to devise various schemes of both interior layouts and color themes for them to choose from. Since designing these schemes is time-consuming and requires professional knowledge and skills, personalized interior design is always expensive, thus limiting its development. For example, a traditional pipeline takes a 3D scene as input, while determining a desired color theme for the given scene is very important for personalized interior design. However, if customers are not satisfied with the current color theme, designers have to manually adjust the colors and/or materials of all involved components and then re-render the 3D scene.

In recent years, several techniques have been proposed to automatically suggest colors or materials for a given 3D indoor scene [CXY*15,LWZ*17,ZGM18]. These approaches mainly aim at providing various plausible color themes, but lack mechanisms to ensure user-desired colors. Moreover, given an automatically suggested color theme, it is still time-consuming to render the 3D scene photo-realistically. On the other hand, benefiting from the real-time rendering techniques [WMG*09] and engines such as Unity3D and Unreal, it is possible to develop real-time material editing systems. However, such systems work under the limited pre-stored materials for user manipulation, and the scenes rendered in real-time are still less vivid compared to those by non-real-time rendering techniques [XCM*14]. Therefore, a natural question arises: could a system efficiently generate high-quality rendering snapshots of indoor scenes directly with their color themes easily editable by users?

This question has been partially addressed by some techniques of image processing, such as color editing [CZZT12] and gray image colorization [ZZI*17]. Since the interior color theme design always requires color variations, single image inputs are not adequate for both color editing and gray image colorization methods in our task. Using color editing methods with color image inputs might bias users' decisions, lead to global color changes, and increase the difficulty to determine proper colors for preparing input images. Therefore, we intend to employ gray image colorization methods with multiple gray image inputs in our system. On the other hand, if we use multiple gray images as input, the amount of required user intervention with a naïve solution would increase sharply, since users have to specify more color points to explicitly

† Joint first author
‡ Corresponding author (email:lixm@bupt.edu.cn)

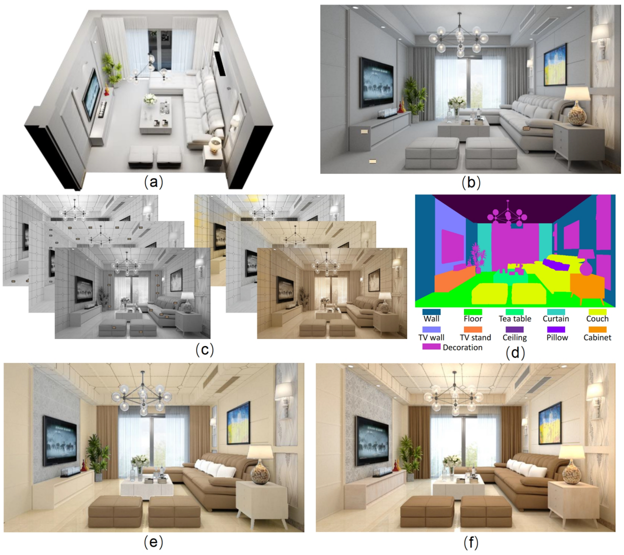

Figure 1: *Given a 3D indoor scene (a) with a small set of user-specified colors (indicated by small color boxes) on its snapshot (b), our system suggests additional harmonious colors and allocates them to multiple layers of pre-rendered gray snapshots of the same scene under the same viewpoint for colorization (c). The colorized multi-layer images are then merged together based on the semantic map (d) to generate a colored snapshot of the given scene respecting the user-designed color theme (e). Once a desired color theme is determined, the corresponding materials and colors can finally be attached to the given 3D scene for high-quality rendering (f).*

identify which gray images should be used for colorization with respect to a certain region [CZG*11, ZZI*17].

To address these problems, we aim at an interactive system which supports color specification on pre-rendered multi-layer gray images of a given 3D scene, and merge them into a single colorized snapshot respecting a user-specified color theme (e.g., Figure 1). To this end, we design and develop an interface that assists users in assigning colors to specific elements (e.g., floor, wall, furniture) of the snapshot of a given 3D indoor scene. We employ material/color priors, which consist of the color collocation expert knowledge and relations between objects, materials, and colors, to prepare the multi-layer images. These priors also guide the users to find the proper color to draw the elements without user-specified colors in order to ensure global color harmony. Afterwards, the selected colors are then assigned to the proper layers corresponding to the pre-rendered grayscale images of the given 3D scene. In each layer, we perform a neural network based colorization method to draw the gray image by the assigned colors. Finally, we use a semantic map, which can be obtained from the given indoor scene with the same view as the snapshot, to merge the colorization results from different layers to create a fast preview of the designed color theme of the given 3D scene. Such a snapshot can provide a vivid colorizing effect for users to evaluate the designed color theme. Once users are satisfied with a specific color theme, our system attaches the associated materials and colors to the given 3D scene, and complete the interior design (e.g., Figure 1-(f)).

We present a series of experimental results to show the usability and effectiveness of our system. We also conduct two user studies to evaluate the user experience of our system, and to compare the visual quality of our method, an automatic material and color method [CXY*15], and real-time rendering software. The results show that by optimizing the time cost and rendering quality, our system makes the task of interactive interior color theme design easier and faster.

In summary, our work has made the following main contributions: 1) a unified framework for interactive interior color theme design with the pre-collected color harmony priors that encode both color-material-object relations and color relations between different object categories; 2) a synergy colorization mechanism that leverages multiple layers of gray images to quickly generate an indoor scene snapshot with respect to a user-specified color theme.

2. Related Work

Image Colorization. Image colorization, which aims at colorizing black-and-white movies/photos, has been a hot topic in the fields of graphics and vision in the past decades. Most of the recent works are data-driven methods that leverage color priors learned from colorful image examples. One feasible solution of the data-driven methods is to transfer the color styles from the reference to the source [FPC*14]. For example, Welsh et al. [WAM02] proposed a general technique for colorizing grayscale images by transferring colors between a source color image and a destination grayscale image. Chia et al. [CZG*11] proposed a colorization system that leverages the rich image content on the internet, and filters them to obtain suitable reference images that are reliable for color transfer to a given grayscale photo. These methods perform well in gray image colorization. However, since the color brightness variation of these methods is limited by the input gray image, we cannot directly apply these methods to color theme design.

Some other works focused on revealing the color relations among well-colored image examples. For example, Huang et al. [HZMH14] proposed a compact representation that summarizes color and geometric features of image regions, and geometric relationships between them, aiming at learning the correlations between color property distributions and geometric features of regions from a database of well-colored photos. Lin et al. [LRFH13] presented a probabilistic factor graph model based on Markov Chain Monte Carlo for automatically coloring 2D patterns. Motivated by these methods, our system leverages expert knowledge to obtain the color relations between different objects and thus facilitate the colorization task.

In recent years, more and more attentions have been paid to the deep learning based approaches, e.g., leveraging Deep Convolution Neural Networks (CNNs) to facilitate image colorization [ZIE16, ISI16], or learning a low-dimensional embedding of color fields using a variational autoencoder (VAE) [DLY*17]. Sangkloy et al. [SLF*17] proposed a deep adversarial image synthesis architecture that allows users to sketch and scribble colors to control

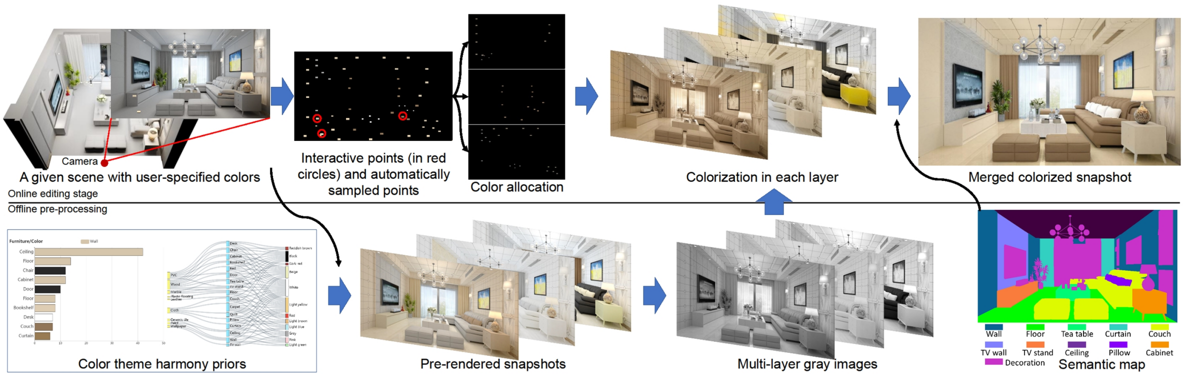

Figure 2: *System overview. The offline pre-processing stage aims at generating multi-layer gray images for a given 3D indoor scene. In the online stage, our system suggests and samples more points based on a small set of user-specified colors, and chooses the proper layers for colorization. Finally, the colorized layers are merged to synthesize the final colorized snapshot, with the help of the semantic map obtained in the offline stage.*

deep image synthesis. Zhang et al. [ZZI*17] proposed a deep learning approach for user-guided image colorization. Rather than using hand-defined rules, their network propagates user edits by fusing low-level cues along with high-level semantic information, learned from large-scale data. In our work, we employ the neural network proposed by [ZZI*17] for colorizing the pre-rendered indoor scene gray images. To enhance the network of indoor scene colorization, we adopt a synergy colorization mechanism which tackles multi-layer gray images of a given indoor scene and merges colorization images into a snapshot of the designed color theme.

Interior Color Theme Design. Most of the currently developed systems for interior design have focused on generating plausible interior layouts (e.g., [YYT*11, YYT16, FCW*17]). Given such layouts, another important task for interior design is to assign proper materials with harmonious colors to indoor objects under certain color themes, while a large number of optional parameters of both material and color would lead to heavy workload for this task. Recently several works have been proposed to suggest materials/colors for indoor scenes. For example, Chajdas et al. [CLS10] proposed an algorithm that assists users in assigning textures to surfaces by propagating a chosen image for a certain surface throughout an entire environment. Chen et al. [CXY*15] developed a system that generates material suggestions for 3D indoor scenes automatically, based on both local material rules and global aesthetic rules. Zhu et al. [ZGM18] presented a data-driven approach that colorizes 3D furniture models and indoor scenes by leveraging indoor images on the internet. Moreover, some works on material labeling, recognition, and synthesis (e.g., [BUSB13, BUSK15, ZWW18]) have been proposed that can be used to facilitate the interior color theme design task.

Although the above-mentioned material suggestion methods are able to relieve the workload of interior color theme design, the outputs of these methods still need commercial tools, such as SketchUp, 3ds Max, or Maya, to render 3D scenes with the suggested materials. On the contrary, we attempt to generate a colored snapshot of an indoor scene with suggested materials/colors quickly. Since our method focuses on interactive interior design, such a mechanism enables interior designers to spend less time waiting for rendered results of indoor scenes after editing their color themes.

Color Harmony. Harmonic colors are sets of colors that are aesthetically pleasing in terms of human visual perception. This topic is closely related to the quality of both colorized images and designed indoor scenes [WYW*10, LE07, WJLC12, LZNH15]. For example, Cohen-Or et al. [COSG*06] presented a method that enhances the harmony among the colors of a given photograph or of a general image, while retaining the original colors as much as possible. Odonovan et al. [OAH11] studied color compatibility theories using large datasets, and developed new tools for choosing colors. Lu et al. [LKPL14] proposed a hierarchical unsupervised learning approach to learn compatible color combinations from a large dataset. Kim et al. [KYKL14] presented a novel method for automatic color assignment based on theories of color perception.

Moreover, some works leverage database examples to establish the harmonious color relations between components of objects and thus to facilitate the object material and color suggestion (e.g., [JTRS12, PRFS18]). Our method also attempts to establish relationships among the materials of different objects. The major difference of our method from the above-mentioned works is that we collect user-defined color theme examples to extract the material/color priors of indoor objects, rather than using indoor scene examples. This is mainly because we expect the priors to reflect the user psychology during the interaction process, thus supporting the heuristic color theme suggestion for objects without user assigned colors.

3. System Overview

Figure 2 illustrates the workflow of our system. The input of our system is a 3D indoor scene whose objects have been segmented with semantic labels. Specifically, users can choose a certain view of the given scene for the snapshots, which will be the targets of all edit operations of our system. The output of our system is a colorized snapshot with a user-designed color theme. We leverage the expert knowledge what we called as the *color theme harmony priors* (Section 4) to boost the system intelligence. Such priors contain two-fold relations, including the color-material-object relations, which are used to estimate the potential range of color variations of a given scene, and the object-object color relations, which are used to predict more colors for the given scene based on a small number of user-specified colors.

The offline pre-processing stage is employed to render a given 3D scene to multi-layer images in different gray levels. These gray images are created via a sampling strategy based on the color-material-object relations, and will be used for the online colored snapshot generation. The variations of the indoor color themes always need lots of gray images to explore the proper one for colorization. For example, a dark gray floor can be colorized in dark red or dark brown easily, but can hardly be colorized in light pink. But we expect to limit the number of pre-rendered gray images to make our system easy to set up. To address this issue, we divide the color space to different gray levels based on the content of a given room. This follows the principle that each kind of object usually has limited color variations, and enables our system to create adequate gray images to support the color theme generation of the given scene (Section 4). The pre-processing stage also needs to generate a semantic map of the given scene with the same user-specified view as the multi-layer gray images.

In the online stage, users may specify colors for certain objects on a snapshot of the given scene. Benefiting from the object-object color relations within our collected priors, our system can explore more colors that are harmonious to the user-specified ones for the unspecified objects to reduce the required user interventions. After that, our system first allocates both the user-specified and the suggested colors to the proper layers, then uses a deep neural network to complete the colorization, and finally merges the per-layer results together based on the semantic map of the given scene (Section 5). This allows designers to easily generate color themes of the given scene and obtain the fast preview of high quality.

4. Color Theme Harmony Priors

A common strategy adopted by the automatic indoor color theme generation methods [CXY*15, ZGM18] is to collect the relations between objects, colors, and materials from indoor scene examples as priors. Our interactive system also needs color theme harmony priors to filter out the improper color combinations and assist users in completing the design. Large-scale dataset of indoor scenes such as indoor images, contain not only examples of scenes with conspicuous color themes, but also scenes with hybrid color styles. In other words, since most of the daily-life indoor scenes do not have well-designed color themes, priors collected from such data could suggest proper color combinations, but might not conform to exact color themes, which are the goals of our interactive design system. To address this concern, we collect the color theme harmony priors adopted in our system by two steps. We first establish the relations between objects, colors, and materials, and then invite interior designers to create examples of well-designed color themes to extract the object-object color relations.

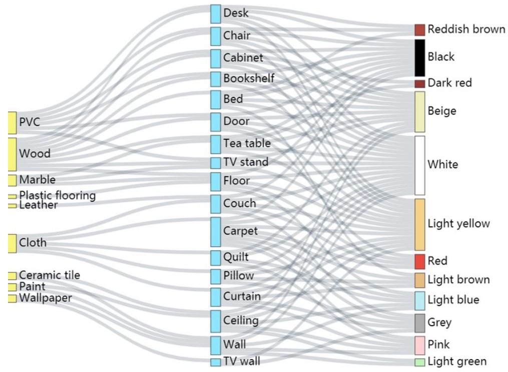

Figure 3: *The visualization of the color-material-object relations.*

In Figure 3, we illustrate the color-material-object relations summarized in our color theme harmony priors. These relations indicate the potential combinations of materials and colors for certain elements, and can be used to determine the distributions of the gray levels involved in our system for colorization. We expect that colors in such relations should be representative within a range of gray levels, in order to reduce the number of the required gray images. To obtain such relations, we first collect a large number of photos for each element category considered in our system. For each kind of element, we cluster the colors, which are extracted from the collected photos, into groups based on their gray levels, and choose a representative color for each group. Besides, for each kind of element, we also manually specify proper materials to each representative color, and relate the texture maps to the associated materials to ensure the pre-rendered gray images and colorized snapshot visually pleasing. In this way, the color-material-object relations can be established. In our implementation, we take 17 kinds of elements, 12 colors, and 9 kinds of materials into consideration, and obtain 75 groups of color-material-object relations. Figuring out the correct combinations of them will help determine how many and what kinds of pre-rendered gray images are required to enable the interactive color theme design. Note that, since some object categories such as decorations always have uniform colors that are largely independent of the other objects of the scene, these categories are not involved in our priors. For these categories, users can directly attach colors/materials to them in the given 3D scene. Regions of such object categories in all snapshots would keep their original colors unchanged after rendering the given scene.

Even though our purpose is to develop an interactive system rather than automatically suggesting various potential color combinations for the given scene, the object-object color relations are still important to reduce the amount of required user intervention and make our system easier to use. Aiming at collecting color theme examples designed by interior designers, we invited 5 participants with the background of digital media technology to create such examples. We collected 74 different indoor scenes from 6 types for the participants to design color themes, including living room, bedroom, dining room, study room, conference room, and office room. For each scene, we asked the participants to select proper color combinations for the elements considered in our system in the

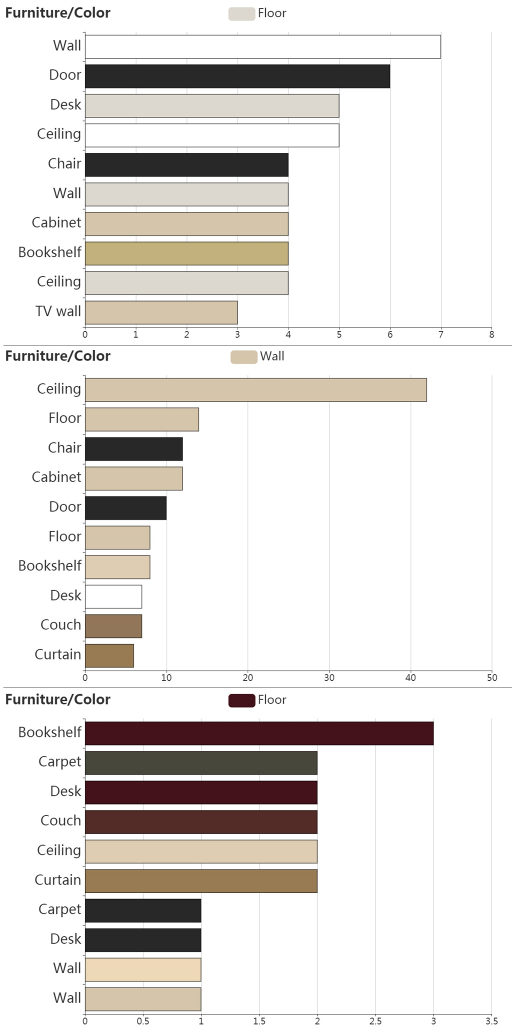

Figure 4: *Three examples of the object-object color relations. In each case, we show the coexistence frequency of the objects with certain colors to a specific colored object (the header of each chart).*

scene, in order to create various color themes. This was conducted through our system to reflect the potential user choices when using our system for color theme design. The participants totally created 322 color theme examples by using the scenes we provided. In this way, we obtain both the labeled color-material-object relations and various color combinations that conform to certain color themes, i.e., the object-object color relations as our color theme harmony priors. In Figure 4, we show three examples of the object-object color relations. We can see that for a certain element with a certain color, the related elements with harmonious colors would have high coexistence frequencies.

5. Interactive Indoor Scene Colorization

By using the collected priors for colorization, our system first employs a pre-processing stage to prepare the multi-layer gray images for a given 3D scene under a specific view. We provide a simple interface allowing users to specify colors for certain areas in a gray snapshot of the given scene. Then an allocation algorithm is performed to assign the user-specified colors to the proper layers (i.e., pre-rendered gray images). Finally, we employ the colorization network to obtain multi-layer colorized images, and use their semantic map to merge them into a colorized snapshot of the given indoor scene.

5.1. Pre-processing

The goal of our pre-processing stage is to render images of the given scene with sufficient grayscale variations. This would ensure that designed color themes could be well presented in the colorized snapshots. This stage is semi-automatic and requires users to assign a proper rendering view and attach materials (associated with textures) and colors to the editable elements of the given 3D scene. A good rendering view could lead to less inter-element occlusion and thus make the colorization process easier, while the attached materials and colors determine the grayscale variations of the pre-rendered images. Our system can assist users in attaching materials/colors to editable elements by providing several *rendering schemes* based on the content of the given scene. Each rendering scheme provides a kind of combination of materials and colors for the editable elements in the given scene, and each pre-rendered image has a unique rendering scheme. Since a single high-quality scene image typically requires 5-10 minutes for offline rendering, and the pre-processing is required when inputting a new 3d scene or changing the viewing position, we expect to limit the maximum number of pre-rendered images under 10 to balance the colorization ability of our system and the time cost of the pre-processing stage. To achieve this, our pre-processing stage leverages the color-material-object relations to explore proper rendering schemes of the given scene for pre-rendering.

The color-material-object relations in our color theme harmony priors contain the commonly-seen coexisting materials and colors for certain elements. To generate rendering schemes, we first enumerate all material and color combinations for each editable element existing in the given scene. For example, a floor can be white marble, light yellow wood, light yellow plastic flooring, etc. Due to our colorization mechanism, the number of rendering schemes is equal to the maximum number of material and color combinations among all editable elements. Then, for the other elements with fewer combinations, we simply repeat the same combinations of such elements in multiple rendering schemes to fill in the vacancy. In this manner, our system generates a series of rendering schemes by picking the combinations in turn. Users can choose one of the themes to render the given 3D scene with the same user-specified viewing and lighting conditions. In our implementation, the pre-processing stage is conducted through 3ds Max. Usually, a given 3D scene needs 8 to 10 pre-rendered images.

Besides, we also create a semantic map by manually segmenting the snapshot into semantic objects, with different labels visualized in different colors, as shown in Figure 2. The generation of this semantic map can be easily automated, since individual objects in the given scene are often represented as separate meshes. This semantic map will be used to merge the colorization results of these gray images together (Section 5.2). Lastly, we render a representative image with the same view as the gray images for user interactions. Since we do not expect the traces of materials and colors left on each pre-rendered gray image to impact users' choices, we render a new image with white color and Lambert material for all editable objects as the representative image.

5.2. Multi-layer Colorization

Given the rendered representative image of a given 3D scene, a user can assign specific colors to a small set of objects using a simple interface provided in our system. Our system then automatically infers colors for the objects without any user-assigned colors. Then, a multi-layer synergy mechanism will allocate these colors to certain gray images based on the gray value similarity, perform the colorization network, and finally merge the colorized images together through the semantic map.

Color Allocation. One core functionality of our multi-layer synergy mechanism is to explore a proper gray image to colorize a certain object or component in terms of its specified color. The first step is to infer the colors of the elements beyond the user-specified ones based on the colors of the input points. Let $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n\}$ ($\mathbf{p}_k = (x_k, y_k)$) be the set of input points drawn by users on the representative image, with a corresponding set of their colors $\mathcal{C} = \{c_1, c_2, \cdots, c_n\}$ specified by users. Benefited from the semantic map, the elements with the user-specified colors in $\mathcal{C}$ can be inquired by the coordinates $\mathcal{P}$. We denote these elements as $e_1, e_2, \cdots, e_n$ while the others (without explicitly specified colors) as $\tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_m$. Since the pre-collected color theme harmony priors also include the object-object color relations, they can be represented as conditional probability functions. Specifically, assume a and b are two kinds of elements (e.g., floor and wall) involved in the color theme harmony priors, and $\theta_a^b(c|\tilde{c})$ is denoted as the conditional probability that element a has the color c when element b has been painted in color $\tilde{c}$. The proper color c for an element $\tilde{e}$ can be explored by the elements with the user-specified colors as follows:

$$\arg\max_c \sum_k \theta_{\tilde{e}}^{e_k}(c|c_k). \tag{1}$$

Note that our color theme harmony priors only involve a limited number of discrete colors. In the case that the user-specified colors are not in the priors, we use the conditional probability functions of their most similar colors involved in our priors instead. Thus, all elements on the snapshot have associated colors.

The next step is to sample more points and explore the proper gray images for colorization. Considering that the elements might have various sizes and possibly disconnected regions on the snapshot, only using the small numbers of input points might fail to colorize the whole element. The elements without user-specified colors also need points with the inferred colors through Equation 1. Moreover, for a certain element, since the gray value of a pre-rendered gray-scale image with the same coordinate as a certain input point might be influenced by texture and/or shadow, it might fail to represent the major gray value of the element, thus making the incorrect gray image suggested for colorization. Therefore, besides the input points, we also sample more points on each element

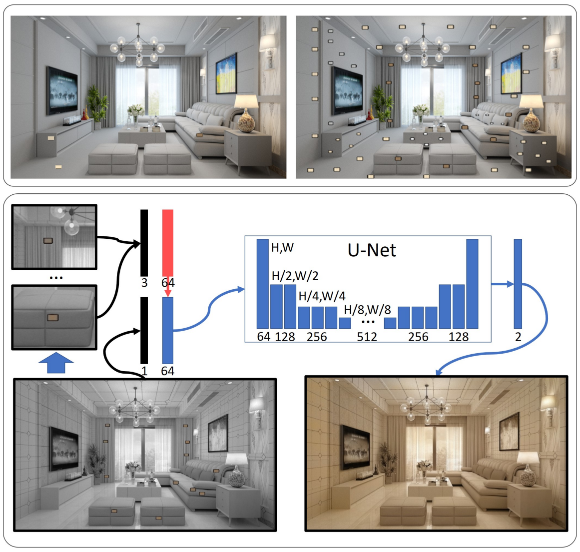

Figure 5: ***Top:*** *A small set of user-specified colors (left) and the sampled points for both user-specified and system-suggested colors (right).* ***Bottom:*** *The architecture of the colorization network employed in our system.*

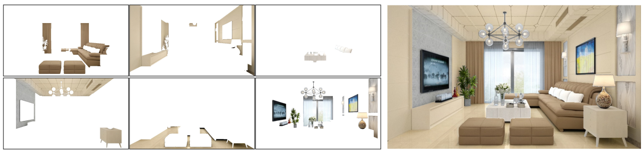

Figure 6: *Colorized regions are segmented from the associated layers (left) and then merged to a colored snapshot (right).*

to minimize the influences of texture and/or shadow on the similarity of the gray values between a user-specified or inferred color and the element of a certain gray image (e.g., Figure 5-Top). The number of sampled points is empirically set based on the element size, which is calculated as the number of covered pixels. A piece-wise constant function that maps the element size to the sampled point number is adopted to ensure each element to have 4 to 12 sampled points. These points are uniformly sampled on the snapshot via a sliding window. After that, we use each point and its $\alpha \times \alpha$ (α is an integer value depending on the element size and $\alpha \in [2,5]$ in our implementation) to explore the proper gray image. More specifically, we define a distance metric that calculates the difference between the gray values of the sampled and input points on a certain gray image and the associated user-specified or inferred color with respect to a certain element as follows:

$$D(g,i) = \sum_j \omega \cdot |g - G_i(\mathbf{p}_j)|, \tag{2}$$

where g is the gray value of the user-specified or inferred color for element e, $G_i(\mathbf{p}_j)$ is the gray value of the j-th sampled or input point for element e on the i-th gray image. Note that these points might have slight gray value differences even on the same gray image for a certain element as a result of the influences of texture and/or shadow. The weight ω is set to 5 for the input points, while for the sampled points $\omega = 1$. Then we choose the layer (i.e., gray image) with the minimum distance as the one to allocate such a user-specified or inferred color.

Colorization Network. After color allocation, all points with colors are assigned to certain gray images for colorization. Our colorization module is based on the network proposed by [ZZI*17], which uses a U-Net architecture [RFB15] as the main branch (Figure 5-Bottom). Since the network takes 256×256 images as input, we resize all gray image snapshots from their original sizes (800×450 in our implementation) to this size and adjust the coordinates of the user-specified and sampled color points correspondingly. For each gray image, the values of its lightness channel in the CIE Lab color space (denoted as $X^{H \times W \times 1}$) are integrated by concatenation with a tensor that encodes the associated color points (denoted as $U^{H \times W \times 3}$). Each element of the tensor is $[a, b, m]$. For the coordinate where a certain color point exists, $[a, b]$ are the values of two color channels in the CIE Lab color space of the color point. For the sampled and the interactive points we set $m = 1$, otherwise $m = 0$ to mask out the unspecified points. The output of the colorization network is denoted as $Y^{H \times W \times 2}$, representing the values of two color channels of the colorized image. Then we upsample Y to the original size of the snapshot, and combine it with the values of the lightness channel of the associated gray image in order to obtain the colorization result of this gray image. We directly use the network parameters of [ZZI*17], which were trained via the ImageNet dataset.

After that, our system uses the semantic map to cut out the colorized elements from each gray image and merges them together as our final output (e.g., Figure 6). To improve the colorization efficiency, the elements specified with gray colors (including white and black) are directly cut out from the allocated gray image(s) without colorization. Note that the user can specify more than one color to a certain element in our UI. In this scenario, the first one is regarded as the primary color for the color theme suggestion, while all these user-specified color points are used in the colorization network to generate the same-class elements but with different colors. Moreover, our system UI also allows users to modify the colorized snapshot by adding/removing input points, or changing the colors of certain input points. The output snapshot can be updated based on the user inputs, thus assisting users in completing the color theme design for the given scene (see our supplementary demo video). Even though our system only presents a colorized snapshot with a fixed view, it can organize the materials and colors based on the designed color theme and attach them to the objects/components of the given 3D indoor scene model. Hence, the design results can be easily extended to other software such as 3ds Max and Unity3D for various applications.

Figure 7: *Galleries of indoor scene color theme design results. Each row shows two different color themes for a given input scene. The user-specified colors are shown as color points in the second and third columns.*

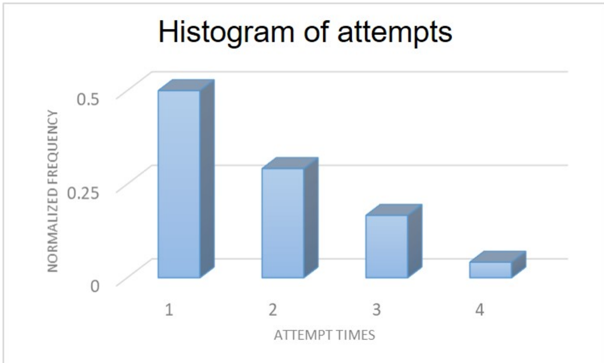

Figure 8: *Histogram of the average attempt times to complete the color theme design results in Figure 7.*

6. Experiment and Discussion

In this section, we first show representative results by our system, and then evaluate the efficiency and effectiveness of our system in indoor color theme design via several experiments. We also compare our system with a state-of-the-art automatic indoor material suggestion method [CXY*15], and a real-time indoor material editing toolkit based on Unity3D.

In Figure 7, we show the snapshot colorization results of six indoor scenes of different types. Each row has the representative image (Left) and the outputs with two different design color themes (Middle and Right). We highlight the input points with user-specified colors on each output colorized snapshot. In Figure 8, we evaluate the user-friendliness of our system with a histogram of attempt times to produce the results in Figure 7. It reflects how many times we tried by adjusting the inputs, until a satisfactory color theme is created. On average, the pre-processing of our system for an indoor scene takes about 25 minutes, the user interaction takes

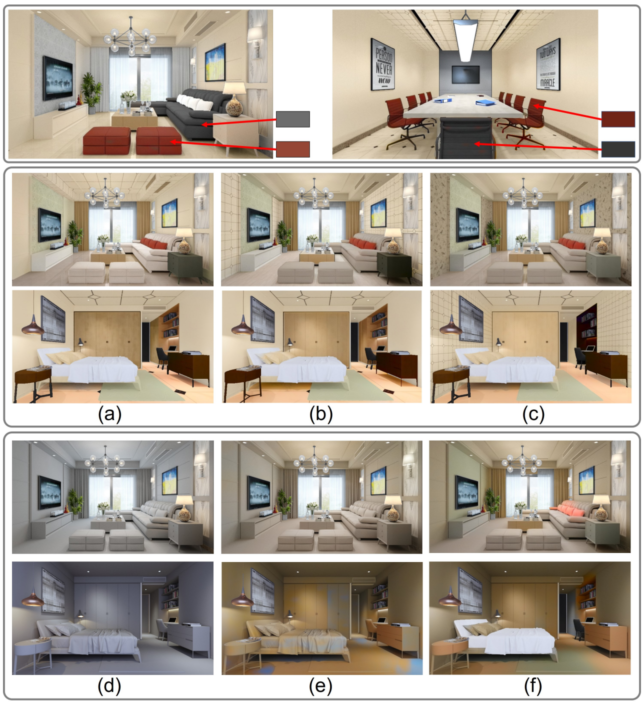

Figure 9: ***Top***: *Our system allows users to specify multiple colors to different instances of the same kind of objects, or different components of an object to enrich the colorization results.* ***Middle***: *We reduce the numbers of pre-rendered gray images from the original ones (9 for (a)-top and 8 for (a)-bottom) to 6 (b) and 3 (c) via uniform sampling, to evaluate the impact of the number of the pre-rendered gray images on colorization.* ***Bottom***: *The colorization results that we directly apply the network of [ZZI* 17] to the representative images (d) without (e) and with (f) automatically inferred color points.*

about 2 minutes, and the snapshot colorization takes about 13 seconds. The designer might try 1 to 4 times to obtain a satisfactory color theme. These tests are performed on a PC with Intel Core I7-7700K 4.2GHz CPU with 32GB RAM and GTX 1080 GPU.

In Figure 9-Top, we illustrate two examples of colorized objects in the same category with different user-specified colors. These examples show that users can modify certain regions or instances of certain kinds of elements flexibly with our system. In Figure 9-Middle, we test our system with reduced numbers of pre-rendered gray images. Snapshots in column (a) are generated with all gray images (i.e., 9 gray images for the top case and 8 for the bottom one). We uniformly sample 6 and 3 pre-rendered gray images and use the same inputs to generate the colorized snapshots in (b) and (c), respectively. We can see that our system can still work even with the reduced numbers of the pre-rendered gray images. However, for the two examples in Figure 9-(a), the numbers of pre-rendered images used for colorization are 5 (top) and 6 (bottom). Considering the examples in Figure 9-(b) have 6 pre-rendered images, while the examples in Figure 9-(c) only have 3 pre-rendered images, lacking pre-rendered gray images might cause colorization

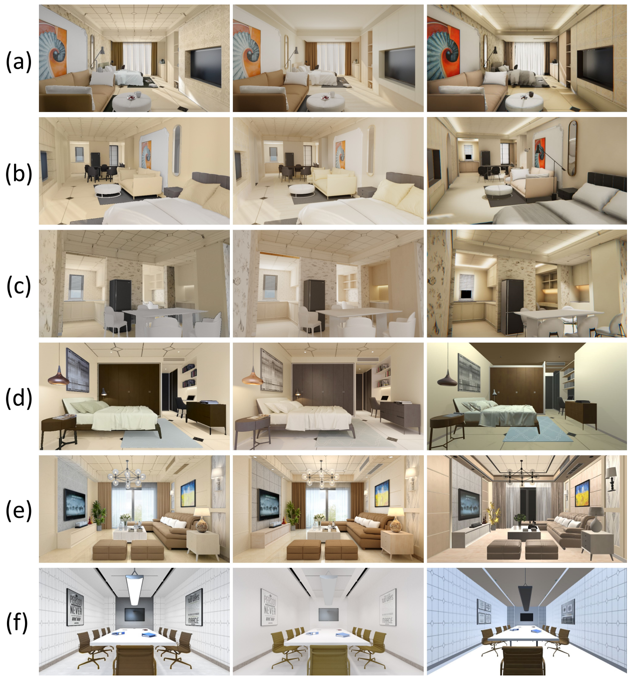

Figure 10: *The comparison between our results (left), similar color themes (middle) designed by [CXY* 15] and rendered by 3ds Max, and the same color themes rendered through Unity3D (right).*

issues. For example, some furniture might fail to be colorized faithfully respecting the user inputs (e.g., cabinet in two cases in Figure 9-(c)). It can be seen that more pre-rendered images could give more proper grayscale images and lead to richer color theme variations and better-colorized snapshots, though the required number of pre-rendered images depends on the complexity of both given scene and a desired color theme. On the other hand, considering the time cost of the pre-rendering process, limiting the number of pre-rendered grayscale images is a cost-effective solution. In Figure 9-Bottom, we validate our multi-layer colorization mechanism by comparing our results to the outputs of [ZZI*17] with the representative image as input (d). The results in (e) are obtained given the user-specified colors as those for generating (a). For the results (f), the same set of both user-specified and automatically sampled colors are used as input to the colorization algorithm in [ZZI*17]. These results show that the image colorization method cannot be directly used for the color theme design of 3D indoor scenes. Even with more color points, the quality of the colorized snapshot largely depends on the pre-defined brightness values in the input gray image. It validates the usability of our multi-layer colorization mechanism.

We also compare our system with a closely related work [CXY*15] and a real-time toolkit developed based on Unity3D, which allows users to interactively edit the materials of objects/components in the given 3D scene. In Figure 10 from (a) to (f), we show 6 groups of comparisons (each with 3 cases) of the indoor scene snapshots with the color themes created by our system (left and right cases) and the method of [CXY*15] (middle case).

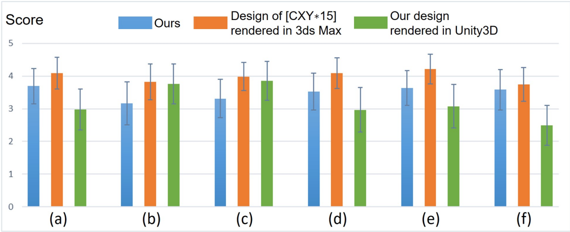

Figure 11: *The user study assessment results about the cases in Figure 10, in terms of the designed color themes and visualization qualities.*

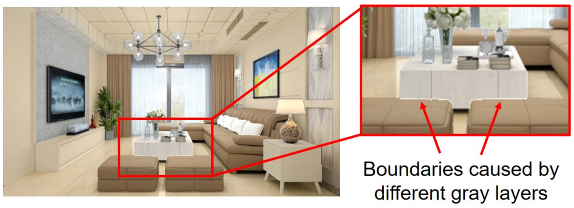

Figure 12: *The merging mechanism might lead to hard boundaries among the regions from different gray layers.*

Images in each group have the same views and light conditions. We illustrate our results (left) and the results with similar color themes suggested by [CXY*15] but rendered through 3ds Max (middle). We also specify the same materials as ours to the same scenes but render the scenes in Unity3D for comparison (right). Note that for the Unity3D results (a)-(c), the rendering parameter configurations are settled by experts and require more time for pre-processing than the results in (d)-(f), which use the default configurations.

To compare the qualities of both the color themes and the visualization effects of these results, we conducted a user study, for which we invited 55 participants (9 graduated and 46 undergraduate students majored in digital media technology) to rate images in Figure 10 between 0 (worst) and 5 (best), indicating the degree of visual quality. The images were presented to the participants in groups to ensure that the scores would reflect the quality differences between the three cases in each group. We summarized the user study results in Figure 11, with the average score for each image. It can be seen that benefited from our colorization mechanism, our system has a better visualization performance comparing to Unity3D, especially for scenarios when non-professional users choose the default rendering parameter configurations. Since the snapshots of [CXY*15] are rendered through 3ds Max, they have higher evaluation scores than ours. However, we can still qualitatively compare these designed color themes between [CXY*15] and ours in Figure 10. We can see that both interactive and automatic approaches can create harmonious indoor color themes. This indicates the effectiveness of our priors, which are not collected from large-scale indoor scene images. Considering that our system can generate a high-quality preview with no need to render the given scene again after the color theme is designed, our system has advantages on the intuitive interactive design paradigm compared to [CXY*15].

Limitations. Even though our system can assist users in previewing indoor scene snapshots with user-designed color themes quickly, there are still several limitations with our current system. First, the pre-processing stage is somewhat intricate. Since these pre-rendered multi-layer gray images are important to the quality of colored snapshots, and we cannot change the viewpoint of the 3D scene after the pre-processing, users should choose proper viewing and lighting conditions in this stage. It takes about an hour on average to prepare the pre-rendered images, thus increasing the time cost of our system. However, since we need to perform the pre-processing stage only once for a given indoor scene, these data can be reused to generate color theme variations to meet different customer demands. Second, since we leverage the semantic map of the snapshot to merge the colorized multi-layer images, when two adjacent elements are colorized in different layers, there might exist some artifacts around their boundaries in the colored snapshot (e.g., Figure 12). To alleviate this problem, some image fusion algorithms (e.g., [PGB03]) might be helpful to generate soft boundaries and improve the quality of the merged snapshot. Moreover, due to such a merging mechanism, lighting effects such as reflection and color bleeding between different objects can hardly be realistically generated by the current system especially when these objects come from different gray layers. Therefore, materials with specular reflection properties such as metal and glass are not suggested for objects in our system. Lastly, slight chromatic aberrations might exist between the user-specified colors and the colored snapshot, as a result of using the colorization network. A slight color adjustment might be still required even after the color theme is determined through our system.

7. Conclusion

In this paper, we presented a novel interactive system for color theme design of indoor scenes, via providing a quick preview of colorized snapshots with user-specified colors. The UI of our system displays a reference image of the given white indoor scene model to assist users in designing a desired color theme by painting certain colors on the image. Benefited from the collected color theme harmony priors, which encode both the color-material-object relations and color relations among different object categories, we perform a synergy colorization mechanism to leverage multiple layers of gray images to generate a colored snapshot of the given indoor scene quickly with the user-designed color theme. We demonstrate that our system can improve the efficiency of interior color theme design through several experiments and user studies that compare the results by our method, a closely related work, and Unity3D.

In the future, we would explore how to reveal the effects of more factors on interior color theme design, such as the relationship between the geometry styles of objects and colors/materials. To enable our system on the specular materials, we are interested in integrating our method into specially-designed rendering engines that could split diffuse and specular effects during rendering. This could make the color of a certain element and its mirror image on other specular surfaces editable simultaneously. Integrating our method into such rendering engines could also speed up the pre-rendering process. Moreover, using some generative deep networks might be more efficient if the preview of a user-designed color theme can be generated quickly through the networks without pre-rendering any image. We believe that this interactive system for the interior color theme will relieve interior designers from the heavy and duplicate workload.

Acknowledgements

We thank the anonymous reviewers for the constructive comments. This work was partially supported by grants from the NSFC (No.61902032), the Fundamental Research Funds for the Central Universities (2020RC17), and the Open Project Program of the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (VRLAB2019B01).

References

[BUSB13] BELL S., UPCHURCH P., SNAVELY N., BALA K.: Opensurfaces: a richly annotated catalog of surface appearance. *Acm Transactions on Graphics 32*, 4 (2013), 111. 3

[BUSK15] BELL S., UPCHURCH P., SNAVELY N., K.BALA: Material recognition in the wild with the materials in context database. In *CVPR* (2015), pp. 3479–3487. 3

[CLS10] CHAJDAS M. G., LEFEBVRE S., STAMMINGER M.: Assisted texture assignment. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (2010), pp. 173–179. 3

[COSG*06] COHEN-OR D., SORKINE O., GAL R., LEYVAND T., XU Y.-Q.: Color harmonization. *Acm Transactions on Graphics* (2006), 624–630. 3

[CXY*15] CHEN K., XU K., YU Y., WANG T., HU S.: Magic decorator: automatic material suggestion for indoor digital scenes. *ACM Transactions on Graphics 34*, 6 (2015), 232. 1, 2, 3, 4, 7, 8, 9

[CZG*11] CHIA A. Y., ZHUO S., GUPTA R. K., TAI Y., CHO S., TAN P., LIN S.: Semantic colorization with internet images. *ACM Transactions on Graphics 30*, 6 (2011), 156. 2

[CZZT12] CHEN X., ZOU D., ZHAO Q., TAN P.: Manifold preserving edit propagation. *Acm Transactions on Graphics 31*, 6 (2012). 1

[DLY*17] DESHPANDE A., LU J., YEH M., CHONG M. J., FORSYTH D.: Learning diverse image colorization. In *CVPR* (2017), pp. 2877–2885. 2

[FCW*17] FU Q., CHEN X., WANG X., WEN S., ZHOU B., FU H.: Adaptive synthesis of indoor scenes via activity-associated object relation graphs. *ACM Transactions on Graphics 36*, 6 (2017), 201. 3

[FPC*14] FARIDUL H. S., POULI T., CHAMARET C., STAUDER J., TREMEAU A., REINHARD E.: A survey of color mapping and its applications. In *Eurographics -State of the Art Reports* (2014), pp. 43–67. 2

[HZMH14] HUANG H. Z., ZHANG S., MARTIN R. R., HU S.: Learning natural colors for image recoloring. *Computer Graphics Forum 33*, 7 (2014), 299–308. 2

[ISI16] IIZUKA S., SIMOSERRA E., ISHIKAWA H.: Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *Acm Transactions on Graphics 35*, 4 (2016), 110. 2

[JTRS12] JAIN A., THORMÄHLEN T., RITSCHEL T., SEIDEL H.-P.: Material memex: Automatic material suggestions for 3d objects. *ACM Transactions on Graphics 31*, 6 (2012). 3

[KYKL14] KIM H., YOO M., KANG H., LEE I.: Perceptually-based color assignment. *Computer Graphics Forum 33*, 7 (2014), 309–318. 3

[LE07] LALONDE J., EFROS A. A.: Using color compatibility for assessing image realism. In *ICCV* (2007), pp. 1–8. 3

[LKPL14] LU P., KUANG Z., PENG X., LI R.: Discovering harmony: A hierarchical colour harmony model for aesthetics assessment. In *ACCV* (2014), pp. 452–467. 3

[LRFH13] LIN S., RITCHIE D., FISHER M., HANRAHAN P.: Probabilistic color-by-numbers: suggesting pattern colorizations using factor graphs. *Acm Transactions on Graphics 32*, 4 (2013), 37. 2

[LWZ*17] LIU T., WEI Y., ZHAO Y., LIU S., WEI S.: Magic-wall: Visualizing room decoration. In *ACM International Conference on Multimedia* (2017), pp. 429–437. 1

[LZNH15] LI X., ZHAO H., NIE G., HUANG H.: Image recoloring using geodesic distance based color harmonization. *Computational Visual Media 1*, 2 (2015), 143–155. 3

[OAH11] O'DONOVAN P., AGARWALA A., HERTZMANN A.: Color compatibility from large datasets. *Acm Transactions on Graphics 30*, 4 (2011), 63. 3

[PGB03] PEREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Transactions on Graphics 22*, 3 (2003), p.313–318. 9

[PRFS18] PARK K., REMATAS K., FARHADI A., SEITZ S. M.: Photoshape: Photorealistic materials for large-scale shape collections. *ACM Transactions on Graphics 37*, 6 (2018). 3

[RFB15] RONNEBERGER O., FISCHER P., BROX T.: U-net: Convolutional networks for biomedical image segmentation. In *MICCAI* (2015), pp. 234–241. 7

[SLF*17] SANGKLOY P., LU J., FANG C., YU F., HAYS J.: Scribbler: Controlling deep image synthesis with sketch and color. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 6836–6845. 2

[WAM02] WELSH T. F., ASHIKHMIN M., MUELLER K.: Transferring color to greyscale images. *Acm Transactions on Graphics 21*, 3 (2002), 277–280. 2

[WJLC12] WANG X., JIA J., LIAO H., CAI L.: Affective image colorization. *Journal of Computer Science and Technology 27*, 6 (2012), 1119–1128. 3

[WMG*09] WALD I., MARK W. R., GÜNTHER J., BOULOS S., IZE T., HUNT W., PARKER S. G., SHIRLEY P.: State of the art in ray tracing animated scenes. *Computer Graphics Forum 28*, 6 (2009), 1691–1722. 1

[WYW*10] WANG B., YU Y., WONG T., CHEN C., XU Y.: Data-driven image color theme enhancement. *Acm Transactions on Graphics 29*, 6 (2010), 146. 3

[XCM*14] XU K., CAO Y. P., MA L. Q., DONG Z., WANG R., HU S. M.: A practical algorithm for rendering interreflections with all-frequency brdfs. *Acm Transactions on Graphics 33*, 1 (2014), 1–16. 1

[YYT*11] YU L.-F., YEUNG S.-K., TANG C.-K., TERZOPOULOS D., CHAN T. F., OSHER S. J.: Make it home: Automatic optimization of furniture arrangement. *ACM Transactions on Graphics 30*, 4 (2011). 3

[YYT16] YU L., YEUNG S., TERZOPOULOS D.: The clutterpalette: An interactive tool for detailing indoor scenes. *IEEE Transactions on Visualization and Computer Graphics 22*, 2 (2016), 1138–1148. 3

[ZGM18] ZHU J., GUO Y., MA H.: A data-driven approach for furniture and indoor scene colorization. *IEEE Transactions on Visualization and Computer Graphics 24*, 9 (2018), 2473–2486. 1, 3, 4

[ZIE16] ZHANG R., ISOLA P., EFROS A. A.: Colorful image colorization. In *ECCV* (2016), pp. 649–666. 2

[ZWW18] ZSOLNAIFEHER K., WONKA P., WIMMER M.: Gaussian material synthesis. *ACM Transactions on Graphics 37*, 4 (2018), 76. 3

[ZZI*17] ZHANG R., ZHU J., ISOLA P., GENG X., LIN A. S., YU T., EFROS A. A.: Real-time user-guided image colorization with learned deep priors. *ACM Transactions on Graphics 36*, 4 (2017), 119. 1, 2, 3, 7, 8