

Studying the Effect of Material and Geometry on Perceptual Outdoor Illumination

Miao Wang, *Member, IEEE*, Jin-Chao Zhou, Wei-Qi Feng, Yu-Zhu Jiang and Ana Serrano, *Member, IEEE*

Abstract—Understanding and modeling perceived properties of sky-dome illumination is an important but challenging problem due to the interplay of several factors such as the materials and geometries of the objects present in the scene being observed. Existing models of sky-dome illumination focus on the physical properties of the sky. However, these parametric models often do not align well with the properties perceived by a human observer. In this work, drawing inspiration from the Hosek-Wilkie sky-dome model, we investigate the perceptual properties of outdoor illumination. For this purpose, we perform a large-scale user study via crowdsourcing to collect a dataset of perceived illumination properties (*scattering*, *glare*, and *brightness*) for different combinations of geometries and materials under a variety of outdoor illuminations, totaling 5,000 distinct images. We perform a thorough statistical analysis of the collected data which reveals several interesting effects. For instance, our analysis shows that when there are objects in the scene made of rough materials, the perceived *scattering* of the sky increases. Furthermore, we utilize our extensive collection of images and their corresponding perceptual attributes to train a predictor. This predictor, when provided with a single image as input, generates an estimation of perceived illumination properties that align with human perceptual judgments. Accurately estimating perceived illumination properties can greatly enhance the overall quality of integrating virtual objects into real scene photographs. Consequently, we showcase various applications of our predictor. For instance, we demonstrate its utility as a luminance editing tool for showcasing virtual objects in outdoor scenes.

Index Terms—Perceptual Outdoor Illumination, Geometry, Material.

1 INTRODUCTION

HUMAN perception of environmental illumination is influenced not only by the surrounding lighting, such as indoor lamps or outdoor natural sunlight, but also by the material reflectance, and surface geometry of the main observed objects in the scene. Different materials have different effects on the reflective properties of lighting, resulting in variations in perceived illumination. For instance, the same light illuminating a piece of wood and a piece of glass of the same shape shows different reflections, while the same wood material with different shapes forms different light transports. Understanding the interplay of material and geometry on perceived illumination is key to improving virtual object insertion and editing in captured real-scene photos. Our work bridges the gap between physical properties of sky-dome illumination and the perceptual traits of outdoor illumination, contributing to creating more realistic results.

Perceptual material appearance is a popular research field that is highly related to the perception of illumination, as lightness can be regarded as a material property. There are many works investigating the perceptual properties of materials [1], [2], [3], [4], most of which focus on only a few appearance properties

such as gloss [5], [6], [7] or translucency [8], [9]. Some works study the impact of both geometry and illumination under a limited variety of material reflectances, geometries and illuminations [10], [11], [12], [13], [14], [15]. Closer to our work, Serrano et al. [16] provide a comprehensive summary of the existing research and systematically organize the study of the impact of geometry and illumination on material appearance perception. Although lightness can be regarded as one of the main material reflectance properties and has been well explored in previous studies, the interactions between lightness properties and perceptual illumination, and the impact of geometry and material reflectance on perceptual illumination are still not fully understood, even for natural outdoor environments with simple lighting conditions.

For outdoor environments with natural scenes, the sun is commonly regarded as the light source. Previous works take efforts toward analytic outdoor illumination modeling [17], [18], which aims to reconstruct the ground-truth environment map from a single captured image by recovering a compact set of sky parameters. Some works build a probability distribution over the sun position and visibility, attempting to estimate the analytical outdoor environment from extracted image statistics through data-driven optimization [19]. Deep learning-based methods [18], [20] use end-to-end image feature extraction and regression pipelines to estimate the outdoor sunlighting by incorporating convolutional neural networks, achieving considerable performance.

However, although these models can provide an accurate estimation of the properties of outdoor environments, there is still a disconnect between purely physical properties and the perception of illumination, which heavily depends on other characteristics of the scene such as the materials and geometries of the objects that are visible. In this work, we use the Hosek-Wilkie analytical model [17] as a conceptual inspiration. We take key parameters from this model as a guide, reinterpreting them as perceptual

- Miao Wang is with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, Beijing 100191, China, and also with Zhongguancun Laboratory, Beijing 100094, China. E-mail: miaow@buaa.edu.cn
- Jin-Chao Zhou, Wei-Qi Feng, and Yu-Zhu Jiang are with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, Beijing 100191, China. E-mail: {zhoujinchao, fengweiqi, 20373779}@buaa.edu.cn
- Ana Serrano is with Universidad de Zaragoza, 50018, Zaragoza, Spain. Email: anase@unizar.es
- Corresponding author: Miao Wang.

Manuscript received Sept., 2023

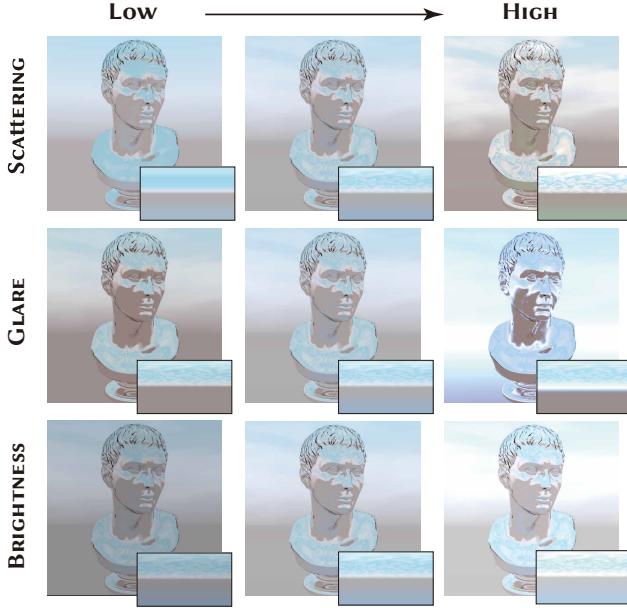


Fig. 1: Illustration of the perceptual attributes *scattering*, *glare*, and *brightness* under changing environments. The top row represents the perceptual scattering, where the surface reflection appears to be increasingly muddy when raising the scattering level. The middle row shows the perceptual glare, where the surface becomes shinier when increasing the glare level. The bottom row represents the perceptual intensity of the sunlight. All the perceptual attributes are controlled by tuning the environment maps (bottom-right), where the environment maps in the middle column are set as identical for better visual comparisons.

attributes to investigate outdoor sky lighting. We focus on outdoor environments and investigate how the human visual system (HVS) perceives such illumination under different combinations of surface geometries and material reflectances [21].

Specifically, in order to collect sufficient data for the study, we synthesize 5,000 images of combinations of geometries and materials under a variety of outdoor illuminations using photo-realistic rendering software [22]. Then, we use crowd-sourcing to collect ratings for these images for three subjective outdoor sky attributes: *scattering*, *glare* and *brightness*, building a comprehensive perceptual outdoor illumination dataset. Fig. 1 illustrates examples of these three perceptual attributes from low (left) to high (right) levels. Further, we conduct a statistical analysis of the effect of geometry and material on perceived outdoor illumination to understand how the HVS is influenced by geometry or material reflectance factors. Leveraging our dataset, we also train a deep neural network to predict such perceptual outdoor illumination attributes from real-world photos. These two components, though distinct, form integral parts of our research, collectively enhancing our understanding and predictive capabilities in the realm of outdoor illumination. Finally, we demonstrate the benefits of our trained perceptual outdoor illumination attribute predictor with several applications. For instance, our predictor can be used as a tool for designers and artists to conveniently select and produce background images that match their perceptual goals. The main contributions of our work are as follows:

- A new perceptual outdoor illumination dataset that includes 5,000 images with distinct combinations of mate-

rial, shape and outdoor environment illumination produced using physically-based rendering, as well as their corresponding crowd-sourced subjective ratings for perceived sky attributes;

- A comprehensive study on the impact of geometry and material on perceived outdoor illumination;
- A novel perceptual outdoor illumination model that predicts the scattering, glare, and brightness properties of an observed scene;
- A demonstration of potential applications of our perceptual outdoor illumination model.

Our collected dataset and predictor are publicly available at: <https://buaavrcg.github.io/PerceptualOutdoorIllumination>

2 RELATED WORK

Our goal is to understand the effect of surface geometry and material reflectance when perceiving outdoor illuminations and propose a model that accounts for these interactions. This is closely related to the fields of environment illumination estimation (Sec. 2.1), overall appearance perception (Sec. 2.2), and in particular lightness perception (Sec. 2.3).

2.1 Illumination estimation

Illumination estimation aims to accurately reconstruct real-world complex environment lighting with daily used equipment or exactitude instruments under controllable costs. Environment lighting can mainly be divided into two categories: outdoor environments and indoor environments. Due to diverse surface geometries and materials with complex visibility situations in indoor scenes, indoor illumination estimation is complex and highly challenging [23], [24], [25]. As a first step towards modeling the effect of geometry and materials in perceived illumination, in this work we focus on outdoor environment lighting, remaining indoor illumination a promising line of future work. For outdoor illumination estimation, Lalonde et al. [19] proposed a generic estimation method, which calculates probability distribution related to sun properties using image statistics extracted from multiple images. Hold-Geoffroy et al. [18] proposed a deep convolutional network trained with a large image dataset capable of regressing the estimated sun position and intensity for a given image, following the Hosek and Wilkie parametric sky representation model [17]. Later, Hold-Geoffroy et al. [20] extended their work by proposing an end-to-end deep neural network, replacing the parametric sky presentation with a latent code, which led to better performance in approximating outdoor illuminations. Recent advancements improve the feature extraction network through attention mechanism [26] or multi-scale feature fusion [27], resulting in significant success across numerous computer vision tasks. We leverage the power of neural networks while building our research on the foundation of Hosek-Wilkie parametric sky model, which has been widely used for real-time sky-dome rendering [28].

2.2 Image-based appearance perception

Since the light entering the human eye is produced by a combination of the surface properties of the object and the light source, the perception of material appearance is strongly influenced by incident illumination. Therefore, illumination and material perception are two closely related topics. Classic methods of image-based appearance perception argued that glossiness and lightness



Fig. 2: Representative sample environment maps from our collected illumination database.

correlate with simple image statistics (e.g. standard deviation, kurtosis, skewness) derived from luminance histograms [12], or the similarity of the subband histograms of the images [29]. Marlow et al. [11] suggested that the perception of gloss relies on a complex heuristic weighting of cues of specular image structures such as highlight coverage, sharpness and contrast. In follow-up work, Fleming et al. [30] pointed out that these methods are not ubiquitous on existing perceptual data and proposed highly nonlinear visual encodings to explain the material perception process. Since complex non-linear encodings have been shown to better explain perceptual impressions in complex images, deep learning has been recently adopted in many works as a unified framework for computationally modeling such non-linear encodings. Deep features [31] have been leveraged to discover sets of representative materials by training material classifiers. Further, Storrs and Fleming [32] showed that deep learning architectures, in particular a variational autoencoder (pixelVAE), can be trained in an unsupervised way to reproduce human gloss judgments for simple scenes with variable bump-maps and different environment illuminations. Recently, Serrano et al. [16] showed that deep neural networks trained with their proposed extensive set of perceptual data can effectively predict perceptual reflectance attributes and are robust to various scenarios. Inspired by these works, we leverage a deep neural network for modeling the perception of illumination properties as well as discovering the complex interactions between geometry and material reflectance on illumination perception.

2.3 Lightness perception

Common objects with shiny surfaces can reflect surrounding lighting information such as sun position or light source intensity, while it is difficult for humans to perceive the environment illumination information from rough surfaces such as wood or paper [13], [33]. Previous work [12] suggested that there are neural mechanisms sensitive to perceptual lightness and image statistics that correlate perceived illumination with the standard deviation of the luminance of the resulting image. Recently, the work of Toscani et al. [34] showed that, while the brightest regions of matte surfaces are good predictors of lightness, these regions have a limited impact on lightness perception for glossy surfaces. The combination of local and global anchoring of lightness values [35] appears to provide a unified account of intrinsic image models and surface patterns. The visible surface region also seems to correlate with perceived lightness [34], indicating that geometry surface properties play an important role in the human visual perception of lightness. Recent studies [13], [21], [36] found that specular highlights may be ignored by human observers while they evaluate the lightness of glossy objects, resulting in

a weak correlation to perceived lightness. Existing studies usually target a particular effect with a limited set of simple stimuli for controlled experiments. In contrast, in our work we aim to study perceived lightness and its interactions with material and geometry in realistic images with real-world illuminations building upon a formal sky lightness model.

3 PERCEPTUAL DATA COLLECTION

This section outlines the experimental stimuli and method used to collect our perceptual responses through crowdsourcing. Further details are available in the supplementary material.

3.1 Stimuli

We use a wide variety of environment light probes under different combinations of geometries and materials, totaling 5,000 distinct rendered images covering 5 materials, 5 geometries, and 200 illuminations represented by lat-long warped environment images. We use Mitsuba 1 [22] for rendering our stimuli into low dynamic range images with resolution 512×512 and gamma correction.

3.1.1 Illuminations

The real world has complex and variable light conditions. For simplification, we only focus on outdoor environments. We collect 312 real-world natural outdoor high dynamic range (HDR) environment maps from online sources^{1,2}, and compute their high-frequency content [37] after normalizing them so that the integral of the luminance is the same for all environment maps [16], [38]. We select 200 of them aiming for variable outdoor illuminations and a balanced distribution of image statistics, and ensuring a variety of sun sky brightness, different types of atmospheres, and different weather conditions. The manual selection we have employed demonstrates strong credibility regarding our perceptual attributes. Furthermore, we have ensured a well-balanced distribution of illumination within the selected samples by incorporating the concept of importance sampling, particularly emphasizing their high-frequency content. Representative images of the selected environment maps are shown in Fig. 2, and a detailed analysis of the illuminations can be found in Sec. 4.

3.1.2 Materials

We represent real-world materials with the bidirectional reflectance distribution function (BRDF) and, inspired by Guo et al. [39], we select 5 representative BRDFs from 520 measured ones used in previous studies [16]. These materials are categorized

1. <https://polyhaven.com>

2. <https://hdremaps.com/hdre/>

as *leather*, *fabric*, *bio*, *metal*, and *natural* based on their physical properties and real-world category. *Bio* stands for materials that have rough and bumpy opaque surface and *natural* stands for materials made of natural substances. These categories are easy for humans to understand intuitively. For example, *leather* is less reflective while *metal* reflects more the surrounding environment. More details on our selection criteria are provided in Section 4.

3.1.3 Geometries

The 5 selected geometries have varied features in order to produce different reflections and perceived effects. We include commonly used geometries in graphics and material perception research (*bunny*, *buddha*, *teapot*). *Bunny* and *teapot* represent relatively simple shapes with low complexity, while *buddha* represents a more complex geometry and has more smoothed normals. Then we add *blob* as it has been shown to be one of the best commonly used shapes for material discrimination tasks [40] and has the most smooth gaussian curvature ($k_G = k_1 k_2$, where k_1, k_2 are two principal curvatures at a point on a surface). Finally, we add the *ghost* which has the most rugged surface with high gaussian curvature and can easily influence the human visual perception since it has been specifically optimized for single image material comparison tasks [41].

3.2 Methods

3.2.1 Perceptual attributes

Hosek and Wilkie [17] proposed an analytic model (Hosek-Wilkie model) to represent outdoor lighting, which parameterizes the HDR sky dome using 4 parameters. For simplicity, we focus on two parameters: ground albedo and turbidity, and we do not take into account the remaining parameters modeling complex solar elevation variations as these do not affect solar radiance. Inspired by the Hosek-Wilkie sky model, we propose a perceptual sky model, which uses three parameters to parameterize the perceived outdoor sky: *scattering*, *glare* and *brightness*. *Scattering* is a simple and intuitive measure of the number of aerosols in the air, which is reflected in the clarity of the sky, *glare* indicates whether the sunlight is focused on the object or is reflected on the object, and *brightness* indicates the intensity of the sunlight. In Fig. 1 we illustrate the perceived visual impact of these parameters.

3.2.2 Participants

A total of 133 participants (50.4% female, average age=21.7, $\sigma = 2.93$ years old) took part in our perceptual survey. Of those participants, 33.9% claimed to have experience in computer graphics, 39.2% claimed to have experience with design and modeling software and 24.7% to claimed have artistic knowledge. Participants completed the study online on regular office displays.

3.2.3 Procedure

To collect our subjective ratings we implemented a web interface that allowed participants to complete the survey on their own displays, similarly to previous works [16], [42]. Please, see Fig. 3 and supplementary material for more details. Before launching this large-scale survey, we run several iterations of a pilot survey on a small range of controllable participants with the goal of refining the interface and explanations, and in order to ensure that the final version was well-understood and the chosen sky parameters were meaningful and descriptive. When these pilot tests were performed with good feedback from the participants, we moved to

the large-scale survey. We performed our study by crowdsourcing, publicizing our web interface on the internet, and inviting a region of participants to access our survey. For ensuring the robustness of the collected data and for anchoring the limits of the rating scales, we included a tutorial and a training stage with examples [43]. In this stage, the task and different attributes were explained to the participants, and they were presented with a short training session with obvious examples (not part of our tested stimuli) to help them get familiar with the web interface and minimize unreliability [44]. We also included control images throughout the survey for detecting unreliable participants [45]. We collected demographic information: gender (male-female-other), knowledge of computer graphics, experience with design or modeling software, and artistic experience (none-basic-intermediate-professional). Each trial consisted of an image in which the participant had to rate our three sky parameters (scattering, glare and brightness). Similarly to previous works [16], [33], we chose a Likert rating task for a good trade-off between the number of trials and difficulty of the task and used a 7-point Likert scale for a good balance between granularity and complexity [46]. The extremes of the scales were labeled as "low" and "high". Each survey consisted of 20 images selected randomly from our image pool and 3 control images: one showing a very muddy image and two showing very clear images, to detect inattentive or malicious participants. Data from a participant is discarded if the responses for the muddy image include low *scattering* and high *brightness* (*scattering* < 3 or *scattering* > 5) or the responses for the clear image include high *scattering* and low *glare* (*scattering* > 5 or *scattering* < 3). As a result, we obtained 20,112 valid responses, from which we ensure that each image had been viewed at least 3 times.

4 PERCEPTUAL DATA ANALYSIS

In this section, we first perform a statistical analysis of the effects of material and geometry on our perceptual outdoor illumination attributes: *scattering*, *glare*, and *brightness*. Then we analyze the pairwise interaction effects of geometry and material. In the supplementary material, we provide more analysis data that further elaborate on the perceptual illumination attributes.

4.1 Statistical analysis

We first investigate whether *material* and *geometry* each have a significant impact on participants' subjective ratings. Our dependent variables (sky attributes) are 7-point Likert items and the collected observations do not obey a normal distribution ($p < 0.05$ for the Kolmogorov-Smirnov test). Therefore we adopt a cumulative link mixed model (CLMM) for analyzing our ordinal-scale observations [47], [48], [49]. Cumulative link models have been used in previous works to analyze Likert data [16] since they are particularly tailored to handle the ordered nature of Likert items by treating these ratings as consecutive instead of quantitative. For each level of the ordinal response, the cumulative link model computes *cut-off points* that separate the levels of the ordinal response as the cumulative probability of being in such level or lower. We employed the Estimated Marginal Means with Bonferroni correction for post-hoc multiple comparisons. The fixed factors of our analysis are the *material* and *geometry* parameters, while *user* (representing individual participant) and *illumination* are considered as mixed effects. This approach enables us to account for residual differences caused by the diversity

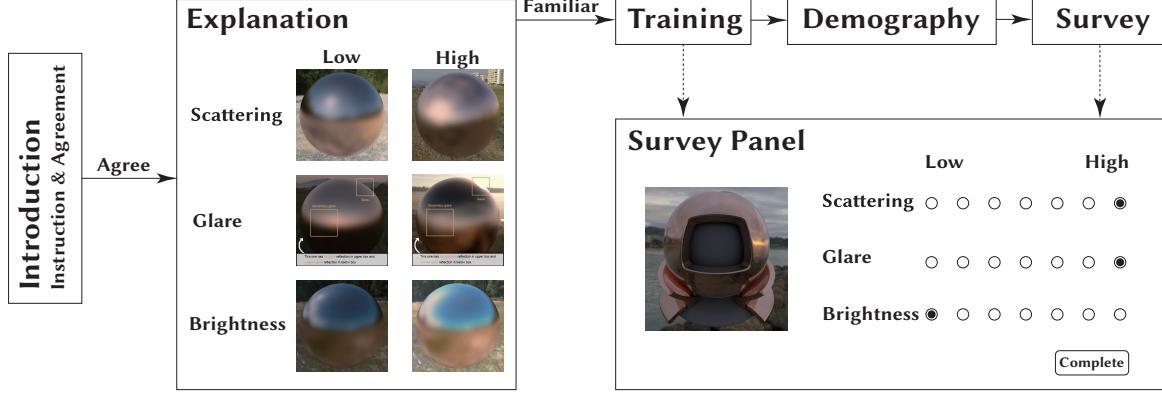


Fig. 3: Procedure for crowdsourcing perceptual data. We represent the procedure with a diagram to simplify the actual web interface for a better understanding of our workflow. The screenshots of our web interface are provided in the supplementary material.

and variations of participants. The statistical significance level α for all tests was set to 0.05.

4.1.1 Effect of material

The study demonstrated a significant effect of material on the perceived illumination attributes ($p < 0.001$). The reflective properties of surfaces determine the visual patterns of illumination perceived, making the selected materials a critical factor in determining the perceived illumination attributes. Fig. 4 presents the mean response of attributes for each level of the factor, adjusted for other variables in the model. Materials with high roughness, such as *leather*, *fabric*, and *bio*, result in a relatively high perceived scattering of illumination. On the other hand, materials with clear reflections, such as *metal* and *natural*, lead to lower perceived scattering of illumination. The roughness of materials results in cloudy reflections that contribute to the perception of a higher scattering environment. In terms of glare perception, a significant difference was observed with *leather* compared to other materials due to its absorption properties, leading to lower perceived glare. For perceived brightness, *fabric* and *natural* materials exhibit the highest ratings. Clear reflections are observed in all materials except for *leather*, which contributes to higher perceived brightness.

4.1.2 Effect of geometry

We discovered a significant impact of geometry on the perceived illumination attributes, with $p < 0.001$. The effect of geometry on the perception of illumination is highly relevant since it plays a vital role in determining the light transport path. Fig. 4 depicts the adjusted impact of geometry on each attribute. Our findings suggest a uniform trend in the effect of geometry on all perceived attributes. Geometries featuring large, smooth surfaces that offer clear reflections (*ghost*, *teapot* and *blob*) produce lower perceived scattering, while geometries with more intricate details (*bunny*, *buddha*) result in higher perceived scattering, potentially due to the lack of a large flat surface area to clearly observe reflections. The results also suggest that complex geometry shapes with steep gradients tend to produce lower glare effects, as exemplified by *buddha*, which produced the lowest perceived glare rating. While the smooth surface of *blob* resulted in stable ratings for all perceived attributes, it is more prone to reflect light with varying normals. A possible reason for this is that the perception of relative lightness among different regions is determined almost entirely by the diffuse component of visible surface [36].

4.2 Interactions of material and geometry

We compute all pairwise post-hoc tests which yield 25 combinations (5 materials \times 5 geometries) for investigating the interaction of material and geometry. For better visualization, we process a subset of the data and visualize it in Fig. 5. Inspired by Guo et al. [39], we compute the statistics of mean albedo (M_α), scatter gloss (M_{sg}) and scatter anisotropy (M_{sa}) for materials as follows:

$$\begin{cases} M_\alpha = \sum_{k=0}^N \frac{\alpha(\frac{k\pi}{2N})}{N} \\ M_{sg} = \sum_{k=0}^N \frac{N \cdot s_3(\frac{k\pi}{2N})}{s_1(\frac{k\pi}{2N})} \\ M_{sa} = \sum_{k=0}^N \frac{N \cdot s_2(\frac{k\pi}{2N})}{s_1(\frac{k\pi}{2N})}, \end{cases} \quad (1)$$

where N denotes sample counts to approximate the integral, $\alpha(\omega_o)$ represents the albedo computed as $\alpha(\omega_o) = \int_{\mathcal{S}^2} \rho(\omega_o, \omega_i) \cos \theta_i d\omega_i$, and $s_i(\omega_o)$ denotes the eigen values of scatter matrix $\mathbf{S}(\omega_o) = \int_{\mathcal{S}^2} \omega_i \omega_i^T \hat{\rho}(\omega_o, \omega_i) d\omega_i$.

The above three variables then have the ability to represent all aspects of material properties in the material space. For geometries, since users can only observe the geometry from a fixed viewpoint based on rendered image settings, we compute the high-frequency content (G_{hfc}) of the visible region to represent the geometry variations [37]. We provide the statistics computed from material and geometry in terms of the perceived *scattering*, *glare* and *brightness* in Fig. 6, in order to enrich our analysis.

4.2.1 Scattering

The changes in the perception of scattering, observed by variations in either material or geometry, exhibit similar trends across different geometries and materials, as demonstrated in Fig. 5 (left column). Our findings suggest that the perception of scattering is strongly influenced by the reflectance of the material (M_α), with an inverse trend observed for material anisotropy (M_{sa}), where the anisotropically reflected light is restricted to a limited range of directions. In addition, surface geometry also plays a significant role in the perceived scattering. Specifically, the *blob* geometry consistently receives the lowest rating due to its smooth and simplified shape, resulting in low values of G_{hfc} . The curvature of the surface varies slowly across the visible region in the rendered

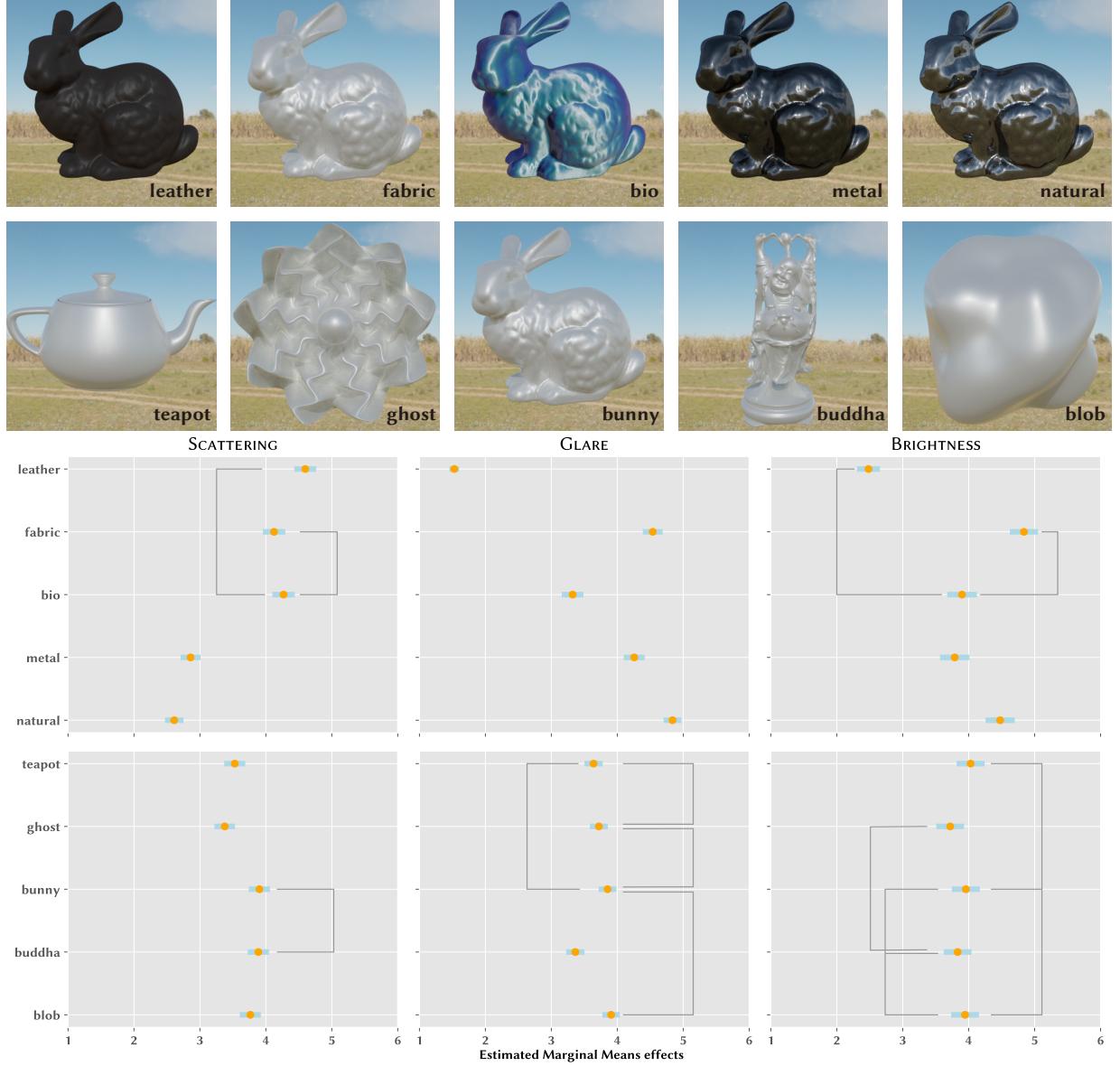


Fig. 4: Top: Example of our five materials (first row) and geometries (second row) with the illumination *hdre-276* from the collected environment maps. Bottom: Effect of material (third row) and geometry (fourth row) in the perceived attributes of illumination (scattering, glare and brightness). Estimated Marginal Means approximate the mean response for each factor, adjusted by *Bonferroni* correction for multiple comparisons. The orange dots indicate the mean values, and the blue bars indicate the 95% confidence intervals. The clusters of factors that do not yield statistically significant differences are connected with gray lines.

TABLE 1: Evaluation results with the *bunny* and *metal* test dataset (Set A) and extrapolation results with the additional validation dataset (Set B). Mean absolute error \pm standard deviation for our two tested architectures and loss functions.

Network structure	Evaluation (Set A)		Extrapolation (Set B)	
	MSE Loss	MSLE Loss	MSE Loss	MSLE Loss
ResNet52	0.1320 ± 0.1065	0.1289 ± 0.1000	0.2394 ± 0.1612	0.2323 ± 0.1541
VGG16	0.1274 ± 0.1010	0.1247 ± 0.0999	0.2238 ± 0.1506	0.2056 ± 0.1330

TABLE 2: Evaluation (set A) and extrapolation (set B) results for each of our three predicted attributes with our selected configuration (VGG16, MSLE loss). Mean absolute error \pm standard deviation for each attribute.

	Scattering	Glare	Brightness	All
Evaluation (Set A)	0.1195 ± 0.0095	0.1209 ± 0.1017	0.0990 ± 0.0450	0.1247 ± 0.0999
Extrapolation (Set B)	0.1859 ± 0.0944	0.3180 ± 0.0532	0.0639 ± 0.0229	0.2056 ± 0.1330

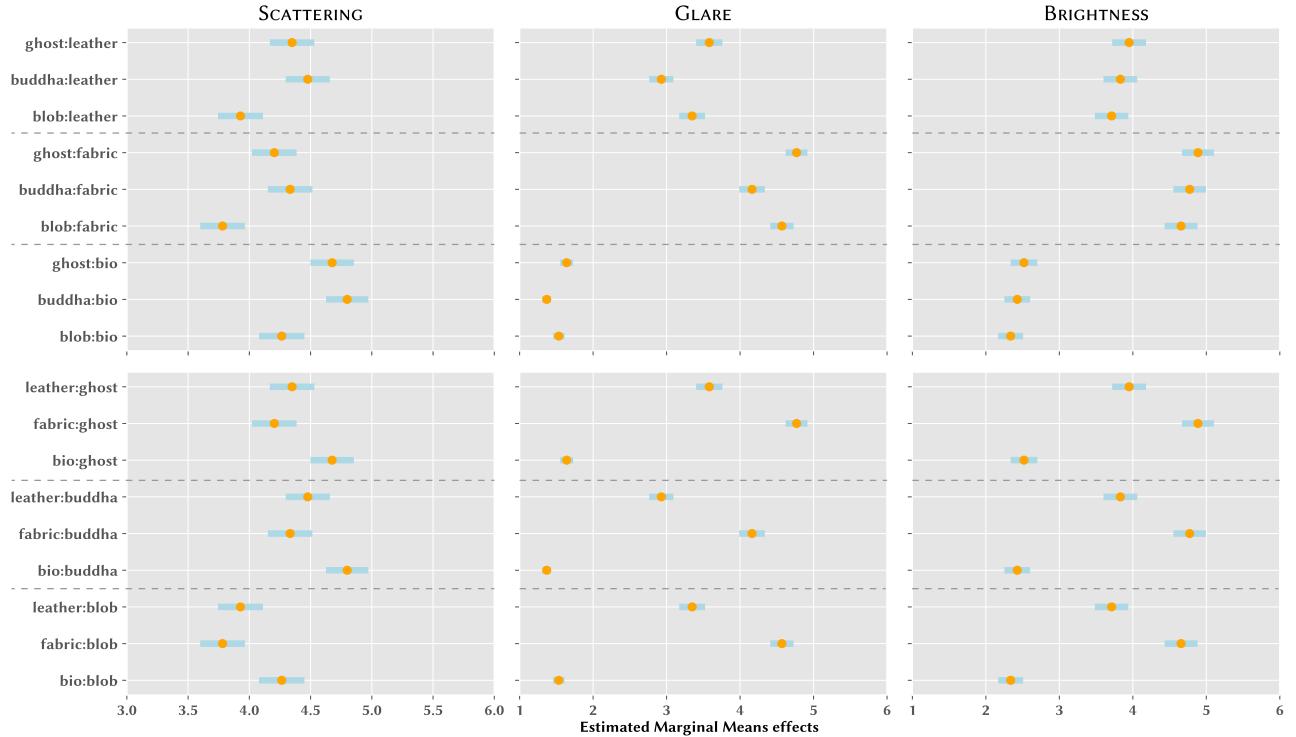


Fig. 5: Estimated marginal means effects of scattering and brightness for the interaction of materials and geometry.

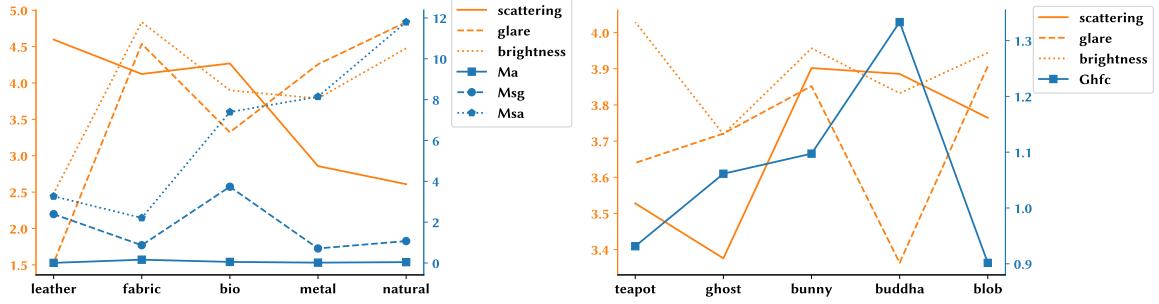


Fig. 6: Estimated marginal effects for scattering, glare and brightness for different materials (left) and geometries (right).

viewpoint, leading to a more regular reflection of illumination and consequently a lower perception of scattering.

4.2.2 Glare

The changes in glare perception resulting from variations in surface material and geometry were investigated in this study (Fig. 5, central column). A similar trend in perceived glare was observed across different geometries when only material variations were applied and vice versa, suggesting the important role of both factors. Specifically, the *natural* material exhibited the highest anisotropy (M_{sa}) and relatively low glossiness (M_{sg}), which resulted in the highest glare perception. In general, the trend of perceived brightness followed the material anisotropy, except for the *fabric* material, where high glare perception corresponded to the highest mean albedo among all materials. Our data also revealed that the perception of glare was strongly impacted by material anisotropy (M_{sa}), and the influence of material reflectance (M_a) should not be ignored. Surface geometry also had a strong impact on perceived glare. The *buddha* had the lowest glare

perception due to the rapidly changing surface curvature, while the *blob* with its most smooth shape resulted in the highest rating. Although the remaining geometries did not yield statistically significant differences, we can view glare perception as tightly related to surface curvature.

4.2.3 Brightness

The effects of material and geometry variations on perceived brightness were investigated, and similar trends were observed for changes in both factors (Fig. 5, right column). Specifically, materials with higher anisotropy (M_{sa}) and lower glossiness (M_{sg}), such as the *natural* material, were associated with higher brightness perception. However, for the *fabric* material, high brightness perception corresponded with low anisotropy. Surface geometry was also found to play a significant role in brightness perception. The *blob* geometry, with slowly changing curvature, and the *ghost* geometry, which always had a perpendicularly reflectable direction with respect to the viewpoint, tended to expose the illumination brightness directly to the viewer.



Fig. 7: The example showcase of our extrapolation dataset: *cylinder* and *elephant* with two new material illuminate by left environment probes.

5 PREDICTING PERCEPTUAL ATTRIBUTES

In this section, we introduce our approach for predicting perceptual illumination attributes from images. To the best of our knowledge, our work is the first that estimates perceived illumination attributes for outdoor scene perception. This section begins with an overview of our neural network and loss functions, followed by validation of its robustness in different scenarios and showcasing potential failure cases. We conclude by evaluating the consistency of the predictor under varying controlled conditions of point of view, geometry, and material.

5.1 The learning-based model

Our approach to estimating perceptual properties leverages our collected data. We train a neural network through supervised learning, using our dataset of images as input and their corresponding crowdsourced subjective attributes for *scattering*, *glare*, and *brightness* as output targets. For each image, we compute the mean rating for each attribute (each image has been rated at least three times). We use the full dataset for training, except images containing the *bunny* geometry and the *metal* material, which we use as a validation set to test the predictor performance. This leads to a training set of 3,200 images ($200 \text{ illuminations} \times 4 \text{ geometries} \times 4 \text{ materials}$), which are further augmented using crops, flips, shifts, rotations, scaling, and Gaussian and Poisson noise. As a result, more than 320,000 augmented examples are generated for training, where 80% of the geometry are guaranteed visible by adjusting the positions. We further normalize the 7-point Likert attribute data to the range of 0 to 1 for each example. We experiment with the VGG16 [50], [51] and ResNet52 [42], [52] architectures which have been widely used for various graphics and vision tasks. We train the networks for 10 epochs with batch size 4 and input resolution 256×256 . The optimization is performed using the Adam optimizer [53] with an initial learning rate set to 10^{-5} . We replace the last layer of the networks with a fully-connected layer that regresses a 3-dimensional attribute vector. The Mean Squared Error (MSE) loss is used for network training:

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{v}_i - v_i)^2, \quad (2)$$

where \hat{v} , v are the 3-dimensional vectors containing the predicted attribute values and ground-truth user ratings, respectively. $N = 3$ is the number of attributes. We additionally experimented with the Mean Squared Logarithmic Error (MSLE) loss for comparisons:

$$\mathcal{L}_{MSLE} = \frac{1}{N} \sum_{i=1}^N (\log(\hat{v}_i + 1) - \log(v_i + 1))^2, \quad (3)$$



Fig. 8: For the case that glare effects are rare to happen on nearly light-absorption material, our predictor is not able to fully learn the image features that characterize the illumination glare: when encountering fully diffuse surfaces, it shows similar perceptual ratings for a non-specular surface.

Note that our statistical analysis provides foundational insights into the influence of material and geometry on perceived sky attributes. This analysis informs the development of the predictor, which uses a neural network to encode the same factors observed in the statistical analysis implicitly. Specifically, the predictor is trained on images and rated attributes collected in the user study, which includes variations in geometry, illumination, and materials. Both are integral components of our work, providing foundational insights and highlighting the usefulness of our collected dataset.

5.2 Validation

As mentioned in Sec. 5.1, we evaluate our model with a subset of our dataset containing the *bunny* geometry and the *metal* material (set A) that have not been used for model training. We also provide an additional evaluation to test the extrapolation abilities of our model, which contains newly collected data (set B) as challenging cases. We include brand-new extrapolation data with 100 additional outdoor illuminations collected from an open-sourced website³, two complex geometries and two edited BRDFs [54] from MERL and UTIA (see Fig. 7). This results in completely new scenes never seen during training. The ground-truth perceptual illumination rating data for this new extrapolation dataset is collected following the same procedure described in Sec. 3.2.3. For the sake of fair comparisons of the loss functions (Eq. 2 and 3) that are already used for training, we employ another Mean Absolute Error (MAE) metric for quantitative evaluations:

$$\mathcal{E}_{MAE} = \frac{1}{N} \sum_{i=1}^N ||\hat{v}_i - v_i||. \quad (4)$$

Tab. 1 shows the results for our two tested architectures and loss functions with the MAE metric. Note that VGG16 architecture with MSLE loss performs slightly better than the other settings, therefore we choose this model as our formal predictor. Tab. 2 illustrates the MAE metrics of the predicted perceptual attributes by our method.

In general, our predictor performs well in most cases. However, we find it may fail in special cases. Glare effects are common with materials that produce reflections, but it is hard to observe glare effects on diffuse surfaces, which often absorb most incoming rays. In these cases, our predictor may yield incorrect predictions (see Fig. 8). Thus we ignore glare perception for the following consistency tests.

3. <https://hdri-skies.com>

5.3 Consistency tests

Our predictor's robustness under varying conditions of object orientation, geometry complexity, and material is evaluated in further detail. As previously mentioned, limited data and image effects can cause perceived glare to be unstable. Therefore, we place particular emphasis on the results of perceptual scattering and brightness. Note that the geometries and materials utilized in this section were not included in the training dataset. Examples of the consistency tests performed can be found in Fig. 9, while complementary examples are available in the supplementary material for reference.

5.3.1 Object orientation

Fig. 9 top row shows the same object with different placing orientations under the same illumination, labeled by our predictions for scattering and brightness. While we rotate the geometry across the Y-axis, the prediction for the above perceptual attributes remains consistent. Different orientations of the geometry can produce reflections at different scales, however, as reflectance can be perceived as the interaction of geometry and materials, variations in object orientation that do not produce significant changes in the observed geometry produce consistent perceptions of scattering and brightness.

5.3.2 Geometry complexity

The middle row of Fig. 9 reveals that when increasing the geometry complexity from a simple sphere to a bumpy sphere, the predicted scattering also increases. The rating is consistent with our previously analyzed perceptual ratings, where the sphere has lower ratings on perceptual scattering with simple geometry. A reasonable explanation is that simple geometry may reflect clearer outdoor illuminations on the same material, while complex shapes will introduce shadows and interact with material reflectance.

5.3.3 Material variations

As illustrated in Fig. 9 bottom row, the predicted scattering increases as the roughness of the material increases from a mirror-like surface to a light-absorbing surface. This is also consistent with our observations: material roughness has a strong impact on perceived illumination.

6 APPLICATIONS

In this section, we demonstrate representative applications that can use our illumination attribute predictor to benefit related fields.

6.1 Luminance editing for outdoor scenes

Since we adopt a neural architecture to predict the perceptual attributes, integrating our predictor into a differentiable renderer is straightforward, which allows us to perform outdoor scene illumination editing in an intuitive manner. For simplicity, we focus on luminance editing, which relates to camera exposure and illumination intensity. We demonstrate this application by adapting our predictor to the differentiable renderer *Mitsuba 3* [55], which is a research-oriented rendering system specialized in differentiable rendering. The input to the system is a scene with a target material and geometry. For the geometry, we choose the *blob*, *bunny* and *teapot* since they are common objects, and for the material, we choose a rough conductor (with surface roughness 0.01) for better visual discrimination. We set the camera



Fig. 9: Changes in perceived scattering and brightness in consistency tests. **Top:** different placing orientations; **Middle:** increasing geometry complexity; **Bottom:** increasing material roughness (bottom). Our prediction is consistent with the visual changes.

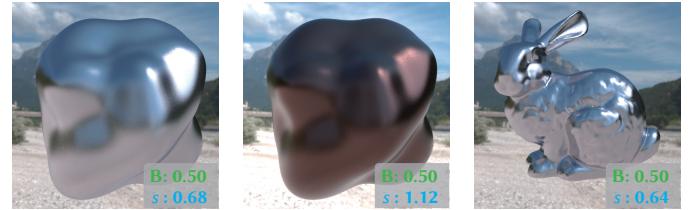


Fig. 10: Example of luminance editing using our predictor under a different material (center) and geometry (right). We set the target perceived brightness of the environment to 0.5. Our predictor takes into account changes in material and geometry and updates the exposure scale s accordingly.

to look at the geometry above the horizon, and render the initial image with sample count 1024 in order to reduce the noise at the resolution of 512×512 . Then we fix all scene parameters except for exposure scale s , which represents the outdoor illumination intensity. After we render the scene, we input the rendered image into our predictor for predicting perceptual brightness. Finally, the scene parameters are optimized by our proposed perceptual loss to match the target perceived brightness:

$$\mathcal{L} = \|A(R(L_x; M_t, G_t)) - T_l\|^2, \quad (5)$$

where R is the differentiable rendering operation under target material M_t and geometry G_t , x are the analytic illumination intensity parameters that we actually want to optimize, and T_l is the target brightness. We simply multiply the scale s to the environment map as input illumination L_x after exposure. Fig. 10 shows the luminance editing results for different materials (left and middle) and geometry (left and right) with the target perceptual brightness set to 0.5. We mark the target brightness as B in our figure. Fig. 11 shows the edited examples of *blob* when we adjust

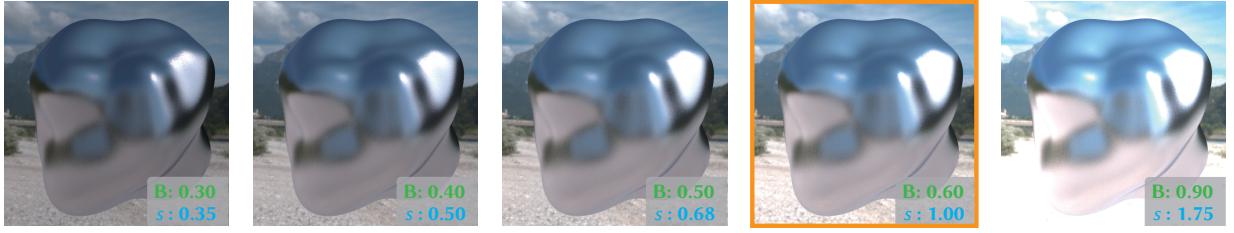


Fig. 11: Luminance editing results. From left to right edited with target brightness: 0.3, 0.4, 0.5, 0.6 (fixed starting parameter), and 0.9.

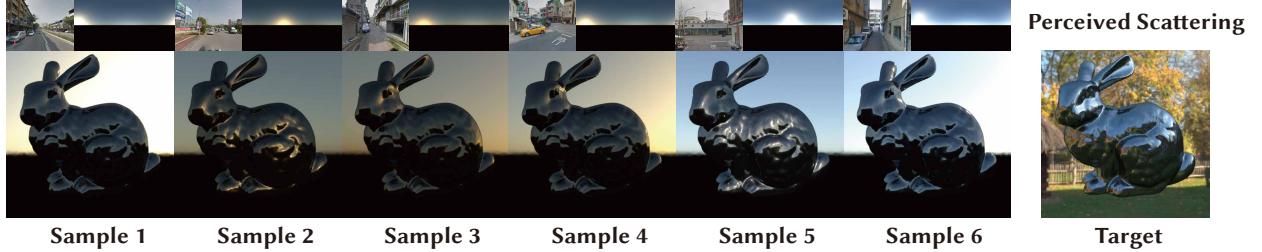


Fig. 12: Top: Collection of captured sample photos with predicted outdoor illumination. Bottom: Rendered sample using the above environment map. Sample 6 is the best approximation of perceived scattering to our target image.

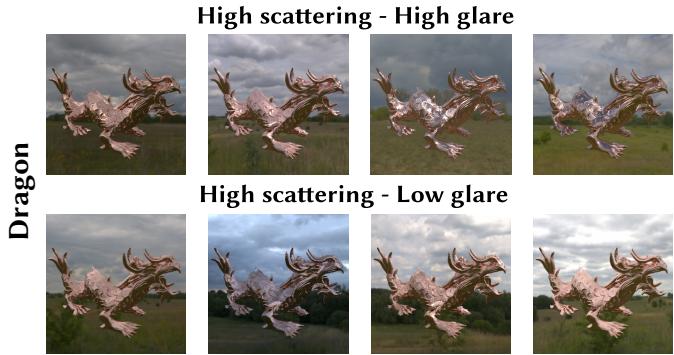


Fig. 13: Top: Recommended scenes with high scattering and high glare. Bottom: Recommended scenes with high scattering and low glare. Although both the top and bottom rows feature hazy skies, the top row shows more glossy and reflective object surfaces, while the bottom row does not reveal intense environmental illuminations.

the target brightness to 0.3, 0.4, 0.5, 0.6, and 0.9 from left to right (the framed image shows the fixed starting parameter at 0.60). For each target brightness, we optimize it in 50 iterations with Adam optimizer with a learning rate of 0.05. Note that, while in this illustrative case the exposure could be easily adjusted manually in an editing software, our setup provides a unique benefit: It aligns the final image with the target perceptual brightness by considering the interplay of illumination, geometry, and material. More results are available in the supplementary material.

6.2 Outdoor illumination reproduction

Reproducing outdoor illuminations from a single regular field-of-view image is a challenging problem due to the missing full information about the original scene. A variety of works [18], [20], [26] have tried to directly estimate the outdoor illumination from captured images by leveraging the power of neural networks.

More sophisticated methods [25], [56] have been proposed for more accurate light estimation results, which have complex neural architectures and even more complex inputs. We start from a perspective of human visual perception, and aim to obtain physically reliable outdoor illumination under a controllable cost, without using complex learning-based methods. For this purpose, first, we captured 1,290 images from the real world, aiming to construct a large-scale reference dataset for reproducing desired outdoor illumination. Then we use a lightweight but robust outdoor sky parameter predictor [18] to estimate a roughly plausible prediction, yielding a collection of available measured outdoor illumination candidates generated using HW model implemented in the Mitsuba renderer [22]. Our goal is to use our predictor to find the desired illumination that best approximates the target perceived scattering of a target image I_g under given material M_t and geometry G_t by finding the closest perceived scattering to a specified target scattering value s_{opt} within the collected measured outdoor illumination candidates \mathbf{s} :

$$s_{opt} = \arg \min_{\mathbf{s}_i \in \mathbf{s}} \|A(R(s_i; M_t, G_t)) - A(I_g)\|_2, \quad (6)$$

where $A(I)$ means applying our predictor to a given image I , R denotes the rendering operation. We can use any given image to specify the target attribute value because the optimization target does not rely on pixel-wise distances. Fig. 12 shows sampled environment maps from the reference dataset and the corresponding rendering results of *bunny*. Our predictor retrieves the rendering result whose perceived scattering most resembles the target image.

6.3 Illumination recommendations for scene design

The selection of suitable environment maps to illuminate a desired scene is often troublesome for artists. Conventional scene illumination design typically requires users to engage in extensive trial-and-error to identify environment illumination that conforms to human aesthetic design. Our predictor can be used to estimate and visualize the perception of a given object with fixed material and geometry under a range of outdoor illuminations. Fig. 13

illustrates a preliminary example of illumination recommendation with high scattering for the *dragon* scene. More examples are provided in the supplementary material.

7 CONCLUSION

In this work, we have studied the effect of material and geometry on perceptual outdoor illumination. We have modeled three perceptual outdoor illumination attributes (scattering, glare, and brightness) and presented a dataset focused on outdoor illumination perception, containing 5,000 images built upon 200 illuminations, 5 geometries and 5 materials. Moreover, we have crowdsourced subjective ratings for the perceptual attributes via a web-based user study. We have further analyzed the effect of material and geometry in outdoor illumination perception and we have discussed their interaction effects with our collected perceptual ratings. Our analysis reveals many interesting trends, and we believe that our data has the potential to support further research. Since fully understanding material reflectance, shape, and illumination directly from an image is a complex task, extracting physically-based parameters from a single image is still an open problem. Various works [16], [31], [32] have tried to use image-level statistics such as luminance to train neural networks to extract better deep features or predict perceptual attributes of material appearance from a given input image. In contrast, we predict perceptual attributes of outdoor illumination by leveraging a neural network that is trained with our collected dataset, and have shown that our trained predictor yields promising performance, and is robust for various scenarios such as shape or material variations. This makes it well-suited for integration into a differentiable rendering pipeline for various applications, allowing users to reproduce real-world outdoor illuminations that match user-specified perception attributes. It has the potential to assist artists in illumination recommendations for scene design tasks.

While our method is restricted to outdoor scenes, an extension of our method to a more comprehensive one using advanced physically-based illumination models is a promising research direction. Moreover, although our model performs well for unseen scenes with new materials, illuminations, and geometries as demonstrated in our validation set B (Sec. 5.2), highly complex scenes with several objects made of different materials would pose a challenge. In addition, improving the current sampling strategy for material and geometry with quantitative description can provide more insights into understanding their interactions.

ACKNOWLEDGMENTS

This work was directly supported by the National Natural Science Foundation of China (Project Number: 62372025 and 61932003) and the Fundamental Research Funds for the Central Universities. Ana Serrano was supported by the Spanish Agencia Estatal de Investigación (project PID2022-141539NB-I00).

REFERENCES

- [1] B. L. Anderson, “Visual perception of materials and surfaces,” *Current biology*, vol. 21, no. 24, pp. R978–R983, 2011.
- [2] R. W. Fleming, S. Nishida, and K. R. Gegenfurtner, “Perception of material properties,” 2015.
- [3] R. W. Fleming, “Material perception,” *Annual review of vision science*, vol. 3, pp. 365–388, 2017.
- [4] P. J. Marlow, J. Kim, and B. L. Anderson, “The perception and misperception of specular surface reflectance,” *Current Biology*, vol. 22, no. 20, pp. 1909–1913, 2012.
- [5] A. C. Chadwick and R. Kentridge, “The perception of gloss: A review,” *Vision research*, vol. 109, pp. 221–235, 2015.
- [6] F. Pellacini, J. A. Ferwerda, and D. P. Greenberg, “Toward a psychophysically-based light reflection model for image synthesis,” in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 55–64.
- [7] B. Chen, M. Piovarči, C. Wang, H.-P. Seidel, P. Didyk, K. Myszkowski, and A. Serrano, “Gloss management for consistent reproduction of real and virtual objects,” in *SIGGRAPH Asia 2022 Conference Papers*, 2022, pp. 1–9.
- [8] R. W. Fleming and H. H. Bülthoff, “Low-level image cues in the perception of translucent materials,” *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 3, pp. 346–382, 2005.
- [9] I. Gkioulekas, B. Walter, E. H. Adelson, K. Bala, and T. Zickler, “On the appearance of translucent edges,” in *Proceedings of the ieee conference on computer vision and pattern recognition*, 2015, pp. 5528–5536.
- [10] W. J. Adams, G. Kucukoglu, M. S. Landy, and R. K. Mantiuk, “Naturally glossy: Gloss perception, illumination statistics, and tone mapping,” *Journal of Vision*, vol. 18, no. 13, pp. 4–4, 2018.
- [11] P. J. Marlow and B. L. Anderson, “Generative constraints on image cues for perceived gloss,” *Journal of vision*, vol. 13, no. 14, pp. 2–2, 2013.
- [12] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, “Image statistics and the perception of surface qualities,” *Nature*, vol. 447, no. 7141, pp. 206–209, 2007.
- [13] M. Olkkonen and D. H. Brainard, “Perceived glossiness and lightness under real-world illumination,” *Journal of vision*, vol. 10, no. 9, pp. 5–5, 2010.
- [14] M. Lagunas, A. Serrano, D. Gutierrez, and B. Masia, “The joint role of geometry and illumination on material recognition,” *Journal of Vision*, vol. 21, no. 2, pp. 2–2, 2021.
- [15] B. Chen, C. Wang, M. Piovarči, H.-P. Seidel, P. Didyk, K. Myszkowski, and A. Serrano, “The effect of geometry and illumination on appearance perception of different material categories,” *The Visual Computer*, vol. 37, no. 12, pp. 2975–2987, 2021.
- [16] A. Serrano, B. Chen, C. Wang, M. Piovarci, H.-P. Seidel, P. Didyk, and K. Myszkowski, “The effect of shape and illumination on material perception: model and applications,” *ACM Trans. on Graph.*, vol. 40, no. 4, 2021.
- [17] L. Hosek and A. Wilkie, “An analytic model for full spectral sky-dome radiance,” *ACM Transactions on Graphics*, vol. 31, no. 4, July 2012.
- [18] Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, and J.-F. Lalonde, “Deep outdoor illumination estimation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7312–7321.
- [19] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan, “Estimating the natural illumination conditions from a single outdoor image,” *International Journal of Computer Vision*, vol. 98, pp. 123–145, 2011.
- [20] Y. Hold-Geoffroy, A. Athawale, and J.-F. Lalonde, “Deep sky modeling for single image outdoor lighting estimation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 6927–6935.
- [21] M. Toscani, M. Valsecchi, and K. R. Gegenfurtner, “Lightness perception for matte and glossy complex shapes,” *Vision research*, vol. 131, pp. 82–95, 2017.
- [22] W. Jakob, “Mitsuba renderer,” 2010, <http://www.mitsuba-renderer.org>.
- [23] M.-A. Gardner, K. Sunkavalli, E. Yumer, X. Shen, E. Gambaretto, C. Gagné, and J.-F. Lalonde, “Learning to predict indoor illumination from a single image,” *arXiv preprint arXiv:1704.00090*, 2017.
- [24] P. P. Srinivasan, B. Mildenhall, M. Tancik, J. T. Barron, R. Tucker, and N. Snavely, “Lighthouse: Predicting lighting volumes for spatially-coherent illumination,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 8080–8089.
- [25] F. Zhan, Y. Yu, C. Zhang, R. Wu, W. Hu, S. Lu, F. Ma, X. Xie, and L. Shao, “Gmlight: Lighting estimation via geometric distribution approximation,” *IEEE Transactions on Image Processing*, vol. 31, pp. 2268–2278, 2022.
- [26] P. Yu, J. Guo, L. Wu, C. Zhou, M. Li, C. Wang, and Y. Guo, “Dual attention autoencoder for all-weather outdoor lighting estimation,” *Sci. China Inf. Sci.*, vol. 64, 2021.
- [27] J. Zhang, F. He, Y. Duan, and S. Yang, “Aidednet: anti-interference and detail enhancement dehazing network for real-world scenes,” *Frontiers of Computer Science*, vol. 17, no. 2, p. 172703, 2022.
- [28] E. Bruneton and F. Neyret, “Precomputed atmospheric scattering,” in *Computer graphics forum*, vol. 27, no. 4. Wiley Online Library, 2008, pp. 1079–1086.

- [29] I. Motoyoshi and H. Matoba, "Variability in constancy of the perceived surface reflectance across different illumination statistics," *Vision Research*, vol. 53, no. 1, pp. 30–39, 2012.
- [30] R. W. Fleming, "Visual perception of materials and their properties," *Vision research*, vol. 94, pp. 62–75, 2014.
- [31] G. Schwartz and K. Nishino, "Recognizing material properties from images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 8, pp. 1981–1995, 2019.
- [32] K. R. Storrs, B. L. Anderson, and R. W. Fleming, "Unsupervised learning predicts human perception and misperception of gloss," *Nature human behaviour*, vol. 5, no. 10, pp. 1402–1417, 2021.
- [33] F. Zhang, H. de Ridder, P. Barla, and S. Pont, "Effects of light map orientation and shape on the visual perception of canonical materials," *Journal of vision*, vol. 20, no. 4, pp. 13–13, 2020.
- [34] M. Toscani and M. Valsecchi, "Lightness discrimination depends more on bright rather than shaded regions of three-dimensional objects," *i-Perception*, vol. 10, no. 6, p. 2041669519884335, 2019.
- [35] A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, and E. Economou, "An anchoring theory of lightness perception," *Psychological review*, vol. 106, no. 4, p. 795, 1999.
- [36] J. T. Todd, J. F. Norman, and E. Mingolla, "Lightness constancy in the presence of specular highlights," *Psychological Science*, vol. 15, no. 1, pp. 33–39, 2004.
- [37] P. Brossier, J. P. Bello, and M. D. Plumbley, "Real-time temporal segmentation of note objects in music signals," in *Proceedings of ICMC 2004, the 30th Annual International Computer Music Conference*. University of Surrey, 2004.
- [38] R. W. Fleming, R. O. Dror, and E. H. Adelson, "Real-world illumination and the perception of surface reflectance properties," *Journal of Vision*, vol. 3, no. 5, pp. 3–3, 07 2003.
- [39] J. Guo, Y. Guo, J. Pan, and W. Lu, "Brdf analysis with directional statistics and its applications," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 3, pp. 1476–1489, 2018.
- [40] P. Vangorp, J. Laurijssen, and Ph. Dutré, "The influence of shape on the perception of material reflectance," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 77:1–77:9, July 2007.
- [41] V. Havran, J. Filip, and K. Myszkowski, "Perceptually motivated brdf comparison using single image," *Comput. Graph. Forum*, vol. 35, no. 4, p. 1–12, jul 2016.
- [42] M. Lagunas, S. Malpica, A. Serrano, E. Garces, D. Gutierrez, and B. Masia, "A similarity measure for material appearance," *ACM Trans. Graph.*, vol. 38, no. 4, jul 2019.
- [43] S. Doroudi, E. Kamar, E. Brunskill, and E. Horvitz, "Toward a learning science for complex crowdsourcing tasks," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 2623–2634.
- [44] P. Welinder, S. Branson, P. Perona, and S. Belongie, "The multidimensional wisdom of crowds," *Advances in neural information processing systems*, vol. 23, 2010.
- [45] D. W. Cunningham and C. Wallraven, *Experimental design: From user studies to psychophysics*. CRC Press, 2011.
- [46] J. Nunnally and I. Bernstein, "Psychometric theory 3rd edition (macgraw-hill, new york)," 1994.
- [47] A. Agresti, *Categorical data analysis*. John Wiley & Sons, 2003.
- [48] R. H. B. Christensen, "Cumulative link models for ordinal regression with the r package ordinal," *Submitted in J. Stat. Software*, vol. 35, 2018.
- [49] P. McCullagh, "Regression models for ordinal data," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 42, no. 2, pp. 109–127, 1980.
- [50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [51] H. Xu, Z. Han, S. Feng, H. Zhou, and Y. Fang, "Foreign object debris material recognition based on convolutional neural networks," *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 1, pp. 1–10, 2018.
- [52] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [54] T. Sun, H. W. Jensen, and R. Ramamoorthi, "Connecting measured brdfs to analytic brdfs by data-driven diffuse-specular separation," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, p. 273, 2018.
- [55] W. Jakob, S. Speierer, N. Roussel, M. Nimier-David, D. Vicini, T. Zeltner, B. Nicolet, M. Crespo, V. Leroy, and Z. Zhang, "Mitsuba 3 renderer," 2022, <https://mitsuba-renderer.org>.
- [56] J.-P. Xu, C. Zuo, F.-L. Zhang, and M. Wang, "Rendering-aware hdr environment map prediction from a single image," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 3, pp. 2857–2865, Jun. 2022.



Miao Wang is an Associate Professor with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, China. He is also an adjunct researcher with Zhongguancun Laboratory, Beijing, China. His research interests include Virtual Reality, Computer Graphics and Visual Computing. He received his Ph.D. degree from Tsinghua University in 2016. He serves as a program committee member of IEEE VR and ISMAR conferences.



Jin-Chao Zhou is a master graduate from Beihang University, Beijing, China. He obtained his bachelor degree from University of Electronic Science and Technology of China in 2020. His major research interest is in virtual reality and computer graphics and now he is working on research topics about perceptual material appearance and foveated rendering.



Wei-Qi Feng is a graduate student with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, China. He received the bachelor degree in Beihang University in 2022. His research interests include Computer Graphics and Visual Computing.



Yu-Zhu Jiang is an undergraduate student with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, China. Her research interests include Computer Graphics and Visual Computing.



Ana Serrano is an Assistant Professor at Universidad de Zaragoza (Spain). Previously she was a Postdoctoral Research Fellow at the Max-Planck-Institute for Informatics. She received her PhD in Computer Science in 2019. She was the recipient of an Adobe Research Fellowship honorable mention in 2017, and a NVIDIA Graduate Fellowship in 2018. Her doctoral thesis was recognized with the Eurographics 2020 PhD award and she was recognized with the Eurographics Young Researcher Award in 2023. Her research focuses on various areas of visual computing, including computational imaging, material appearance perception and editing, and virtual reality. She is particularly interested in using perceptually-driven approaches to improve user experiences and develop tools to assist content creation.