Visual Informatics 3 (2019) 59–68

Contents lists available at ScienceDirect

Visual Informatics

journal homepage: www.elsevier.com/locate/visinf

Deep appearance modeling: A survey

Yue Dong

Microsoft Research Asia, T2 Microsoft Building, 5th Danling Street, Beijing, China

ARTICLE INFO

Article history:
Received 30 June 2019
Accepted 1 July 2019
Available online 22 July 2019

Keywords:
Appearance modeling
Surface reflectance
Reflectance field
Light field

ABSTRACT

Appearance modeling is an essential task in computer graphics for capturing and reproducing rich appearance of real world materials under different lighting and viewing conditions. With recent advances of deep learning techniques, a set of deep learning based approaches have been proposed for improving the efficiency and result quality of appearance modeling. In this paper, we provide a survey of these deep appearance modeling techniques from both graphics and machine learning perspectives, and discuss the challenges and opportunities along this direction.

Contents

1. Introduction

Modeling the appearance of an object is a fundamental problem in computer graphics (Dorsey et al., 2008; Weinmann and Klein, 2015). In computer graphics, the appearance represents how the light interacts with the surface, works together with geometry shape and lighting condition to determine the final rendering results. One major challenge of appearance modeling is the high dimensionality of the appearance data. In its raw form, the appearance usually represented as high dimensional functions, such as 8D light field (Levoy and Hanrahan, 1996; Gortler et al., 1996), 6D spatially varying surface reflectance function (Nicodemus et al., 1992), 4D light transport (Ng et al., 2003;

E-mail address: yuedong@microsoft.com.

Peer review under responsibility of Zhejiang University and Zhejiang University Press.

https://doi.org/10.1016/j.visinf.2019.07.003

Peers et al., 2009), etc. Fortunately, the high dimensional appearance data usually exhibits strong coherency, which can be utilized for efficient appearance modeling.

Brute-force capture the high dimensional function requires long measurement and processing time with dedicated capture devices. Data-driven approaches either design a compact model following heuristics and fit the measurements into the model (Matusik et al., 2003; Lawrence et al., 2006; Dong et al., 2010), or take data coherency as a constraint (Peers et al., 2006; Chen et al., 2014; Zhou et al., 2016). However, due to the limited number of real measured appearance data and the limited tools that support learning from large scale datasets, classical data-driven approaches still follow the ad-hoc design of the models and data-fitting algorithms. Many classical data-driven methods only exploit one specific type of data coherency or data coherency within the specific input material sample, limiting their efficiency.

With the recent advance of deep learning (Goodfellow et al., 2016), efficient feature representations can be learned directly from a large collection of data; as a result, deep learning became an efficient tool for appearance modeling. Both the representation model and the data-fitting algorithm can be learned from data; the representation model is generic, sample independent, utilizing the data coherency across the whole training dataset.

Recently, there are several works that successfully apply deep learning to surface reflection estimation (Li et al., 2017; Deschaintre et al., 2018; Li et al., 2018a,b; Kang et al., 2018; Gao et al., 2019; Deschaintre et al., 2019), light transport modeling (Ren et al., 2013, 2015; Xu et al., 2018), face modeling (Lombardi et al., 2018) material synthesis (Zsolnai-Fehér et al., 2018) and BTF compression (Rainer et al., 2019). Among those topics, surface reflectance modeling has most works embracing the trend of deep learning. In recent years, researchers developed a series of works using deep learning techniques in different perspectives for surface appearance modeling (some experimental results can be found in Fig. 1). To better discuss and compare different design choices of applying deep learning techniques, this survey will discuss those deep surface reflectance modeling works from both the computer graphics and machine learning perspectives.

In this paper, we will first provide preliminary background knowledge and briefly revisit data-driven appearance modeling works, and discuss the generic scheme for deep learning based appearance modeling (Section 2). Then we summarize recent deep appearance modeling works based on their tasks and methodology, from both the computer graphics (Section 3) and machine learning (Section 4) perspective. Finally, we will discuss research opportunities and future directions for developing new deep appearance modeling solutions (Section 5).

2. Scope and overview

Appearance modeling research focuses on a wide range of optical phenomena, such as surface reflection, subsurface scattering, and diffraction materials. Among those appearance modeling areas, surface reflectance modeling has most works embracing the trend of deep learning. In recent years, researchers developed a series of works using deep learning techniques in different perspectives for surface appearance modeling. As a result, this survey will mainly focus on works that leveraging deep learning for surface appearance modeling. Besides having the deep surface reflectance as the main topic, we will also cover closely related works such as deep learning based light transport modeling.

There are other topics in appearance modeling such as translucent materials (Nicodemus et al., 1992; Jensen and Buhler, 2002; Wang et al., 2007), multi-layered material (Jakob et al., 2014a; Yan et al., 2016; Zeltner and Jakob, 2018; Belcour, 2018), bi-scale materials (Wu et al., 2009, 2011, 2013; Lan et al., 2013), diffraction (Toisoul and Ghosh, 2017; Holzschuch and Pacanowski, 2017; Werner et al., 2017) and polarization effects (Riviere et al., 2017), glint reflections (Yan et al., 2014; Jakob et al., 2014b; Werner et al., 2017), light-transport decomposition (Dong et al., 2015), etc. As far as our knowledge, deep learning has not been widely applied in those directions, however, modeling those effects also shares similarities to surface reflectance modeling, thus adapting deep learning on those areas would lead to potential future works.

2.1. Preliminary

The surface reflectance models how the light reflects on an opaque surface point. Specifically, surface reflectance at a point x can be described by the Bidirectional Reflectance Distribution Function (BRDF) (Nicodemus et al., 1977), a 4D function that relates incident irradiance to outgoing radiance: A common approximation is the dichromatic BRDF which decompose the reflection into a diffuse component and a specular component:

$$f_r(\omega_i, \omega_o; x) = \frac{\rho_d(x)}{\pi} + \rho_s(x) f_s(\omega_i, \omega_o; x), \tag{1}$$

where ω_i and ω_o are the incident and outgoing directions respectively, ρ_d and ρ_s are the diffuse and specular coefficient or albedo. The diffuse component is constant over different incident and outgoing direction, and f_s is an angular function that models the specular surface reflectance.

Many physical based models have been proposed to further approximate and simplify the BRDF, including Ward model (Ward, 1992), Cook–Torrance model (Cook and Torrance, 1982), generalized microfacet model (Ashikmin et al., 2000), etc. Here we introduce some basic BRDF terminology using the Ward BRDF model as an example. At each surface point x the Ward model (Ward, 1992) can be determined by its diffuse albedo ρ_d, specular albedo ρ_s and specular roughness parameter α follows:

$$f_r(\omega_i, \omega_o, x) = \frac{\rho_d(x)}{\pi} + \rho_s \frac{e^{-\tan^2 \delta/\alpha^2}}{4\pi\alpha^2 \sqrt{(\omega_i \cdot n(x))(\omega_o \cdot n(x))}}, \tag{2}$$

where δ is the angle between the halfway vector $\omega_h = (\omega_i + \omega_o)/\|\omega_i + \omega_o\|$ and the normal $n(x)$. The diffuse ρ_d and specular albedo ρ_s determine the overall energy scale of the diffuse and specular reflection, while the roughness α affects the shape of the specular highlights. Smaller roughness represents shiny surfaces with mirror liked specular highlight, larger roughness means glossy or matte surfaces with blurred highlight. Aside from those parameters, the surface normal n affects the local coordinate that defines the incident and outgoing direction, as a result, surface normal is also regarded as a part of surface appearance in many surface appearance modeling works.

2.2. Data-driven surface appearance modeling

To better understand the recent works on deep surface reflectance modeling, here we first give a brief review of classical data-driven surface appearance modeling works.

Surface reflectance can be reconstructed independently at each point, for example using a gonioreflectometer (Mcallister, 2002; Lawrence et al., 2006) for brute-force sampling the complete 4D function. With a sufficient number of measurements, one can fit a low parametric model at each surface point individually (Gardner et al., 2003; Aittala et al., 2013)

Realworld materials often exhibit spatial coherency, different spatial points may have similar BRDFs. Following this assumption, for spatially smooth varying materials, reflectance sharing (Zickler et al., 2005) formulates SVBRDF reconstruction as

Fig. 1. Examples rendering results of deep appearance modeling. Rendering results of SVBRDF of planar surface captured by Li et al. (2017) (a), Gao et al. (2019) (b) and Kang et al. (2018) (c). (d) Rendering results of geometry shape and SVBRDF estimated by Li et al. (2018b).

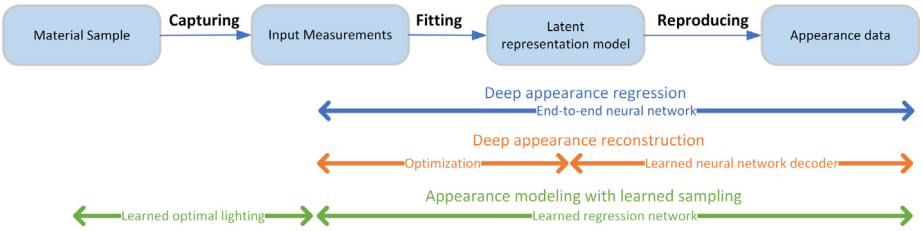

Fig. 2. A general data-driven appearance modeling pipeline. Based on this pipeline, we classify recent deep appearance modeling works into three categories: the deep appearance regression that models the fitting and reproducing process as an end-to-end network; the deep appearance reconstruction that explicitly learns a latent space with learned reproducing process and optimization as the fitting process; and appearance modeling works with learned lighting for the capturing process.

a scattered interpolation problem, and fits both its angular and spatial distribution with basis functions. For piecewise constant materials, Lombardi and Nishino (2012) assumes pixels with similar chromaticity represent the same material and estimate the SVBRDF of a curved object from a single image under unknown point illumination. (Wang et al., 2008) proposed an anisotropic SVBRDF acquisition method that finds surface points with similar BRDF by comparing partial NDF computed from densely sample BRDF slice.

With separately captured basis BRDF (Dong et al., 2010) or basis BRDF reference (Ren et al., 2011), the local linear blending weights of those basis can be estimated from sparsely captured images. (Lensch et al., 2003) proposed method estimating a Lafortune model basis by progressive clustering over surface points. Low-rank optimization can be used to jointly reconstruct the basis BRDF and spatially varying weights (Chen et al., 2014; Zhou et al., 2016).

Traditional data-driven surface appearance modeling relies on handcrafted heuristics, such as spatially smooth distribution, piece-wise constant, linear basis blending, etc. Each heuristic only works efficiently for one specific type of materials and tied to one specific kind of reconstruction algorithms. With the recent advances of deep learning, it is possible to learn heuristic from large scale datasets, enable more efficient and general surface appearance modeling.

2.3. The deep appearance modeling pipeline

As illustrated in Fig. 2, a general scheme for data-driven appearance modeling can be divided into three major steps: first the *capturing* process gets the input measurements from the exemplar object; then the input measurement is *fitted* into latent representation model; which can be used to *reproduce* the compete high dimensional appearance data. In the following discussion, We summarize the three main steps as the *capturing process*, *fitting process* and *reproducing process*. In the following discussion, we will discuss recent deep appearance modeling works in this framework, and how to use deep learning techniques to model those processes.

3. Discuss of deep appearance modeling from a computer graphics perspective

In this section, we will discuss deep surface appearance modeling works from the appearance modeling pipeline's perspective. We classify recent works into three categories: The deep appearance regression, which represents the appearance modeling problem as an image to image regression problem, representing the *fitting* and *reproducing* process as an end-to-end learned system; The deep appearance reconstruction, which has an explicitly designed latent space, the *fitting* and *reproducing* process are represented by separated learned part; The appearance modeling with learned lighting conditions methods that represent the *capturing* process with a trainable part and learn the optimal lighting conditions together with the appearance modeling problem. A brief summary of recent works from the computer graphics perspective is presented in Table 1.

3.1. Deep appearance regression

For the most appearance modeling scenarios, the user wants a simplified input, either a single image that collected from the Internet or one flash lit photograph of the material sample, which can be easily captured via a mobile phone. For those scenarios, the *capturing process* is pre-determined by the scenario, the *fitting* and *reproducing* process can be represented by a regression model which learned the mapping from the input photograph to the output appearance data.

Li et al. (2017) first proposed a learning based solution that estimates surface reflectance from a single image lit by natural environment map. By assuming the input image is a planar material sample lit by environment map, which holds true for many online images, such method allows the user to select input images directly from the Internet. This also allows collecting large amount of valid input images from online image collections, which can be utilized by their specialized training scheme. (discussed in Section 4). On the other hand, rely on environment map lit input means less control about the illumination and there is no guarantee that every surface pixel exhibits specular reflection

effects, leading unstable specular estimation. As a result, they limited their output surface reflectance model with a homogeneous specular component and spatially varying albedo and normal variations.

Using active lighting, such as a flashlight on the mobile phone can reveal more appearance information especially for specularity. By assuming the input image is a planar material sample lit by a unknown environment map and a dominating flashlight collocated with the camera, Li et al. (2018a) learned a regression network that estimates surface appearance with spatially varying albedo, roughness, normal variations, together with a homogeneous specular coefficient determined by the automatically classified material type. Deschaintre et al. (2018) shared a similar assumption of the input, while outputting spatially varying albedo, specular, roughness and normal maps. One benefit of the known dominating flash light is the known dominating specular highlight, specifically, the local lit flash light produces one strong highlight lobe due to the spatially varying incident light direction for different spatial pixels. In order to better analyze those spatial varying information, different mechanisms are designed to aid the network. Li et al. (2018a) included a radial coordinate image that correlated to the spatially varying incident light direction as an auxiliary input to the neural network. Deschaintre et al. (2018) designed a global features network that gathers the spatially varying information, processed into global features, then feedback to the regular convolutional network on every spatial pixel. The knowledge of the collocated lighting condition also enables using the rendering process as a supervision. Both methods leverage the rendering in their training. We will discuss the rendering loss details in Section 4.

Li et al. (2018b) further extended such regression approach for estimating both the spatially varying surface reflectance and the geometry shape from a single input image lit by environment map and dominating collocated flashlight. To better handle spatially varying global features, Li et al. (2018b) designed their neural network and ensured the final layer of their encoder has a receptive field covering the whole input image. To further enhance the capability of global reasoning and also increase the capacity of the neural networks, Li et al. (2018b) also introduced a cascaded estimation scheme. We will have an in-depth discussion on the network structure design in Section 4.

The global illumination also has a stronger influence on objects with complex geometry shapes rather than planar surfaces. In order to account for global illumination effects during the training and evaluation while maintaining a reasonable computation cost, Li et al. (2018b) trained a neural network to predict the global illumination effects from direct illumination renderings.

All those deep appearance regression methods take input under predefined lighting condition, thus the *capturing process* is fixed and determined by the assumption of the input. The learned neural network represents both the *fitting process* and the *reproducing process* and trained in an end-to-end manner, without explicitly modeling of the latent space.

Kim et al. (2017) proposed a solution to estimate homogeneous BRDF from multiple view observations. Although their method supports arbitrary measurements, all the measurements are projected onto a unit sphere, and the neural network takes this projected measurement as input. Thus in the neural network's perspective, the *capturing process* is still fixed and the fully connected network is performing the fitting process to the parametric BRDF.

3.2. Deep appearance reconstruction

Previously discussed deep appearance regression methods can produce plausible results even with a single image input, however, in many cases, multiple measurements are needed to get accurate reconstruction or distinguish between ambiguity cases. The end-to-end trained encoder–decoder networks from those regression based solutions are trained for one specific task with predetermined input condition. Different numbers of input images essentially correspond to different regression tasks, it is not practical to train a separate regression network for every possible number of inputs. The key to resolve the problem is separating the *fitting* and *reproducing process* and explicitly modeling the latent space.

Following this idea, Gao et al. (2019) proposed a deep inverse rendering framework, which supports an arbitrary number of inputs. The precision of the estimated results varies from plausible when the input images fail to capture all the reflectance information, to accurate for large input sets. They design the *fitting process* as an optimization process in the explicitly modeled latent space. Instead of casting manually designed constraints, the latent space, which is trained by an adapted autoencoder, works as a learned constraint that exploits the prior distribution of large SVBRDF dataset. In Fig. 3 we briefly compare the results of deep appearance regression (Li et al., 2018a; Deschaintre et al., 2018) and deep appearance reconstruction (Gao et al., 2019). Showing the benefit of latent space optimization and additional measurements.

Deschaintre et al. (2019) also proposed a learning based SVBRDF estimation system that supports arbitrary number of input. Instead of representing the *fitting* process with optimization, they followed the regression approach that trains a neural network to extract latent features from each individual input image, and assemble them via an max pooling. The advantage of this design is the neural network does not need light and view point information, however, at the same time limited its capability to leverage the light and view information to form a more constrictive optimization system.

For surfaces with stationary textures, Aittala et al. (2016) proposed a method that models the SVBRDF and normal variations from a single flash lit image. Instead of explicitly formulating a latent space for optimization, they directly optimize the material maps with a CNN based constraint. They avoid explicit point-to-point correspondence but trained a CNN-based texture descriptor, and use the descriptor as a reference to guide the material map optimization.

The Deep Lambertian Network (Tang et al., 2012) combines deep belief network and Lambertian surface reflection yielding a multilayer generative model with the albedo, surface normal and light source direction as latent variables. By explicitly including the Lambertian shading model, Deep Lambertian Network enables the generation of images that follow the Lambertian reflection from a trainable network.

3.3. Appearance modeling with learned lighting conditions

All the works discussed above have a predefined *capturing process* in their deep appearance modeling pipeline, since those methods are designed for casual user capture, with minimal requirements on the acquisition setup. For professional acquisition with dedicated devices, it is possible to find the optimal capture setup for one specific task.

Xu et al. (2018) proposed an image-based relighting method that can synthesize scene appearance under novel illuminations. Similar to deep surface appearance regression, their deep relighting network represents the *fitting* and *reproducing process*, which takes five images captured under predefined directional lights, together with the output light direction as inputs, then output the final relit image.

The predefined light directions of inputs are optimized for their relighting network. To find the optimal lighting setup, Xu

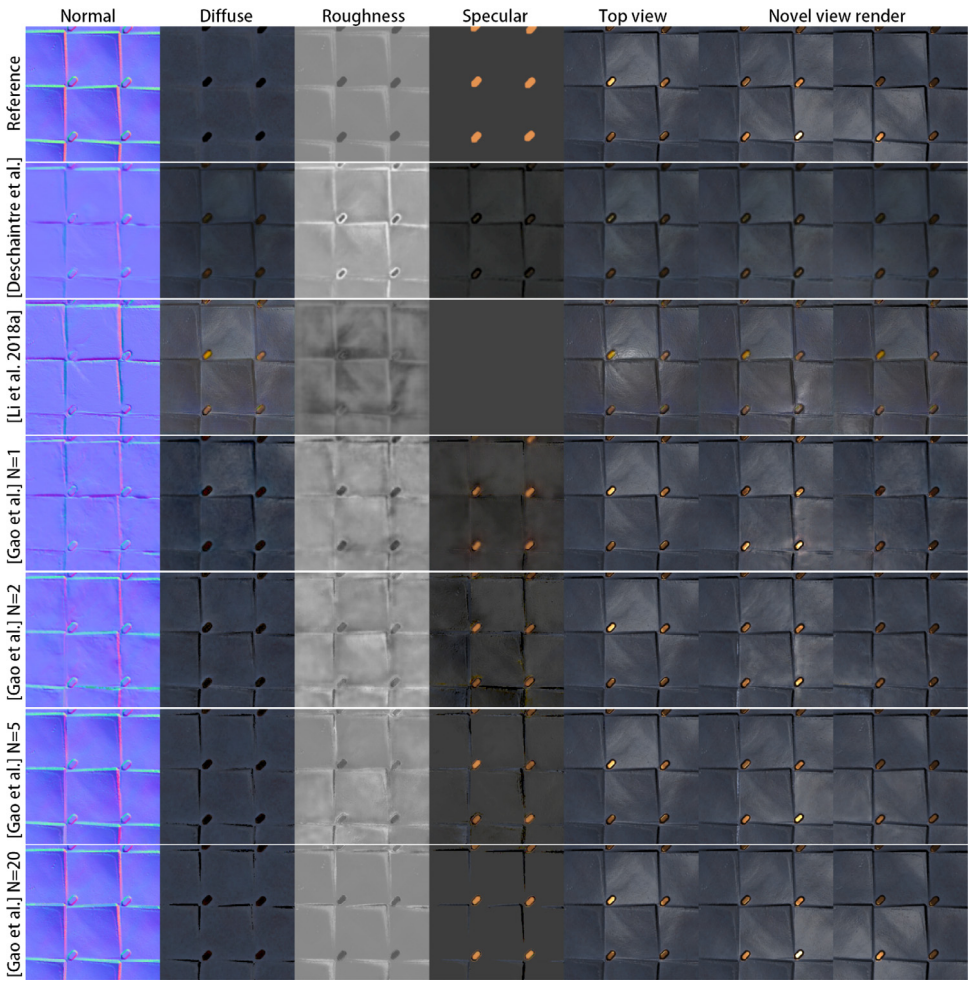

Fig. 3. Comparing deep surface reflectance modeling with similar setup. Deep appearance regression methods Deschaintre et al. (2018) and Li et al. (2018a) produces plausible results with only a single input image. Optimization based method Gao et al. (2019) can further improve the quality of the results by optimizing in the explicitly modeled latent space, given exactly the same single input. With an increasing number of input photographs Gao et al. (2019) show an improvement in accuracy, yielding a more accurate reconstruction given sufficient number of inputs. Gao et al. (2019).

et al. (2018) designed a trainable component called "sample-net" which represents the *capturing process*. The essence of the sample-net is a projection matrix that projects the light transport of a scene into relit results under certain lighting patterns. If one can find the optimal sample-net, the corresponding light pattern of the optimal sample-net will be the optimal lighting condition.

The "sample-net" is trained together with the relighting network, where the goal for the relighting net is to accurately reproduce the relit results based on the input provided by the sample-net; at the same time, the sample-net is also trained to provide the optimal input for the relighting net. They also cast additional constraints to the sample-net, thus after the training, the sample-net will represent a projection matrix that corresponding to directional light sources, thus the resulting optimal lighting pattern is easy to be realized by their capture setup.

For surface reflectance acquisition, Kang et al. (2018) proposed a learning based framework that automatically learns the optimal lighting patterns for efficient anisotropic SVBRDF acquisition. They model the whole deep appearance modeling pipeline with a trainable asymmetric deep autoencoder. The encoder part is a non-negative, linear encoder corresponding to the *capturing process*, their optimized lighting patterns used in physical acquisition directly project the BRDF at each surface point into a latent code. The latent code can then be decoded via a non-linear neural network with complete BRDF information, and fitted into BRDFs.

Kang et al. (2018) designed the encoder as a trainable projection matrix. The acquisition device supports arbitrary complex lighting patterns that can be displayed by its LED panel, as a result, unlike (Xu et al., 2018) that require the matrix corresponding to a set of directional light sources, in Kang et al. (2018) the only constraint for the projection matrix is the non-negative weight should be within the range that their LEDs can display. The decoder is designed as a series of fully connected layers with leakyReLU activations, the input of the decoder is the measured 'projected' latent code from one surface point under predefined LED patterns, and the decoded output is a complete set of measurement for every LED of their capture device. The decoded compete measurement can be fitted to a physical based BRDF model

4. Discuss of deep appearance modeling from a machine learning perspective

In the previous section, we discussed deep surface appearance modeling works from the appearance modeling pipeline's perspective. In this section, we will discuss several important machine learning topics that essential for deep appearance. A taxonomy of those works following the machine learning perspective can be found in Table 2.

Table 1
Taxonomy of deep appearance modeling works from a computer graphics perspective, details for each work are explained in Section 3.

Method	Modeling target	Num. input	Lighting setup
Li et al. (2017)	SVBRDF (planar)	1	Env.
Deschaintre et al. (2018)	SVBRDF (planar)	1	Env.+Flash
Li et al. (2018a)	SVBRDF (planar)	1	Env.+Flash
Kang et al. (2018)	Anisotropic SVBRDF (planar)	Fixed	LED panel
Ye et al. (2018)	SVBRDF (planar)	1	Env.
Li et al. (2018b)	SVBRDF + Shape	1	Env.+Flash
Gao et al. (2019)	SVBRDF (planar)	Arbitrary	Env.+Flash
Deschaintre et al. (2019)	SVBRDF (planar)	Arbitrary	Env.+Flash

Table 2
Taxonomy of deep appearance modeling works from a machine learning perspective, details for each work are explained in Section 4.

Method	Network design	Loss function	Training scheme
Li et al. (2017)	U-Net	L_2 map	Self-augmented
Deschaintre et al. (2018)	U-Net + global feature net	Render	Supervised
Li et al. (2018a)	U-Net + axillary input + CRF	Render + L_2 map	Supervised
Kang et al. (2018)	Linear encoder + FC decoder	L_2 output value	Supervised
Xu et al. (2018)	U-Net	L_2 output value	Supervised
Ye et al. (2018)	U-Net	L_2 map	Self-augmented
Li et al. (2018b)	Cascaded CNN	Render + L_2 map	Per stage, Supervised
Gao et al. (2019)	Encoder–decoder	Render + L_1 map	Supervised
Deschaintre et al. (2019)	U-Net+CNN decoder	Render + L_1 map	Supervised

4.1. Network design

For most appearance regression tasks, the input and output are in the same spatial domain with pixel-to-pixel correspondence, as a result, encoder–decoder CNN with skip connections is a common choice (Li et al., 2017; Ye et al., 2018; Li et al., 2018a; Deschaintre et al., 2018; Li et al., 2018b; Xu et al., 2018). The encoding and decoding process of the encoder–decoder structure provides a large reception field for the latent feature maps, while the skip link connections preserve rich spatial details.

Deschaintre et al. (2018) designed a *Global Features Network* which pools features from the whole image and computes a global feature set, then the global feature set is added back onto each feature map as basis. Such design provides a very large receptive field, leads to better preservation of the global information and removes artifacts caused by the strong spot highlight in the input image. Fig. 4 shows the effect of the global feature network design.

The encoder–decoder structure in Kang et al. (2018) is not symmetrical, its encoder is realized via physical measurements, and the mapped latent space does not have any correspondence to the output, as a result, Kang et al. (2018) designed a series of fully connected network to decode from its latent space.

To find an optimal lighting condition, the mapping from lighting patterns to the final rendered results needs to be modeled as a trainable operation. Thank to the linear properties of the light transport, such process can be modeled as a matrix multiplication (Xu et al., 2018; Kang et al., 2018), and any physical constraints can be directly applied to the matrix element values as a loss function.

Batch normalization or instant normalization is often used to stable the training process. Gao et al. (2019) in their paper discussed how different choices of batch normalization affect the latent space. As illustrated in Fig. 5, their result shows that, without batch normalization, the reconstruction result will be noisy; and using batch normalization will lead to blurred results. They designed a latent space smooth constraint to further constraint the latent space, resulting sharp reconstruction without noise.

4.2. Loss function

Aside from traditional L_1 and L_2 loss over the output, since each different component of the output material maps has unique physical meanings, one can render an image based on those material maps and use the rendering error as a loss function (Deschaintre et al., 2018; Li et al., 2018a). Deschaintre et al. (2018) in their paper compared the neural network trained with and without rendering loss. The neural network trained with rendering loss produces sharper results, especially on the normal map. More

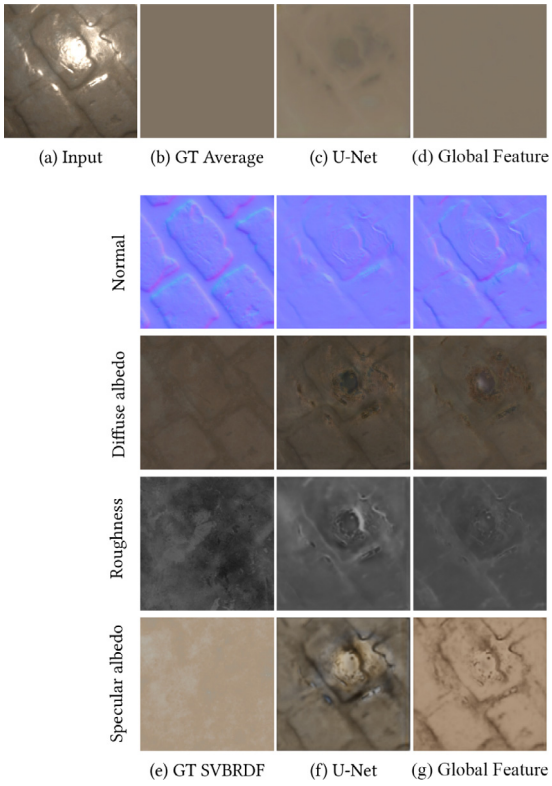

Fig. 4. Effect of the global feature network. Deschaintre et al. (2018) trained U-Net neural network without the global feature network and compared with their global feature network design. The U-Net failed to produce a constant image (c) when trained to output the average color of the input (b), and producing artifacts when estimating SVBRDFs (f), while global feature network design solved such problem, producing artifacts free results (d, g) (Deschaintre et al., 2018).

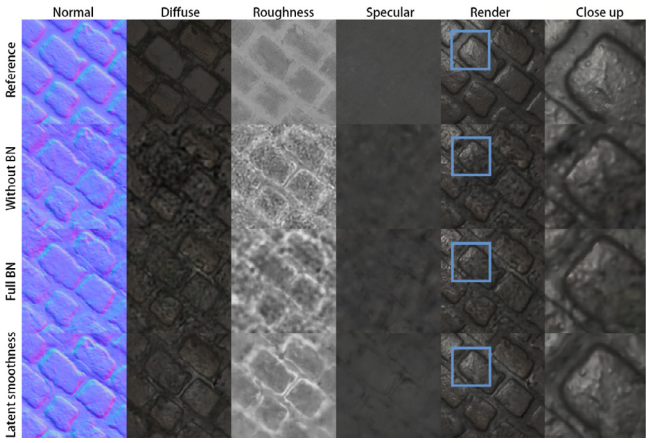

Fig. 5. Effect of batch normalization. Gao et al. (2019) test the effect of batch normalization in their auto-encoder design. The auto-encoder without batch normalization produces noisy results (2nd row), on the contrary, the auto-encoder with batch normalization after every activation layer generates smoothed results (3nd row). Their designed auto-encoder with single batch normalization layer together with the space smoothness design produce the most plausible result (Gao et al., 2019).

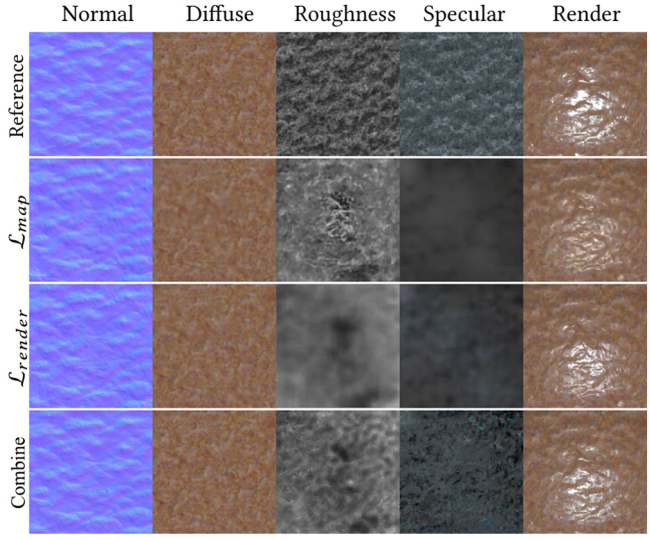

Fig. 6. Effect of render loss and map loss. The neural network trained only with map loss produces sharper material maps, however the rendering results do not fit the expected appearance. Training with the rendering loss solves the problem and the final appearance matches the target well, but produces blurred maps. Combining the map loss and rendering loss produces balanced between map details and render accuracy (Gao et al., 2019).

importantly, L_1 loss produces maps that do not reproduce the appearance of the ground truth, while the rendering loss produces a more faithful reproduction of the ground truth appearance.

Gao et al. (2019) also confirm the effect of render loss, and propose a combination of map loss and rendering loss which can preserve details in the rendering results as well as each individual material maps. Fig. 6 shows the results of optimizing an SVBRDF using an auto-encoder trained with different loss functions (Gao et al., 2019).

4.3. Training data

For single point BRDF, random sampling over parametric BRDFs could act as a good training set. Kang et al. (2018) randomly generate Anisotropic GGX BRDFs for their training dataset.

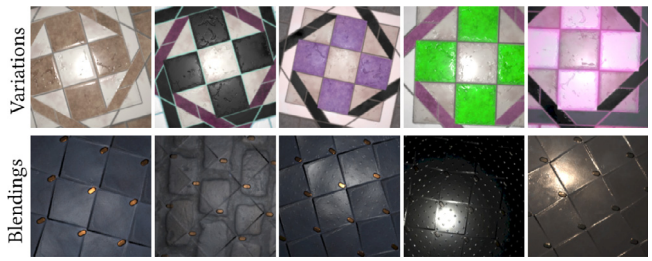

Fig. 7. Example of data augmentation. Perturbing parameters to generate different variations (top row); Blending different SVBRDFs (bottom row) (Deschaintre et al., 2018).

Li et al. (2017) also performed a synthetic test on isotropic Ward models.

For spatially varying materials, Li et al. (2017) collected a small scale manually crafted SVBRDF to bootstrap their self-augmented training. Deschaintre et al. (2018) collected their training data from Algorithmic Substance Share (Allegorithmic, 2018), a large collection of artist manually crafted SVBRDFs. Li et al. (2018a,b) collect their SVBRDF data from Adobe Stock (Allegorithmic, 2019). The original data of Li et al. (2017) and augmented training and test data of Deschaintre et al. (2018) are publicly available for future research, while the adobe SVBRDF dataset requires a commercial license.

Augmentation is essential for enhancing the coverage and diversity of the training data. Deschaintre et al. (2018) introduced four types of augmentation (see Fig. 7). They first randomly perturb the important parameters of the procedural SVBRDFs. Second, they generate convex combinations of random pairs of SVBRDFs. They then include random scaling and orientations, and finally apply a random crop to get the final training data. Li et al. (2018a) also augment their training data with random crop, rotation, and flip operations.

For geometry shapes, Li et al. (2018b) and Xu et al. (2018) use a procedural way to generate complex shapes and scenes. They first generate random primitive shapes (cube, ellipsoid, cylinder, box, and L-shape) enhanced by randomly generated height maps. Then for each scene, they randomly select 1 to 5 primitive shapes and randomly combine them. The augmented SVBRDF patches (Li et al., 2018a) are randomly applied onto those primitive shapes.

4.4. Training scheme

Unlike other machine learning tasks, the underlying image formation process is well known for appearance modeling. Utilizing the known rendering equation leads to many efficient training schemes.

Li et al. (2017) proposed a self-augmentation training strategy that combines a small labeled training set of measured SVBRDFs and a large unlabeled set of regular photographs of spatially-varying materials for learning a regression network for surface appearance estimation. Such scheme is further improved by Ye et al. (2018), remove the requirement of labeled training data, replace it with synthetically generated SVBRDFs based on neural texture synthesis.

Based on the fact that the re-rendered image from the estimated SVBRDF and lighting should match the input, Li et al. (2018b) designed a cascading network. The render error between the input and the image rendered by the previous stage estimations can be used as additional input for the next stage network, indicating the distribution of the error, guiding the refinement on those regions, progressively improving the estimation results.

5. Research opportunities and future directions

From the computer graphics perspective, there are many open opportunities on applying deep learning to complex lighting effects currently solved with traditional data-driven approaches; Learning based appearance model for rendering is also another interesting future direction. From the machine learning point of view, finding a generic latent space, creating large scale training datasets and developing semi-supervised and unsupervised training schemes for appearance modeling would be important future research directions.

5.1. Modeling complex lighting effects

Deep learning has shown its capability in many appearance modeling tasks, especially on surface reflectance modeling that covered in this survey. On the other hand, there are still many appearance modeling tasks that have not been tackled in a deep learning way, which opens new opportunists.

Light transport decomposition has many applications in computer graphics, such as scene editing and inverse global illumination. Xu et al. made some early attempts on this direction in their learning based relighting work (Xu et al., 2018). They designed two neural networks for the relighting task, one of the designed networks do have the capability to predict different components (visibility map, direct and global illumination components) of the relighted image, archiving marginal improvement on the prediction results. It is possible to extend this approach for light transport decomposition, or support advanced light transport editing (Dong et al., 2015).

Materials exhibit strong diffraction effects such as holographic surfaces (Toisoul et al., 2018) that could be one potential topic for future deep appearance modeling. Similar to traditional surface reflectance modeling, the key to inferring the underlying holographic pattern is finding corresponding repeated patterns on the measured input. Which can be archived by deep analysis (Aittala et al., 2016).

Current works on multi-layer material modeling still follow a traditional optimization based approach, with more in-depth understanding of the multi-layer light transport (Jakob et al., 2014a; Zeltner and Jakob, 2018) and efficient rendering (Belcour, 2018), large collection of training dataset can be generated, laying down the foundation for future deep learning based solutions.

5.2. Learning based model for rendering

Currently, the final output of the deep learning based surface reflectance estimation methods is mainly classical physical based appearance models, such as Cook–Torrance (Cook and Torrance, 1982) or GGX (Walter et al., 2007) BRDF models. Those physical based models render friendly and represent a wide range of real-world appearances. On the other hand, it is unclear whether the parameter settings of those physical based appearance models are ideal output for a deep learning system. Learning based appearance model for rendering, at the same time support back-propagation, can seamlessly integrate with existing deep learning framework. Such learned rendering model, optimized end-to-end with other learning based components, potentially can further improve the performance of the deep appearance modeling works.

The decoder part of the relight-net proposed in Xu et al. (2018) can be regarded as an early attempt in this direction. Their encoder can synthesize one relit image with the lighting direction as an input parameter, thus their decoder can be regarded as a neural renderer. However, due to their network design, such as the skip link structures, the output relit image is not fully determined by the latent code, thus one cannot render a new image with only a latent code, forbidding using only the decoder as a separated neural render model.

On searching better parametric BRDF models, Brady et al. (2014) proposed an evolution programming scheme and found several learned parametric models. It is also possible to explore in this direction that designs a learning friendly appearance model.

5.3. Generic latent space

Latent space is one of the key components in data-driven appearance modeling. The latent space determines the range of appearance that can be well modeled and affects the encoder and decoder design. In classical data-driven appearance modeling works, the latent space is always explicitly defined usually task independent. For example, the linear basis model of SVBRDF can be applied to SVBRDF acquisition, editing or even efficient rendering.

However, the majority of existing deep appearance modeling works learn a latent space specific for their task. Furthermore, deep appearance regression works do not explicitly model the latent space, with the *fitting* and *reproducing process* represented by a single neural network. In order to ensure the encoder can be realized by physical measurements, Kang et al. (2018) designed a special latent space that tailored for their device setup. Recently, Gao et al. (2019) begin to explore explicitly learning a latent space that supports arbitrary number of input measurements with a collocated flashlight. Along this track, learning a generic latent space that supports many appearance modeling tasks would be an interesting and useful application.

In the computer vision research communities, deep neural networks have been widely used to learn a generic latent space. The backbone feature extraction networks trained on large scale datasets such as ImageNet (Russakovsky et al., 2015) or MS-COCO (Lin et al., 2014) can be used for many computer vision task, without retraining. Whether there is an equivalent approach in appearance modeling still remains an open question.

5.4. Large scale appearance dataset

Data is always at the center of deep appearance modeling. However, there is no well designed large scale public dataset for deep appearance modeling. Different works are trained, tested and validated on different datasets, usually generated by the author following their own way of augmentation. The difference in the dataset setup and augmentation makes it difficult to compare different deep appearance learning approaches.

The open surfaces dataset (Bell et al., 2013, 2015) and the dataset prepared by Li et al. (2017), Deschaintre et al. (2018) could be a good starting point on this direction. In the future, building a large scale high quality public dataset, such as ImageNet (Russakovsky et al., 2015) in computer vision research field, would lay down a very important foundation for the appearance modeling field, and such dataset will serve as a powerful engine drives generations after generations of new research works on deep appearance modeling.

5.5. Semi-supervised and unsupervised training

Like any other learning based solution, deep appearance modeling is craving for large scale high quality datasets, insufficient training data will lead to reduced generality or over-fitting. Creating high quality large scale dataset would definitely pave the way for better learning based solutions. At the same time, designing novel training schemes that support semi-supervised learning or fully unsupervised learning would be also beneficial.

Li et al. (2017) proposed the self-augmented training scheme, which combines a small set of labeled SVBRDF data and a large collection of unlabeled rendered image to train a surface reflectance regression network. This scheme can be further extended to be fully unsupervised (Ye et al., 2018), where the labeled SVBRDF dataset can be replaced by synthetically augmented textures.

Combining the existing wisdom of data augmentation (Xu et al., 2018; Deschaintre et al., 2018) and self-supervised training would potentially lead to more efficient semi-supervised or even fully unsupervised training schemes that can benefit many deep appearance modeling applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author want to thank Xin Tong for the discussion and suggestion. The author also want to thank Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, Adrien Bousseau, Kaizhang Kang, Zimin Chen, Jiaping Wang, Kun Zhou, Hongzhi Wu, Zhengqin Li, Zexiang Xu, Ravi Ramamoorthi, Kalyan Sunkavalli and Manmohan Chandraker for providing their images shown in this paper.

References

Aittala, M., Aila, T., Lehtinen, J., 2016. Reflectance modeling by neural texture synthesis. ACM Trans. Graph. 35 (4), 65:1–65:13. http://dx.doi.org/10.1145/2897824.2925917, http://doi.acm.org/10.1145/2897824.2925917.

Aittala, M., Weyrich, T., Lehtinen, J., 2013. Practical SVBRDF capture in the frequency domain. ACM Trans. Graph. 32 (4), 110:1–110:12. http://dx.doi.org/10.1145/2461912.2461978, http://doi.acm.org/10.1145/2461912.2461978.

Allegorithmic, 2018. Substance share. https://share.allegorithmic.com/.

Allegorithmic, 2019. Adobe stock. https://stock.adobe.com/.

Ashikmin, M., Premože, S., Shirley, P., 2000. A microfacet-based brdf generator. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. In: SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, pp. 65–74. http://dx.doi.org/10.1145/344779.344814, http://dx.doi.org/10.1145/344779.344814.

Belcour, L., 2018. Efficient rendering of layered materials using an atomic decomposition with statistical operators. ACM Trans. Graph. 37 (4), 73:1–73:15. http://dx.doi.org/10.1145/3197517.3201289, http://doi.acm.org/10.1145/3197517.3201289.

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2013. Opensurfaces: a richly annotated catalog of surface appearance. ACM Trans. Graph. 32 (4).

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database. In: Computer Vision and Pattern Recognition (CVPR).

Brady, A., Lawrence, J., Peers, P., Weimer, W., 2014. genBRDF: discovering new analytic BRDFs with genetic programming. ACM Trans. Graph. 33 (4), http://doi.acm.org/10.1145/2601097.2601193.

Chen, G., Dong, Y., Peers, P., Zhang, J., Tong, X., 2014. Reflectance scanning: estimating shading frame and BRDF with generalized linear light sources. ACM Trans. Graph. 33 (4), http://doi.acm.org/10.1145/2601097.2601180.

Cook, R.L., Torrance, K.E., 1982. A reflectance model for computer graphics. ACM Trans. Graph. 1 (1), 7–24. http://dx.doi.org/10.1145/357290.357293, http://doi.acm.org/10.1145/357290.357293.

Deschaintre, V., Aittala, M., Durand, F., Drettakis, G., Bousseau, A., 2018. Single-image SVBRDF capture with a rendering-aware deep network. ACM Trans. Graph. 37 (4), 128:1–128:15. http://dx.doi.org/10.1145/3197517.3201378, http://doi.acm.org/10.1145/3197517.3201378.

Deschaintre, V., Aittala, M., Durand, F., Drettakis, G., Bousseau, A., 2019. Flexible SVBRDF capture with a multi-image deep network. Comput. Graph. Forum 38 (4).

Dong, B., Dong, Y., Tong, X., Peers, P., 2015. Measurement-based editing of diffuse albedo with consistent interreflections. ACM Trans. Graph. 34 (4), http://dx.doi.org/10.1145/2766979.

Dong, Y., Wang, J., Tong, X., Snyder, J., Lan, Y., Ben-Ezra, M., Guo, B., 2010. Manifold bootstrapping for SVBRDF capture. ACM Trans. Graph. 29 (4), 98:1–98:10. http://dx.doi.org/10.1145/1778765.1778835, http://doi.acm.org/10.1145/1778765.1778835.

Dorsey, J., Rushmeier, H., Sillion, F., 2008. Digital Modeling of Material Appearance. Morgan Kaufmann Publishers Inc.

Gao, D., Li, X., Dong, Y., Peers, P., Xu, K., Tong, X., 2019. Deep inverse rendering for high-resolution SVBRDF estimation from an arbitrary number of images. ACM Trans. Graph. 38 (4).

Gardner, A., Tchou, C., Hawkins, T., Debevec, P., 2003. Linear light source reflectometry. ACM Trans. Graph. 22 (3), 749–758. http://dx.doi.org/10.1145/882262.882342, http://doi.acm.org/10.1145/882262.882342.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, http://www.deeplearningbook.org.

Gortler, S.J., Grzeszczuk, R., Szeliski, R., Cohen, M.F., 1996. The lumigraph. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. In: SIGGRAPH '96, ACM, New York, NY, USA, pp. 43–54. http://dx.doi.org/10.1145/237170.237200, http://doi.acm.org/10.1145/237170.237200.

Holzschuch, N., Pacanowski, R., 2017. A two-scale microfacet reflectance model combining reflection and diffraction. ACM Trans. Graph. 36 (4), 66:1–66:12. http://dx.doi.org/10.1145/3072959.3073621, http://doi.acm.org/10.1145/3072959.3073621.

Jakob, W., d'Eon, E., Jakob, O., Marschner, S., 2014a. A comprehensive framework for rendering layered materials. ACM Trans. Graph. 33 (4), 118:1–118:14. http://dx.doi.org/10.1145/2601097.2601139, http://doi.acm.org/10.1145/2601097.2601139.

Jakob, W., Hašan, M., Yan, L.-Q., Lawrence, J., Ramamoorthi, R., Marschner, S., 2014b. Discrete stochastic microfacet models. ACM Trans. Graph. 33 (4), 115:1–115:10. http://dx.doi.org/10.1145/2601097.2601186, http://doi.acm.org/10.1145/2601097.2601186.

Jensen, H.W., Buhler, J., 2002. A rapid hierarchical rendering technique for translucent materials. ACM Trans. Graph. 21 (3), 576–581. http://dx.doi.org/10.1145/566654.566619, http://doi.acm.org/10.1145/566654.566619.

Kang, K., Chen, Z., Wang, J., Zhou, K., Wu, H., 2018. Efficient reflectance capture using an autoencoder. ACM Trans. Graph. 37 (4), 127:1–127:10. http://dx.doi.org/10.1145/3197517.3201279, http://doi.acm.org/10.1145/3197517.3201279.

Kim, K., Gu, J., Tyree, S., Molchanov, P., Nießner, M., Kautz, J., 2017. A lightweight approach for on-the-fly reflectance estimation. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 20–28. http://dx.doi.org/10.1109/ICCV.2017.12.

Lan, Y., Dong, Y., Pellacini, F., Tong, X., 2013. Bi-scale appearance fabrication. ACM Trans. Graph. 32 (4), 145:1–145:12. http://dx.doi.org/10.1145/2461912.2461989, http://doi.acm.org/10.1145/2461912.2461989.

Lawrence, J., Ben-Artzi, A., DeCoro, C., Matusik, W., Pfister, H., Ramamoorthi, R., Rusinkiewicz, S., 2006. Inverse shade trees for non-parametric material representation and editing. ACM Trans. Graph. 25 (3), 735–745. http://dx.doi.org/10.1145/1141911.1141949, http://doi.acm.org/10.1145/1141911.1141949.

Lensch, H.P.A., Kautz, J., Goesele, M., Heidrich, W., Seidel, H.-P., 2003. Image-based reconstruction of spatial appearance and geometric detail. ACM Trans. Graph. 22 (2), 234–257. http://dx.doi.org/10.1145/636886.636891, http://doi.acm.org/10.1145/636886.636891.

Levoy, M., Hanrahan, P., 1996. Light field rendering. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. In: SIGGRAPH '96, ACM, New York, NY, USA, pp. 31–42. http://dx.doi.org/10.1145/237170.237199, http://doi.acm.org/10.1145/237170.237199.

Li, X., Dong, Y., Peers, P., Tong, X., 2017. Modeling surface appearance from a single photograph using self-augmented convolutional neural networks. ACM Trans. Graph. 36 (4), 45:1–45:11. http://dx.doi.org/10.1145/3072959.3073641, http://doi.acm.org/10.1145/3072959.3073641.

Li, Z., Sunkavalli, K., Chandraker, M.K., 2018a. Materials for masses: SVBRDF acquisition with a single mobile phone image. In: ECCV.

Li, Z., Xu, Z., Ramamoorthi, R., Sunkavalli, K., Chandraker, M., 2018b. Learning to reconstruct shape and spatially-varying reflectance from a single image. ACM Trans. Graph. 37 (6), 269:1–269:11. http://dx.doi.org/10.1145/3272127.3275055, http://doi.acm.org/10.1145/3272127.3275055.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: common objects in context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014. Springer International Publishing, Cham, pp. 740–755.

Lombardi, S., Nishino, K., 2012. Single image multimaterial estimation. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). In: CVPR '12, IEEE Computer Society, Washington, DC, USA, pp. 238–245, http://dl.acm.org/citation.cfm?id=2354409.2354759.

Lombardi, S., Saragih, J., Simon, T., Sheikh, Y., 2018. Deep appearance models for face rendering. ACM Trans. Graph. 37 (4), 68:1–68:13. http://dx.doi.org/10.1145/3197517.3201401, http://doi.acm.org/10.1145/3197517.3201401.

Matusik, W., Pfister, H., Brand, M., McMillan, L., 2003. A data-driven reflectance model. ACM Trans. Graph. 22 (3), 759–769. http://dx.doi.org/10.1145/882262.882343, http://doi.acm.org/10.1145/882262.882343.

Mcallister, D.K., 2002. A generalized surface appearance representation for computer graphics (Ph.D. thesis). AAI3061704.

Ng, R., Ramamoorthi, R., Hanrahan, P., 2003. All-frequency shadows using non-linear wavelet lighting approximation. ACM Trans. Graph. 22 (3), 376–381. http://dx.doi.org/10.1145/882262.882280, http://doi.acm.org/10.1145/882262.882280.

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W., Limperis, T., 1977. Geometric considerations and nomenclature for reflectance. In: Monograph, vol. 161. National Bureau of Standards (US).

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W., Limperis, T., 1992. Radiometry. In: Wolff, L.B., Shafer, S.A., Healey, G. (Eds.), Jones and Bartlett Publishers, Inc, USA, pp. 94–145, http://dl.acm.org/citation.cfm?id=136913.136929.

Peers, P., vom Berge, K., Matusik, W., Ramamoorthi, R., Lawrence, J., Rusinkiewicz, S., Dutré, P., 2006. A compact factored representation of heterogeneous subsurface scattering. ACM Trans. Graph. 25 (3), 746–753. http://dx.doi.org/10.1145/1141911.1141950, http://doi.acm.org/10.1145/1141911.1141950.

Peers, P., Mahajan, D.K., Lamond, B., Ghosh, A., Matusik, W., Ramamoorthi, R., Debevec, P., 2009. Compressive light transport sensing. ACM Trans. Graph. 28 (1), 3:1–3:18. http://dx.doi.org/10.1145/1477926.1477929, http://doi.acm.org/10.1145/1477926.1477929.

Rainer, G., Jakob, W., Ghosh, A., Weyrich, T., 2019. Neural BTF compression and interpolation. In: Proc. Eurographics. Comput. Graph. Forum 38 (2), 1–10.

Ren, P., Dong, Y., Lin, S., Tong, X., Guo, B., 2015. Image based relighting using neural networks. ACM Trans. Graph. 34 (4), 111:1–111:12. http://dx.doi.org/10.1145/2766899, http://doi.acm.org/10.1145/2766899.

Ren, P., Wang, J., Gong, M., Lin, S., Tong, X., Guo, B., 2013. Global illumination with radiance regression functions. ACM Trans. Graph. 32 (4), 130:1–130:12. http://dx.doi.org/10.1145/2461912.2462009, http://doi.acm.org/10.1145/2461912.2462009.

Ren, P., Wang, J., Snyder, J., Tong, X., Guo, B., 2011. Pocket reflectometry. ACM Trans. Graph. 30 (4), 45:1–45:10. http://dx.doi.org/10.1145/2010324.1964940, http://doi.acm.org/10.1145/2010324.1964940.

Riviere, J., Reshetouski, I., Filipi, L., Ghosh, A., 2017. Polarization imaging reflectometry in the wild. ACM Trans. Graph. 36 (6), 206:1–206:14. http://dx.doi.org/10.1145/3130800.3130894, http://doi.acm.org/10.1145/3130800.3130894.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. 115 (3), 211–252. http://dx.doi.org/10.1007/s11263-015-0816-y.

Tang, Y., Salakhutdinov, R., Hinton, G., 2012. Deep lambertian networks. In: Langford, J., Pineau, J. (Eds.), Proceedings of the 29th International Conference on Machine Learning (ICML-12). In: ICML '12, Omnipress, New York, NY, USA, ISBN: 978-1-4503-1285-1, pp. 1623–1630.

Toisoul, A., Dhillon, D.S., Ghosh, A., 2018. Acquiring spatially varying appearance of printed holographic surfaces. ACM Trans. Graph. 37 (6), 272:1–272:16. http://dx.doi.org/10.1145/3272127.3275077, http://doi.acm.org/10.1145/3272127.3275077.

Toisoul, A., Ghosh, A., 2017. Practical acquisition and rendering of diffraction effects in surface reflectance. ACM Trans. Graph. 36 (5), http://dx.doi.org/10.1145/3012001, http://doi.acm.org/10.1145/3012001.

Walter, B., Marschner, S.R., Li, H., Torrance, K.E., 2007. Microfacet models for refraction through rough surfaces. In: Proceedings of the 18th Eurographics Conference on Rendering Techniques. In: EGSR'07, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 195–206. http://dx.doi.org/10.2312/EGWR/EGSR07/195-206, http://dx.doi.org/10.2312/EGWR/EGSR07/195-206.

Wang, J., Zhao, S., Tong, X., Lin, S., Lin, Z., Dong, Y., Guo, B., Shum, H., 2007. Modeling and rendering of heterogeneous translucent materials using the diffusion equation. ACM Trans. Graph. 27.

Wang, J., Zhao, S., Tong, X., Snyder, J., Guo, B., 2008. Modeling anisotropic surface reflectance with example-based microfacet synthesis. ACM Trans. Graph. 27 (3), 41:1–41:9. http://dx.doi.org/10.1145/1360612.1360640, http://doi.acm.org/10.1145/1360612.1360640.

Ward, G.J., 1992. Measuring and modeling anisotropic reflection. SIGGRAPH Comput. Graph. 26 (2), 265–272.

Weinmann, M., Klein, R., 2015. Advances in geometry and reflectance acquisition. In: ACM SIGGRAPH Asia, Course Notes.

Werner, S., Velinov, Z., Jakob, W., Hullin, M.B., 2017. Scratch iridescence: wave-optical rendering of diffractive surface structure. ACM Trans. Graph. 36 (6), 207:1–207:14. http://dx.doi.org/10.1145/3130800.3130840, http://doi.acm.org/10.1145/3130800.3130840.

Wu, H., Dorsey, J., Rushmeier, H., 2009. Characteristic point maps. Comput. Graph. Forum 28 (4), 1227–1236.

Wu, H., Dorsey, J., Rushmeier, H., 2011. Physically-based interactive bi-scale material design. ACM Trans. Graph. 30, 145:1–145:10, http://doi.acm.org/10.1145/2070781.2024179.

Wu, H., Dorsey, J., Rushmeier, H., 2013. Inverse bi-scale material design. ACM Trans. Graph. 32 (6), 163:1–163:10. http://dx.doi.org/10.1145/2508363.2508394, http://doi.acm.org/10.1145/2508363.2508394.

Xu, Z., Sunkavalli, K., Hadap, S., Ramamoorthi, R., 2018. Deep image-based relighting from optimal sparse samples. ACM Trans. Graph. 37 (4), 126:1–126:13. http://dx.doi.org/10.1145/3197517.3201313, http://doi.acm.org/10.1145/3197517.3201313.

Yan, L.-Q., Hašan, M., Jakob, W., Lawrence, J., Marschner, S., Ramamoorthi, R., 2014. Rendering glints on high-resolution normal-mapped specular surfaces. ACM Trans. Graph. 33 (4), 116:1–116:9. http://dx.doi.org/10.1145/2601097.2601155, http://doi.acm.org/10.1145/2601097.2601155.

Yan, L.-Q., Hašan, M., Marschner, S., Ramamoorthi, R., 2016. Position-normal distributions for efficient rendering of specular microstructure. ACM Trans. Graph. 35 (4), 56:1–56:9. http://dx.doi.org/10.1145/2897824.2925915, http://doi.acm.org/10.1145/2897824.2925915.

Ye, W., Li, X., Dong, Y., Peers, P., Tong, X., 2018. Single photograph surface appearance modeling with self-augmented CNNs and inexact supervision. Comput. Graph. Forum 37 (7).

Zeltner, T., Jakob, W., 2018. The layer laboratory: a calculus for additive and subtractive composition of anisotropic surface reflectance. ACM Trans. Graph. 37 (4), 74:1–74:14. http://dx.doi.org/10.1145/3197517.3201321, http://doi.acm.org/10.1145/3197517.3201321.

Zhou, Z., Chen, G., Dong, Y., Wipf, D., Yu, Y., Snyder, J., Tong, X., 2016. Sparse-as-possible SVBRDF acquisition. ACM Trans. Graph. 35 (6), 189:1–189:12. http://dx.doi.org/10.1145/2980179.2980247, http://doi.acm.org/10.1145/2980179.2980247.

Zickler, T., Enrique, S., Ramamoorthi, R., Belhumeur, P., 2005. Reflectance sharing: image-based rendering from a sparse set of images. In: Proceedings of the Sixteenth Eurographics conference on Rendering Techniques. In: EGSR'05, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 253–264. http://dx.doi.org/10.2312/EGWR/EGSR05/253-264.

Zsolnai-Fehér, K., Wonka, P., Wimmer, M., 2018. Gaussian material synthesis. ACM Trans. Graph. 37 (4), 76:1–76:14. http://dx.doi.org/10.1145/3197517.3201307, http://doi.acm.org/10.1145/3197517.3201307.