Computing the Bidirectional Scattering of a Microstructure Using Scalar Diffraction Theory and Path Tracing

V. Falster¹, A. Jarabo²,³, and J. R. Frisvad¹

¹Technical University of Denmark, Denmark
²Universidad de Zaragoza, Spain
³Centro Universitario de la Defensa Zaragoza, Spain

Abstract
Most models for bidirectional surface scattering by arbitrary explicitly defined microgeometry are either based on geometric optics and include multiple scattering but no diffraction effects or based on wave optics and include diffraction but no multiple scattering effects. The few exceptions to this tendency are based on rigorous solution of Maxwell’s equations and are computationally intractable for surface microgeometries that are tens or hundreds of microns wide. We set up a measurement equation for combining results from single scattering scalar diffraction theory with multiple scattering geometric optics using Monte Carlo integration. Since we consider an arbitrary surface microgeometry, our method enables us to compute expected bidirectional scattering of the metasurfaces with increasingly smaller details seen more and more often in production. In addition, we can take a measured microstructure as input and, for example, compute the difference in bidirectional scattering between a desired surface and a produced surface. In effect, our model can account for both diffraction colors due to wavelength-sized features in the microgeometry and brightening due to multiple scattering. We include scalar diffraction for refraction, and we verify that our model is reasonable by comparing with the rigorous solution for a microsurface with half ellipsoids.

CCS Concepts
• Computing methodologies → Reflectance modeling;

1. Introduction
Scalar diffraction theory is frequently used in the modeling of reflectance functions [Kaj85; HTSG91; Sta99; DWMG15; HP17; WVJH17; YHW*18; KHZ*19]. This is especially the case when it comes to rendering of scratched or brushed metallic surfaces or glinty surfaces in general. Wave optics is important for these material types because the surface microgeometry has features of a size comparable to the wavelength of visible light. Another important observation in recent work on the modeling of reflectance functions is the necessity to account for multiple scattering between microfacets [HHdD16; LJ*18; XH18; CCM19]. This is important to avoid loss of energy that is not due to absorption but simply due...
to the limitation of single scattering. While single scattering models can be renormalized to avoid loss of energy, this renormalization leads to an inaccurate distribution of the scattered light because multiple scattering was not accounted for. The wave optics models from scalar diffraction theory are single scattering models and thus suffer from the same issue.

One way to include both multiple scattering and diffraction effects is using a rigorous solution of Maxwell’s equations [MMRO13; AHB18]. Use of rigorous solvers is however a computationally demanding problem that becomes intractable when we consider a patch of microgeometry that can be tens or hundreds of microns in size. We therefore aim at a more practical solution, where we combine single scattering scalar diffraction theory with multiple scattering from geometric optics. Although this is an approximation, we find that this approach has the ability to capture the important visual effects: color variation due to diffraction and brightening due to multiple scattering.

We consider explicitly defined microgeometry instead of the popular normal distribution function maps [DWMG15; YHM16; YHW*18; KHZ*19]. This enables us to account for shadowing and masking without the simplifications imposed by an analytic geometric attenuation term like the often used Smith approximation [Smi67]. Information about surface microgeometry is more and more commonly available. We can measure it with a profilometer [DWMG15] or model it based on a desired surface structure [WDR11; MMRO13; LFD*17; AHB18; LFD*20]. It is even an option to find an explicit microgeometry representative of a normal distribution function [RBSM19]. We thus find it advantageous to base our model directly on the microgeometry. This means that we can support a wide variety of different surface types including optical functional surfaces with engineered microstructure.

The surface microstructures in Figure 1 are examples for which both diffraction and multiple scattering effects are visually significant. Without the combined model that we suggest, we would have to choose between multiple scattering or diffraction or almost intractable rigorous numerical evaluation of Maxwell’s equations. We find that the combination of path tracing and scalar diffraction theory, both well known and often employed tools in graphics, is a very practical method for computing the scattering properties of this kind of surface. We also find this approach an excellent tool for analyzing the differences between scalar diffraction and geometric optics approximations and for making decisions on the adequateness of analytic models.

2. Related Work

Rendering of material appearance using light scattering models is a multiscale problem. We use surface scattering models to cope with the complexity of light-matter interaction at a more microscopic scale. Early analytic models [TS67; Bu77] rely on single scattering geometric optics, V-groove geometry, and a distribution of microfacet orientations (random roughness). This significantly limits the types of surfaces that one can faithfully model. An extensive body of work provides extension of this outset and models more advanced surface types and light scattering phenomena [FJM*20]. An important extension is the use of explicitly defined microgeometry based on Kirchhoff theory [Kaj85]. This approach includes diffraction effects, but is challenged by the limitations of Kirchhoff theory, which is valid only if shadowing, masking, and multiple scattering are negligible.

We can deal with shadowing, masking, and multiple scattering in a geometric optics setting by ray tracing a patch of microgeometry to compute a bidirectional reflectance distribution function (BRDF) [CMS87; WAT92] If we account for only shadowing and masking, the effect of modifying the microgeometry can be computed interactively using rasterization and shadow mapping [WDR11]. These are very flexible approaches in terms of types of microgeometry that one can faithfully model, but they do not account for the fact that geometric features have a size where diffraction effects become important. We can include information about the light waves in the ray tracing and perform wavefront tracing [GMN94; SML*12]. This enables us to account for interference effects when an outgoing ray arrives in a solid angle bin, but diffraction effects (bending of light around geometric edges) are not accounted for by this method. Using the Wigner distribution function, ray tracing can be extended to include diffraction effects [OKG*10; CHB*12]. This method however relies on analytic solutions for regular structures, which makes it hard to use it for an arbitrary surface microstructure. Other work incorporate wave effects into reflectance based on geometric optics by accounting for thin coatings on the microgeometry [BB17; GCM*20], for example, but diffraction is not accounted for in these works.

Extension of the Kirchhoff model (or Harvey-Shack) to account for shadowing and masking while retaining diffraction effects has been proposed as well [HTSG91; Sta99; DWMG15; HP17]. However, to employ appropriate analytic functions for shadowing and masking, these models revert to use of modified normal distribution functions instead of explicitly defined microgeometry. The analytic functions used for shadowing and masking are the same in these models based on scalar diffraction theory as in the models based on single scattering geometric optics [WM17]. This is clear in the work of Holzschuch and Pacanowski [HP17] where these two single scattering solutions are combined with each model addressing reflectance due to geometry at different scales. However, only models based on geometric optics [TS67; WML70] have been extended to multiple scattering. This has been done using a Monte Carlo approach based on sampling of the normal distribution function [HHG16] and using analytic solutions based on an assumption of V-groove microgeometry [LJ*18; XH18]. Recently, the sampling of multiple scattering in normal distribution functions was extended to the spatially varying BRDFs of glinty surfaces [CCM19].

The accuracy of the Kirchhoff approximation is based on a small-angle assumption and has limited accuracy for wide-angle scattering and grazing angles of incidence. Various modifications of the theory therefore exist to improve its accuracy [HKV07]. Rayleigh-Rice theory is another single scattering scalar diffraction approximation that can be used [LYKU12] and the generalized Harvey-Shack (GHS) theory was introduced to have a model with good accuracy at arbitrary angles of incidence and scattering [HKS11]. In a sense, one could say that the GHS theory is a wave optics approach to dealing with the missing shadowing and masking in the Kirchhoff approximation. GHS theory is con-
sidered in the single scattering BRDF models by Holzschuch and Pacanowski [HP17] and Yan et al. [YHW+18]. The term from wave optics corresponding to the shadowing and masking from geometric optics is similar but not exactly the same [BNM15; WYH+18; KHZ+19]. To the best of our knowledge, it is unknown to what extent ray traced shadowing and masking in combination with the original Kirchhoff approximation (as suggested by Sancer [San69]) compares with results obtained with GHS theory. In any case, GHS is a single scattering theory and uses renormalization to ensure energy conservation [KHC11; HP17]. This means that the resulting energy distribution will be increasingly incorrect as multiple scattering effects increase in significance.

In the following, we use radiometry and Monte Carlo integration to combine multiple scattering based on geometric optics with scalar diffraction theory aiming at a technique that can capture both diffraction effects and brightening due to multiple scattering. To the best of our knowledge, we are the first to present a technique for computing a full BSDF from an arbitrary explicitly defined patch of microgeometry that includes both these effects. Holzschuch and Pacanowski [HP17] suggest a concept similar to ours as future work, namely combination of their work with multiple scattering [HHdD16]. The theoretical framework that we present could be used for such a combination if one is willing to accept the limitations of normal distribution functions. One would then use the scalar diffraction theory of Holzschuch and Pacanowski [HP17] for single scattering and the method of Heitz et al. [HHdD16] for secondary bounces. Taking a first step, we decided to keep our theory general and applicable to explicitly defined microgeometry.

3. Combining Geometric and Scalar Wave Optics

To combine a multiple scattering geometric optics approach with wave optics, we need to set up a measurement equation for computing a bidirectional scattering distribution function (BSDF) using Monte Carlo integration. We separate the integral in this measurement equation into a sum of two terms and use scalar diffraction theory for one term and regular path tracing without single scattering for the other term. These are two different approximations, and we have to accept an error since light will be partially coherent after the first bounce while geometric optics disregards coherence. We can compute the amount of coherent and incoherent light after the first bounce and estimate the significance of this error, but we intentionally do not use the result from scalar diffraction when computing the secondary bounces. Since we keep the evaluation of first and secondary bounces separate, use of different approximations for the two terms is valid.

3.1. Measurement Equation for the BSDF

The BSDF at a surface location \( \mathbf{x} \) for directions of incidence and observation \( \hat{o}_i \) and \( \hat{o}_o \) is defined by [BDW81]

\[
fs(x, \hat{o}_i, \hat{o}_o) = \frac{dl_{oo}(x, \hat{o}_o)}{dE(x, \hat{o}_i)},
\]

where \( l_{oo} \) is outgoing radiance and \( E \) is irradiance. The differential element of irradiance \( dE \) incident at \( x \) from a differential element of solid angle around \( \hat{o}_i \) is

\[
dE(x, \hat{o}_i) = L_o(x, \hat{o}_i) |\cos \Theta_i| d\Omega_i.
\]

Here, \( L_o \) is incident radiance and \( \Theta_i \) is the angle of incidence (the angle between \( \hat{o}_i \) and the surface normal \( \hat{n} \)). Using the definition of radiance [Nic63], we have

\[
L_o = \frac{d^2J_o}{\cos\Theta_o d\Omega d\omega},
\]

where \( \Theta_o \) is the angle of reflection or transmission, while \( d\Omega \) is a differential element of surface area around \( x \). To obtain \( L_o \), we measure the radiant flux \( \Phi_o \) scattered by a small patch of area \( \Omega_o \) centered around \( \hat{o}_o \). As a consequence of Eqs. (1) and (3), we can measure the BSDF using

\[
fs(x, \Omega_i, \Omega_o) \approx \frac{\Phi_o(A, \Omega_o)}{\cos \Theta_o} \frac{d\Omega_o}{d\Omega_i} \frac{d\omega}{d\Omega},
\]

where \( \Theta_o \) is the angle between the surface normal at \( x \) and the direction in the centre of the \( \Omega_o \) solid angle. Using Eq. (3), the outgoing radiant flux in \( \Omega_o \) is

\[
\Phi_o(A, \Omega_o) = \int_{\Omega_o} L_o(x, \hat{o}_o) |\cos \Theta_i| d\omega, \quad \text{(5)}
\]

while it follows from Eqs. (1) and (2) that

\[
L_o(x, \hat{o}_o) = \frac{1}{2\pi} \int_{\Omega_o} d\omega \int_{\Omega_i} d\omega' \Phi_o(A, \hat{o}_o, \hat{o}_i) \cos \Theta_i |\cos \Theta_i| d\omega_d \omega_i \quad \text{(6)}
\]

but now \( \omega_d \) is a microfacet BSDF and \( \omega_i \) is a position within the microgeometry of the patch of area \( A \) centered at \( x \). Considering Eq. (2), the irradiance is

\[
E(x, \Omega_i) = \int_{\Omega_o} L_o(x, \hat{o}_o) |\cos \Theta_i| d\omega.
\]

We can thus solve the measurement equation (4) by Monte Carlo integration and store the resulting BSDF values in \( \Phi_o(A, \Omega_o) \)-bins.

The reflected radiance equation (6) is recursive just like the rendering equation. Let us split it into single and multiple scattering contributions: \( L_o = L_o^s + L_o^m \). The single scattering contribution arrives directly from \( \Omega_i \) and is reflected directly into \( \Omega_o \), whereas the multiple scattering contribution involves at least one other position in the patch microgeometry. The two terms thus correspond to direct and indirect illumination in conventional path tracing. Inserting this split of the equation for outgoing radiance into Eq. (4) and propagating the superscript, we have

\[
fs(x, \Omega_i, \Omega_o) \approx \frac{\Phi_o^s(A, \Omega_o) + \Phi_o^m(A, \Omega_o)}{\cos \Theta_o} \frac{d\Omega_o}{d\Omega_i} \frac{d\omega}{d\Omega},
\]

In the following, we develop Monte Carlo estimators for evaluation of these two terms. First, we discuss geometric optics path tracing for the indirect illumination term (Sec. 3.2), and then scalar diffraction theory for the direct illumination term (Sec. 3.3).

3.2. Progressive Path Tracing

Let us set up an estimator for Eq. (4). This is easily modified to compute only one of the terms in Eq. (8) when we do path tracing. Sampling a direction of incidence \( \hat{o}_i \) uniformly in a bin \( \Omega_i \) with \( k \) as sample index, the probability density function is \( \text{pdf}(\hat{o}_i) = 1/\Omega_i \). Assuming unit incident radiance \( L_o = 1 \), we
have the irradiance estimator

\[ E_k(x, \Omega) = \frac{1}{K} \sum_{j=1}^{K} \frac{L_j(x, \mathbf{\theta}_j)}{pdf(\theta_j)} \cos \theta_j \]  

where \( \mathbf{\theta} \) is the normal of the macroscopic surface.

The interaction of a light wave with geometric features significantly smaller than the wavelength is insignificant (this is the Rayleigh criterion of optical smoothness). At the micro scale, we therefore assume that the surface is perfectly smooth and use the BSDF of a perfectly specular material for \( \Phi_{sk} \). The expression for this BSDF is available from Walter et al. [WMILT07]. It basically tells us that we can evaluate Eq. (6) by path tracing with a Russian roulette to select reflection or refraction at each path vertex within the microgeometry. Fresnel reflectance \( F_r \) is used as the probability of reflection and \( F_t \) is the probability of transmission.

Each path is then fully deterministic and the integral over outgoing directions in Eq. (5) becomes one if \( \mathbf{\theta}_o \in \Omega_o \) and otherwise zero. We can now write up an estimator for Eq. (5) by uniformly sampling positions \( x_{m,j} \) in the microgeometry, \( j \) is the sample index and \( pdf(x_{m,j}) = 1/A \). Considering spectral radiance, where Fresnel reflectance or transmittance \( (F_r \) or \( F_t) \) cancels perfectly in every Russian roulette, we have that the outgoing radiance is the light source visibility of the sampled point \( L_{o,j,k} = V(x_{m,j}, \mathbf{\theta}_o) \) when the path exits the microgeometry in a direction \( \mathbf{\theta}_o,j,k \). Then

\[ \Phi_o(A, \Omega_o) = \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{N} \frac{L_{o,j,k}}{pdf(x_{m,j})} [\mathbf{\theta}_o,j,k \in \Omega_o] \]  

where \( m \) is the microfacet normal at \( x_{m,j} \) and \( x \) is an Iverson bracket, which is 1 if the condition \( x \) is true and 0 otherwise. The origin of a sampled path to be traced through the microgeometry is \( x_{m,j} + r \mathbf{\theta}_o \), where \( r \) is the radius of the bounding sphere of the microgeometry, and the initial direction of the path is \( -\mathbf{\theta}_o \). We would have to discard rays not reaching \( x_{m,j} \) to account for the visibility term (or sample the illuminated area only, see Sec. 5).

Suppose we set up a path tracer to use an orthographic camera with resolution \( W \times H \), and we let the camera observe the square \([-1,1] \times [-1,1]\). Conveniently, we can store the orthographic projection of a hemispherical function in the image produced by this path tracer using each pixel as a projected solid angle bin. Orienting the microgeometry so that the \( z \)-axis represents the normal \( \mathbf{\hat{n}} \) of the macro surface, the \( x \) and \( y \)-coordinates of \( \mathbf{\theta}_o,j,k \) identify the \( \Omega_o \) bin that a sampled path arrives in. We then have the interesting construction that the area of one pixel \( A_p \) corresponds to the projected solid angle of a bin:

\[ A_p = \frac{4}{WH} \Omega_o \cos \Theta_o. \]  

By insertion of these different results [Eqs. (9)–(11)] in Eq. (4), our collective Monte Carlo estimator for the BSDF becomes

\[ \Phi_{sk}(x, \Omega) = \frac{WH}{N} \sum_{k=1}^{N} \frac{V(x_{m,j}, \mathbf{\theta}_o)}{pdf(x_{m,j})} \frac{[\mathbf{\theta}_o,j,k \in \Omega_o]}{4 \sum_{k=1}^{N} \Omega_o / \Omega_k \cdot \mathbf{\hat{n}} \cdot \mathbf{\theta}_o,j,k}. \]  

Out of convenience, we can choose one area sample in the microgeometry per pixel and use progressive path tracing. Then we have \( N = WH \) and only \( K = 1 \) direction of incidence per \( \Omega_o \) bin per progressive update. A simple check on the trace depth in the path tracing is all we need to evaluate one or the other term in Eq. (8).

### 3.3. Scalar Diffraction Theory

As scalar diffraction theory is a single scattering model, we can use it for computing \( \Phi_{sk} \) in Eq. (8). Modeling the incident field as a scalar plane wave and using Beckmann’s version of the Kirchhoff approximation [BS65], Kajiya [Kaj85] provided a formula for the complex amplitude of the reflected wave

\[ \psi_r(A) = -\frac{e^{i kr}}{4\pi r} \int_A \mathbf{\hat{n}} \cdot |k_1 - k_2| R - (k_1 + k_2)| e^{i(k_1 - k_2) \cdot x} \, dA, \]  

where \( m \) is the microfacet normal at the position \( x_m \) in the microgeometry, \( r \) is the distance to the observer, and \( k_1 = -k_2 \mathbf{\hat{n}} \) and \( k_2 = k_2 \mathbf{\hat{n}} \) are the wave vectors of the incident and the reflected fields (with wave numbers \( k_1 = 2\pi n_1 / \lambda \) and \( k_2 = 2\pi n_2 / \lambda \), where \( n_1 \) and \( n_2 \) are refractive indices of the media the waves propagate in). The factor \( R \) is a complex reflection coefficient given by the Fresnel equations before taking the squared absolute value.

The Kirchhoff approximation is based on the assumptions that all points in the microgeometry \( x_m \) are visible and that a plane wave of unit amplitude is incident in all points. Shadowing and masking effects are in other words neglected. By explicitly including the amplitude of the incident wave in the derivation, Sancer [San69] showed that we can account for shadowing and masking using ray tracing. If we evaluate the integral by Monte Carlo integration, we can use ray tracing to check for shadowing or masking and only include a sample if visible from both light source and detector. Inclusion of shadowing and masking in the Kirchhoff integral was also explained in the very useful appendix to the paper by He et al. [HTSG91]. The standard approach (as also outlined in this appendix) would now be to simplify Eq. (13) until analytic approximation becomes manageable.

The standard simplification of Eq. (13) is to assume that the surface geometry has a large area and is slowly varying. We can then collect the two terms in the integrand and consider the Fresnel term independent of the microfacet normal \( m \). With this simplification, we can pull the Fresnel term outside the integral and make it manageable without computerization [Bec67]. A good description of this simplification is provided by Ishimaru [Ish78]. As described by Walter et al. [WHYH18], the resulting simplified integral has a form that with variation of a few terms can describe several different commonly used models from scalar diffraction theory. We can thus use Eq. (13) or one of the other models from previous work [DWMG15; WVIH17; VHY18; KHZ19] for computing the complex amplitude of the reflected wave.

In the case of transmission, Eq. (13) must be modified to specify the complex amplitude of the transmitted wave. This modification is available from Caron et al. [CLA02]:

\[ \psi_t(A) = -\frac{e^{i kr}}{4\pi r} \int_A \mathbf{\hat{n}} \cdot (|k_1 + k_2| T) e^{i(k_1 + k_2) \cdot x} \, dA, \]  

© 2020 The Author(s)

Computer Graphics Forum © 2020 The Eurographics Association and John Wiley & Sons Ltd.
where $T$ is the complex transmission coefficient from the Fresnel equations (expressions for $R$ and $T$ are available in the appendix of Kajiya’s paper [Kaj85]) and

$$k_1 \cdot \vec{m} = -\sqrt{k_2^2 - k_1^2} + (\vec{k}_1 \cdot \vec{m})^2.$$  

(15)

To have reflectance and transmittance factors, we need the ratios $S$ and $T$ of the reflected and transmitted wave amplitudes to the incident wave amplitude.

The amplitudes of the scattered waves ($\psi_s$ and $\psi_t$) depend on the distance to the observer, which is impractical when we are looking for a bidirectional function. This is resolved by going to the far field, where we collect the scattered energy in a small sensor perpendicula to $\vec{\omega}_0$ at the distance $r$ and take the limit of $r$ going to infinity [WYH18]. This limit is taken in way so that $r$ approaches infinity in steps of a full period of oscillation. Incidentally, the same result can be obtained in the reflection case through normalization using the amplitude of the wave that would be reflected by a perfectly smooth perfect conductor [BS63; Kaj85]. Using $R = 1, \vec{m} = \vec{n}$, and $\vec{n} \cdot \vec{k}_1 = \vec{n} \cdot \vec{k}_2$ in Eq. (13), this amplitude is

$$\psi_{s0}(A) = \frac{i e^{i k r}}{4 \pi r} \left(2 |\vec{n} \cdot \vec{k}_1| \right)^{2}.$$  

(16)

To modify Eq. (13) with Sancer’s ray traced shadowing and masking, we insert visibility terms $V(x, \vec{\omega}_0)$ that are 0 if the ray of origin $x$ and direction $\vec{\omega}$ intersects geometry, 1 if not. Then we have

$$\mathcal{S}(A) = \frac{\psi_{s}(A)}{\psi_{s0}(A)} = \frac{1}{2 |\vec{n} \cdot \vec{k}_1|} \int_{A} V(x_m, \vec{\omega}_0) V(x_m, \vec{\omega}_0) \frac{\vec{m} \cdot (\vec{k}_1 - \vec{k}_2) R - (\vec{k}_1 + \vec{k}_2) e^{i(\vec{k}_1 - \vec{k}_2) \cdot x_m}}{\sqrt{\Theta}} dA,$$  

(17)

where we can evaluate by Monte Carlo integration using uniform sampling of the microgeometry (pdf$(x_m) = 1/A$) or sampling of the visible area (see Sec. 5). Oscillatory functions of this kind can be difficult to integrate using Monte Carlo. Aided by the computational power of modern graphics hardware, we deal with this issue by using a very large number of samples.

Extending Beckmann’s normalization and Sancer’s ray traced shadowing and masking to the transmission case, we use $T = 1, \vec{m} = \vec{n}$, and $\vec{k}_2 = \vec{k}_1$ in Eq. (14) to find $\psi_{t0}$ and obtain

$$\mathcal{T}(A) = \frac{\psi_{t}(A)}{\psi_{t0}(A)} = \frac{1}{2 |\vec{n} \cdot \vec{k}_1|} \int_{A} V(x_m, \vec{\omega}_0) V(x_m, \vec{\omega}_0) \frac{\vec{m} \cdot (\vec{k}_1 + \vec{k}_2) T e^{i(\vec{k}_1 - \vec{k}_2) \cdot x_m}}{\sqrt{\Theta}} dA.$$  

(18)

We now have separate far field expressions for $\mathcal{S}$ and $\mathcal{T}$, but a BSDF based on these would not be normalized due to the visibility terms and the separate normalization. This is not an issue as we work with in-surface scattering only and have no absorption. We can thus normalize the resulting BSDF in a post process.

The remaining challenge is to connect the ratios $\mathcal{S}$ and $\mathcal{T}$ with our measurement equation. Let us use $\mathcal{S}$ to denote $\mathcal{R}$ or $\mathcal{T}$ depending on whether the detector is observing reflected or transmitted light. Due to Poynting’s theorem, the energy transfer in electromagnetic waves is given by the absolute square of the complex amplitude [BW99, §8.4]. We thus have $|\mathcal{S}|^2 = d\Phi_{s0}/d\Phi_{s}$. This carries a connection to our measurement equation (8) because we by definition have $E = d\Phi_{s}/dA$. Then

$$f_s(x, \Omega_i, \Omega_o) \approx \frac{|\mathcal{S}(A, \Omega_i, \Omega_o)|^2}{\cos \Theta_{\Omega_i} \Omega_o} + f_{N,K}(x, \Omega_i, \Omega_o),$$  

(19)

where we compute the first term using scalar diffraction theory and the second term using path tracing (see Appendix A). It should be noted that the first term implicitly involves integration of Eqs. (17) and (18) over the solid angle bins $\Omega_i$ and $\Omega_o$.

4. Coherence and Energy Conservation

Evaluation of Eq. (19) results in a full anisotropic BSDF (where $x$ is $R$, $T$, or $S$), which we tabulate for use in rendering. For each bin of incident directions $\Omega_i$, we uniformly sample a direction of incidence $\vec{\omega}_0$ and compute the outgoing reflected field in two steps. In the first step, we compute single scattering by evaluating the first term of Eq. (19) using a uniformly sampled direction of observation $\vec{\omega}_0$, for each orthogonally projected $\Omega_i$ bin. In the second step, we compute multiple scattering (second term of Eq. (19)) by path tracing the microgeometry but including only paths with trace depth larger than 1. We bucket the exiting rays in the orthogonally projected $\Omega_i$ bins that they arrive in.

To enable concurrent progressive updates of the two terms, we represent them differently: the estimate of $\mathcal{S}$ as a complex number and $f_{N,K}$ as a real number. The complex number of each bin estimates a phasor (time-invariant representation of phase and amplitude) of the scattered wave in the far field. The real number represents radiant energy that we approximate as being incoherent due to multiple scattering. The primary difference between the two representations is that the scalar may immediately represent the reflected radiance, whereas the phasor must allow superposition of both phase and amplitude. Superpositioning of wave amplitudes is only valid as long as the incident light can be treated as coherent [GMN94; BW99]. Light is spatially coherent when all phasors on a wavefront are synchronized constituents in a wavetrain with distinct fringes in-between, where fringes are the edges between the wave amplitudes. Coherence is not a binary construct, it is directly related to the fringe intensity falloff [Hen06].

The coherence area is the extent of surface area in which we can reasonably assume that the incident light is spatially coherent. The cross sectional area of the microgeometry that we use for computing a bidirectional scattering function should thus be smaller than the coherence area but still large enough to capture the important features in the microstructure. Depending on the spatial dimensions of the considered microgeometry, this may result in a conflict. To account for the fringe intensity falloff, and following previous work [WVH17; YHW18], we use a spatial Gaussian filter with standard deviation $\sigma$ in order to consider microgeometries larger than the coherence area. We use this kernel together with the visibility terms as a factor under the integrals in Eqs. (17) and (18). This underlines the need for normalization in a postprocess.

The size of the coherence area can be estimated using the van Cittert–Zernike theorem. This relates the degree of coherence at a fixed point and a variable point illuminated by a quasi-monochromatic incoherent light source to the amplitude of a diffraction pattern centered at the fixed point [BW99]. We can use
it to obtain the coherence length: the maximum spatial distance between two phasors in the incident plane wave considered to have a correlated phase. As an example, sunlight has a worst case coherence length of $\delta_c = 50 \, \mu m$ [Hec17], which we can use to set $\sigma = \delta_c/6$ [WVJH17]. In general, the coherence length depends on the solid angle subtended by the light source and the ratio between the wavelength and the width of the emission spectrum.

In cases where significant features in the microstructure lie outside the coherence area, the simulation can be split into multiple passes. Alternatively, we can use path tracing only and neglect the wave effects. As seen in Figure 2, the overall distribution of energy due to significant features changes with the coherence area. This is an important point if we consider a so-called metasurface with microfeatures at different orders of magnitude. Figure 3 demonstrates the effect of an overestimated coherence area. The microgeometry is a plane displaced by noise with amplitude and frequency so that we would not expect diffraction effects (Figure 4). However, if too large a coherence area is assumed, we see a significant change in the spectral composition of the BRDF (see Figures 3b and 3c). This is due to diffraction effects where there should be none. The coherence length ($\delta_c$) is thus an important parameter.

Coherent light reflected (or transmitted) in the perfectly specular direction retains its coherence. We can compute this light using the path tracing approach (Sec. 3.2) or by taking the absolute square of the integrands in Eqs. (17) and (18) (each individual term in the Monte Carlo integration) when computing $\mathcal{S}$. The remaining light accounted for by the scalar diffraction theory is incoherent diffracted light [CLA02]. We can use this information to separate a computed BSDF into coherent and incoherent light, which can be valuable information in a coherence-aware renderer. The part of the incoherent diffracted light that is masked by the microgeometry is not accounted for in the multiple scattering part of our method. We can compute the magnitude of this part of the BSDF using

$$f_{\text{incoh}}(x, \Omega_i, \Omega_o) \approx \frac{|\mathcal{J}_{\text{msk}}|^2 - \Phi_{\text{msk}}(A, \Omega_o)}{\cos \Theta_o},$$

where the subscript msk is short for masked and means that $V(x_m, \hat{\omega}_o)$ is replaced by $1 - V(x_m, \hat{\omega}_o)$. If we had used the path tracing approach on its own (without the scalar diffraction theory), none of this incoherent diffracted light would have been included. In secondary bounces, we accept such an omission of wave effects, but we include the wave effects in the first bounce.

5. Sampling the Illuminated Microgeometry

As an optimization for both path tracing and scalar diffraction theory, we can sample the illuminated area only instead of uniformly sampling the full microgeometry ($pdf(x_m) = 1/A$). This requires a change of variables in the integrals over surface area. Instead of sampling the area of the microgeometry $A$ directly, we would like to sample the orthographic projection of the microgeometry into the reference plane and trace a ray toward this point from outside the microgeometry to the first point $x_m$ that it meets. Let us refer to this orthographic projection of the microgeometry as $A'$. This is usually a square or a disk. We can use sampling of this illuminated area by considering that

$$dA' = |\vec{n} \cdot \vec{n}| V(x_m, \hat{\omega}_o) dA.$$

With uniform sampling of $A'$, we then have $pdf(x_m) = 1/A'$ and the Monte Carlo estimator in Eq. (12) becomes

$$f_{x,N,K}(x, \Omega_i, \Omega_o) = \frac{WH}{N^2} \frac{A'}{A} \frac{1}{4\pi} \sum_{k=1}^{K} \sum_{j=1}^{N} \frac{|\vec{n} \cdot \hat{\omega}_o|}{|\vec{n} \cdot \hat{\omega}_o|} \left[ \hat{\omega}_{o,j,k} \in \Omega_o \right].$$

We have a similar exchange of the visibility term $V(x_m, \hat{\omega}_o)$ for $1/|\vec{n} \cdot \vec{n}|$ and integration over $A'$ instead of $A$ in Eqs. (17) and (18).

6. Results

In the following, we exemplify the use of our model and highlight the main differences in BSDFs computed for explicit microgeo-

\[Figures\ 2\ and\ 3:\ Log-\ transformed\ false\ color\ BRDF\ slices\ (color\ scale\ in\ Figure\ 5)\ for\ normal\ incidence\ with\ two\ different\ standard\ deviations\ for\ the\ Gaussian\ limiting\ the\ coherence\ area.\]

\[Figures\ 4\ and\ 5:\ Height\ map\ (left)\ of\ a\ planar\ surface\ displaced\ by\ a\ low-\ frequency\ noise\ (\text{S}_{\text{q}} = 281.8\ \text{nm})\ and\ the\ effect\ of\ filtering\ the\ surface\ by\ a\ Gaussian\ with\ \sigma = 10\ \mu m\ (right).\]
try when including indirect illumination or not and when including wave effects or not. The models being compared are:
- GS Geometric optics with Single scattering.
- GM Geometric optics with Multiple scattering.
- WS Wave optics with Single scattering.
- WM Wave optics with Multiple scattering (our model).

Our model is concerned only with the scattering in the surface of a medium, so we assume non-absorbing BSDFs and normalize all BSDFs to one. This means that energy loss in the single scattering-only models (GS and WS) becomes an incorrect distribution of the scattered light.

To visualize slices for a given direction of incidence $\vec{\omega}$ of the directional BSDF functions, we use orthogonal projections of the dependency on $\vec{\omega}$ (hemispheres) onto the disk spanned by the direction cosines $\sin \theta \cos \phi$, and $\sin \theta \sin \phi$ (the first two coordinates of $\vec{\omega}$, we only need two as the vector is of unit length). We transform spectral results to RGB space using normalized CIE RGB color matching functions. To ease visual comparison, we display spectral BSDF slices without scaling. We also integrate the spectral BSDF values to a scalar and show false color slices of these integrated values in logarithmic scale (Figure 5).

Implementation details. We implemented our BSDF generator on the GPU using OptiX [PBD*10] following the recipe in Appendix A. For the single scattering part, we computed one complex phasor per spectral sample for each ray. For the multiple scattering part, we assumed a wavelength-independent index of refraction and traced a single real scalar. Precomputation time and memory usage depend heavily on the choice of resolution. For an anisotropic BSDF, we discretize the hemisphere on 75 polar and 300 azimuthal angles (a total of 22,500 bins). We sample the visible spectrum uniformly using eight samples, which we transform into RGB for rendering. Each 2D slice of the BSDF is generated using five million samples, which we found enough even for the highly-oscillating integral in Equations (17) and (18). Each slice takes a few seconds to compute on a NVIDIA GTX 1080 TI graphics card. By exploiting reciprocity, the complete RGB anisotropic BRDF/BTDF is about 3 GB, which we do not compress.

For rendering, our model is similar to other methods that employ precomputed or measured anisotropic BRDFs. We define the microgeometry coordinates system in which the BSDF is computed, and transform the BSDF to the shading local coordinates for evaluation. While in our results we do not apply any importance sampling, this could be implemented trivially by sampling the tabulated BSDF as a discrete distribution. All renderings were computed using a path tracer implemented in OptiX. We used 35 thousand samples per pixel or more. Rendering with a BSDF took two to three hours for a resolution of 2048 × 2048. With a BSDF and 5 million samples per pixel, the rendering time was 23 to 24 hours.

Random surface. We first investigate the energy loss in single scattering on a slowly varying random surface of about 20 by 20 $\mu$m$^2$ with a root mean square height of $S_q = 132$ nm (Figure 6). The height variation was generated using sparse convolution noise [FW07; LFD*20], which we also use when adding ground noise to a modeled microgeometry. We use $S_q$ to denote the amplitude of ground noise added to a nonplanar surface, as it is then no longer the actual root mean square height of the microgeometry. Results at normal incidence for the slowly varying microsurface are in Figure 6. We observe that the energy distribution is similar between geometric and wave optics, with a small difference in the sharpness of the peak values. However, note that even for a surface with relatively low multiple scattering, our model (WM) is able to predict light that is ignored by single scattering wave optics alone (WS). We have experimented with similar surfaces of low curvature and without milli-scale features, and have found that our results align well with results in previous work [DWMG15; YHW*18]. In the next two paragraphs, we demonstrate the ability of our model to estimate the distribution of radiant energy from mixed-scale microgeometry (metasurfaces).

Reflective metasurfaces. Manufacturing processes used for metasurfaces often lead to surfaces that cannot be represented by height maps. Unfortunately, most recent scattering models assume that a given microsurface can be described as a height field [DWMG15; HP17; WVJH17; YHW*18; KHZ*19]. Since we work with explicit geometry, our work does not suffer this limitation. We demonstrate this with the OVERHANG microgeometry shown in Figure 1.
further study this microgeometry by adding ground noise [FW07; al. [HFM16], the effect on the final appearance is significant. We scene for perceptual evaluation of BRDFs suggested by Havran et
which compares the appearance rendered by each model using the LFD*20] to the base geometry. The results are in Figures 8 and
each model (GS, GM, WS and our WM). As also seen in Figure7,
shadowing/masking and multiple scattering dominate). nanoscale (where diffraction dominates) and microscale (where
9 and show that our model is able to deal with features at both
crossscale are represented in the simulated BSDF, as one can also see in Figures11b and 11e. Adding noise to the hemisphere microstructure,
we can investigate the effect of increased roughness. Despite the fact that added noise results in more features in the WS result,
the energy loss due to masking and shadowing is still dominant on the energy distribution. In contrast, our model (WM) includes the
sharp coloured features from the WS result and avoids the energy loss through use of multiple scattering.

Refractive metasurface. We now demonstrate our model for a dielectric surface. To the best of our knowledge, no prior work in graphics has demonstrated BSDFs based on scalar diffraction theory. We build a dielectric RIDGED metasurface (Figure 12, top),
that creates an apparent double refraction, and acts as a diffraction grating which creates visible colored patterns. Figure 13 shows a
dielectric disk with RIDGED microgeometry and index of refraction based on BK7 glass [Sch17]. The disk is slanted above the ground
floor while intersecting it at the bottom. This figure clearly shows the visible color shift on transmission predicted by scalar diffraction theory (bottom), as well as the importance of multiple scattering for energy conservation, especially in the case of our WM.

Comparison against a rigorous solution. In shaping of beams, a diffraction grating with so-called elliptic axicons is of particular interest [TJF03]. This is a surface with elongated half ellipsoids arranged in a regular pattern. We use this kind of microgeometry
as an example for comparison of our method with a rigorous wave solver. Specifically, we computed the BSDF of the surface illustrated in Figure 12 (bottom) using a finite difference time domain (FDTD) solver by Lumerical. Comparison of our results with the rigorous solution are in Figure 14. While the rigorous solver is in most cases most accurate, it has too low resolution in its simulation volume to capture the specific shaping of the beam that is the key characteristic of an elliptic axicon diffraction grating. We clearly capture this effect as illustrated in Figure 15. As we seem to be the first in graphics to use scalar diffraction theory for transmission, we find it comforting that the transmission outward through the elliptic axicon surface matches the rigorous solution rather well. We also find it interesting that the characteristic caustic is only captured if secondary bounces are included in the model.

Figure 9: BRDF slices for the overhang microsurface with added ground noise of different amplitudes: (a–f) $S^a_\nu = 1.283\, nm$, and (g–l) $S^a_\nu = 13.82\, nm$. The surface is illuminated at $\theta_i = 55^\circ$. We display spectral BRDF slices transformed to RGB (a–c and g–i) and integrated across the spectrum (d–f and j–l). The importance of multiple scattering is obvious for this surface (compare the middle column with the other two). The need for wave optics effects is discernible around the specular peak.

Figure 10: Close-ups of the specular peak to the right in Figure 9(h–i) for three different angles of incidence ($\theta_i$). Scalar diffraction models (WS) exhibit problems in the spectral composition when not accounting for shadowing and masking. We use ray tracing to account for shadowing and masking in our single scattering component ($f^1_s$), and we therefore do not experience this problem.

7. Conclusion

We have presented a model for computing the bidirectional scattering properties of a surface with a known microstructure. Our model is not limited to height maps. The microstructure can be chosen arbitrarily. We combine multiple scattering from geometric optics with single scattering from scalar diffraction theory to obtain improved energy distribution in the computed bidirectional functions. We presented a practical technique for the computation of a BSDF using path tracing for the geometric optics part and Monte Carlo integration for the wave optics part. As opposed to previous work, our single scattering contribution from wave optics is based on the Kirchhoff approximation but includes ray traced shadowing and masking. This approach was the key for us to avoid a height field assumption and to include the transmission mode. Our results show a clear advantage in terms of estimating a plausible energy distribution in reflected and transmitted light while we can also faithfully capture diffraction colors and brightening due to multiple scattering in the surface microgeometry.

Limitations and future work. The main limitation of our work is that we neglect diffractive interference in near-field multiple scattering. As a part of this limitation, the incoherent diffracted light that does not escape the microgeometry is not included in our multiple scattering. Whether this is a reasonable assumption can be assessed using the measure we provide for evaluating the magnitude of the problem (Eq. 20). Our current implementation is based on brute-force Monte Carlo integration for computing the full BSDF. This is likely suboptimal for the oscillatory integrals from scalar diffraction theory and could be improved by leveraging integration in the Fourier domain. Finally, since we precompute the BSDF, it is not trivial to support spatially varying BSDFs. Finding a suitable
Figure 11: BRDF slices for the HEMISPHERES microsurface at $\theta_i = 55^\circ$ with no added noise (a–f) and with added ground noise of $S_g^* = 13.57$ nm (g–l). We display BRDF slices in true color (a–c and g–i) and in false color with logarithmic scale (d–f and j–l).

Figure 12: Top: RIDGED microgeometry used to generate the surface BSDF of the dielectric in Figure 13. Bottom: AXICON microgeometry with regularly arranged elongated half ellipsoids of radii 0.4, 0.45, and 0.3 $\mu$m, used for the comparison of our theory with a rigorous wave solver in Figures 14 and 15.

Figure 13: A planar dielectric disk with RIDGED microgeometry (Figure 12, top). For comparison, we show geometric optics in the top row with single scattering (GS, top-left) and multiple scattering (GM, top-right), single scattering wave optics (WS, bottom-left), and our method (WM, top-right). Insets are BRDF and BTDF slices (left and right, respectively) at normal incidence.

Figure 14: Reference BSDF slices (top row) for the AXICON surface in Figure 12 (index of refraction is 2) computed using Lumerical FDTD for normally incident light of 850 nm linearly polarized in the vertical direction. Results are displayed using a logarithmic color scale from $10^{-7}$ to $10^{-2.5}$. As expected, we do not have a perfect match, but our results (middle row) are significantly closer than a geometric optics approach (bottom row). From left to right: outward transmission, inward reflection, inward transmission, and outward reflection. Some deviation is due to the fact that we assume unpolarized light. We may conjecture that the inward reflection is off because we do not account for diffraction in secondary bounces.
analytical model or basis function representation for our computed BSDFs would allow their use on spatially varying surfaces.

Acknowledgments. This work has been funded by Innovation Fund Denmark (5163-00001B), the European Research Council (ERC) under the EU’s Horizon 2020 research and innovation programme (project REVEAL, HR0011-16-C-0025), and the Spanish Ministry of Economy and Competitiveness (TIN2016-78753-P and PID2019-105004GB-I00).

References


© 2020 The Author(s)
Computer Graphics © 2020 The Eurographics Association and John Wiley & Sons Ltd.
Appendix A: Algorithmic description of our method.

Initialization

- Find the radius \( r \) of the bounding sphere of the microgeometry.
- Allocate a complex number and a real number for each \( (\Omega_x, \Omega_y, \Omega_z) \) bin and set them to zero.

Main computation (repeat until the desired spectrum and the desired resolution of \( \Omega_z \) bins have been covered).

- Sample a wavelength \( \lambda \) and a direction of incidence \( \omega_i \) and find the associated spectral bin and \( \Omega_z \) bin.
- Compute a BSDF slice for the sampled \( \omega_i \) and \( \Omega_z \), using progressive updates (see below) until the result is satisfactory.

Progressive update of a BSDF slice for given \( \lambda \) and \( \omega_i \)

- For each \( \Omega_z \) bin, sample a point \( x_i \) in the projected visible area \( A' \) modulated by the coherence area Gaussian. Evaluate \( B \) or \( \mathcal{F} \) and update the complex number of the \( \Omega_z \) bin accordingly.
- Allocate an accumulation buffer with a scalar value initialized to zero for each \( \Omega_z \) bin. This is for bucketing of path tracing results.
- Repeat \( W \times H \) times: Sample a point \( x_i \) in the projected visible area \( A' \). Trace a path from \( x_i + \delta \) in the direction \( -\omega_i \). When the path exits the microgeometry with direction \( \omega_i \), find the corresponding \( \Omega_z \) bin and bucket a term from the sum of Eq. (12) in the accumulation buffer.
- Combine direct and indirect illumination using Eq. (19) and normalize the BSDF slice.