A Comprehensive Physical Model for Light Reflection

Xiao Dong He, Kenneth E. Torrance, François X. Sillion, Donald P. Greenberg

A Comprehensive Physical Model
for Light Reflection

Xiao D. He
Kenneth E. Torrance
François X. Sillion
Donald P. Greenberg

Program of Computer Graphics
Cornell University
Ithaca, NY 14853

Abstract

A new general reflectance model for computer graphics is presented. The model is based on physical optics and describes specular, directional diffuse, and uniform diffuse reflection by a surface. The reflected light pattern depends on wavelength, incidence angle, two surface roughness parameters, and surface refractive index. The formulation is self consistent in terms of polarization, surface roughness, masking/shadowing, and energy. The model applies to a wide range of materials and surface finishes and provides a smooth transition from diffuse-like to specular reflection as the wavelength and incidence angle are increased or the surface roughness is decreased. The model is analytic and suitable for Computer Graphics applications. Predicted reflectance distributions compare favorably with experiment. The model is applied to metallic, nonmetallic, and plastic materials, with smooth and rough surfaces.

CR Categories and Subject Descriptors: I.3.7—[Computer Graphics]: Three-Dimensional Graphics and Realism; I.3.3—[Computer Graphics]: Picture/Image Generation; J.2—[Physical Sciences and Engineering]: Physics.

Additional Key Words and Phrases: reflectance model, specular and diffuse reflection, comparison with experiment.

1 Introduction

Photorealistic image generation is an active research area in Computer Graphics. Ray-tracing and Radiosity have been developed to obtain realistic images for specular and diffuse environments, respectively. However, applications of these methods to general environments have been hindered by the lack of a broadly-applicable local light reflection model. To obtain a true global illumination solution of a general environment, a physically based reflection model of general applicability is needed.

A comprehensive light reflection model is presented in this paper. The model compares favorably with experiment and describes specular, directional diffuse, uniform diffuse and combined types of

reflection behavior. The model is analytic and provides a smooth transition from specular to diffuse-like behavior as a function of wavelength, incidence angle and surface roughness.

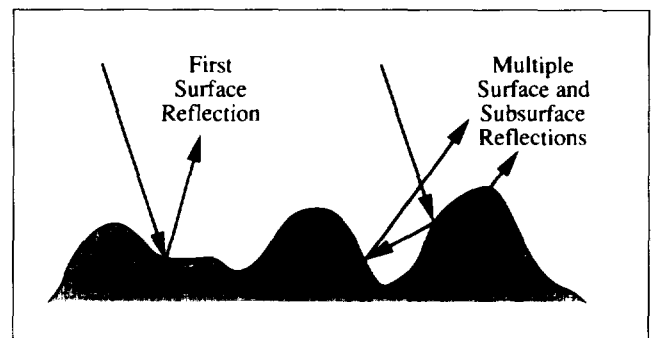As illustrated in Figure 1, we classify the reflection process from

Figure 1: Reflection processes at a surface.

an arbitrary surface as consisting of first-surface reflections and multiple surface and/or subsurface reflections. The first-surface reflection process is described by physical optics and is strongly directional. As the surface becomes smooth this part evolves toward specular or mirror-like behavior. As the surface becomes rough, a diffuse-like behavior due to diffraction and interference effects becomes more important and, at larger roughnesses, it controls the directional distribution of the first-surface reflected light. The model partitions energy into specular and diffuse-like components according to the roughness of the surface. The multiple surface and subsurface reflections sketched in Figure 1 are geometrically complex, but may be expected to be less strongly directional than the first-surface reflected light. Hence, they are approximated as uniform diffuse. Our model leads to analytic expressions suitable for the full range of surface roughnesses and thus is useful for implementation in computer graphics.

The present model builds on, and extends, existing models from optics [3] [5]. It allows for polarization and masking/shadowing effects. The model extends the geometric optics model of Cook [8] to the physical optics region, and correctly includes specular reflection as the surface roughness is decreased. The model is physically based in contrast to empirical approaches [13].

The following sections provide a conceptual introduction, the model, a comparison with physical experiments, and example implementations. The mathematical derivation of the model appears in Appendix A. For unpolarized incident light, the reflectance model is summarized in Appendix B.

A	=	projected area of the surface (Figure 5)
$BRDF$	=	bidirectional reflectance distribution function
$C(r)$	=	correlation coefficient, equation (48)
c	=	complex coefficient of polarization state
D	=	distribution function, equation (78)
$E_s, \vec{E}_s$	=	scalar and vector electric fields
F	=	Fresnel reflection coefficient, equation (44)
$|F|^2$	=	Fresnel reflectivity
$\bar{F}$	=	Fresnel matrix, equation (44)
G	=	geometrical factor, equation (76)
G'	=	Green's function, equation (2)
g	=	surface roughness function, equation (9)
I	=	intensity
$\bar{I}$	=	unit tensor
i	=	unit imaginary number, i.e., $i = \sqrt{-1}$
k	=	wave number, i.e., $k = 2\pi/\lambda$
$\vec{k}$	=	wave vector
$\hat{k}$	=	unit vector in wave direction
L	=	length
L_x, L_y	=	length dimensions of the surface
m	=	summation index
$\bar{n}$	=	refractive index
$\hat{n}$	=	local surface normal, unit vector
$\hat{n}_b$	=	bisecting unit vector, equation (51)
$\mathbf{p}$	=	incident polarization state vector, equation (34)
$p(z)$	=	Gaussian distribution function, equation (3)
R	=	distance from origin to field point
$\vec{R}$	=	positional vector to field point
$\vec{r}$	=	positional vector of a surface point
S	=	shadowing function, equation (23)
$\hat{s}, \hat{p}$	=	s and p polarization unit vectors
$\bar{T}$	=	transformation matrix, equation (39)
$\vec{v}$	=	wave vector change, equation (20)

v_{xy}	=	$\sqrt{v_x^2 + v_y^2}$
$\hat{x}, \hat{y}, \hat{z}$	=	unit vectors in Cartesian coordinates
z	=	surface height
Γ	=	area of bounding surface, Figure 2
Δ	=	delta function
$\vec{\eta}$	=	horizontal distance vector, equation (28)
θ, ϕ	=	polar and azimuthal angles (Figure 5)
λ	=	wavelength
$\xi(x, y)$	=	Gaussian distributed random function
ρ_{bd}	=	bidirectional reflectivity, equation (4)
ρ_{dh}	=	directional-hemispherical reflectivity
ρ_{hd}	=	hemispherical-directional reflectivity
σ^2	=	apparent variance of $z = \xi(x, y)$
σ_0^2	=	variance of $z = \xi(x, y)$
τ	=	autocorrelation length, equation (48)
ω	=	solid angle

Subscripts

a	=	ambient
b	=	bisecting
bd	=	bidirectional
dd	=	directional-diffuse
i	=	incident
p	=	p polarization
r	=	reflected
s	=	s polarization
sp	=	specular
ud	=	uniform-diffuse
x, y, z	=	Cartesian coordinates
$1, 2$	=	surface points

Superscripts

n	=	local plane
$*$	=	complex conjugate

Table 1: Nomenclature

2 Theory of light reflection

This section introduces the principal techniques often used to analyze the reflection of an electromagnetic wave by a general surface [3] [5]. The improved model presented later in this paper uses all of these techniques.

2.1 Kirchhoff theory

Consider the geometry sketched in Figure 2. According to classical electromagnetic theory, the scalar electromagnetic field $E(\vec{R})$ at an arbitrary point in space can be expressed as a function of the scalar field E_s and its normal derivative $\partial E_s/\partial n$ on any enclosing surface Γ. The governing equation is [5]

$$E(\vec{R}) = \frac{1}{4\pi} \int_\Gamma \left(E_s(\vec{r}) \frac{\partial G'(\vec{R}, \vec{r})}{\partial n} - G'(\vec{R}, \vec{r}) \frac{\partial E_s(\vec{r})}{\partial n} \right) d\Gamma \quad (1)$$

where G' is the free space Green's function given by [12]

$$G'(\vec{R}, \vec{r}) = \frac{e^{ik|\vec{R} - \vec{r}|}}{|\vec{R} - \vec{r}|} \quad (2)$$

Equation (1) is an integral representation of the wave equation and is known as the Kirchhoff integral of scalar diffraction theory.

For a single reflecting surface, the domain of integration Γ reduces to the area of the reflecting surface. This has allowed a class

of surface reflection models, known as "physical or wave optics" models, to be derived [5]. "Physical optics" uses a complete physical or wave description of the reflection process, thus allowing for diffraction and interference effects. Wave effects must be included if a reflection model is to describe both specular and diffuse-like reflection from a surface.
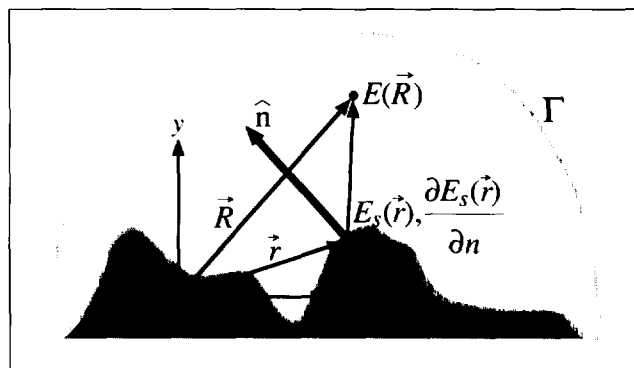

Figure 2: Geometry for application of the Kirchhoff integral. $\hat{n}$ is the local surface normal.

2.2 Tangent plane approximation

For reflection processes, the Kirchhoff formulation reduces the general problem of computing the field everywhere in space to the simpler one of determining the field on the reflecting surface. However, even this is a complex task, and the so-called "tangent plane approximation" is often used. This is done by setting the value of the field at a given point on the surface to be the value that would exist if the surface were replaced by its local tangent plane. This is sketched in Figure 3 where E_i and E_s are the incident and scattered fields, respectively, and $F(\theta)$ is the local Fresnel (electric field) reflection coefficient. The approximation is valid when the local radius of curvature of the surface is large compared to the wavelength. The reflected field depends on the Fresnel reflection coefficients for horizontal and vertical polarizations, as well as on the local slope and position of the reflecting point.
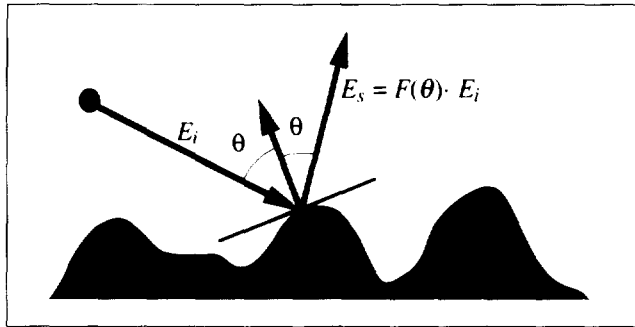

Figure 3: Tangent plane approximation for a reflecting surface. The statistical parameters σ and τ for the surface are indicated schematically.

2.3 Statistical surfaces

The complete geometrical specification of a reflecting surface is rarely known, but information at length scales comparable to the radiation wavelength is required when the Kirchhoff theory is used. However, small scale variations of the electromagnetic field on the surface are averaged out when viewed from a distance. This averaging over points on a surface is statistically equivalent to averaging over an entire class of surfaces with the same statistical description. Interesting quantities, such as the reflected intensity in a given direction, can then be obtained by a weighted average of the Kirchhoff integral.

Frequently, the height distribution on a surface (Figure 3) is assumed to be Gaussian and spatially isotropic. Under such conditions, the probability that a surface point falls in the height range z to $z + dz$ is given by $p(z)dz$, with a probability distribution

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(z^2/2\sigma_0^2)} \qquad (3)$$

A mean value of $z = 0$ is assumed and σ_0 is the rms roughness of the surface. To fully specify an isotropic surface a horizontal length measure is also needed. One such measure is the autocorrelation length τ (defined in equation (48)), which is a measure of the spacing between surface peaks. The rms slope of the surface is proportional to σ_0/τ.

2.4 Shadowing and masking

The effect of self-shadowing and self-masking by a rough surface (Figure 4) was introduced in computer graphics by Blinn [6]

and Cook [8]. This effect manifests itself at large angles of incidence or reflection, where parts of the surface are shadowed and/or masked by other parts, reducing the amount of reflection. Beckmann [4] argued that to first order, the effect of shadow-
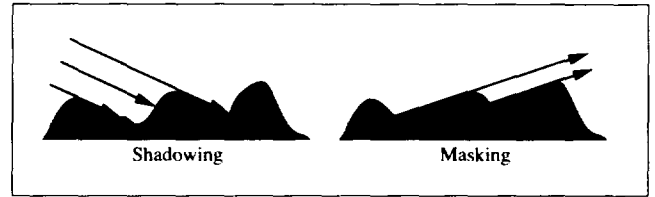

Figure 4: Shadowing and masking.

ing/masking can be obtained by using a multiplicative factor which accounts for the fraction of the surface that is visible both to the source and the receiver. Such a concept was used by both Blinn and Cook in their geometrical optics approaches, but the V-groove shadowing/masking factor they used [20] is first-derivative discontinuous. Many other shadowing/masking factors have appeared in the literature. Of these, the one due to Smith [16] is continuous in all derivatives and has been found to agree with statistical numerical simulations of a Gaussian rough surface [7].

2.5 Discussion

An early comprehensive model of light reflection from a rough surface, using physical optics, was introduced by Beckmann [5]. Beckmann applied the scalar form of the Kirchhoff theory, used the tangent plane approximation, and performed a statistical average over the distribution of heights to get the reflected intensity. The Beckmann distribution function was used by Blinn and Cook for their computer graphics applications.

Stogryn applied a more general, vector form of the Kirchhoff theory, thus taking polarization effects and the correct dependency of the Fresnel reflectivity into account [18]. Furthermore, he used a more complete statistical averaging scheme that averages over both height and slope. However, shadowing/masking was not considered, and the derivation of the reflected intensity was limited to special cases of incident polarization. A more general model, which accounts for polarization, Fresnel, and shadowing/masking effects, has been described by Bahar [1] [2]. However, it is difficult to implement because it relies on the solution of a set of coupled integro-differential equations.

Finally, it should be noted that these models were very rarely compared with experimental results.

3 An improved model

This section presents an improved light reflection model of broad applicability. Section 3.1 summarizes the techniques and key assumptions; Section 3.2 presents the improved model. Details of the mathematical derivation appear in Appendix A and a full set of equations for unpolarized incident light in Appendix B.

3.1 Techniques and key assumptions

To develop a general reflection model which avoids many of the limitations of previous models, the overall formulation of Beckmann was used, but with the following improvements:

- The vector form of the Kirchhoff diffraction theory is used. This allows, for the first time, a complete treatment of polarization and directional Fresnel effects to be included. Such

effects are required for a comprehensive formulation. The model permits abitrary incident polarization states (e.g., plane, circular, unpolarized, partially polarized, etc.) and includes effects like depolarization and cross-polarization.

- The surface averaging scheme of Stogryn [18] is employed with its improved representation of the effects of surface height and slope. Averaging of the Kirchhoff integral is over a four-fold joint probability function (i.e., height, slope, and two spatial points).

- The scheme of Stogryn [18] is extended to average only over the *illuminated* (unshadowed/unmasked) parts of the surface. This requires a modified probability function with an *effective roughness*, σ, given by equation (53). When roughness valleys are shadowed/masked (Figure 4), the effective surface roughness can be significantly smaller than the rms roughness, σ_0, especially at grazing angles of incidence or reflection. For the first time, the concept of an effective roughness, which depends on the angles of illumination and reflection, is applied.

- The geometrical shadowing/masking factor of Smith [16] is introduced as a multiplicative factor. The function has appropriate smoothness and symmetry.

With the above, the model leads to a fairly-complex integral formulation. Simplifications result by making the local tangent-plane approximation and assuming gentle roughness slopes. These assumptions should be realistic for many surfaces over a wide range of radiation wavelengths. Significantly, the assumptions lead to an analytical form for the light reflection model.

3.2 The improved light-reflection model

The light reflection model is presented in terms of the bidirectional reflectivity ρ_{bd}, also called the bidirectional reflectance distribution function (BRDF). The coordinates are shown in Figure 5, together with the propagation unit vectors $(\hat{k}_i, \hat{k}_r)$ and the polarization unit vectors $(\hat{s}, \hat{p})$ for the polarization components perpendicular $(\hat{s})$ and parallel $(\hat{p})$ to the incident and reflecting planes (i.e., the $(\hat{k}, \hat{z})$ planes). The total BRDF is defined as the ratio of the total reflected
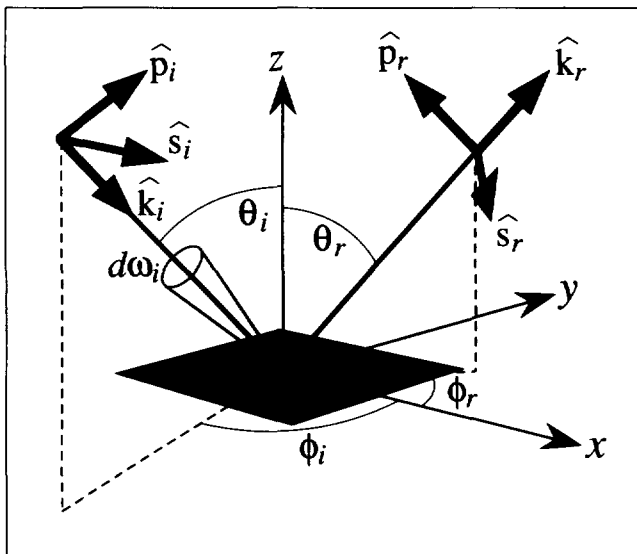
Figure 5: Coordinates of illumination and reflection.

intensity (i.e., the sum of reflected s and p intensities) in the direction (θ_r, ϕ_r) to the energy incident per unit time and per unit area onto the surface from the direction (θ_i, ϕ_i) [14]. The incident energy flux may be expressed in terms of the incident intensity I_i and the incident solid angle $d\omega_i$:

$$\rho_{bd}(\theta_r, \phi_r, \theta_i, \phi_i) = \frac{dI_r(\theta_r, \phi_r; \theta_i, \phi_i)}{I_i(\theta_i, \phi_i) \cos\theta_i d\omega_i} \qquad (4)$$

The BRDF may also be defined for each polarization component of the reflected intensity (see Appendix A). Equation (4) gives the frequently-used total BRDF.

We propose a bidirectional reflectivity consisting of three components:

$$\rho_{bd} = \rho_{bd,sp} + \rho_{bd,dd} + \rho_{bd,ud} \qquad (5)$$

The additional subscripts correspond to specular (sp), directional-diffuse (dd), and uniform-diffuse (ud) reflection. The first two components in (5) result from the first-surface reflection process (see Figure 1) and are respectively due to specular reflection by the mean surface and diffraction scattering by the surface roughness. The third component, taken as uniform diffuse, is attributed to multiple surface and/or subsurface reflections.

An example of a light intensity distribution corresponding to equation (5) is shown in Figure 6. A general reflecting surface is
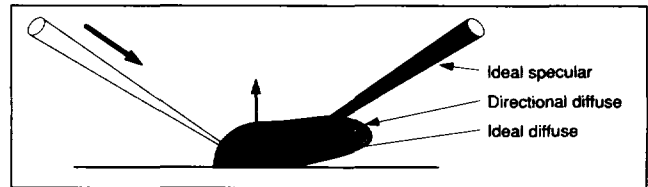
Figure 6: Example of a light intensity distribution.

assumed, with some specular reflection, some diffraction scattering due to roughness, and some multiple or subsurface scattering. The specularly-reflected part is contained within the specular cone of reflection. The diffraction-scattered part shows a directional distribution which is far from ideal diffuse. The last part is uniform diffuse (Lambertian).

An analytic form for the first two terms in (5) is derived in Appendix A. With the local-tangent-plane and gentle-slope assumptions for the first-surface reflection process, and for arbitrary incident polarization, we have:

$$\rho_{bd,sp} = \frac{\rho_s}{\cos\theta_i d\omega_i} \cdot \Delta = \frac{|F|^2 \cdot e^{-g} \cdot S}{\cos\theta_i d\omega_i} \cdot \Delta \qquad (6)$$

$$\rho_{bd,dd} = \frac{\mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p}) \cdot S}{\cos\theta_i \cdot \cos\theta_r} \cdot \frac{\tau^2}{16\pi} \cdot \sum_{m=1}^{\infty} \frac{g^m e^{-g}}{m! \cdot m} \cdot \exp(-\frac{v_{xy}^2 \tau^2}{4m}) \qquad (7)$$

$$\rho_{bd,ud} = a(\lambda) \qquad (8)$$

where ρ_s is the specular reflectivity of the surface, Δ is a delta function which is unity in the specular cone of reflection and zero otherwise, $|F|^2$ is the Fresnel reflectivity which depends on the index of refraction $(\bar{n}(\lambda))$ of the surface material [14, p.100], g is a function of the effective surface roughness given by

$$g = [(2\pi\sigma/\lambda)(\cos\theta_i + \cos\theta_r)]^2, \qquad (9)$$

S is the shadowing function (see equation (23)), $\mathcal{F}$ is a function involving the Fresnel reflection coefficients (see equations (68) and

(59), (60)), p is the polarization state vector of the incident light (see equation (34)), v_{xy} is a function which depends on the illumination and reflection angles (see equation (20)), and $a(\lambda)$ is a parameter to be discussed later.

For convenience and for the special case of incident unpolarized light, the governing equations are gathered together and presented in Appendix B. The directional-diffuse term in this appendix (equation (71)) uses nomenclature to permit comparison with the geometric optics model of Cook-Torrance [8].

The physical basis of the three reflection components in (5) is discussed in the following subsections. Before proceeding, we note that the dependence of the specular component on $d\omega_i$ drops out if equation (5) is converted to an intensity basis by multiplying by $I_i \cos\theta_i d\omega_i$. From (6), the specular term becomes $\rho_s I_i \Delta$, which is the well known form used in Ray-tracing. The specular intensity is then independent of $d\omega_i$, but the directional-diffuse and uniform-diffuse intensities are proportional to $d\omega_i$.

3.2.1 Specular contribution: $\rho_{bd,sp}$

The specular term accounts for mirror-like reflection from the *mean* plane of the reflecting surface. The term is proportional to the Fresnel or mirror reflectivity, $|F|^2$. For rough surfaces, the specular term is reduced by the roughness and shadowing factors e^{-g} and S, respectively.

For a smooth surface, as the wavelength of the incident light becomes large relative to the projected surface roughness, i.e., $\lambda \gg \sigma\cos\theta_i$, the specular term is not attenuated since $g \rightarrow 0$ and $S \rightarrow 1$. Also in this limit, the specular component dominates the first-surface reflection process, since the contribution from equation (7) diminishes as $g \rightarrow 0$. For smooth surfaces, equation (6) reduces to

$$|F|^2/\cos\theta_i d\omega_i, \qquad (10)$$

which is the usual form of the bidirectional reflectivity for a specular surface.

3.2.2 Directional diffuse contribution: $\rho_{bd,dd}$

When the wavelength of the incident light is comparable to or smaller than the projected size of surface roughness elements (i.e., $\lambda \sim \sigma\cos\theta_i$), the first-surface reflection process introduces diffraction and interference effects. The reflected field is spread out to the hemisphere above the reflecting surface. We call this *directional diffuse*, to indicate that the field is diffused to the hemisphere but may have a directional, nonuniform character.

The reflected light pattern given by equation (7) depends on surface statistics through the effective roughness σ and the autocorrelation length τ. For smooth surfaces, as σ/λ or g approach zero, the bidirectional reflectivity given by equation (7) diminishes to zero. For rough surfaces, with σ/λ or g large, equation (7) describes the directional distribution of the first-surface reflected light. The reflected pattern can be complex with maximal values in the specular direction for slightly rough surfaces, at off-specular angles for intermediate roughnesses, or at grazing reflection angles for very rough surfaces.

3.2.3 Uniform diffuse contribution: $\rho_{bd,ud}$

The light reflected by multiple surface reflections or by subsurface reflections is generally more difficult to describe analytically than light reflected by the first-surface reflection process. This contribution is small for metallic (opaque) surfaces with shallow roughness slopes. However, the contribution can be important for surfaces with large slopes, or for nonmetals if significant radiation crosses the

first surface and is reflected by subsurface scattering centers (e.g., paints, ceramics, plastics).

Estimates of the multiple-reflection process within surface V-grooves, based on geometrical optics, have been carried out [10] [17]. Also, estimates of the subsurface scattering are available [14]. The analytical results often suggest that the reflected field due to these two processes may be approximated as nearly directionally uniform. Therefore, the multiply-reflected and/or subsurface scattered light is approximated as *uniform-diffuse* (i.e., Lambertian), and we denote it by $a(\lambda)$.

The coefficient $a(\lambda)$ can be estimated theoretically if the V-groove geometry is applicable, or if the subsurface scattering parameters are known. Alternatively, $a(\lambda)$ can be estimated experimentally if equation (5) is integrated over the reflecting hemisphere, and the results are compared with measured values of the directional-hemispherical reflectivity, ρ_{dh}. This reflectivity is equal to the hemispherical-directional reflectivity ρ_{hd} (for the case of uniform incident intensity [14]), and which can be easily measured using an integrating sphere reflectometer. For the present paper, in the absence of additional surface or subsurface scattering parameters, or experimental measurements, we will treat $a(\lambda)$ as a constrained, but otherwise free, parameter. The constraint is based on energy conservation and gives an upper bound for $a(\lambda)$.

3.3 Discussion

The theoretical model described by equation (5) allows specular, directional-diffuse, and uniform-diffuse reflection behavior as sketched in Figure 6. The governing equations in general form are given in equations (5) to (8) and Appendix A, or for unpolarized incident light in Appendix B. The actual reflection patterns depend on wavelength, incidence angle, surface roughness and subsurface parameters, and index of refraction. The model provides a unified approach for a wide range of materials and surface finishes, and is in a form suitable for use in computer graphics.

4 Comparison with experiments

In this section we compare the reflection model with experimental measurements. Appropriate comparison experiments appear only infrequently in the literature, since well-characterized surfaces as well as good wavelength and directional resolution are required. The measurements selected for comparison consist of BRDF's for roughened aluminum [19], roughened magnesium oxide ceramic [19], sandpaper [9], and smooth plastic [11]. The comparisons cover a wide range of materials (metallic, nonmetallic) and reflection behavior (specular, directional diffuse, uniform diffuse).

Polar comparisons are presented in Figures 7 to 10. Results are shown in the plane of incidence; the polar angle is θ_r and the curve parameter is the angle of incidence θ_i. Theoretical predictions are shown with solid lines and experimental measurements with dashed lines. The polar radius is the BRDF normalized with respect to the specular reflecting ray direction, i.e.,

$$\frac{\rho_{bd}(\theta_i, 0; \theta_r, \phi_r)}{\rho_{bd}(\theta_i, 0; \theta_i, 0)} \qquad (11)$$

Results for an aluminum surface (very pure; measured roughness: $\sigma_0 = 0.28\mu m$) are shown in Figures 7 and 8, respectively, for wavelengths of $\lambda = 2.0\mu m$ and $0.5\mu m$. These figures illustrate the effects of wavelength and incidence angle. The autocorrelation length and measured hemispherical reflectances were not reported. Therefore, values of $\tau = 1.77\mu m$ and $a(\lambda) = 0$ were selected as best fits at both wavelengths. Several points can be noted.
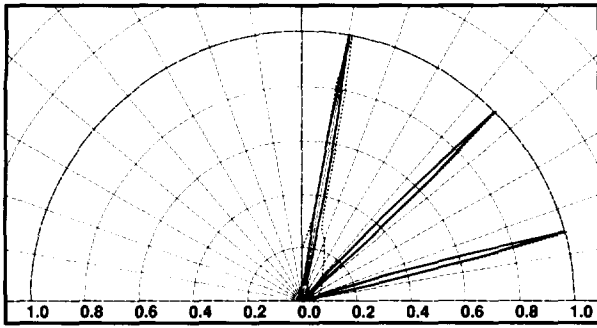
Figure 7: Normalized BRDF's of roughened aluminum as obtained from theory (solid lines) and experiment (dashed lines) for incidence angles of θ_i = 10°, 45°, and 75°. λ = 2.0μm. This is the same surface as in Figure 8. The surface shows strong specular reflection at this wavelength.
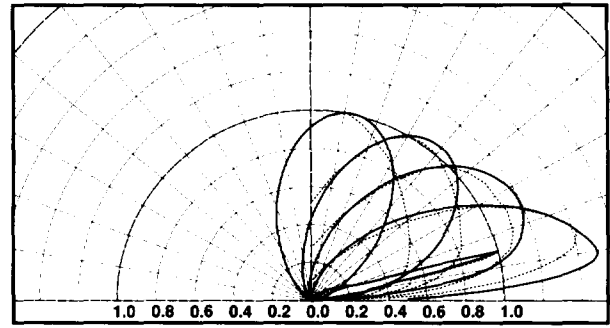

Figure 8: Normalized BRDF's of roughened aluminum as obtained from theory (solid lines) and experiment (dashed lines) for incidence angles of θ_i = 10°, 30°, 45°, 60°, and 75°. λ = 0.5μm. This is the same surface as in Figure 7. The surface shows strong directional diffuse and emerging specular reflection at this wavelength.
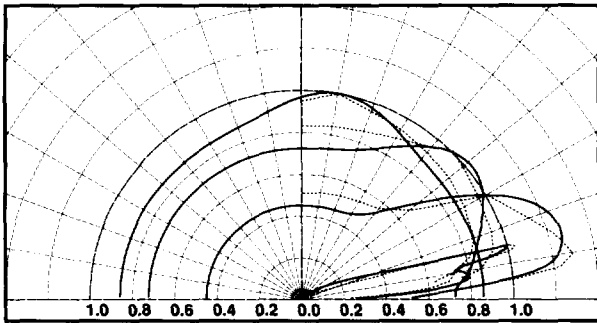

Figure 9: Normalized BRDF's of roughened magnesium oxide ceramic as obtained from theory (solid lines) and experiment (dashed lines) for incidence angles of θ_i = 10°, 45°, 60°, and 75°. λ = 0.5μm. The surface shows strong uniform diffuse and emerging specular reflection.

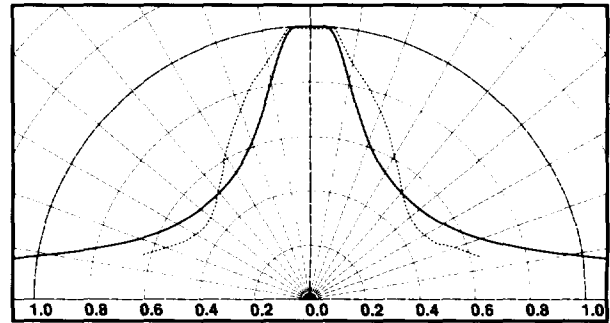

Figure 10: Normalized BRDF's of sandpaper as obtained from theory (solid lines) and experiment (dashed lines) for normal incidence, θ_i = 0°. λ = 0.5μm. The surface shows a large reflectance at grazing reflection angles.

When σ_0 is small compared to λ, as in Figure 7, strong specular reflection occurs. The angular width of the measured specular peak is determined by the solid angles of incident and received light in the experiments ($d\omega_i = d\omega_r = \pi/1024$). To allow comparisons, the theoretical peaks have been averaged over the same solid angles. For incidence at θ_i = 10°, the reflected pattern displays both specular and directional diffuse components. In Figure 8, when the roughness is more comparable to the wavelength, a strong directional diffuse pattern appears, and for θ_i = 10°, 30°, 45°, and 60°, the reflected intensity is maximal at larger-than-specular angles. For θ_i = 75°, a specular peak emerges as the surface appears somewhat smoother to the incident radiation.

A comparison with a magnesium oxide ceramic (very pure; measured roughness: σ_0 = 1.90μm, but model best fit σ_0 = 1.45μm) at λ = 0.5μm is displayed in Figure 9. This surface shows nearly uniform diffuse behavior at θ_i = 10° and an emerging specular peak for larger values of θ_i. The model employed best-fit parameters of τ = 13.2μm and $a(\lambda)$ = 0.9, the latter expressing the relatively stronger role of subsurface scattering as compared to the aluminum surface. Significantly, the experimental and theoretical trends in Figures 7 to 9 for both the metal and the nonmetal are in qualitative accord. Importantly, both materials display an emerging specular peak as the angle of incidence is increased, and, for the metal, as the wavelength is increased. Further, the metal shows a strong directional diffuse pattern, and the nonmetal a strong uniform diffuse

pattern, both of which are in accord with the model.

A dramatically different reflection pattern is displayed in Figure 10, corresponding to 220 grit sandpaper at θ_i = 0° and λ = 0.55μm. Parameters used for the comparison are σ_0/τ = 4.4 and $a(\lambda)$ = 0. For very rough surfaces, only the ratio σ_0/τ is required, not σ_0 and τ separately [5]. Although the large ratio of σ_0/τ challenges the gentle slope assumption of the model, the agreement between experiment and theory is striking as both display large reflected intensities at grazing angles of reflection.

A comparison of experiment and theory in terms of absolute BRDF's is shown in semilog form in Figure 11 for a smooth blue plastic at λ = 0.46μm. The shape of the specular spikes is determined by the geometry of the incident and receiving optical systems. The distributions for four incidence angles reveal a linear combination of specular and uniform diffuse behavior. This is consistent with the model (equations (5) to (8)). For a smooth surface with σ_0 = 0, the directional-diffuse term drops out and the specular term reduces to equation (10). The directional-hemispherical reflectivity at θ_i = 0° and λ = 0.46μm was measured (ρ_{dh} = 0.195) and yields the value $a(\lambda)$ = 0.15 used for the uniform diffuse term in the model. The agreement between experiment and theory in Figure 11 in terms of shape and absolute magnitude is encouraging.

In conclusion, the experimentally-measured directional distributions in Figures 7 to 11 show a wide range of behavior and complexity. The present model describes the major features of the dis-
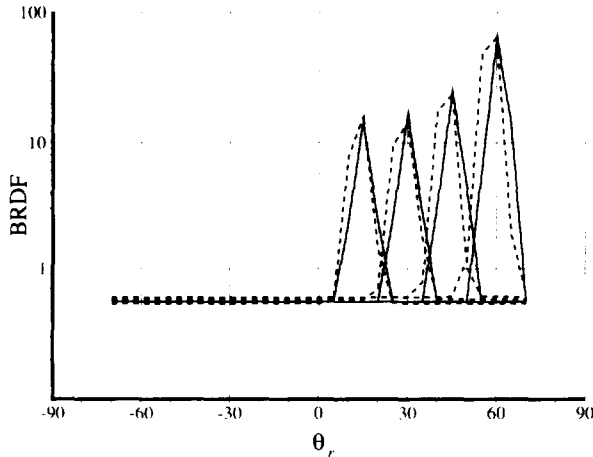
Figure 11: Absolute BRDF's for smooth blue plastic as obtained from theory (solid lines) and experiment (dashed lines) for incidence angles of θ_i = 15°, 30°, 45°, and 60°. λ = 0.46μm. This surface shows a typical smooth plastic reflection pattern with combined specular and uniform diffuse behavior.

tributions.

5 Example scenes

The reflection model described by equations (5) to (8) can be incorporated in ray-tracing or extended radiosity [15] methods. We have employed ray tracing. A single reflected ray is used together with ambient and point source illumination. The reflected intensity is given by

$$I_r(\lambda) = \sum_{i=1}^{N_l} \left\{ |F(\theta_i)|^2 \cdot e^{-g_i} \cdot S \cdot \Delta + \left[(\rho_{bd,dd})_i + a(\lambda) \right] \right. $$
$$\left. \cdot \cos\theta_i \cdot d\omega_i \right\} \cdot I_i(\lambda) + \rho_{hd}(\lambda) \cdot I_a(\lambda)$$

(12)

where N_l is the number of light sources, subscript i denotes the ith light source, the terms inside the braces respectively correspond to the three terms in equation (5), $\rho_{hd}(\lambda)$ is the hemispherical-directional reflectivity of the surface (taken as a function of λ only, and found from experiment or by integrating (5) over the incident hemisphere), and I_a is the uniform ambient illumination. The directional-diffuse term is included only for light sources. To include a directional-diffuse term from the environment, a distributed ray-tracer or an extended radiosity method [15] must be employed.

Figure 12 displays six aluminum cylinders in front of a brick wall. Each cylinder is rendered in isolation. Cylinders (a) to (f) are in order of increasing surface roughness. Other parameters are τ = 3.0μm for cylinders (a) to (e) and τ = 16.0μm for cylinder (f), and $a(\lambda)$ = 0. Note that the sharp specular image in the top faces of the cylinders diminishes, but is not blurred, with increasing surface roughness, and the image of the light source on the front vertical face spreads out. These are characteristics, respectively, of the specular and directional diffuse terms in the reflection model that are derived from physical optics. Note also that the apparent roughness of a given cylinder varies with viewing angle. The top and lateral edges can appear specular or nearly specular at grazing angles, even when the vertical face on the front side appears to be rough. A slight color shift is also apparent for a given rough surface (i.e., as λ in σ_0/λ varies). For visible light, this is most apparent in the blue shift on the front faces of the cylinders. The enhanced red shift

of the specular images is not so apparent. Clearly, the specular and directional diffuse terms of the model vary with wavelength, incidence angle, and roughness, and are responsible for the realism of the cylinders in Figure 12.

The aluminum cylinders (a) to (c) in Figure 13 illustrate limiting cases of each of the three terms in the reflection model. Cylinder (a) in Figure 13 is the same as cylinder (f) in Figure 12. Cylinder (b) is a smooth cylinder described by the specular term, in which the reflectance is a function of incidence angle according to the Fresnel reflectivity. Specular images are apparent on the top and lateral edges. (To emphasize the specular images, we have set the ambient illumination term to zero in rendering cylinder (b).) Cylinder (a) represents the directional diffuse term in the limit of $\sigma_0/\lambda \rightarrow \infty$ with σ_0/τ fixed at 0.16 (i.e., a limiting form for very rough surfaces). Cylinder (c) is ideal diffuse and is described by the uniform diffuse term. Note the striking differences between the three cylinders.

Figure 14 illustrates a scene consisting of a rough aluminum cylinder (σ_0 = 0.18μm, τ = 3.0μm, $a(\lambda)$ = 0), a rough copper sphere (σ_0 = 0.13μm, τ = 1.2μm, $a(\lambda)$ = 0), and a smooth plastic cube (σ_0 = 0, τ = 2.0μm, $a(\lambda)$ = 0.55μm) = 0.28), all resting on a rough plastic table (σ_0 = 0.20μm, τ = 2.0μm, $a(\lambda)$ = 0.55μm) = 0.28). The cube and table have the same Fresnel reflectivity.

Several effects can be noted in Figure 14. On the faces of the cube, the specular image varies with reflection angle, an effect caused solely by the Fresnel reflectivity $|F|^2$ in equation (6). The specular images on the table top also vary with reflection angle (and disappear), but this is caused mainly by roughness effects (i.e., e^{-g}) in equation (6). The cylinder in Figure 14 corresponds to cylinder (a) in Figure 12 and displays some of the specular and directional diffuse characteristics of that image.

Figure 14 gives a hint of the comprehensiveness of the light reflection model derived in this paper. Several materials of different roughnesses appear. A given surface can display specular or diffuse-like behavior depending on reflection angles and surface properties. Specular images appear or disappear based on correct physical principles. The high level of realism in Figure 14 is due to a physically-correct treatment of specular, directional diffuse, and uniform diffuse effects by the reflection model.

6 Conclusions

1. The general reflection model given by equations (5) to (8), in a single formulation, describes specular, directional diffuse, and uniform diffuse behavior. For unpolarized incident light, the model reduces to the form given in Appendix B. All of the parameters of the model are physically based.

2. The model compares favorably with experimental measurements of reflected radiation for metals, nonmetals, and plastics, with smooth and rough surfaces.

3. The model accurately predicts the emergence of specular reflection with increasing wavelength or angle of incidence, or decreasing surface roughness.

4. The model predicts a directional-diffuse pattern which can have maximal values at specular, off-specular, or grazing angles, depending on surface roughness.

5. The model is in analytical form and can improve the realism of synthetic images.

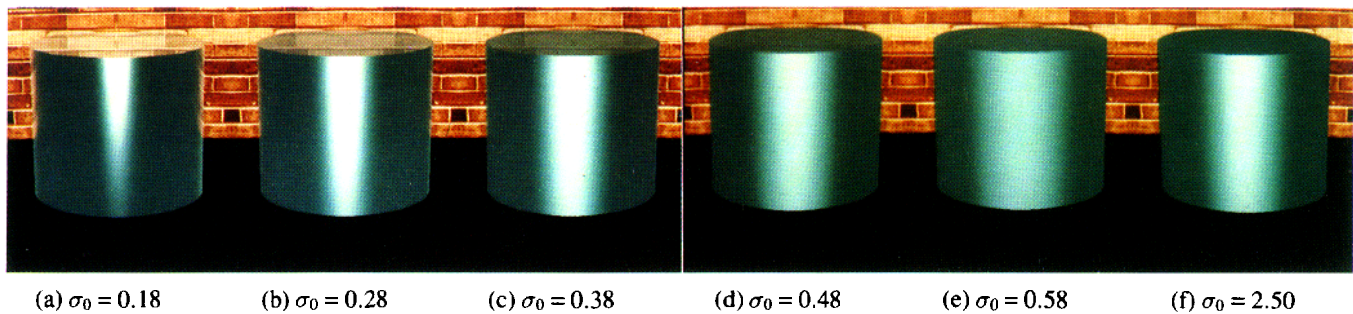6. The model can be employed for ray-tracing or extended radiosity [15] methods.

(a) $\sigma_0 = 0.18$ (b) $\sigma_0 = 0.28$ (c) $\sigma_0 = 0.38$ (d) $\sigma_0 = 0.48$ (e) $\sigma_0 = 0.58$ (f) $\sigma_0 = 2.50$

Figure 12: Aluminum cylinders with different surface roughnesses. σ_0 is in μm. $\tau = 3.0\mu m$ for cylinders (a) to (e) and $\tau = 16.0\mu m$ for cylinder (f). Note that the specular and directional-diffuse reflection characteristics vary with reflection angle and roughness.

7. The model highlights the need for tabulated databases of parameterized bidirectional reflectivities. The parameters include two surface roughness parameters (σ_0, τ), the index of refraction (as a function of wavelength), and the constrained parameter $a(\lambda)$. The latter can be inferred from measured hemispherical reflectivities.

In conclusion, the reflection model is comprehensive, physically-based, and provides an accurate transition from specular to diffuse-like reflection. Further, the model is computable and thus useful for graphics applications.

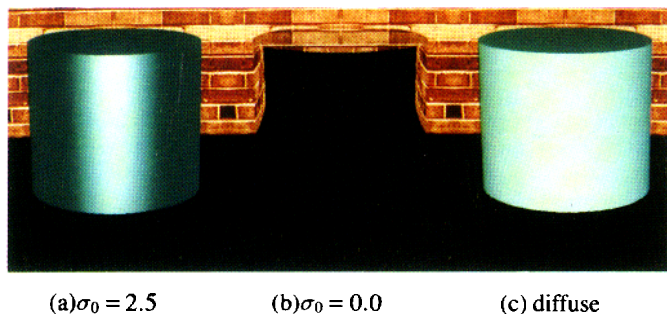(a)$\sigma_0 = 2.5$ (b)$\sigma_0 = 0.0$ (c) diffuse

Figure 13: Aluminum cylinders in extreme limiting cases. Each cylinder corresponds to one of the three terms in the reflection model. σ_0 is in μm. (a) Directional diffuse reflection; (b) Ideal specular reflection; (c) Uniform diffuse (Lambertian) reflection.
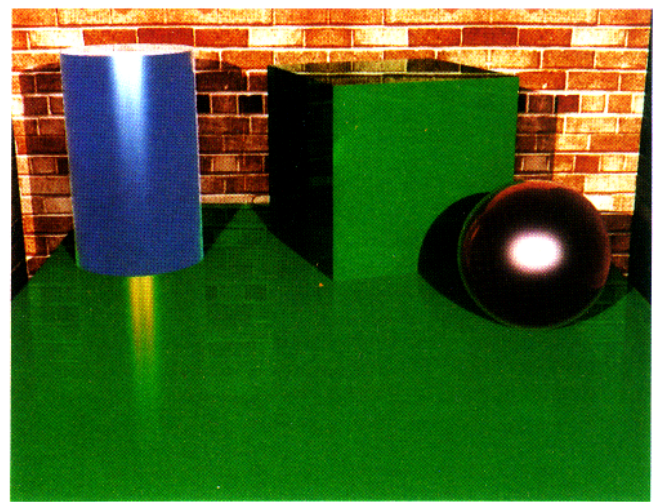

Figure 14: A general scene with metallic and plastic objects in the foreground, with smooth and rough surfaces. The specular images in the smooth plastic box vary with incidence angle due to the Fresnel effect. In the table top, the decay of the specular images with reflection angle is due to roughness. In the rough metallic surfaces, the glossy highlights result from directional diffuse reflection.

discussions, and to Emil Ghinger for photographing the raster images. We thank the reviewers for their extensive and constructive comments which have helped to clarify a difficult paper.

Acknowledgments

References

[1] Bahar, E. and S. Chakrabarti. "Full wave theory applied to computer-aided graphics for 3-D objects," *IEEE Computer Graphics and Applications*, 7(7), 1987, pages 46–60.

[2] Bahar, Ezekiel. "Review of the full wave solutions for rough surface scattering and depolarization," *Journal of Geophysical Research*, May 1987, pages 5209–5227.

[3] Bass, F.G. and I.M. Fuks. *Wave Scattering from Statistically Rough Surfaces*, Pergamon Press, 1979.

[4] Beckmann, Petr. "Shadowing of Random Rough Surfaces," *IEEE Transactions on Antennas and Propagation*, May 1965, pages 384–388.

[5] Beckmann, Petr and André Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, 1963.

[6] Blinn, James F. "Models of Light Reflection for Computer Synthesized Pictures," *Computer Graphics*, 11, 1977, pages 192–198. (Proceedings SIGGRAPH '77.)

[7] Brockelman, R. A. and T. Hagfors. "Note on the Effect of Shadowing on the Backscattering of Waves from a Random Rough Surfaces," *IEEE Transactions on Antennas and Propagation*, AP-14(5), September 1966, pages 621–629.

[8] Cook, Robert L. and Kenneth E. Torrance. "A Reflectance Model for Computer Graphics," *ACM Transactions on Graphics*, 1, 1982, pages 7–24.

[9] Feng, Xiaofen. *Comparison of methods for generation of absolute reflectance factors for BRDF studies*, Master's thesis, Rochester Institute of Technology, 1990.

[10] Hering, R.G. and T.F. Smith. "Apparent radiation properties of a rough surface," *Application to Thermal Design of Spacecraft*, 23, 1970, pages 337–361.

[11] Himlan, Theodore H., Michael C. Monks, Stephan H. Westin, Donald P. Greenberg, and Kenneth E. Torrance. "Physical measurement Techniques for Improving and Evaluating Computer Graphic Simulations.," 1991. (To be published.)

[12] Jackson, John D. *Classical Electrodynamics*, John Wiley & Son Inc., 1975.

[13] Phong, Bui Tuong. "Illumination for Computer Generated Pictures," *Communications of the ACM*, 18(6), June 1975, pages 311–317.

[14] Siegel, Robert and John R. Howell. *Thermal Radiation Heat Transfer*, McGraw-Hill book Company, 2nd edition, 1981.

[15] Sillion, François, James Arvo, Stephen Westin, and Donald P. Greenberg. "A Global Illumination Solution for General Reflectance Distributions," *Computer Graphics*, 25(4), August 1991. (Proceedings SIGGRAPH '91 in Las Vegas.)

[16] Smith, Bruce G. "Geometrical Shadowing of a Random Rough Surface," *IEEE Transactions on Antennas and Propagation*, AP-15(5), September 1967, pages 668–671.

[17] Smith, T.F. and K.E. Nichols. "Effects of polarization on bidirectional reflectance of a one-dimensional randomly rough surface," *Spacecraft Radiative Transfer and Temperature Control*, 83, 1981, pages 3–21.

[18] Stogryn, Alex. "Electromagnetic Scattering From Rough, Finitely Conducting Surfaces," *Radio Science*, 2(4), 1967, pages 415–428.

[19] Torrance, K.E. and E.M. Sparrow. "Off-Specular Peaks in the Directional Distribution of Reflected Thermal Radiation," *Journal of Heat Transfer – Transactions of the ASME*, May 1966, pages 223–230.

[20] Torrance, K.E. and E.M. Sparrow. "Theory for Off-Specular Reflection from Roughened Surfaces," *Journal of the Optical Society of America*, 57(9), September 1967, pages 1105–1114.

A   Appendix: Derivations

A.1   Reflected intensities

The reflected intensities for the s and p components of polarizations are given by [14][18]

$$dI_r(\theta_r, \phi_r; \theta_i, \phi_i)_s = \frac{R^2}{A \cdot \cos\theta_r} < |\hat{s}_r \cdot \vec{E}_r(\vec{R})|^2 >$$

$$dI_r(\theta_r, \phi_r; \theta_i, \phi_i)_p = \frac{R^2}{A \cdot \cos\theta_r} < |\hat{p}_r \cdot \vec{E}_r(\vec{R})|^2 > \quad (13)$$

where the coordinates are as shown in Figure 5, $\vec{E}_r(\vec{R})$ is the reflected field in vector form, R is the distance from the origin to an arbitrary point in space, A is the area of the reflecting surface projected on the x-y plane, and $\hat{s}_r, \hat{p}_r$ are unit polarization vectors, given by

$$\hat{s}_r = \frac{\hat{k}_r \times \hat{z}}{|\hat{k}_r \times \hat{z}|}$$

$$\hat{p}_r = \hat{s}_r \times \hat{k}_r \quad (14)$$

which are normal and parallel, respectively, to the plane formed by the viewing direction and the mean surface normal. The symbol $<>$ denotes an average over the joint probability distribution function of the random rough surface characterized by

$$z = \xi(x, y). \quad (15)$$

The reflected field can be expressed in terms of the scattered field on the surface by using the vector form of the Kirchhoff diffraction theory [12]:

$$\vec{E}_r(\vec{R}) = \frac{e^{ikR}}{4\pi R} \left(\bar{I} - \hat{k}_r \hat{k}_r \right) \cdot$$
$$\int_\Gamma e^{-i\vec{k}_r \cdot \vec{r}} \left\{ -i\vec{k}_r \times (\vec{E}_s \times \hat{n}) - (\nabla \times \vec{E}_s) \times \hat{n} \right\} d\Gamma$$

$$(16)$$

where $\vec{k}_i, \vec{k}_r$ are wave vectors in the incident and reflection directions, $|k| = 2\pi/\lambda$ is the wave number, $\vec{r}$ is the position vector for a point on the surface, and the tensor $\bar{I} - \hat{k}_r \hat{k}_r = \hat{s}_r \hat{s}_r + \hat{p}_r \hat{p}_r$ is introduced to to make the reflected field transverse.

Substituting (16) into (13), we have

$$dI_s = \frac{1}{A \cos\theta_r (4\pi)^2} \cdot < | \int_\Gamma e^{-i\vec{k}_r \cdot \vec{r}} \cdot$$
$$\left\{ ik\hat{p}_r \cdot (\vec{E}_s \times \hat{n}) + \hat{s}_r \cdot [(\nabla \times \vec{E}_s) \times \hat{n}] \right\} d\Gamma |^2 >$$

$$dI_p = \frac{1}{A \cos\theta_r (4\pi)^2} \cdot < | \int_\Gamma e^{-i\vec{k}_r \cdot \vec{r}} \cdot$$
$$\left\{ ik\hat{s}_r \cdot (\vec{E}_s \times \hat{n}) - \hat{p}_r \cdot [(\nabla \times \vec{E}_s) \times \hat{n}] \right\} d\Gamma |^2 >$$

$$(17)$$

To evaluate the right side of (17), the surface element $d\Gamma$ is expressed in terms of the planar surface area $dA = dx \cdot dy$ by

$$d\Gamma = dA/(\hat{n} \cdot \hat{z}) \quad (18)$$

Further, the squares of the absolute values of the integrals in (17) can be expanded in terms of double surface integrals. We find

$$< | \int_\Gamma e^{-i\vec{k}_r \cdot \vec{r}} \{ \} d\Gamma |^2 > =$$

$$< \int_A dA_1 \int_A dA_2 e^{-i\vec{v}\cdot(\vec{r}_1-\vec{r}_2)} \cdot$$

$$\left(e^{-i\vec{k}_i\cdot\vec{r}_1}\{\}_1 \right)\left(e^{-i\vec{k}_i\cdot\vec{r}_2}\{\}_2 \right)^* /(\hat{n}_1\cdot\hat{z})(\hat{n}_2\cdot\hat{z}) > \tag{19}$$

where v is the wave vector change

$$\vec{v} = k(\hat{k}_r - \hat{k}_i), \tag{20}$$

$*$ denotes a complex conjugate, $\{\}$ refers to the terms in braces in (17), and the subscripts refer to points on area elements dA_1 and dA_2.

The $<>$ in (19) commutes with the surface integral and a term of the form

$$< e^{-i\vec{v}\cdot\hat{z}(\xi_1-\xi_2)}\{\}_1\{\}_2^* /(\hat{n}_1\cdot\hat{z})(\hat{n}_2\cdot\hat{z}) > \tag{21}$$

results. Since the surface is assumed to be isotropic and stationary, (21) is a function only of $x_1 - x_2$ and $y_1 - y_2$. Thus, by making the change of variables

$$x' = x_1 - x_2 \qquad x'' = x_2$$
$$y' = y_1 - y_2 \qquad y'' = y_2 \tag{22}$$

the integrals over x'' and y'' may be carried out separately to give a factor $S \cdot A$, where S is the fraction of the surface that is both illuminated and viewed and represents the shadowing function given by [16]:

$$S = S_i(\theta_i) \cdot S_r(\theta_r) \tag{23}$$

where

$$S_i(\theta_i) = (1 - \frac{1}{2}\text{erfc}(\frac{\tau\cot\theta_i}{2\sigma_0}))/(\Lambda(\cot\theta_i) + 1)$$

$$S_r(\theta_r) = (1 - \frac{1}{2}\text{erfc}(\frac{\tau\cot\theta_r}{2\sigma_0}))/(\Lambda(\cot\theta_r) + 1) \tag{24}$$

and

$$\Lambda(\cot\theta) = \frac{1}{2}\left(\frac{2}{\pi^{1/2}} \cdot \frac{\sigma_0}{\tau\cot\theta} - \text{erfc}(\frac{\tau\cot\theta}{2\sigma_0}) \right) \tag{25}$$

Hence, the reflected intensities in (17) are

$$dI_s = \frac{S}{\cos\theta_r(4\pi)^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} dx'dy'\, e^{-i\vec{v}\cdot\vec{\eta}}B_s \tag{26}$$

$$dI_p = \frac{S}{\cos\theta_r(4\pi)^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} dx'dy'\, e^{-i\vec{v}\cdot\vec{\eta}}B_p \tag{27}$$

where

$$\vec{\eta} = x'\hat{x} + y'\hat{y} \tag{28}$$

and

$$B_s = < e^{-i\vec{v}\cdot\hat{z}(\xi_1-\xi_2)}\mathcal{F}(\hat{n}_1,\hat{n}_2)_s >$$
$$B_p = < e^{-i\vec{v}\cdot\hat{z}(\xi_1-\xi_2)}\mathcal{F}(\hat{n}_1,\hat{n}_2)_p > \tag{29}$$

where

$$\mathcal{F}(\hat{n}_1,\hat{n}_2)_s = e^{-i\vec{k}_i\cdot(\vec{r}_1-\vec{r}_2)}/(\hat{n}_1\cdot\hat{z})(\hat{n}_2\cdot\hat{z})$$
$$\cdot \left(\{ik\hat{p}_r\cdot(\vec{E}_s\times\hat{n}) + \hat{s}_r\cdot[(\nabla\times\vec{E}_s)\times\hat{n}]\} \right)_1$$
$$\cdot \left(\{ik\hat{p}_r\cdot(\vec{E}_s\times\hat{n}) + \hat{s}_r\cdot[(\nabla\times\vec{E}_s)\times\hat{n}]\} \right)_2^* \tag{30}$$

$$\mathcal{F}(\hat{n}_1,\hat{n}_2)_p = e^{-i\vec{k}_i\cdot(\vec{r}_1-\vec{r}_2)}/(\hat{n}_1\cdot\hat{z})(\hat{n}_2\cdot\hat{z})$$
$$\cdot \left(\{ik\hat{s}_r\cdot(\vec{E}_s\times\hat{n}) - \hat{p}_r\cdot[(\nabla\times\vec{E}_s)\times\hat{n}]\} \right)_1$$
$$\cdot \left(\{ik\hat{s}_r\cdot(\vec{E}_s\times\hat{n}) - \hat{p}_r\cdot[(\nabla\times\vec{E}_s)\times\hat{n}]\} \right)_2^* \tag{31}$$

The functions B_s and B_p in (29) depend only on x' and y'. Notice that dI_s and dI_p are the s and p polarized reflected intensities, respectively. The total reflected intensity, as used in equation (4), is given by

$$dI_r = dI_s + dI_p \tag{32}$$

A.2 Tangent plane approximation

The reflected intensities in (26) and (27) are expressed in terms of the scattered field $\vec{E}_s$ on the surface. In turn, $\vec{E}_s$ depends on the incident field, and may be related to the incident field by using the local tangent plane approximation.

For the case of a unidirectional incident field, we have

$$\vec{E}_i = E_0 e^{i\vec{k}_i\cdot\vec{r}}\mathbf{p} \tag{33}$$
$$\mathbf{p} = c_s\hat{s}_i + c_p\hat{p}_i \tag{34}$$

where E_0 is the wave amplitude, $\mathbf{p}$ is the polarization state vector of the incident radiation, c_s, c_p are called the polarization coefficients[1], and $\hat{s}_i$, $\hat{p}_i$ are unit polarization vectors with respect to the plane of incidence ($\hat{k}_i, \hat{z}$). The unit vectors are given by

$$\hat{s}_i = \frac{\hat{k}_i\times\hat{z}}{|\hat{k}_i\times\hat{z}|}$$
$$\hat{p}_i = \hat{s}_i\times\hat{k}_i \tag{35}$$

Equation (33) can be written in the more compact matrix form

$$\vec{E}_i = E_0 e^{i\vec{k}_i\cdot\vec{r}}\left(c_s, c_p\right)\cdot\left(\begin{array}{c}\hat{s}_i\\\hat{p}_i\end{array}\right) \tag{36}$$

$\hat{s}_i, \hat{p}_i$ decompose into incident local polarization unit vectors $\hat{s}_i^n$, $\hat{p}_i^n$ with respect to the local incident plane ($\hat{k}_i, \hat{n}$), given by

$$\hat{s}_i^n = \frac{\hat{k}_i\times\hat{n}}{|\hat{k}_i\times\hat{n}|}$$
$$\hat{p}_i^n = \hat{s}_i^n\times\hat{k}_i \tag{37}$$

Therefore,

$$\left(\begin{array}{c}\hat{s}_i\\\hat{p}_i\end{array}\right) = T_{in}\cdot\left(\begin{array}{c}\hat{s}_i^n\\\hat{p}_i^n\end{array}\right) \tag{38}$$

where T_{in} is the transformation matrix from incident coordinates to local coordinates

$$\bar{T}_{in} = \left(\begin{array}{cc}\hat{s}_i\cdot\hat{s}_i^n & \hat{s}_i\cdot\hat{p}_i^n\\\hat{p}_i\cdot\hat{s}_i^n & \hat{p}_i\cdot\hat{p}_i^n\end{array}\right) \tag{39}$$

Substituting (38) into (36), we have the incident field in terms of $\hat{s}_i^n, \hat{p}_i^n$ as

$$\vec{E}_i = E_0 e^{i\vec{k}_i\cdot\vec{r}}\left(c_s, c_p\right)\cdot\bar{T}_{in}\left(\begin{array}{c}\hat{s}_i^n\\\hat{p}_i^n\end{array}\right) \tag{40}$$

Reflections of the $\hat{s}_i^n, \hat{p}_i^n$ fields are found from the local Fresnel reflection coefficients for each component of polarization, i.e.,

$$\hat{s}_i^n \longrightarrow F_s\cdot\hat{s}_r^n \tag{41}$$
$$\hat{p}_i^n \longrightarrow F_p\cdot\hat{p}_r^n$$

---

[1] For example: for s polarization, $c_s = 1.0$, $c_p = 0$; for p polarization, $c_s = 0$, $c_p = 1$.

where F_s and F_p are the Fresnel reflection coefficients for s and p polarizations, respectively [14, p.100]. The unit vectors $\hat{s}_r^n, \hat{p}_r^n$ are the local polarization unit vectors for reflection from the tangent plane:

$$\hat{s}_r^n = \frac{\hat{k}_r \times \hat{n}}{|\hat{k}_r \times \hat{n}|} \tag{42}$$

$$\hat{p}_r^n = \hat{s}_r^n \times \hat{k}_r$$

where $\hat{k}_r$ is the unit vector in the specular direction from the tangent plane, given by

$$\hat{k}_r = \hat{k}_i - 2(\hat{k}_i \cdot \hat{n})\hat{n} \tag{43}$$

Using the Fresnel matrix

$$\bar{F} = \begin{pmatrix} F_s & 0 \\ 0 & F_p \end{pmatrix} \tag{44}$$

we have in more compact form

$$\begin{pmatrix} \hat{s}_i^n \\ \hat{p}_i^n \end{pmatrix} \longrightarrow \bar{F} \begin{pmatrix} \hat{s}_r^n \\ \hat{p}_r^n \end{pmatrix} \tag{45}$$

From equations (40) and (45), the scattered field on the surface can be expressed as a linear combination of the Fresnel reflection coefficients

$$\vec{E}_s = E_0 e^{i\hat{k}_i \cdot \vec{r}} (c_s, c_p) \cdot \begin{pmatrix} \hat{s}_i \cdot \hat{s}_i^n & \hat{s}_i \cdot \hat{p}_i^n \\ \hat{p}_i \cdot \hat{s}_i^n & \hat{p}_i \cdot \hat{p}_i^n \end{pmatrix} \cdot$$
$$\begin{pmatrix} F_s & 0 \\ 0 & F_p \end{pmatrix} \cdot \begin{pmatrix} \hat{s}_r^n \\ \hat{p}_r^n \end{pmatrix} \tag{46}$$

The scattered field is a function of the incident polarization state, the local surface normal $\hat{n}$, the Fresnel reflection coefficients F_s and F_p of the surface, and the incident and reflection directions $\hat{k}_i, \hat{k}_r$.

A.3   Representation of the surface

Specification of the surface topography is required to carry out the surface integrals and surface averages appearing in equations (26), (27) and (29). Without losing generality, we assume the surface to be Gaussian distributed [5], i.e., we assume the surface height in (15) to be a stationary normally distributed random process whose mean value is zero. In addition we assume the surface to be directionally isotropic. An appropriate two-point joint probability function is given by

$$P(z_1, z_2) = \frac{exp\left[-(z_1^2 + z_2^2 - 2C(r)z_1 z_2)/2\sigma_0^2(1 - C(r)^2)\right]}{2\pi\sigma_0^2\sqrt{1 - C(r)^2}} \tag{47}$$

where $r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$, σ_0^2 is the variance of $z_1 = \xi(x_1, y_1)$ and $z_2 = \xi(x_2, y_2)$, and $C(r)$ is the correlation coefficient, which is assumed to be [5]

$$C(r) = e^{-\frac{r^2}{\tau^2}} \tag{48}$$

where τ is the autocorrelation length.

The parameters σ_0 and τ are the only two surface parameters required for the surface integrations.

A.4   Analytic evaluation of the integrals

Substituting (46) into (29) to (31), B_s and B_p are expressed in terms of known quantities and depend on the surface only through the normals n_1 and n_2 at two surface points. Further, the integrals in equations (26) and (27) can be written as:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i\vec{v}\cdot\vec{\eta}} < e^{-i\vec{v}\cdot\hat{z}(\xi_1 - \xi_2)} \mathcal{F}(n_1, n_2, \mathbf{p}) > dxdy \tag{49}$$

Stogryn [18] has shown that an integral and average of the form in (49) can be approximatly evaluated under either of the following two conditions:

- the surface is very rough (i.e., $(v_z\sigma)^2 \gg 1$)

- the surface has gentle slopes (i.e. $(\frac{\sigma}{\tau}) \ll 1$)

As a result, (49) reduces to

$$\mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p}) \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i\vec{v}\cdot\vec{\eta}} < e^{-i\vec{v}\cdot\hat{z}(\xi_1 - \xi_2)} > dxdy \tag{50}$$

where $\mathcal{F}$ is evaluated at $\hat{n}_b$, which is the unit vector bisecting $\hat{k}_i$ and $\hat{k}_r$, given by

$$\hat{n}_b = \frac{\hat{k}_r - \hat{k}_i}{|\hat{k}_r - \hat{k}_i|} \tag{51}$$

Furthermore, the $<>$ in (50) can be shown to be [5]:

$$< e^{-i\vec{v}\cdot\hat{z}(\xi_1 - \xi_2)} > = e^{-(v_z\sigma)^2[1 - C(\eta)]} \tag{52}$$

where $C(\eta)$ is given by (48).

Note that σ in (52) is the effective surface roughness, not σ_0. This is because the surface averaging is carried over illuminated and visible parts only. σ is given by [4]:

$$\sigma = \frac{\sigma_0}{\sqrt{1 + \frac{z_0^2}{\sigma_0^2}}} \tag{53}$$

where z_0 depends on θ_i and θ_r and is the root of the following equation

$$\sqrt{\frac{\pi}{2}}z = \sigma_0 K \cdot \exp(-\frac{z^2}{2\sigma_0^2}) \tag{54}$$

and

$$K = K_i + K_r$$
$$K_i = \frac{1}{4}\tan\theta_i \cdot erfc(\frac{\tau}{2\sigma_0}\cot\theta_i)$$
$$K_r = \frac{1}{4}\tan\theta_r \cdot erfc(\frac{\tau}{2\sigma_0}\cot\theta_r) \tag{55}$$

The double integral in (50) can be evaluated analytically [5]:

$$N = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i\vec{v}\cdot\vec{\eta}} < e^{-i\vec{v}\cdot\hat{z}(\xi_1 - \xi_2)} > dxdy$$
$$= e^{-g} \cdot A \cdot sinc^2(v_x L_x) sinc^2(v_y L_y) +$$
$$\pi\tau^2 \sum_{m=1}^{\infty} \frac{g^m e^{-g}}{m! \cdot m} \cdot \exp(-v_{xy}^2\tau^2/4m) \tag{56}$$

where L_x, L_y are the dimensions of the reflecting surface. Since we are only interested in cases when $L_x, L_y \gg \lambda$, the first term is nonzero only in the specular direction and zero otherwise. For the case of unidirectional incidence with solid angle $d\omega_i$ and $L_x, L_y \gg \lambda$, the averaged form of the first term in (56) is

$$A \cdot sinc^2(v_x L_x) sinc^2(v_y L_y) \rightarrow (2\pi\lambda)^2 \cdot A/(d\omega_i \cdot \cos\theta_r) \quad (57)$$

Hence, (56) becomes

$$N = e^{-g} \cdot (2\pi\lambda)^2 \cdot A/(d\omega_i \cdot \cos\theta_r) +$$
$$\pi\tau^2 \sum_{m=1}^{\infty} \frac{g^m e^{-g}}{m! \cdot m} \cdot \exp(-v_{xy}^2 \tau^2/4m) \quad (58)$$

Next, $\mathcal{F}_s$ and $\mathcal{F}_p$ in (30) and (31) are evaluated. First, $\hat{n}_1, \hat{n}_2$ are replaced by $\hat{n}_b$ defined in (51). Then they are substituted into (30) and (31). After lengthy vector manipulations, we find

$$\mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_s = \delta \cdot |c_s M_{ss} + c_p M_{sp}|^2 \quad (59)$$

$$\mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_p = \delta \cdot |c_s M_{ps} + c_p M_{pp}|^2 \quad (60)$$

where

$$M_{ss} = \left(F_s(\hat{p}_i \cdot \hat{k}_r)(\hat{p}_r \cdot \hat{k}_i) + F_p(\hat{s}_i \cdot \hat{k}_r)(\hat{s}_r \cdot \hat{k}_i)\right) \quad (61)$$

$$M_{sp} = -\left(F_s(\hat{s}_i \cdot \hat{k}_r)(\hat{p}_r \cdot \hat{k}_i) - F_p(\hat{p}_i \cdot \hat{k}_r)(\hat{s}_r \cdot \hat{k}_i)\right) \quad (62)$$

$$M_{pp} = \left(F_s(\hat{s}_i \cdot \hat{k}_r)(\hat{s}_r \cdot \hat{k}_i) + F_p(\hat{p}_i \cdot \hat{k}_r)(\hat{p}_r \cdot \hat{k}_i)\right) \quad (63)$$

$$M_{ps} = \left(F_s(\hat{p}_i \cdot \hat{k}_r)(\hat{s}_r \cdot \hat{k}_i) - F_p(\hat{s}_i \cdot \hat{k}_r)(\hat{p}_r \cdot \hat{k}_i)\right) \quad (64)$$

$$\delta = (\frac{2\pi}{\lambda})^2 \cdot \frac{|\hat{k}_r - \hat{k}_i|^4}{|\hat{k}_r \times \hat{k}_i|^4 \left(\hat{z} \cdot (\hat{k}_r - \hat{k}_i)\right)^2} \quad (65)$$

The Fresnel reflection coefficients F_s and F_p in (61) to (64) are evaluated at the bisecting angle given by $\cos^{-1}(|\hat{k}_r - \hat{k}_i|/2)$.

Using (59)-(65) and (58) in (26) and (27), we find an analytical expression for the reflected intensity

$$dI_s = \frac{|E_0|^2}{\cos\theta_r (4\pi)^2} \mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_s \cdot N$$

$$dI_p = \frac{|E_0|^2}{\cos\theta_r (4\pi)^2} \mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_p \cdot N \quad (66)$$

where the square of the absolute value of the incident field amplitude, $|E_0|^2$, is related to the incident intensity I_i by

$$|E_0|^2 = I_i d\omega_i \quad (67)$$

Note that the right side of (66) has the correct dimensions of intensity since N has dimension $[L^2]$ whereas the $\mathcal{F}$'s have dimension $[L^{-2}]$.

Finally, substituting (67) into (66) and using (4) and (32), we get exactly the first two terms in (5), given that

$$\mathcal{F}(\hat{k}_i, \hat{k}_r, \mathbf{p}) = \mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_s + \mathcal{F}(\hat{n}_b, \hat{n}_b, \mathbf{p})_p \quad (68)$$

since the BRDF defined in (4) is the total BRDF, which is the sum of the BRDF's for the reflected s and p components.

B Appendix: Governing equations of the reflectance model for unpolarized incident light

Equations (5) to (8) together with the defining equations for all the symbols in (5) to (8) completely define the general BRDF for

arbitrarily-polarized incident light. In most applications, however, we are only interested in the BRDF for unpolarized incident light. The expressions for the BRDF are greatly simplified for this special but useful case. For convenience, the BRDF equations for unpolarized incident light are presented in this appendix. The reader should refer to Figure 5 and the nomenclature list in Table 1 for the angular coordinates and other physical parameters that appear in the reflectance model:

$$\rho_{bd} = \rho_{bd}(\lambda, \sigma_0, \tau, \bar{n}(\lambda), a(\lambda))$$
$$= \rho_{bd,sp} + \rho_{bd,dd} + \rho_{bd,ud} \quad (69)$$

$$\rho_{bd,sp} = \frac{\rho_s}{\cos\theta_i d\omega_i} \cdot \Delta \quad (70)$$

$$\rho_{bd,dd} = \frac{|F|^2}{\pi} \cdot \frac{G \cdot S \cdot D}{\cos\theta_i \cos\theta_r} \quad (71)$$

$$\rho_{bd,ud} = a(\lambda) \quad (72)$$

$$\rho_s = |F|^2 \cdot e^{-g} \cdot S \quad (73)$$

$$\Delta = \begin{cases} 1 & \text{if in specular cone} \\ 0 & \text{otherwise} \end{cases} \quad (74)$$

$$|F|^2 = \frac{1}{2}(F_s^2 + F_p^2) = f(\theta_i, \theta_r, \bar{n}(\lambda)) \quad (75)$$

$$G = \left(\frac{\vec{v} \cdot \vec{v}}{v_z}\right)^2 \cdot \frac{1}{|\hat{k}_r \times \hat{k}_i|^4} \cdot [(\hat{s}_r \cdot \hat{k}_i)^2 + (\hat{p}_r \cdot \hat{k}_i)^2] \cdot$$
$$[(\hat{s}_i \cdot \hat{k}_r)^2 + (\hat{p}_i \cdot \hat{k}_r)^2] \quad (76)$$

$$S = S(\theta_i, \theta_r, \sigma_0/\tau) \quad (77)$$

$$D = \frac{\pi^2 \tau^2}{4\lambda^2} \cdot \sum_{m=1}^{\infty} \frac{g^m e^{-g}}{m! \cdot m} \cdot \exp(-v_{xy}^2 \tau^2/4m) \quad (78)$$

$$g = [(2\pi\sigma/\lambda)(\cos\theta_i + \cos\theta_r)]^2 \quad (79)$$

$$\sigma = \sigma_0 \cdot [1 + (\frac{z_0}{\sigma_0})^2]^{-1/2} \quad (80)$$

$$\sqrt{\frac{\pi}{2}} z_0 = \frac{\sigma_0}{4}(K_i + K_r) \cdot \exp(-\frac{z_0^2}{2\sigma_0^2}) \quad (81)$$

$$K_i = \tan\theta_i \cdot erfc(\frac{\tau}{2\sigma_0} \cot\theta_i) \quad (82)$$

$$K_r = \tan\theta_r \cdot erfc(\frac{\tau}{2\sigma_0} \cot\theta_r) \quad (83)$$

$$\vec{v} = \hat{k}_r - \hat{k}_i, \quad v_{xy} = \sqrt{v_x^2 + v_y^2} \quad (84)$$

$$\hat{s}_i = \frac{\hat{k}_i \times \hat{n}}{|\hat{k}_i \times \hat{n}|}, \quad \hat{p}_i = \hat{s}_i \times \hat{k}_i \quad (85)$$

$$\hat{s}_r = \frac{\hat{k}_r \times \hat{n}}{|\hat{k}_r \times \hat{n}|}, \quad \hat{p}_r = \hat{s}_r \times \hat{k}_r \quad (86)$$

where $\bar{n}$ is the index of refraction, ρ_s is the specular reflectivity, Δ is a delta function, $|F|^2$ is the Fresnel reflectivity for unpolarized light [14, p.100] evaluated at the bisecting angle given by $\cos^{-1}(|\hat{k}_r - \hat{k}_i|/2)$, G is a geometrical factor, S is the shadowing/masking factor given in equation (23), and D is a distribution function for the directional diffuse reflection term.