A Unified Model of Specular and Diffuse Reflectance for Rough, Glossy Surfaces

William A. P. Smith and Edwin R. Hancock
Department of Computer Science, University of York, UK
{wsmith,erh}@cs.york.ac.uk

Abstract

In this paper we consider diffuse and specular reflectance from surfaces modeled as distributions of glossy microfacets. In contrast to previous work, we describe the relative contribution of both of these components in the same terms, namely with recourse to Fresnel theory. This results in a more highly constrained model with a reduced number of parameters. Also, the need for ad hoc and physically meaningless specular and diffuse reflectance coefficients is removed. This ensures that the conservation of energy is obeyed and only physically plausible mixtures of the two components are allowed. In our model, both specular and diffuse reflectance are related to the roughness and refractive index of the surface. We show how physically meaningful parameters of a surface can be measured from uncalibrated imagery and that our model fits observed BRDF data more accurately than comparable existing models.

1. Introduction

Accurate models of the reflectance of light from surfaces have many applications, both in computer graphics (for the creation of realistic images) and in computer vision (for the estimation of surface shape and physical properties from images). Existing models of reflectance have tended to be either phenomenological or physics-based by design. Phenomenological models [6] are attractive because of their simplicity and ability to visually capture some observable feature of surface reflectance. However, such models bear no relation to the underlying physics of light reflection and their parameters are not related in a meaningful way to physical quantities in the real world. Moreover, they are not constrained to obey physical principles of reflectance and can therefore produce reflectance values that are physically impossible. In computer vision, where commonly we aim to tackle highly underconstrained problems (e.g. shape-from-shading [12]), additional constraint is useful in reducing the space of possible solutions. For this reason, physics-based models of reflectance are attractive. Moreover, they allow us to estimate parameters which have meaning in the real world (for example surface roughness) which may be useful for surface classification.

Many real surfaces are macroscopically rough. Physics-based models of such surfaces can be derived in two ways. If the surface perturbations are on a similar scale to the wavelength of the incident radiation, then a physical optics model must be used, such as that of Beckmann and Spizzichino [1]. If, on the other hand, surface perturbations are much larger than the wavelength of the incident radiation, then significantly simpler geometrical optics models can be used. In this paradigm, a popular and analytically simple approach has been to assume that the macrostructure of the surface under study comprises microfacets whose orientation varies from that of the mean surface normal according to a probability distribution. Such a model requires a number of further ingredients. The first is a model of how each microfacet reflects light locally. The second is a model of how the geometry of the microfacets causes shadowing and masking of other parts of the surface. The third is a model of how light may be interreflected between facets. Oren and Nayar [5] assume that each facet reflects light according to Lambert’s law. Torrance and Sparrow [7], on the other hand, assume that each facet is a perfect mirror reflector and derive a model for specular reflectance from rough surfaces. In both cases they assume that the surface is composed of long, symmetric v-grooves and account for shadowing and masking of adjacent facets. Oren and Nayar [5] consider two-bounce interreflections between adjacent facets while Torrance and Sparrow [7] discount the effect of interreflections.

There have been a number of attempts to combine these models to allow for facets that are ‘glossy’ (i.e. light is reflected both diffusely and specularly). For example, Blinn [2] and Cook and Torrance [3] append a Lambertian model to the specular model of Torrance and Sparrow [7]. They weight the relative contribution of the two models using specular and diffuse coefficients. There are two key weaknesses to this approach. The first is that the standard Lambertian model does not account for the effects of microfacets. It is possible to overcome this problem by using the Oren and Nayar [5] model instead, this has been con-
sidered by e.g. van Ginneken et al. [9]. However, this does not address the second weakness which is that the specular and diffuse coefficients have no physical meaning and their relative values may result in physically impossible reflectance. The values of these coefficients may be arbitrarily chosen in order to obtain pleasing visual results or, in the case of computer vision, to fit observed data. However, there is an implicit assumption that all of the incident light contributes to the diffuse reflectance, which is subsequently downweighted by the diffuse coefficient in order to achieve a realistic mixture of diffuse and specular reflectance. This problem was considered superficially by van Ginneken et al. [9] who suggested using Fresnel terms calculated from the mean surface normal as specular and diffuse coefficients. This approach is justified for a perfectly smooth surface, but does not account for the effects of surface roughness.

Wolff et al. [11] incorporate Fresnel theory in a model that attempts to unify diffuse reflectance from rough and smooth surfaces. Walter et al. [10] on the other hand develop a microfacet model for refraction and transmission through rough surfaces, such as etched glass. The model is compared with real world data, and the effect of the microfacet distribution is investigated. In order to implement the method for ray-tracing, an efficient importance sampling scheme is presented.

In this paper, we seek to develop a unified model of specular and diffuse reflectance from microfacet models by describing both components of reflectance in terms of Fresnel theory and the same model of surface roughness. The result is that we remove the need for specular and diffuse coefficients completely. The directional dependence and relative strength of the two components is described entirely in terms of physically meaningful parameters. Our model is more highly constrained than the models mentioned above. For example, it does not reduce to Lambertian reflectance for any choice of parameters. For this reason, fitting our model to observed data is also more constrained. The hope is that this reduces the search space and hence the likelihood of an accurate solution. In addition, we obtain parameter estimates that all have physical meaning. Our model also has the property that the proportion of light available for diffuse reflectance depends on the angle of incidence. For frontal illumination, this exaggerates the effect captured by Lambert’s law by further reducing diffuse reflectance at large angles of incidence.

We begin by introducing some preliminaries of reflectance from rough surfaces. We then derive our unified model of specular and diffuse reflectance and show some of the characteristics of the model. In our experimental section we fit the model to data from uncalibrated images and compare the results with those obtained using existing models. Finally, we present our conclusions and suggest some directions for future work.

2. Reflectance from Surfaces

When incident light strikes a surface, it may interact with it in a number of ways. The light may be absorbed, scattered or transmitted or a combination of these effects. We make the assumption that reflectance effects are local. That is, light leaving a point is due only to light arriving at this point. We wish to describe the way in which light interacts with a surface patch $dA$ whose mean surface normal is $N$. We use a local coordinate system in which the mean surface normal coincides with the $z$ axis. We assume the surface is isotropic and therefore, without loss of generality, may locate the light source vector, $L$, in the $xz$ plane and specify it using only the angle $\theta_s$. The vector describing the direction of the viewer, $V$, is specified in spherical coordinates as $(\theta_v, \alpha_v)$. Because of surface isotropy, $0 \leq \alpha_v \leq \pi$. These vectors and angles are shown in Figure 1. Since we are concerned with rough surfaces, we also specify the local surface normal of a microfacet, $N_f$, which may not coincide with the mean surface normal.

If the illumination, viewer and normal direction vectors are specified in global coordinates, we may convert these to angular representations in the local coordinate system as follows:

$$\theta_s = \arccos(N \cdot L), \quad \theta_v = \arccos(N \cdot V)$$

$$\theta_{s_f} = \arccos(L \cdot V), \quad \cos \alpha_f = \frac{\cos \theta_{s_f} - \cos \theta_s \cos \theta_v}{\sin \theta_s \sin \theta_v}. \quad (1)$$

The formula for $\alpha_f$ follows from applying spherical trigonometry to the spherical triangle formed by $N$, $L$ and
V.

In addition, we can compute the angle of incidence on a microfacet from the microfacet orientation, \((\theta_f, \alpha_f)\), and angle of incidence with respect to the mean surface normal, \(\theta_i\). This is done by applying the spherical law of cosines to the spherical triangle formed by \(N_f\), \(N\) and \(L\):

\[
\theta_i' = \arccos \left[ \cos \theta_i \cos \theta_f + \sqrt{1 - \cos^2 \theta_i \sin \theta_f \cos \alpha_f} \right].
\]

2.1. Surface Roughness Model

Models of surface roughness can be divided into those based on a distribution of surface heights \([1, 9]\) and of surface orientations \([7, 8]\). The former is preferable from a theoretical standpoint since it corresponds to physically realisable geometry. However, such a model proves hard to work with from an analytical point of view. On the other hand, the latter are easier to use but in general are based on an analysis of inconsistent geometry corresponding to a surface that cannot exist. For the purposes of our model, the choice between the two is not important and, for analytical simplicity, we chose to model surface roughness as distributions of surface orientations.

The surface roughness model we use is that of Torrance and Sparrow \([7]\). This model assumes that the surface is comprised of long, symmetric v-grooves and that the direction of the cavities has no preferred direction (i.e., the surface is isotropic). The surface normal of a facet is denoted \(N_f\), or in spherical coordinates, \((\theta_f, \alpha_f)\). The orientation of the facets deviate from the mean surface normal according to a distribution function. Two representations of facet orientation distributions have been considered in the literature. Assuming surface roughness is isotropic, the distribution is a function of only \(\theta_f\). Torrance and Sparrow \([7]\) use a distribution, \(N(\theta_f, \nu)\), which represents the number of facets per unit surface area with slope angle \(\theta_f\). The parameter \(\nu\) is related to surface roughness. This proves easier to use for their model of specular reflectance. Oren and Nayar \([5]\) on the other hand use the slope-area distribution, \(P(\theta_f, \nu)\), which represents the fraction of the surface area occupied by facets with slope angle \(\theta_f\). We make use of both functions in this paper. The two are related by:

\[
P(\theta_f, \nu) = N(\theta_f, \nu)da \cos \theta_f,
\]

where \(da\) is the surface area of a microfacet which is assumed equal for all facets.

A number of facet distribution functions have been proposed. In this paper, we use the Beckmann distribution \([1]\):

\[
P(\theta_f, \nu) = C \frac{1}{\nu^2 \cos^4 \theta_f} e^{-\left[ \frac{\sin^2 \theta_f}{\nu^2} \right]^2}.
\]

Since \(\theta_f\) lies in the closed interval \([0, \pi]\), it holds that \(\int_0^{\pi} P(\theta_f, \nu)da \cos \theta_f = 1\). Hence, the constant \(C\) is given by:

\[
C = \frac{4\nu}{\sqrt{\pi(2 + \nu^2)}}.
\]

The parameter \(\nu\) is the root mean square slope of the microfacets. For real surfaces, \(\nu\) rarely exceeds 0.5. For perfectly smooth surfaces \((\nu = 0)\), the function becomes a Dirac delta function (i.e., all microfacets are oriented in the same direction as the local average surface normal).

3. A Unified Model

Our model (see Figure 2) is based on the following assumptions:

1. The surface is comprised of isotropically distributed, symmetric v-cavities, each facet of which behaves as a dielectric, reflecting some light in the mirror direction, the quantity of which is determined by Fresnel’s equations.

2. Light not reflected specularly enters the surface and is internally scattered. Some light is absorbed, the remainder is remitted and it is assumed that this light is completely diffused by the internal scattering processes.

3. We assume that light does not exit a different facet to the one it entered. In other words, we assume that the subsurface transport length is much smaller than the characteristic scale of the surface roughness.

4. Both the specular and diffuse reflectance are subject to shadowing and masking constraints imposed by the geometry of the facets.

The model requires only three physically meaningful parameters: the refractive index of the surface \((n)\), the surface roughness \((\nu)\) and the albedo of the surface \((\rho_a)\). In the following subsections we describe the various components of this model. We start by describing how shadowing and
masking is modelled. We then describe the specular and diffuse components of reflectance before finally showing the complete model.

3.1. Masking and Shadowing of Microfacets

We employ the same masking and shadowing model as Oren and Nayar [see (5) for derivation]. If the orientation of a facet is specified by \((\theta_f, \alpha_f)\), the proportion of the facet that is shadowed from a light source from direction \((\theta_r, \alpha_r)\) by its adjacent facet is given by: \(G_s(\theta_r, \alpha_r, \theta_f, \alpha_f) = \max(0, \min(1, GAF))\), where

\[
GAF = \frac{2 \cos \theta_r \cos \theta_f}{\cos \theta_r \cos \theta_f + \sin \theta_r \sin \theta_f \cos(\alpha_r - \alpha_f)}.
\]

(6)

This term includes self shadowing, i.e. when \(\theta'_r > \frac{\pi}{2}\). A similar expression can be obtained for masking. We denote the proportion of a facet that is both illuminated and visible by \(G(\theta_l, \theta_r, \alpha_r, \theta_f, \alpha_f)\).

3.2. Fresnel Reflectance

If the refractive index of the surface is \(n\) and the incident light is unpolarised, then the total fraction of light that is reflected when a ray strikes a facet with incident angle \(\theta'_i\) is given by Fresnel’s equation:

\[
R(\theta'_i, \eta) = \frac{(c - \eta g)^2 + (g - \eta c)^2}{2(c + \eta g)^2},
\]

where \(c = \cos(\theta'_i)\) and

\[
g = \sqrt{1 - \frac{\sin(\theta'_i)^2}{\eta^2}}.
\]

(7)

(8)

We use this term to compute the proportion of light that is specularly reflected from the surface and that which enters the surface and is subsequently diffusely reflected.

3.3. Specular Reflectance

Only those microfacets whose normal vectors lie within a solid angle \(d\omega'\) around the vector \(\mathbf{H}\), halfway between the light source and viewer (i.e. where \(\theta'_r = \theta'_r\)), are capable of specularly reflecting light into the solid angle \(d\omega\) around the view direction vector. The vector \(\mathbf{H}\) is computed as:

\[
\mathbf{H} = \frac{\mathbf{L} + \mathbf{V}}{||\mathbf{L} + \mathbf{V}||},
\]

(9)

which we refer to in spherical coordinates as: \((\theta_f^{\text{spec}}, \alpha_f^{\text{spec}})\). These are computed from \(\theta_f^{\text{spec}} = \arccos(N \cdot H)\) and

\[
\cos \alpha_f^{\text{spec}} = \csc \theta_f^{\text{spec}} \csc \theta_i \left( \cos \frac{\theta_f^{\text{spec}}}{2} - \cos \theta_f^{\text{spec}} \cos \theta_i \right).
\]

(10)

The angle of incidence on this microfacet is given simply by: \(\theta'_i^{\text{spec}} = \frac{\theta_i}{\eta}\).

The radiance in the direction \((\theta_r, \alpha_r)\) due to specular reflectance from a surface patch illuminated by a light source with incident angle \(\theta_i\) and radiance \(L_i\), accounting for masking and shadowing of facets, is given by:

\[
L_{r,s} = \frac{R(\theta'_i^{\text{spec}}, \eta)G(\theta_i, \theta_r, \alpha_r, \theta_f^{\text{spec}}, \alpha_f^{\text{spec}}) \times L_i d\omega_i d\omega(\theta_f^{\text{spec}}, \nu) d\omega' \cos \theta'_i^{\text{spec}}}{d\omega_r \cos \theta_r}.
\]

(11)

This is the Torrance and Sparrow [7] model of specular reflectance. The solid angle terms \(d\omega'\) and \(d\omega_r\) are related by:

\[
d\omega' = \frac{d\omega_r}{4 \cos \theta_i^{\text{spec}}}. \quad (12)
\]

The development of our diffuse model is made simpler if roughness is described using the slope area probability distribution. We may describe specular reflectance in terms of the same function. Substituting (3) and (12) into (11) gives:

\[
\frac{L_{r,s}}{K} = \frac{R(\theta'_i^{\text{spec}}, \eta)G(\theta_i, \theta_r, \alpha_r, \theta_f^{\text{spec}}, \alpha_f^{\text{spec}}) \times P(\theta_f^{\text{spec}}, \nu)}{4 \cos \theta_r \cos \theta_i^{\text{spec}}},
\]

(13)

where \(K = L_i d\omega_i\) is a constant determined by the illuminant.

3.4. Diffuse Reflectance

The key contribution of this paper is to extend the formulation of specular reflectance from microfacets given above to consider the effect of microfacets on diffuse reflectance. The Lambertian BRDF assumes that all incoming light on a surface patch is scattered and hence diffusely reflected. However, from our discussion above, we know that a proportion of the light striking each microfacet is reflected specularly. Therefore, we must downweight the diffuse reflectance by a factor that uses Fresnel theory to compute the proportion of light that is transmitted into the surface and subsequently diffusely reflected, as opposed to that which is reflected specularly (either towards the viewer or in a different direction). The resulting expression for radiance due to diffuse reflectance is dependent on the same physical values as the specular expression.

3.4.1 Radiance from a Lambertian patch

We begin by considering a surface patch in which all incident light penetrates the surface and is scattered equally in all directions. If the patch is illuminated by a source with radiance \(L_i\) and solid angle \(d\omega_i\), the surface radiance is given by:

\[
L_{r,L}(\theta_i, \rho_d) = \frac{P_d}{\pi} L_i d\omega_i \cos \theta_i,
\]

(14)
where \( \rho_d \in [0, 1] \) is the intrinsic reflectivity or albedo of the surface. This is the proportion of light that is absorbed by the surface.

### 3.4.2 Diffuse radiance from a rough surface composed of shiny microfacets

If some of the incident light on a microfacet is reflected specularly, the remaining proportion of the incident light that penetrates the surface and is subsequently diffusely scattered is dependent on the refractive index of the surface and the angle of incidence:

\[
T(\theta_i', \eta) = 1 - R(\theta_i', \eta). \tag{15}
\]

We may now define the quantity \( p_T \) as the average proportion of light that penetrates the surface of a facet according to our rough surface model. This term takes into account shadowing of the light source and Fresnel transmission into the surface:

\[
p_T(\theta_i, \eta, \nu) = \frac{1}{2\pi} \int_0^{2\pi} P(\theta_f, \nu) \times \\
\int_0^{2\pi} G_s(\theta_i, \theta_f, \alpha_f) T(\theta_i', \theta_i', \theta_f, \alpha_f, \eta) \, d\alpha_f \, d\theta_f.
\]

Note that this term is independent of the viewing direction.

The radiance due to diffuse reflectance from a surface patch in our rough surface model is therefore the average radiance from each facet over the hemisphere of possible facet orientations, weighted by the facet distribution function. The radiance from each facet is simply the Lambertian radiance, (14) weighted by the proportion of incident light transmitted into the surface, (16). We also account for masking with respect to the viewer and shadowing with respect to the illumination. The radiance from a surface patch is therefore given by:

\[
L_{r,d}(\theta_i, \theta_r, \alpha_r, \nu, \eta, \rho_d) = \frac{K \rho_d}{2\pi} \int_0^{2\pi} P(\theta_f, \nu) \int_0^{2\pi} G(\theta_i, \theta_r, \alpha_r, \theta_f, \alpha_f) \times \\
T(\theta_i', \theta_f, \alpha_f, \eta) \cos(\theta_i', \theta_f, \alpha_f) \, d\alpha_f \, d\theta_f.
\]

where \( K = L_0 \omega_1 \) is the same constant as in (13).

### 3.4.3 Characteristics of the diffuse model

Before presenting the combined specular and diffuse model, we briefly illustrate the characteristics of the diffuse component of the model. We begin by examining the effect of the downweighting term \( p_T \). In the left hand column of Figure 3, we plot \( p_T \) against \( \theta_i \) for various values of \( \eta \) and \( \nu \). In the top left panel, \( \eta = 1.5 \) and \( \nu = \{0.0, 0.2, 0.4, 0.6, 0.8, 1\} \), in the bottom left panel \( \nu = 0.25 \) and \( \eta = \{1 + \epsilon, 1.01, 1.2, 1.5, 2, 4\} \), where \( \epsilon = 2.22 \times 10^{-16} \). We use \( \nu = \epsilon \) and \( \eta = 1 + \epsilon \) as approximations to \( \nu = 0 \) and \( \eta = 1 \), since \( p_T \) cannot be evaluated for these values.

The first feature to note about the plots is that less light enters the surface as the angle of incidence increases. The incident angle at which this effect becomes noticeable increases as surface roughness decreases. As the index of refraction increases, the proportion of light entering the surface decreases except for large angles of incidence. As one would expect, this will result in a greater proportion of the reflectance being specular. In the right hand column of Figure 3 we show diffuse radiance versus incident angle for the same parameters as in the left hand column. Here it is clear that the overall magnitude of diffuse radiance decreases as either surface roughness or refractive index increases. In Figures 4 and 5 we show the diffuse radiance from spheres with the same surface properties as in the plots. In both cases, the first row is frontal illumination while in the second, the light source is from 45° to left.

### 3.5 The combined model

The radiance predicted by our unified model of specular and diffuse reflectance is given simply by:

\[
L_r(\theta_i, \theta_r, \alpha_r, \nu, \eta, \rho_d) = \\
L_{r,s}(\theta_i, \theta_r, \alpha_r, \nu, \eta) + L_{r,d}(\theta_i, \theta_r, \alpha_r, \nu, \eta, \rho_d). \tag{17}
\]

Note that: 1. both specular and diffuse reflectance are described in terms of the same parameters, 2. there is no specular or diffuse coefficient to weight the relative contributions of the two terms.
3.6. Implementation

The diffuse component of our model given in (16) requires solution of a double integral. Finding an analytical solution to this may be impossible. An approximation with a simpler form is likely to exist, however in this paper we wish to evaluate the model precisely without introducing additional sources of error and hence solve the integral using numerical methods. For efficient implementation, precomputed values for the diffuse component of the model can be stored in a lookup table. Because the diffuse reflectance function is smooth, this lookup table can be sparse and intermediate values interpolated with no significant loss of accuracy. The specular component can be computed directly and so does not require the use of a lookup table.

4. Experiments

In this section we present an experimental evaluation of our reflectance model on real world data. We begin by comparing the performance of our model against that of a comparable existing model when fitting to measured reflectance data. We then show how our model can be used to estimate surface properties that correspond to physically meaningful quantities from uncalibrated imagery.

4.1. Fitting to BRDF data

We begin by evaluating the performance of our model when fitted to measured reflectance data. We use data from the Columbia-Utrecht Reflectance and Texture (CURET) Database [4] which contains reflectance measurements for 61 materials with widely varying reflectance properties. This is a challenging dataset as the range of viewing and illumination directions is broad and includes near retro-reflection, grazing angles and views coincident with specular spike. Further, the samples are non-homogenous. The distribution of viewing and illumination directions is shown in Figure 6. We show results for an illustrative subset of these materials, focussing on those which exhibit both substantial roughness and specular reflection.

We perform a constrained non-linear optimisation over the parameters $\nu$, $\eta$ and $\rho_d$ (the constraints are to ensure physical plausibility of the parameters, such as $0 \leq \rho_d \leq$...
1. We aim to minimise the sum of squared errors between the observed and fitted radiance values over the 205 reflectance measurements. To provide comparison, we also fit a combination of the Oren and Nayar and Torrance and Sparrow models (ON+TS) as described in [5]. This model has 5 free parameters. Because our model requires less parameters and is more highly constrained, the fitting process typically converges in less than 10 iterations. The ON+TS model typically requires 50-100 iterations to converge.

The results of this experiment are shown in Table 1. For all of the samples, our model outperforms the ON+TS model, even though it has fewer degrees of freedom. We also include results of using a simple, one parameter Lambertian model for comparison. To observe the distribution of errors, in Figure 8 we show the observed versus predicted normalised radiance values for our model (top row) and the ON+TS model (bottom row). A perfect fit would see all the points distributed along the diagonal.

### 4.2. Measuring surface properties from images

In this experiment we demonstrate how surface properties which correspond to physically meaningful quantities can be estimated from uncalibrated imagery. For this study, we focus on cylindrical rolls of sandpaper of varying grade. The approximate particle size of the grains are known for each sample and we assume that larger average grain sizes will result in larger deviations for the mean normal direction and hence higher surface roughness. As the rolls are cylindrical, normal direction can be inferred from pixel position. We use a camera with linear response. As in our previous experiment, we assume nothing about the strength of the light source or the camera properties, treating them as an unknown linear scale factor. We find optimal values of refractive index, roughness and diffuse albedo by fitting our model to the measured brightness values using the estimated normal directions to infer the reflectance geometry. We use approximately frontal lighting provided by a flash. The input images are shown in Figure 9. Note that the backing paper and particle material differs between samples so some are more specular than others.

In Table 2 we show our results. For each sample we show the ISO 6344 sandpaper grade along with the approximate particle diameters. As predicted, our estimated roughness parameter reduces as the particle diameter decreases.
5. Conclusions

We have presented a model of diffuse and specular reflectance from rough surfaces composed of glossy microfacets. The two components of our model are both expressed in terms of Fresnel theory and use the same distribution of facet orientation. The relative contribution of the two components is therefore entirely determined by three physically meaningful parameters. The resulting model is more highly constrained than previous adhoc combinations of specular and diffuse reflectance models, yet still fits measured reflectance data more accurately. We have shown how the model can be used to estimate real world macroscopic surface properties from uncalibrated imagery.

In future work, we intend to seek a more efficient, analytical expression for the diffuse reflectance. The model could also combine naturally with any subsurface spectral reflectance model which takes as input the wavelength-dependent quantity of light penetrating the surface.

References


