

Physically Based Shading

物理原理

光学现象

在物理世界中，光主要会与三种物质交互，他们产生的光学现象构成了人们眼中观察到的世界。

粒子 (Particles) 主要指气体，分子在其中的位置分布是随机且不相关的。当粒子为大气分子时，光在其中传播会发生瑞利散射 (Rayleigh Scattering)，而当粒子为固体时，发生的散射叫做丁达尔散射 (Tyndall Scattering)。

介质 (Media) 主要指均匀质 (Homogeneous Medium)，光在其中传播会发生吸收和散射，吸收会决定介质的颜色，散射会决定介质的浑浊程度。

表面 (Surfaces) 即各种物体表面，光与其交互会发生反射和折射现象。而折射光进入介质中，又会进一步与介质进行交互。

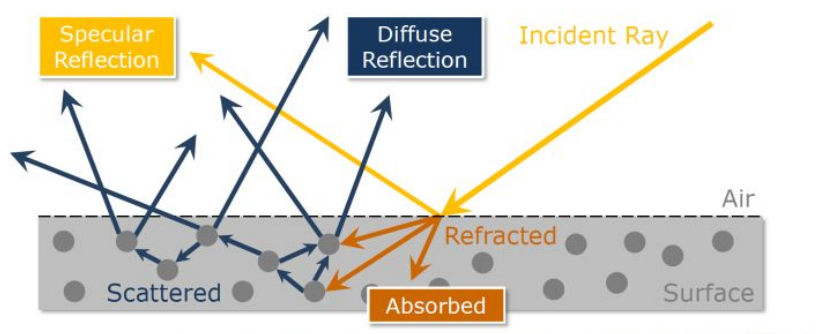


图1 入射光与表面交互图示。

光与表面交互

光与非光学平坦表面 (Non-Optically-Flat Surfaces) 交互时，可以看成光与无数个微小的光学平坦表面交互的集合。物体越粗糙，微小平面越起伏，从而反射光也越模糊。在渲染中，模拟这个物理现象的模型称为微表面模型 (Microfacet Theory)。

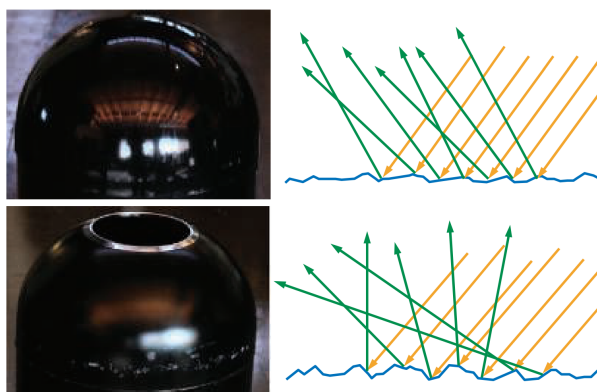


图2 上图模型光滑，微表面起伏较小，反射清晰。下图模型粗糙，微表面起伏较大，反射模糊。

以上理论描述了光与表面交互的反射现象，而折射现象与具体的物体类型有关。

在金属（导体，Conductor）中，折射光的能量会立即被自由电子吸收；在非金属（电介质，Dielectric）中，折射与光在介质中的现象相同，会发生散射和吸收。

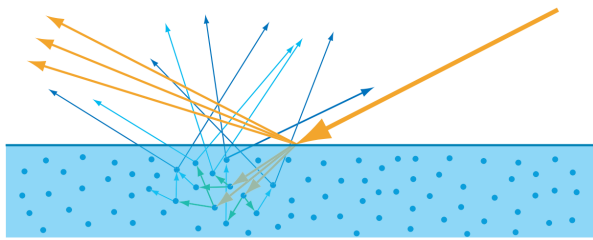


图3 入射光在表面发生折射与反射，折射光在介质内部经过散射与吸收后从表面重新出射。

折射光与物体内部介质交互，经过散射和吸收后，重新从表面出射的现象，称为次表面散射（Subsurface Scattering）。漫反射（Diffuse）和次表面散射在微观本质上指的是同一种物理现象。但从宏观上，当次表面散射的出射光位置与入射光位置的距离差相对于像素大小可以忽略时，一般称之为漫反射，或局部次表面散射（Local Subsurface Scattering），反之称之为全局次表面散射（Global Subsurface Scattering）。

本文不考虑光线透射和全局次表面散射，总的来说，我们可以把着色拆分成衡量微表面散射的高光项（Specular term）和衡量局部次表面散射的漫反射项（Diffuse Term）。

基于物理的渲染理论

渲染方程与BRDF

Kajiya给出完整的渲染方程（Rendering Equation）如下：

$$L_o(\mathbf{p}, \mathbf{v}) = L_e(\mathbf{p}, \mathbf{v}) + \int_{\mathbf{l} \in \Omega} f(\mathbf{l}, \mathbf{v}) L_i(\mathbf{p}, \mathbf{l}) (\mathbf{n} \cdot \mathbf{l})^+ d\mathbf{l}$$

等式左边为表面位置 \mathbf{p} 在观察方向 \mathbf{v} 的辐亮度（Radiance）。等式右边第一项为 \mathbf{p} 点在 \mathbf{v} 方向的自发光辐亮度，第二项为入射辐亮度 L_i 反射至 \mathbf{v} 方向的辐亮度，对入射方向 \mathbf{l} 作半球积分（ Ω 为半球面）。

如果我们不考虑自发光，就可以将渲染方程简化为反射率方程（Reflectance Equation）：

$$L_o(\mathbf{p}, \mathbf{v}) = \int_{\mathbf{l} \in \Omega} f(\mathbf{l}, \mathbf{v}) L_i(\mathbf{p}, \mathbf{l}) (\mathbf{n} \cdot \mathbf{l})^+ d\mathbf{l}$$

方程中 $(\mathbf{n} \cdot \mathbf{l})^+$ 为入射方向 \mathbf{l} 与表面法向 \mathbf{n} 夹角余弦值，它描述了入射角带来的入射辐亮度衰减。可以类比于太阳光直射赤道时，纬度越高，地球上单位面积接收到的辐射通量越少。

方程中 $f(\mathbf{l}, \mathbf{v})$ 描述了如何从 \mathbf{l} 方向入射辐亮度得到 \mathbf{v} 方向出射辐亮度，它叫作双向反射分布函数（Bidirectional Reflectance Distribution Function, BRDF）。在以 \mathbf{p} 点为原点的局部坐标系下， \mathbf{l} 与 \mathbf{v} 分别需要两个参数决定：与法向 \mathbf{n} 的夹角 θ ，与切线 \mathbf{t} 的夹角 ϕ 。

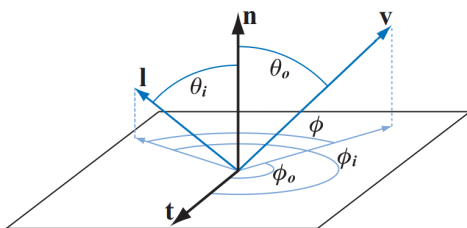


图4 局部坐标系下的入射方向与观察方向。

对立体角 Ω 的半球积分可以具体写成：

$$\int_{\Omega} d\Omega = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\frac{\pi}{2}} \sin \theta d\theta d\phi$$

记 $\mu_i = \cos \theta_i$ 与 $\mu_o = \cos \theta_o$ ，反射率方程可以写作：

$$L_o(\mu_o, \phi_o) = \int_{\phi_i=0}^{2\pi} \int_{\mu_i=0}^1 f(\mu_i, \phi_i, \mu_o, \phi_o) L(\mu_i, \phi_i) \mu_i d\mu_i d\phi_i$$

显然其中的BRDF方程有4个参数，而如果物体材质是各向同性的，那么BRDF与方位角 ϕ_i ， ϕ_o 的绝对大小是无关的，仅与它们的相对大小有关，即方位角之差 ϕ 。所以各向同性的BRDF仅3个参数，这也是绝大多数物体的属性。

BRDF并不是一个随意的函数，物理规律使得它需要满足两条原则。

第一，赫尔姆霍兹互换原则（Helmholtz Reciprocity），即入射与出射方向互换后，函数值不变，数学公式如下：

$$f(\mathbf{l}, \mathbf{v}) = f(\mathbf{v}, \mathbf{l})$$

第二，能量守恒原则（Conservation of Energy），即出射能量不能大于入射能量。这也是基于物理的着色模型与基于经验的着色模型的主要区别之一。定向半球反射率（Directional-Hemispherical Reflectance）反应了BRDF的能量损失，数学公式如下：

$$R(\mathbf{l}) = \int_{\mathbf{v} \in \Omega} f(\mathbf{l}, \mathbf{v})(\mathbf{n} \cdot \mathbf{v}) d\mathbf{v}$$

它的含义是在方向 \mathbf{l} 入射辐照度为1时，BRDF在各个方向出射辐照度的积分。由于能量守恒， $R(\mathbf{l})$ 函数的值域为 $[0, 1]$ ， $R(\mathbf{l}) = 0$ 表示入射光能量被全部吸收。

菲涅尔效应

菲涅尔效应（Fresnel Effect）指的是，光线照射某种材质表面时被反射与折射的比例，随着入射角变化而变化。入射角越大，反射越明显。



图5 菲涅尔效应。直接向下看水面时（垂直入射），被反射的比例小，所以可以直接看到水底；

而向远处看水面时（掠射），被反射的比例高，可以看到天空的倒影。

菲涅尔函数 $F(\theta_i)$ 定义为入射光被反射的比例，当材质一定时，它仅于入射角有关。当 $\theta_i = 0^\circ$ 时，方程反应了垂直入射光被反射的比例，它是材质的一个属性，记为 F_0 ；当 $\theta_i = 90^\circ$ 时，光线入射任何物质都被完全反射，即 $F_{90} = 1$ 。

同一材质，对不同波长的入射光反射比例也是不同的，所以 $F(\theta_i)$ 是一个RGB的矢量函数。对于非金属， $F(\theta_i)$ 对波长的变化非常不明显，且 F_0 值较小，通常在0.06以下；而对于金属， $F(\theta_i)$ 对波长的变化非常明显，且 F_0 值往往在0.5以上。具体实现上，非金属 F_0 通常保存为一个标量，金属 F_0 则保存为一个RGB值。

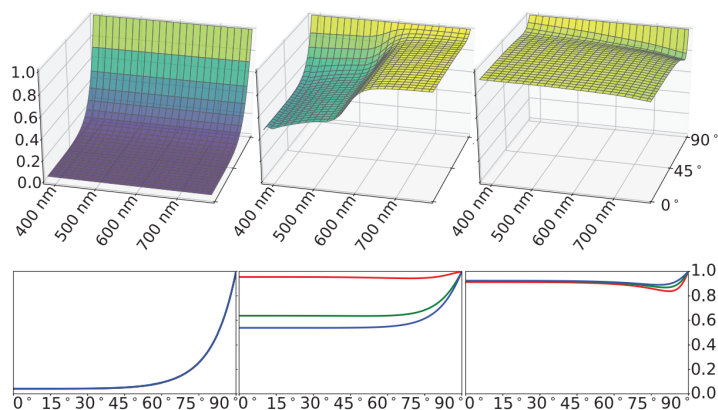


图6 三种材质在不同入射角对不同波长的菲涅尔值。

从左至右分别为玻璃、铜、铝。第一行为不同波长菲涅尔值，第二行为转换为RGB的菲涅尔值。

对于菲涅尔函数，Schlick给出了近似：

$$F(\mathbf{n}, \mathbf{l}) \approx F_0 + (F_{90} - F_0)(1 - (\mathbf{n} \cdot \mathbf{l})^+)^5$$

上式也可以写作：

$$F_c = (1 - (\mathbf{n} \cdot \mathbf{l})^+)^5$$

$$F(\mathbf{n}, \mathbf{l}) \approx F_c F_{90} + (1 - F_c) F_0$$

这样可以将Schlick近似理解为用 F_c 插值 F_0 与 F_{90} ，其中 $F_{90} = 1$ 。

微表面理论

物体表面的微小起伏使得表面粗糙度不同，不过这种起伏远小于一个像素，所以物体模型不可能如此精细。而这些起伏会影响光在物体表面的反射，所以可以使用BRDF模型从统计学上模拟这些微小平面整体对光照着色的影响。微表面理论起源于光学界，由Blinn于1977年以及Cook和Torrance在1981年引入计算机图形学^{[4][5]}。

法向分布函数

在微观上，记每个微平面的法向为 \mathbf{m} ，且它们都遵循一个微观的BRDF $f_\mu(\mathbf{l}, \mathbf{v}, \mathbf{m})$ 来反射光线，一般将可以这个BRDF直接简化为理想的菲涅尔镜面反射。将所有微观BRDF加起来，就得到了宏观上整个表面的BRDF。

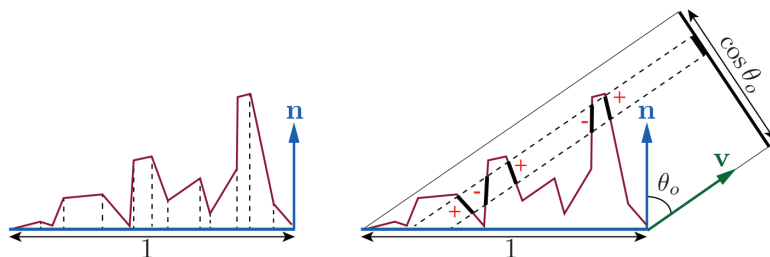


图7 将微表面投影至宏观表面或其他平面。

微表面模型BRDF的一个重要属性就是微平面的法向统计分布，我们使用法向分布函数（Normal Distribution Function, NDF）来衡量，记作 $D(\mathbf{m})$ [6]。它的具体含义是：法向为 \mathbf{m} 的微表面面积。所以对 $D(\mathbf{m})$ 积分将得到微表面总面积，而对 $D(\mathbf{m})$ 在宏观表面上的投影积分将得到宏观表面片元的面积，即 $D(\mathbf{m})$ 需要满足：

$$\int_{\mathbf{m} \in \Theta} D(\mathbf{m})(\mathbf{n} \cdot \mathbf{m}) d\mathbf{m} = 1$$

其中， Θ 表示对整个球面积分， \mathbf{n} 为宏观表面法向。简便起见，规定宏观表面片元面积为1。

同时，将微表面和宏观表面都投影至与观察方向垂直的表面上，得到的面积是相等的，即：

$$\int_{\mathbf{m} \in \Theta} D(\mathbf{m})(\mathbf{v} \cdot \mathbf{m}) d\mathbf{m} = \mathbf{v} \cdot \mathbf{n}$$

几何函数

遮蔽函数

公式x中， $(\mathbf{v} \cdot \mathbf{m})$ 有可能小于零，如图7所示，正是 $(\mathbf{v} \cdot \mathbf{m})$ 项正负抵消，才得到了正确的结果。而如果我们只考虑从观察方向 \mathbf{v} 可见的微平面，就需要引入一个遮蔽函数（Masking Function），记为 $G_1(\mathbf{m}, \mathbf{v})$ 。它描述了法向为 \mathbf{m} 的微表面中，从 \mathbf{v} 方向可见的比例。所以 $G_1(\mathbf{m}, \mathbf{v})D(\mathbf{m})$ 就描述了可见的微表面法向分布。同时，公式x可以写为：

$$\int_{\mathbf{m} \in \Theta} G_1(\mathbf{m}, \mathbf{v}) D(\mathbf{m})(\mathbf{v} \cdot \mathbf{m})^+ d\mathbf{m} = \mathbf{v} \cdot \mathbf{n}$$

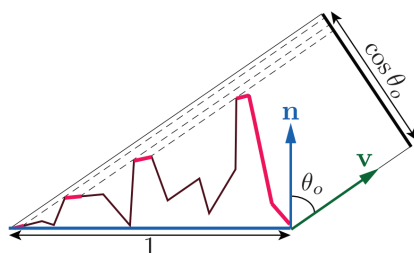


图8 几何遮蔽函数作用示意图。

NDF描述了不同朝向的微表面的统计分布，而遮蔽函数 G_1 则隐含了这些不同朝向的微表面在空间上如何排列的信息。Smith遮蔽函数[7]适用于任意NDF。Heitz证明了它具有良好的性质[6]，一方面，它满足公式x，另一方面，它与微表面具体的法向 \mathbf{m} 是无关的。公式如下：

$$G_1(\mathbf{m}, \mathbf{v}) = \frac{\chi^+(\mathbf{m} \cdot \mathbf{v})}{1 + \Lambda(\mathbf{v})}$$

其中 $\chi^+(x)$ 为正特征函数。 Λ 函数需要从特定的NDF中推导，Walter等人[8]与Heitz[6]给出了方法。在微表面朝向和遮蔽情况无关（Normal-Masking Independence）的假设下，Smith G_1 是精确的，而当这点假设不成立时（比如布料纤维等材质），它的精确性就会下降。

遮蔽阴影函数

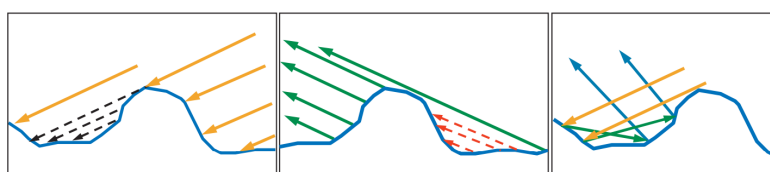


图9 阴影，遮蔽，多次弹射

如图九所示， G_1 函数仅考虑了从观察方向产生的微表面遮蔽情况，而关系在照射微表面时，会产生类似的现象，称之为阴影（Shadowing）。同时考虑两种遮挡的函数称之为遮蔽阴影函数（Masking-Shadowing Function），记作 $G_2(\mathbf{l}, \mathbf{v}, \mathbf{m})$ 。最简单的 G_2 函数假设两种遮挡是无关的，称为分离形 G_2 （Separable Form）公式如下：

$$G_2(\mathbf{l}, \mathbf{v}, \mathbf{m}) = G_1(\mathbf{m}, \mathbf{v})G_1(\mathbf{m}, \mathbf{l})$$

这种假设与现实不符，会造成渲染结果过暗。加入高度相关的因素，从Smith G_1 中可以推导出Smith高度相关的遮挡阴影函数（Smith Height-Correlated Masking-Shadowing Function）：

$$G_2(\mathbf{l}, \mathbf{v}, \mathbf{m}) = \frac{\chi^+(\mathbf{m} \cdot \mathbf{v})\chi^+(\mathbf{m} \cdot \mathbf{l})}{1 + \Lambda(\mathbf{v}) + \Lambda(\mathbf{l})}$$

Heitz还给出了一个结合方向相关（Direction-Correlated）与高度相关的Smith G_2 函数：

$$G_2(\mathbf{l}, \mathbf{v}, \mathbf{m}) = \frac{\chi^+(\mathbf{m} \cdot \mathbf{v})\chi^+(\mathbf{m} \cdot \mathbf{l})}{1 + \max(\Lambda(\mathbf{v}), \Lambda(\mathbf{l})) + \lambda(\mathbf{v}, \mathbf{l}) \min(\Lambda(\mathbf{v}), \Lambda(\mathbf{l}))}$$

$$\lambda(\phi) = 1 - e^{-7.3\phi^2}$$

由于公式x复杂度与Smith G_1 相似，所以广泛应用于实践中^{[9][10][11]}。

多次弹射

从图九中可以看到，除了简单的遮挡，光线还会在微表面之间进行多次弹射后出射。而我们所讨论的几何遮蔽函数都忽略了这一点，所以最终的渲染结果会偏暗，并且材质粗糙度越高，多次弹射现象越明显，能量损失也越严重。Kulla与Conty提出了多次弹射的BRDF项^[10]，达到了能量补偿的效果。

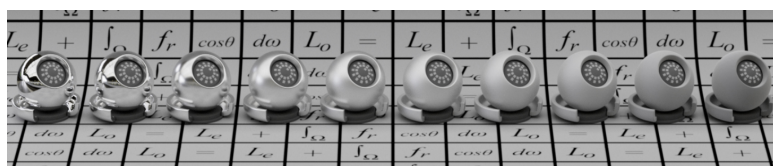


图10 多次弹射能量损失

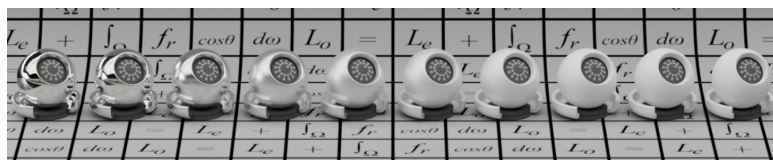


图11 Kulla-Conty多次弹射能量补偿

高光项BRDF

法向分布函数

多次弹射能量补偿

漫反射项BRDF

基于物理的环境光照

漫反射项光照预计算

高光项光照预计算

BRDF预计算

采样

低差异序列

半球面上均匀采样

半球面上余弦重要性采样

GGX重要性采样

参考文献

1. *Real-Time Rendering*
2. James T. Kajiya. 1986. The rendering equation. *SIGGRAPH Comput. Graph.* 20, 4 (Aug. 1986), 143–150. <https://doi.org/10.1145/15886.15902>
3. Schlick, C. (1994), An Inexpensive BRDF Model for Physically-based Rendering. *Computer Graphics Forum*, 13: 233-246. <https://doi.org/10.1111/1467-8659.1330233>
4. James F. Blinn. 1977. Models of light reflection for computer synthesized pictures. *SIGGRAPH Comput. Graph.* 11, 2 (Summer 1977), 192–198. <https://doi.org/10.1145/965141.563893>
5. Robert L. Cook and Kenneth E. Torrance. 1981. A reflectance model for computer graphics. *SIGGRAPH Comput. Graph.* 15, 3 (August 1981), 307–316. <https://doi.org/10.1145/965161.806819>
6. Heitz, Eric. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. *Journal of Computer Graphics Techniques (JCGT)*. 3.
7. B. Smith, Geometrical shadowing of a random rough surface, in *IEEE Transactions on Antennas and Propagation*, vol. 15, no. 5, pp. 668-671, September 1967, doi: 10.1109/TAP.1967.1138991.
8. Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. 2007. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques (EGSR'07)*. Eurographics Association, Goslar, DEU, 195–206.
9. Karis, Brian. 2013. Real Shading in Unreal Engine 4. Physically based shading in theory and practice. In *ACM SIGGRAPH 2013 Courses (SIGGRAPH '13)*. Association for Computing Machinery, New York, NY, USA, Article 22, 1–8. <https://doi.org/10.1145/2504435.2504457>
10. Kulla, Christopher, and Alejandro Conty. 2017. Revisiting Physically Based Shading at Imageworks. Physically based shading in theory and practice. In *ACM SIGGRAPH 2017 Courses (SIGGRAPH '17)*. Association for Computing Machinery, New York, NY, USA, Article 7, 1–8. <https://doi.org/10.1145/3084873.3084893>
11. Lagarde, S´ebastien, and Charles de Rousiers. 2014. Moving Frostbite to Physically Based Rendering. Physically based shading in theory and practice. In *ACM SIGGRAPH 2014 Courses (SIGGRAPH '14)*. Association for Computing Machinery, New York, NY, USA, Article 23, 1–8. <https://doi.org/10.1145/2614028.2615431>
- 12.

