第三部分

反射、折射和阴影

任何光线追踪渲染都会导致关于效果是否支持以及如何支持的一些设计决策。光线追踪的主要历史诉求是它能够很好地处理阴影、反射和折射。然而，当你真的坐下来实现了一个支持这些效果的系统，你将面临一些不明显的设计决策。本书的这一部分描述了几个具体的达成这些设计决策的方法。

虽然对一个清晰的玻璃球进行光线追踪是一蹴而就的，但面对更复杂的模型，就有点令人头疼了。例如，一个简单的装有水的玻璃杯就呈现了光线在三种不同类型的材质之间传递的行为：水和空气，玻璃和空气以及玻璃和水。为了使折射正确，被追踪的光线不仅需要知道自己穿过了哪个接触面，还得知道是从哪种材质穿入哪种材质的。指望美术工作者用这些接触面建模是有问题的；想象一下在玻璃杯中装满水。第[11](#第11章)章《自动处理嵌套体积的材质》中描述了一个巧妙并且经过实战考验的方法来处理这个问题。

困扰几乎所有光线追踪程序的一个问题是当法线贴图产生了物理上不可能的表面法向量时应该怎么处理。显然，用硬编码的条件分支方案忽略这些不可能的法向量会导致不和谐的颜色不连续。每个光线追踪器都有自己的小技巧来处理这个问题。第[12](#第12章)章《用基于微表面的阴影函数解决凹凸不连续的问题》中，提出了一个基于简单的统计学方法的简洁实现。

光线追踪的屏幕空间方法使其在生成精确的屏幕空间阴影方面特别强大，而不会有阴影贴图方法带来的所有锯齿问题。然而光线追踪的计算能快到可以进行实时交互的程度吗？第[13](#第13章)章《光线追踪阴影：维持实时帧率》提供了关于如何做到这一点的详细解释。

大部分简单的光线追踪器都会从眼睛发出光线。通常这些程序几乎不可能产生焦散——就像我们用聚焦光图案制作装有液体的玻璃杯、游泳池或者湖泊。用“几乎”来形容是因为计算结果伴随了太多的噪点。但是，将光线从光源发射到环境中是一种产生焦散的可行方法。事实上，这甚至可以通过实时计算完成，正如第[14](#第14章)章所描述的《使用DXR在单一散射介质中实现光线引导的体积水焦散》。

总之，一个基础的光线跟踪器是非常简单的。把光线跟踪器部署到生产环境需要对基本效果进行一些小心处理，本书的第三部分提供了处理这些基本效果的几种有用的方法。

Peter Shirley

**117**

第12章

通过基于微表面的阴影函数解决凹凸(Bump)突变的问题

Alejandro Conty Estevez, Pascal Lecocq, and Clifford Stein

Sony Pictures Imageworks

摘要

我们提出了一种技术来隐藏在使用强烈的凹凸(Bump)或法线(Normal)贴图模拟微观几何时出现的突兀的阴影终止线。我们的方法是基于微表面阴影函数的，简单而高效。我们并非渲染细节丰富而昂贵的高度场阴影，而是应用统计解决方案假设法线遵循接近正态的随机分布。我们也为GGX提供了有用的近似方差测量，否则它是未在解析（几何）中定义的。

12.1 介绍

凹凸(Bump)贴图被广泛用于游戏的实时渲染和电影的批量渲染中。它会增加表面上的高频细节而不像使用真实几何体或位移贴图进行渲染那么得费时。它还负责给表面添加最终细粒度的细微缺陷（译者注：非光滑物体表面的极小的坑坑洼洼）。

凹凸(Bump)贴图按法线方向扰动而非从底层几何体产生的方式工作，是从一张纹理贴图或者某些程序图案（译者注：程序纹理）得来。但是，与任何其他简化方法一样，它可能会产生不符合预期的人工痕迹（失真）——特别是像图[12.1](#图12_1)所示的众所周知的（阴影）硬边。这是因为当表面突然遮挡入射光线时，原本预期由于法线变化导致的光滑度衰减被中断。这个问题在法线没有扰动时不会出现，因为这种情况出现时辐照度已经降到0了。但是，凹凸(Bump)贴图在将法线朝入射光的方向倾斜时会产生将入射光线延伸（至阴影区域）的效果，使照亮区域穿过和阴影区域的硬边。



图12.1：使用了强凹凸贴图的布料模型的对比。原始结果（左）显示在阴影区边缘处光照突然降低，而我们的阴影技术（右）用更自然和视觉上令人愉悦的平滑过渡替代了这种效果（阴影硬边）。

**127**

我们通过应用受到微表面理论启发的阴影函数来解决这个问题。凹凸(Bump)贴图可以被认为是大规模的正态分布，通过对其属性进行假设，我们可以使用与广泛使用的GGX 微表面分布相同的阴影项。即使这些假设在许多情况下是错误的，但这些阴影项在实践中甚至当凹凸或法线贴图表现出非随机结构时也仍然有效。

12.2 之前的工作

据我们所知，尚未发布针对此硬边问题的具体解决方案。Max [5]的相关工作是在特写镜头中如何计算凹凸与凹凸互相产生的阴影细节，它基于在每个点上找到水平基准线的方法。Onoue[7]等人将这种技术扩展到弯曲表面。但这些方法虽然对于点到点产生的阴影是准确的，却需要辅助表和更多的查找。它们不适用于只是想解决硬边问题而非产生细节阴影的高频凹凸贴图。

然而，阴影硬边问题在几乎每个渲染引擎中都是一个问题，并且提供的解决方案通常只是缓和Bump（凹凸）贴图的强度或者使用Displacement（置换）贴图。我们的解决方案快速而简单，无需任何额外的数据或预计算。

另一方面，自从Cook和Torrance[2]提出微表面理论及其阴影术语以来，Heitz [4]、Walter等人[8]以及其他人对此进行了广泛的研究。我们从他们的工作中获取灵感，以获得合理的信息解决本文档中讨论的问题。

12.3 方法

问题的原因是扭曲法线会改变光辐照度的自然余弦衰减，使得被照亮的区域向阴影区域中侵入过多。由于用贴图模拟的表面（细节）只是虚构的，因此渲染器是不知道任何高度场阴影（没有实际的几何多边形存在），因此光线突然消失，如图[12.2](#图12_2)所示。尽管这些缺陷的存在是符合预期的，但可能会分散注意力，产生不必要的卡通外观（不真实）。艺术家们希望从照亮到阴影(的过渡)要平滑。

在图[12.3](#图12_3)中，我们展示了凹凸贴图如何模拟表面不存在的法线从而导致照亮区域过多侵入阴影区域。这是因为形成阴影的因素（如图所示）完全被忽略。在微表面理论中这个因素的术语被称为阴影或者遮罩，它是值域为[0,1]，根据光线和观察方向计算的值，用于维护双向散射分布（BSDF）的互换性。

**128**

|  |  |
| --- | --- |
|  |  |
|  |

图12.2：插图展示了使用强凹凸贴图看到的阴影硬边的类型。

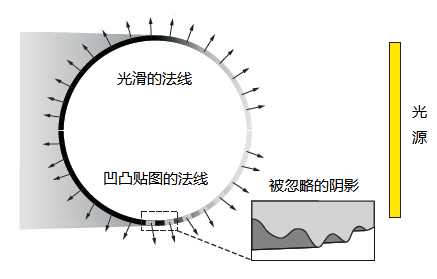


图12.3：在球体的上半部分，按照实际表面的平滑法线没有任何阴影硬边的问题。但是，下半部分引入的扭曲可能会使法线朝向光源倾斜，从而导致在非常接近应该被完全遮挡的地方产生了明亮的区域。这将忽略掉这样一个假想的表面应该会接收到的阴影。

**129**

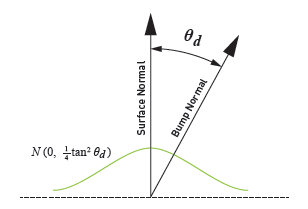


图12.4：基于凹凸法线的差异，我们想象一个以切线方向为极限的正态分布的两个标准差。这使得94％的其他凹凸法线更接近实际表面的方向。（译者注：根据正态分布，两个标准差之内的比率合起来为95%）

我们还使用Smith阴影近似方法进行凹凸贴图。它只降低从入射角到达的散射能量，这将使硬边优雅地变暗并融合照亮和黑暗区域而不改变其余的外观。它的推导需要知道是正态分布，这对于任意凹凸或法线贴图来说是未知的，我们会假设它是随机和正态分布的。这种假设在大部分情况下都不成立，但是它能够很好地进行阴影的渲染。

12.3.1 正态分布

我们选择GGX分布（曲线）是因为它简单并且高效的实现。和大多数分布（曲线）一样，它有一个用于调节微表面斜坡的蔓延（译者注：粗糙表面的凹凸不平）的粗糙度参数α。一个不易察觉的凹凸效果将对应于从低粗糙度到强烈凹凸的高粗糙度的α，主要是不知道如何找到这个α参数。

我们排除了从纹理贴图计算这个属性，有时他们是程序化生成的、不可预测的，我们希望避免任何预先计算过程。我们的想法是不需要额外信息仅根据我们在照明计算时获取到的凹凸法线来猜测α的值。也就是说，我们的猜测仅基于这个点本身来计算而不需要依赖邻近点的信息。

我们观察凹凸法线和真实表面法线的夹角的正切值，为了计算涵盖此法线的阴影项的合理概率，如图[12.4](#图12_4)所示，我们让这个正切值等于正态分布的两个标准差。然后，我们可以用GGX替换它并应用众所周知的阴影项

, (12.1)

其中是入射光方向与真实表面法线的夹角。

但是，这就产生了如何从分布方差中计算GGX的α的问题，GGX基于具有未定义的均值和方差的柯西（Cauchy）分布，Conty等人在数值上发现，

**130**

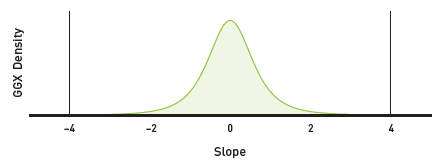


图12.5：如果我们截断GGX分布只存在于[-4α; 4α]区间，我们保留了94％的质量和斜率方差的数值结果始终如一地收敛到2。我们发现这个统计测量很好地表示了分布的视觉影响，否则会有未定义的动量。

[1]如果忽略长尾以保留大部分分布质量，则=2是GGX的方差的一个很好的近似值。见图[12.5](#图12_5)。因此，我们使用

, (12.2)

但我们将结果限制在[0,1]区间，这个测量反映了一个事实就是GGX显示的表面粗糙度高于贝克曼（Beckmann），其正切方差为。因为这样的关系这个等式大致上是。

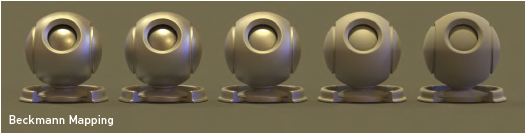
我们通过在一个发生大范围凹凸正态分布扰动的GGX表面上运行一个综合的视觉研究，验证了GGX方差的近似。我们使用了Olano和Dupy等人[3,6]的对mipmap纹理中的凹凸斜率分布的第一、第二时刻编码的法线过滤抗锯齿技术。对于每个像素，我们通过获取该像素所选择的mipmap过滤级别并相应地扩大GGX粗糙度来估算正态分布的方差。我们将我们的GGX方差关系与天真的贝克曼方差映射和以高采样率对未过滤的凹凸法线进行光线追踪作为参考的方法进行了比较。在所有场景中，我们的映射显示出更好地保持了由凹凸正态分布引起的感知GGX粗糙度，如图[12.6](#图12_6)所示。

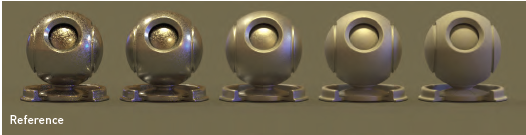
12.3.2 阴影函数

在一个典型的微表面BSDF中，对光和观察方向都要计算阴影/掩蔽项，以保持互换性。在我们的实现中，我们仅将我们的凹凸阴影应用于光线方向以尽可能地保留原始外观，因此会稍微打破这个属性。不像那些无阴影微表面的BSDF，凹凸贴图不会在观察角度上产生能量峰值，因此将公式[12.1](#公式12_1)应用于观察方向会使边缘变暗很多，如图[12.7](#图12_7)所示。如果这种效果造成问题，则可以使用完全可互换的阴影/遮蔽代替所有非主要光线。尽管如此，根据我们的经验，即使使用双向积分器，我们也没有发现任何问题。

我们根据入射角对入射光应用标量乘法。如果着色模型包含具有不同凹凸法线的多个BSDF，

**131**





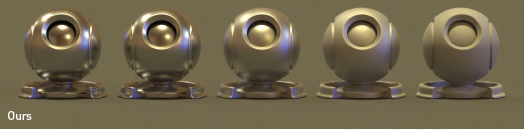


图12.6：根据过滤抗锯齿正态分布对GGX材质进行粗糙度扩大，使用常见的Beckmann方差映射（最上面那张图）和使用我们的GGX的方差近似（底下那张图），将两者与未过滤的参考（中间那张图）进行对比。在这个测试案例中，GGX基础表面粗糙度从0.01（左）到0.8（右）变化并显示我们的近似值更好地保留了由底层引起的整体感知粗糙度正态分布。

图12.7：左边：当网格呈现不规则的细分时，伪影会变得特别分散注意力，甚至显示出底层三角形。中间：应用阴影函数可以使硬边变得平滑并隐藏它们的伪影。右边：但如果我们试图使着色反转，我们就会使边缘产生非必要的变暗，特别是靠近脑袋的右上方。我们选择了中间的非可互换版本用于产品。

**132**

1. // 返回从法线（传入参数）夹角获取的alpha参数的平方
2. **float** bump\_alpha2(float3 N, float3 Nbump)
3. {
4. **float** cos\_d = min(fabsf(dot(N, Nbump)), 1.0f);
5. **float** tan2\_d = (1 - cos\_d \* cos\_d) / (cos\_d \* cos\_d);
6. **return** clamp(0.125f \* tan2\_d, 0.0f, 1.0f);
7. }
8. // 阴影因子
9. **float** bump\_shadowing\_function(float3 N, float3 Ld, **float** alpha2)
10. {
11. **float** cos\_i = max(fabsf(dot(N, Ld)), 1e-6f);
12. **float** tan2\_i = (1 - cos\_i \* cos\_i) / (cos\_i \* cos\_i);
13. **return** 2.0f / (1 + sqrtf (1 + alpha2 \* tan2\_i));
14. }

代码清单12.1：这两个函数足以实现硬边的修复。第二个可以用作入射光或BSDF求值的倍增器。

则每个BSDF将获得不同的缩放比例并且应该被单独计算。清单[12.1](#代码清单12_1)显示了执行调整所需的所有代码，以演示我们的方法的简单性。

由于每个着色点都是获得不同的α值，因此该提案可能看似违反直觉。这意味着凹凸法线与表面对齐的方向将几乎不会受到收到阴影，而不同的方向将会收到明显的阴影。但是，事实证明，这正是解决问题所需的行为。

12.4 结果

我们的方法设法使突变的硬边变得平滑而对其剩余部分的样子影响很小。我们想强调一些允许无缝集成到生产渲染器中的功能：

* 在没有凹凸的情况下，外观保持不变。注意在公式[12.2](#公式12_2)中，对于无失真的情况，计算得到的粗糙度为0，因此将不会有阴影。可以绕过整个函数。
* 由于α的估值很低，微小的凹凸将导致难以察觉的变化。这种情况不会受到伪影的影响，也不需要修复。
* 只有掠射光线方向受到阴影函数的影响，与典型的微表面模型一样，更直接面向表面的角度的入射光将不受影响。

虽然我们的推导是基于与现实脱节的正态分布，但我们证明了这种分布为结构化模式产生了合理的结果，如图[12.8](#图12_8)所示。在左列中的低振幅凹凸的情况下，我们的阴影项仅最小程度地改变不需要校正的图像。如硬边变得更加显眼，我们的技术发挥更强烈的作用，并且使过渡区域变得平滑。这种方法对强烈的凹凸特别有用。

**133**



图12.8：从左到右，结构化的织物的凹凸图案的幅度逐渐增强。顶行显示未校正的凹凸渲染结果，以及底行用平滑的边缘过渡演示我们的阴影版本。

感谢

这项工作是在索尼影业的Arnold渲染器的核心开发工作中与Christopher Kulla和Larry Gritz合作开发的。

参考

[1] Conty Estevez, A., and Lecocq, P. Fast Product Importance Sampling of

Environment Maps. In ACM SIGGRAPH 2018 Talks (2018), pp. 69:1-69:2.

[2] Cook, R. L., and Torrance, K. E. A Reectance Model for Computer

Graphics. ACM Transactions on Graphics 1, 1 (Jan. 1982), 7-24.

[3] Dupuy, J., Heitz, E., Iehl, J.-C., Pierre, P., Neyret, F., and Ostromoukhov,

V. Linear E\_cient Antialiased Displacement and Reectance

Mapping. ACM Transactions on Graphics 32, 6 (Sept. 2013), 211:1{211:11.

[4] Heitz, E. Understanding the Masking-Shadowing Function in Microfacet-Based

BRDFs. Journal of Computer Graphics Techniques 3, 2 (June 2014), 48-107.

[5] Max, N. L. Horizon Mapping: Shadows for Bump-Mapped Surfaces. The

Visual Computer 4, 2 (Mar 1988), 109-117.

**134**

[6] Olano, M., and Baker, D. Lean Mapping. In Symposium on Interactive 3D Graphics and Games (2010), pp. 181-188.

[7] Onoue, K., Max, N., and Nishita, T. Real-Time Rendering of Bumpmap Shadows Taking Account of Surface Curvature. In International Conference on Cyberworlds (Nov 2004), pp. 312-318.

[8] Walter, B., Marschner, S. R., Li, H., and Torrance, K. E. Microfacet Models for Refraction Through Rough Surfaces. In Eurographics Symposium on Rendering (2007), pp. 195-206.

**135**