

# 硬阴影绘制难点介绍

21721097, DrinkingCoder

## 硬阴影绘制难点介绍

前言

介绍

基础知识

误差分析

误差类型

简化的采样误差分析

准确的采样误差分析

参考文献

## 前言

阴影的位置与形状在人类的视觉系统中具有相当重要的意义，它能增加人对图像中物体的空间感知。而在计算机图形学中，关于阴影的绘制自1978年推出以来(7)，阴影映射一直是生成快速硬阴影的主流方法之一。与许多绘制算法一样，它一开始被应用于离线的电影制作，之后随着算法的改进与硬件计算能力的增强，被用于实时的图形绘制中。阴影映射在游戏行业、VR和其他许多应用中被广泛使用，并且它也是许多软阴影算法的基础。这篇文章简单介绍了当前实现硬阴影算法的难点，并尝试从信号采样的角度对它们进行误差分析。

## 介绍

光线从光源出发，经过环境的反射折射之后进入传感器或者人眼中形成图像，而由于环境中物体之间存在遮挡关系，光线在最终成像中某些区域的能量较低，我们称之为阴影。阴影是场景中光线传输的重要结果。如图1所示，他们给出了视觉线索来澄清物体之间以及物体和光源之间的几何关系，在人类的感知系统中阴影显得十分重要，在绘制中阴影的质量直接决定了其用户的体验感。

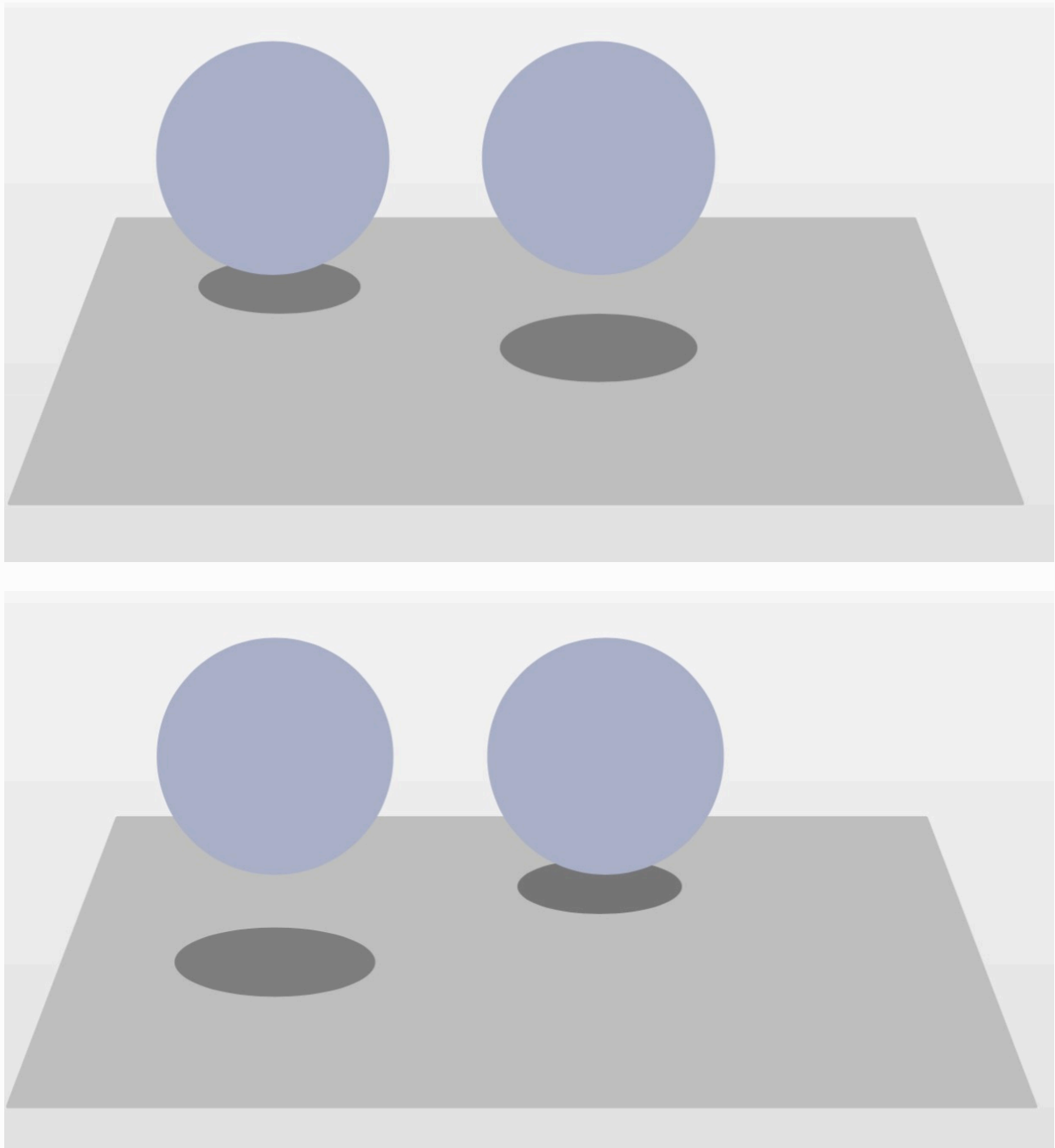


图1

阴影的类型大体可以分为两种，一种是硬阴影（图2上），一种是软阴影（图2下）。其中硬阴影主要由点光源形成，而软阴影则主要由面光源形成。尽管面光源引起的软阴影在游戏等应用中越来越流行，但许多应用仍使用由点光源或定向光源引起的硬阴影。由于使用过滤算法可以使硬阴影或阴影轻微柔化，能够很好地模拟许多光源，即使应用中的某些光源使用软阴影硬阴影算法也能够胜任这一角色（由太阳造成的阴影就是一个十分典型的例子）。这篇文将集中介绍硬阴影算法，因为它们是最广泛使用的阴影算法。

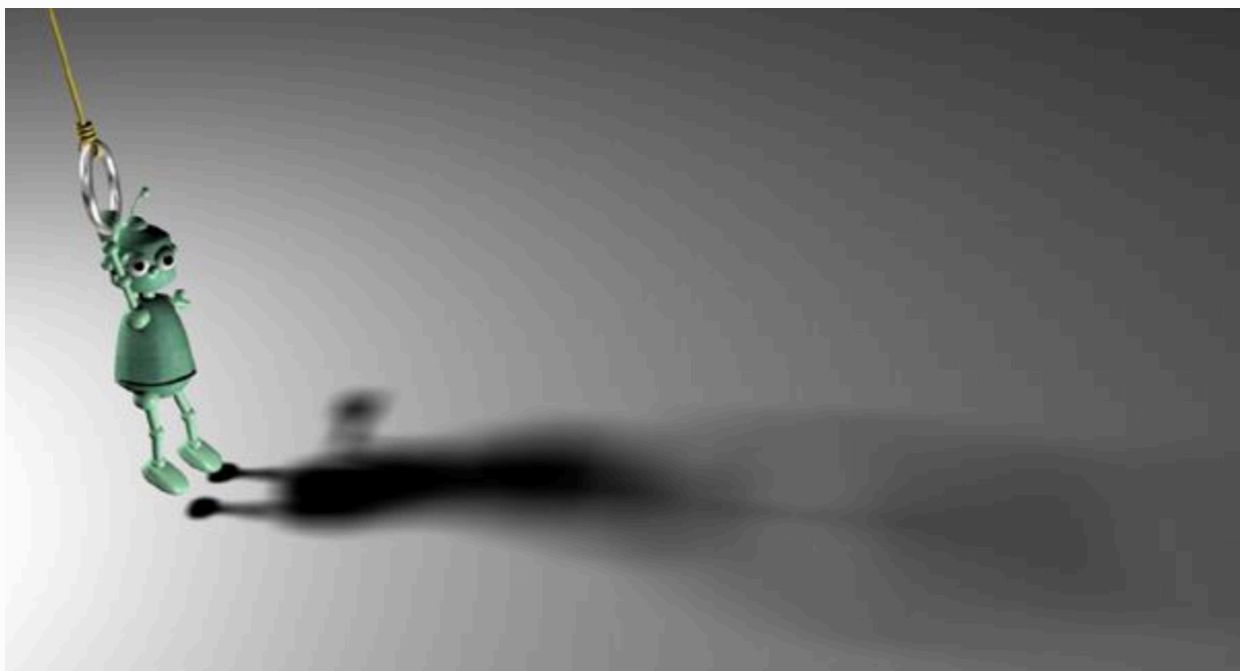
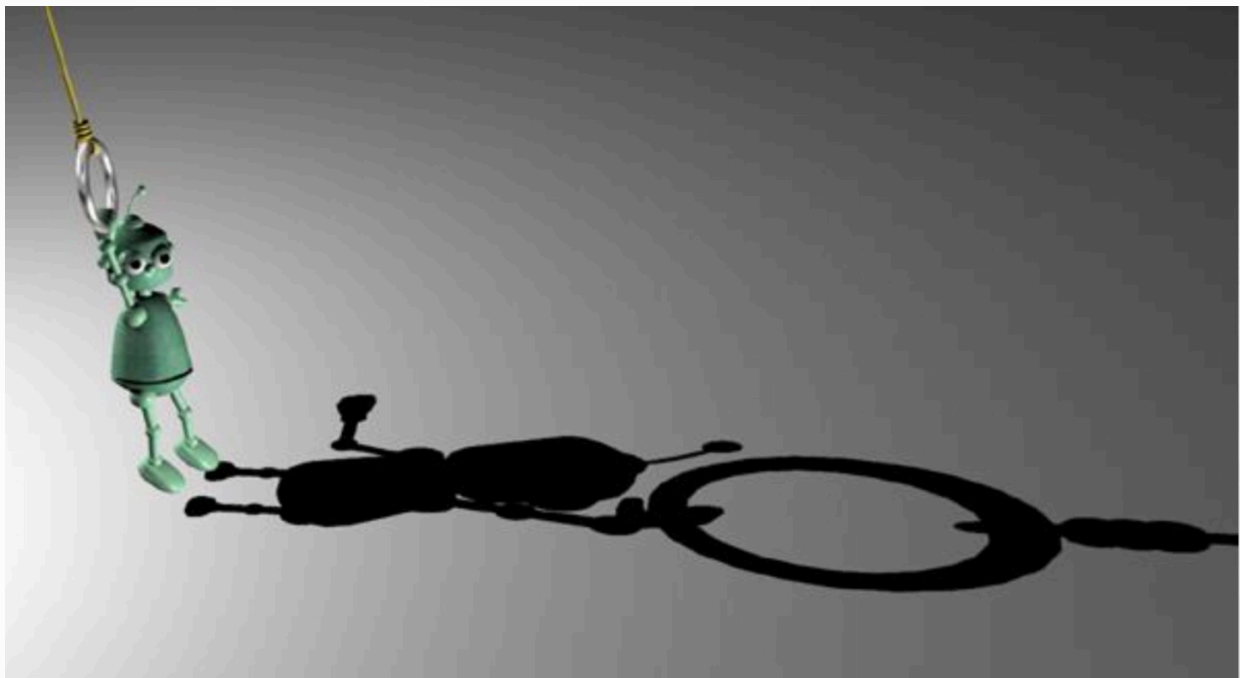


图2

当我们从光源的角度出发，某些点被遮挡时，这些点点将落在在阴影中。阻挡光线到达此点的物体称为阴影施加物，遮挡物或阻挡物。阴影中的点所在的对象称为阴影接收物。存在实时硬阴影的两种主要方法：基于几何的和基于图像的。 尽管阴影计算算法一直是计算机图形学中的一个重要问题，但强大而有效的硬阴影生成仍然是一个没有解决的问题。基于几何的算法在像素级上产生了很好的结果，但它们在不同的观察者视点下仍然存在鲁棒性问题（参见图3）。几乎所有基于几何的算法的研究都集中在阴影体上(1) (8) (9)，这种算法的主要缺点是渲染所有阴影体积所需要的填充比相当大。由于它必须进行轮廓检测，在多边形丰富的场景下耗时十分严重。而且因为需要一种检测边的简单方法，这种算法只能处理多边形数据。

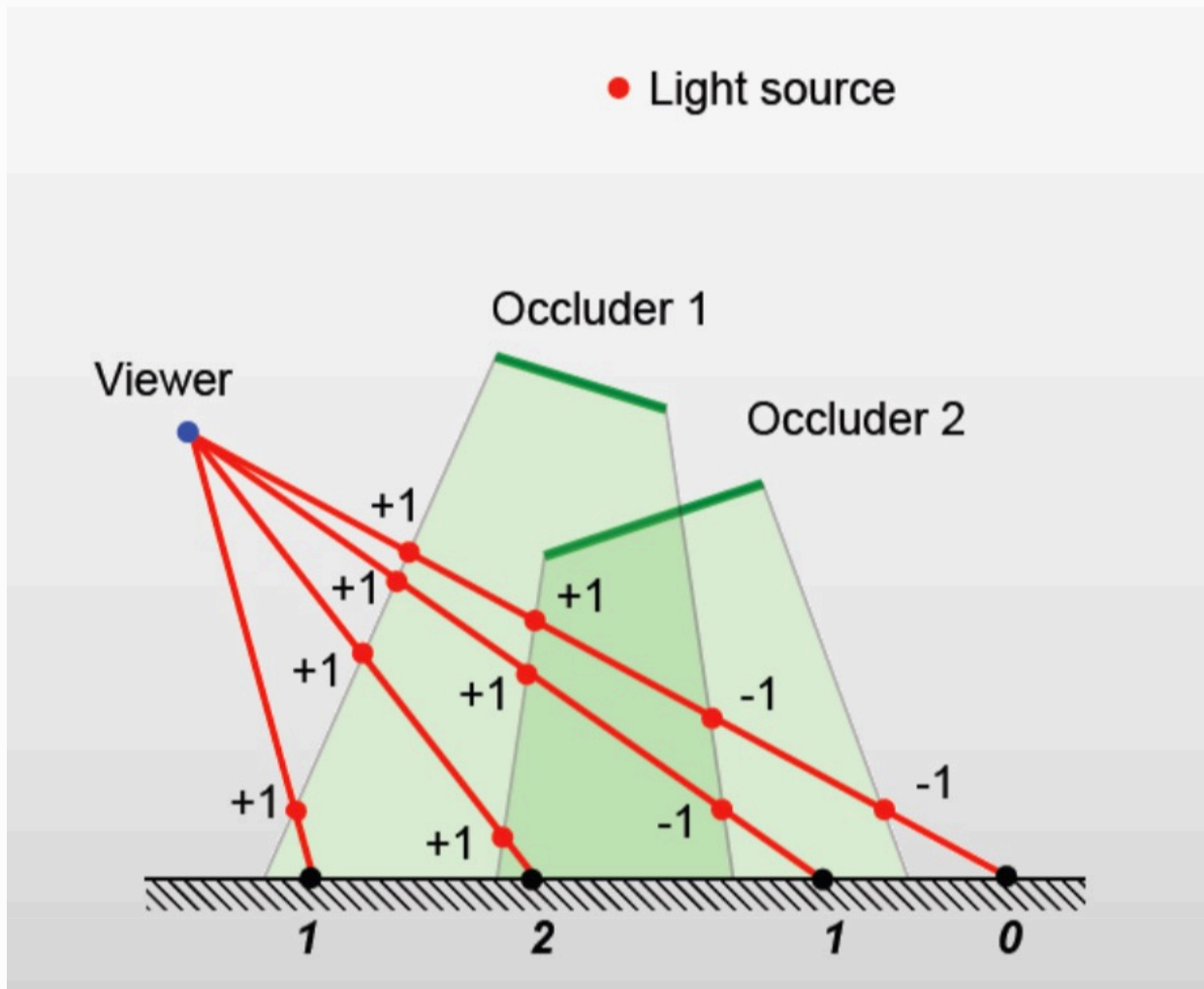
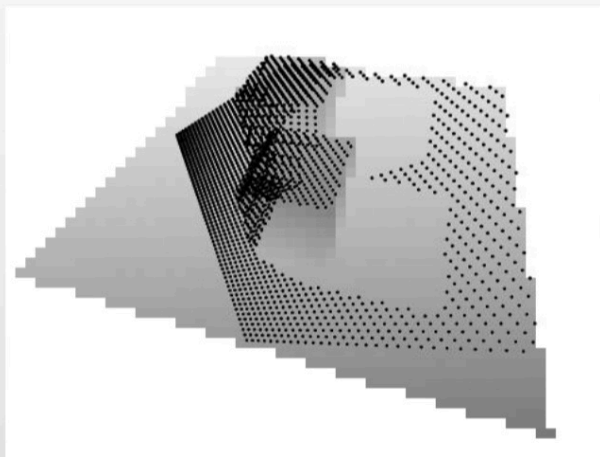


图3

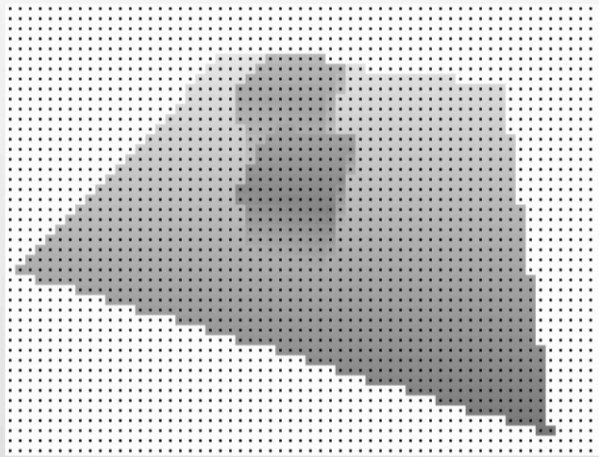
另一方面，基于图像的算法速度则非常快，因为它们的复杂度与标准场景渲染相似。阴影映射 (shadow mapping)(2) 是一种基于图像的算法，可以处理任意的光源和观察者视点，可以计算物体造成的自阴影，甚至可以处理非多边形输入。第2节介绍了基本阴影算法。阴影映射也有许多缺点：

首先，它无法使用单个平截头体来捕捉全向灯，这就导致可能需要构建多张阴影图来解决阴影问题。这个问题在第2节讨论。第二个也是更严重的问题是由于阴影图的采样和投影到阴影图中的图像像素的采样通常不匹配而产生的走样（如图4、5所示）。在第3节中，我们将分析这些走样，并从该分析中导出若干不同类型的误差，而第4部分概述了可减少抽样误差的方法。第三个问题是物体产生的不正确的自阴影，这是由于存储在每个纹理元素的阴影贴图深度信息的欠采样和不精确性造成的。这就需要对深度测试进行整体处理以提供可靠的结果。我们将在第5节讨论这个问题的各种方法。



Points where the depth  
is queried

$\neq$



Points where the depth  
is computed/known

图4





第6节介绍了应用采样理论的滤波算法，以更好地重构第二遍中存储在阴影图中的信息。最后，第7节给出了如何为给定应用程序选择最佳算法的指南。

## 基础知识

在阴影映射中，阴影计算分两步进行：首先渲染并存储从光源（在光照空间中）看到的当前场景（阴影图）的深度图像（请参见图4左）。该图像包含每个纹理元素最近物体到光源的深度。这个想法是，光源背后无法看到隐藏在这些深处的一切，因此处于阴影之中。在第二遍中，从视点（在视图空间中）渲染场景，并且将每个3D片段重新投影到阴影图中。如果重投影的点深度比存储在阴影图（深度测试）中的深度更远，则碎片处于阴影中并相应地进行着色（请参见图4右）。由于没有一个台体能够独立的描述一个球形的视图，那么如果光源是一个全方位灯，我们必须使用多个缓冲区进行计算，因此必须建立多个阴影图和台体来划分这个球形视图。最常用的方法是将从光源出发的全景图绘制到一个立方体上，那么我们就可以用六个截锥（立方体贴图的每一面都有一个）来描述这个光源了，但是这会导致对这种灯光性能的巨大损失。采用抛物线映射(3)可以实现更快的解决方案。这只会导致两个绘制环境，每个半球都对应一个绘制环境，但也会产生如何在图形硬件上高效模拟抛物线映射（线变为曲线）的问题。最简单的解决方案是假定场景细分得足够精细，以至于只有顶点的抛物线映射就足够了。

## 误差分析

绘制阴影贴图时，我们会对给定的场景进行离散采样，之后的阴影映射使用采样的结果来重建每个物体的表面对于该光源的可见性。阴影映射可以看作类似于纹理映射，它有以下步骤：

1. 第一步采样输入函数（即生成阴影贴图）。从信号处理的角度来看，由于在初始采样阶段没有频带限制以避免走样，因此采样频率理想情况下应高于信号的奈奎斯特频率。但是由于我们希望生成的硬阴影有锐利的边缘，因此频率的信号其实上限是不存在的。  
对采样信号进行插值和重建。对于阴影映射，就是将片段投影到阴影贴图中。由于深度测试的非线性，对于直接的阴影映射我们一般使用最近邻
2. 对重建信号进行滤波以避免输出分辨率的锐度过高。由于深度测试的非线性，这对于直接的阴影映射来说是不可能的。
3. 在最后的像素位置重新采样重建信号。

在进行离散采样时，重建只发生在最终的像素位置，因此重建和重采样是合并在一起的。为了通过上述步骤将阴影映射看做信号重建过程，我们可以将阴影映射操作视为具有透视变换的纹理的纹理映射，该映射是对某些已经处理好的物体进行阴影测试，即如果纹理元素暴露在光线中被填充1，而在阴影中填充0。通过这种方式，标准阴影映射可以被视为信号重建和重采样，我们使用最近邻查找或双线性插值完成信号重建，使用mipmapping完成带通滤波，并且重采样只是最终像素处函数的评估位置。

## 误差类型

误差产生的主要原因在于：

- 欠采样，当投影到屏幕上的阴影图样本的采样率低于屏幕像素时发生。这是由于初始采样频率太低所致。
- 过采样，当投影到屏幕的阴影图样本具有比屏幕像素更高的采样率时发生。在这种情况下与纹理采样中的经典走样问题类似。这可以通过调整初始采样率，或通过上面讨论的带通滤波来解决。
- 重建误差或阶梯状走样，这是由于最近邻重建造成的。这可以通过更好的重建滤波器来解决。
- 由于阴影图在每一帧中的光栅化都会不一样，这有可能在时间尺度上出现锯齿和闪烁的走样。如果发生欠采样，并且使用了非最佳重构方式这种情况可能会十分严重。

信号处理或纹理映射最重要的区别在于，阴影映射过程可以对初始采样步骤进行一些控制。改变初始采样可以减少所有类型错误的影响，因此大多数关于阴影映射的出版物都采用这种方法。在3.2节和3.3节中，我们将更详细地讨论抽样误差，并讨论两种描述这种误差类型的方法。还有一些方法专门用于处理阴影映射环境中的过采样和重构，我们将在第6节中对其进行描述。

## 简化的采样误差分析

大多数减少阴影图采样误差算法的根源是分析场景中误差的分布。Stamminger和Drettakis (4)首先介绍了透视阴影地图的简化误差分析，并在之后的许多方法中使用了相同的公式。该分析假设了一个头顶定向光源，并查看位于视锥体z轴某处的表面元素。图6显示了一条边的相关参量的描述，便于我们之后形式化误差分析。阴影图中的一个像素代表穿过它的光线轴，并且在阴影图的局部参数化中具有 $d_s \times d_s$ 的大小。我们局部参数化阴影图在观察者的近平面和远平面之间的为0到1，这已经假设阴影图适当地集中到视锥体上，而不是浪费任何分辨率在视觉的不可见场景部分上。在世界空间中，以均匀阴影图为例，光线的轴的长度为 $d_z = (z_f - z_n)d_s$ 。

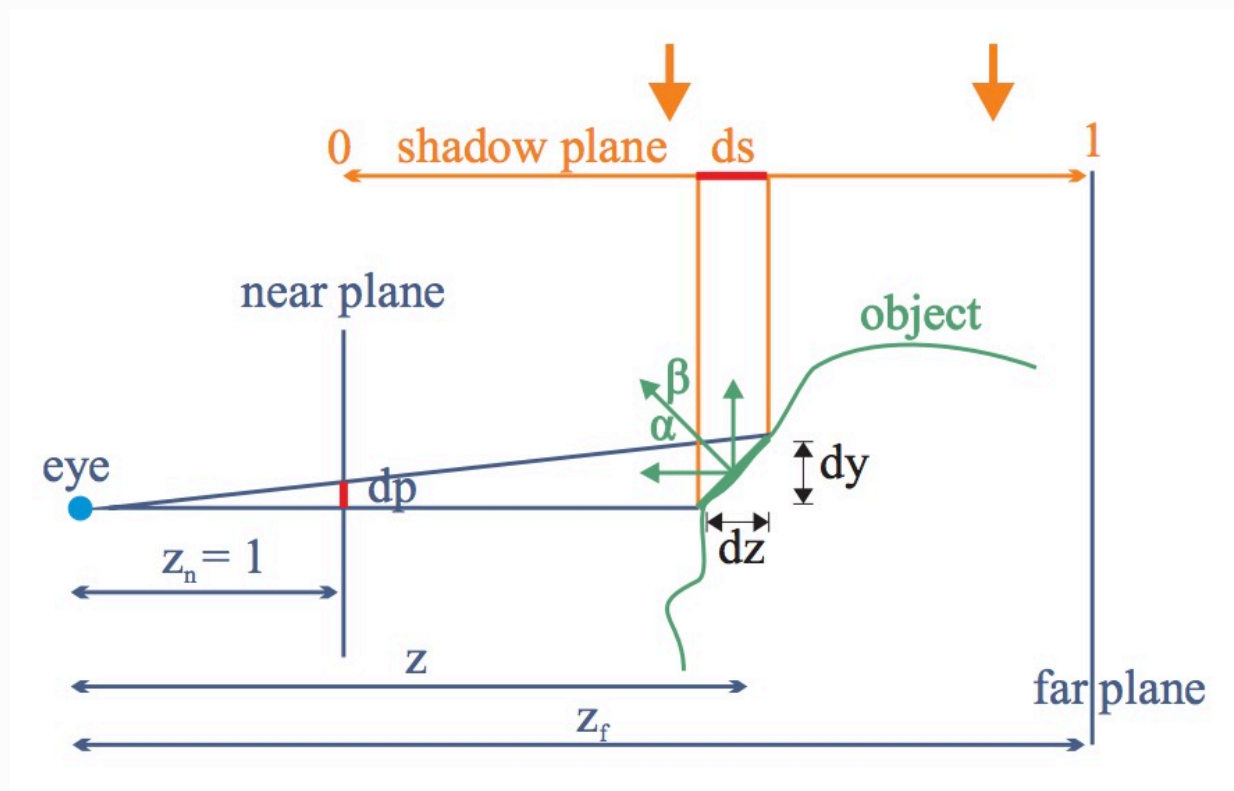


图6

当在世界坐标系中光线沿着  $d_z/\cos\beta$  的长度接触到一条边时。这表示眼睛空间中的长度  $d_y = d_z \frac{\cos\alpha}{\cos\beta}$ , 在屏幕上显示  $d_p = d_y/z$  (假设近平面距离为1)。请注意, 我们假设边可以沿着z轴平移, 即z在进行分析时是自由参数。那么阴影图走样误差  $d_p/d_s$  就是

$$\frac{d_p}{d_s} = \frac{1}{z} \frac{d_z}{d_s} \frac{\cos\alpha}{\cos\beta}$$

当  $d_p$  比一个像素大或者对于存在一个观察点在近平面上 (此时  $d_p/d_s$  大于  $res_{shadowmap}/res_{screen}$ ) 时会发生阴影图欠采样。Stamminger和Drettakis (4) 在文章中解释这种情况可能发生的原因有两个: 当  $\frac{d_z}{zd_s}$  很大时的会产生透视走样, 以及当  $\cos\alpha/\cos\beta$  大时会产生投影走样。

投影走样是一种局部现象, 发生在几乎平行于光线方向的表面上 (参见图7左侧)。减少这种误差需要在这些区域采用更高的采样密度。只有基于场景分析才能利用自适应性地局部采样算法可以实现这一点。

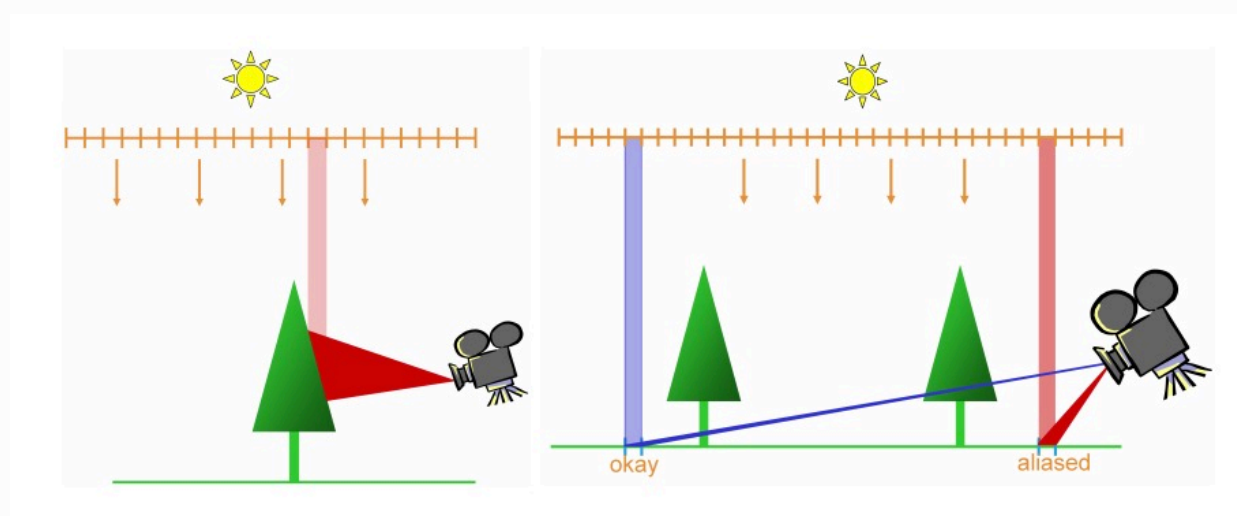


图7

另一方面, 透视走样是由观看者的透视投影引起的 (见图7右)。如果沿着阴影图的一个轴发生透视缩短 (透视投影会使得较远的物体被缩短) 的效果, 不同的阴影图参数化方式会产生不同的效果。如果选择不同的参数化, 这将导致阴影图上的采样密度分布不同。如果使用均匀分布来参数化, 那么就会使得  $d_z/d_s$  变为常数, 于是采样误差  $d_p/d_s$  会随着  $1/z$  变大而变大, 在靠近近平面的时候会产生这种效应。这会导致在相机附近的物体上只会有较少的采样, 而这会导致透视走样十分之明显 (如图8所示)。我们可以通过使用不同的参数化或将阴影图分割成更小的部分 (图9所示) 来减少透视走样, 在观察者附近进行更多的阴影图采样。



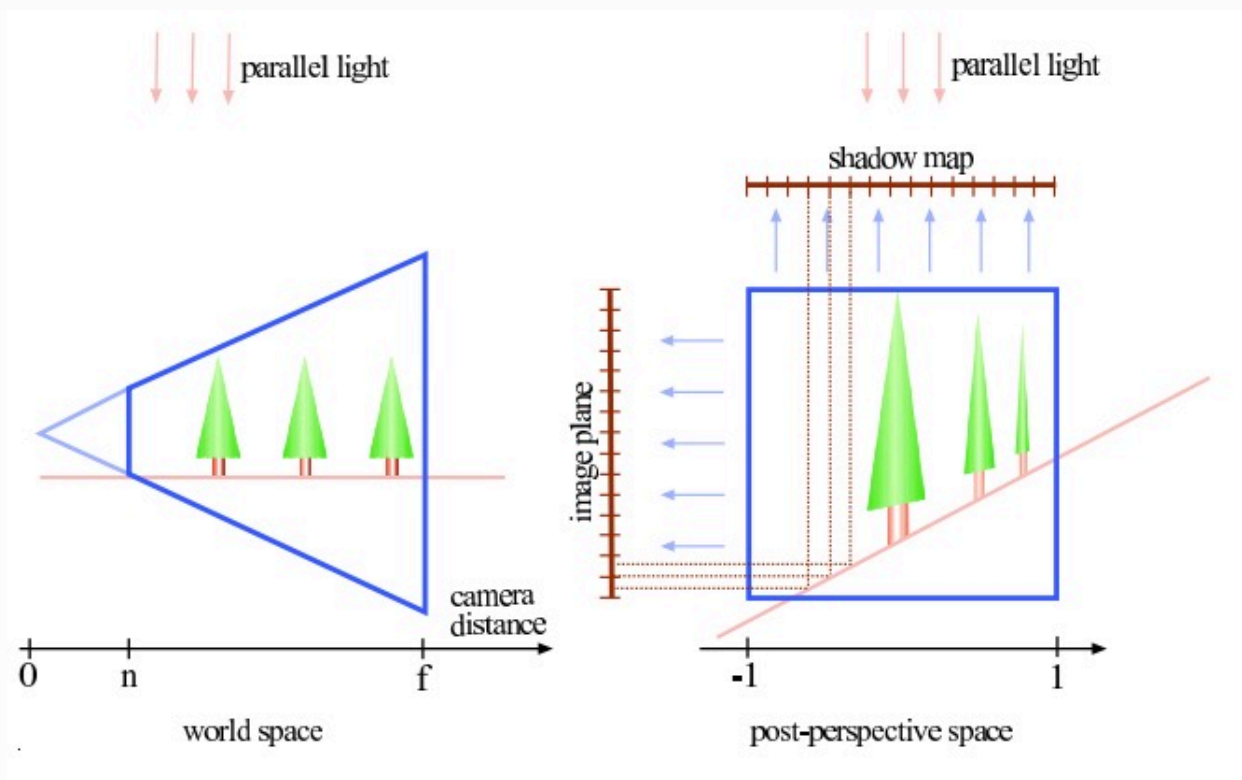


图8

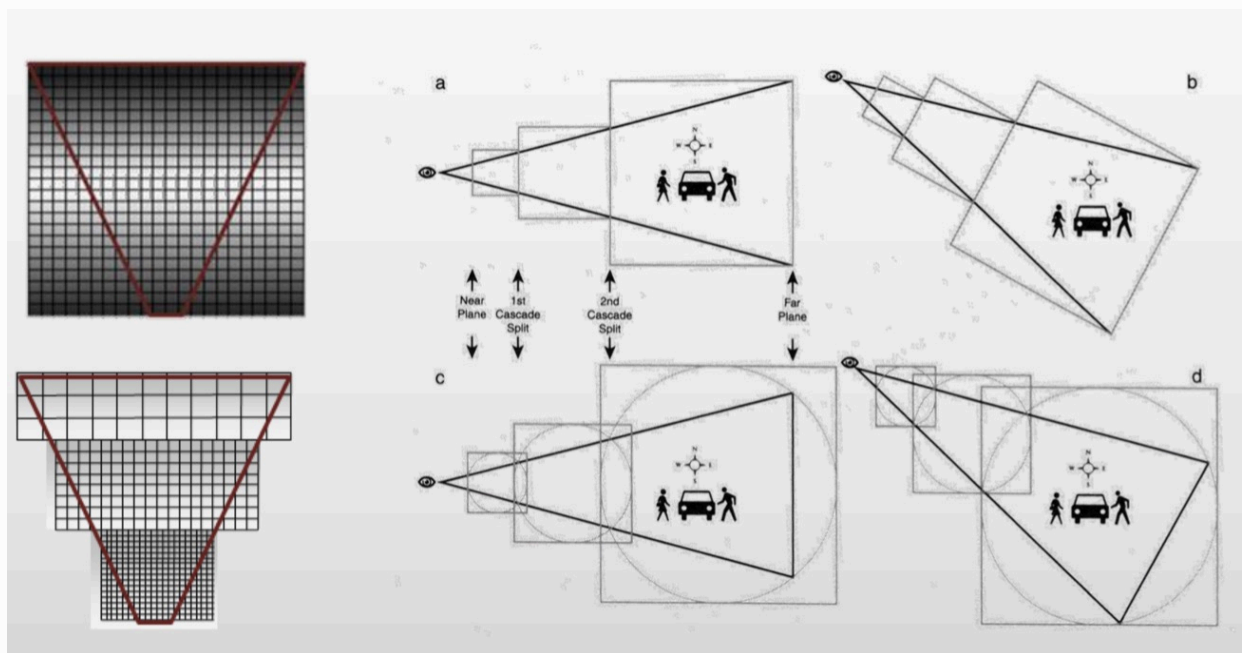
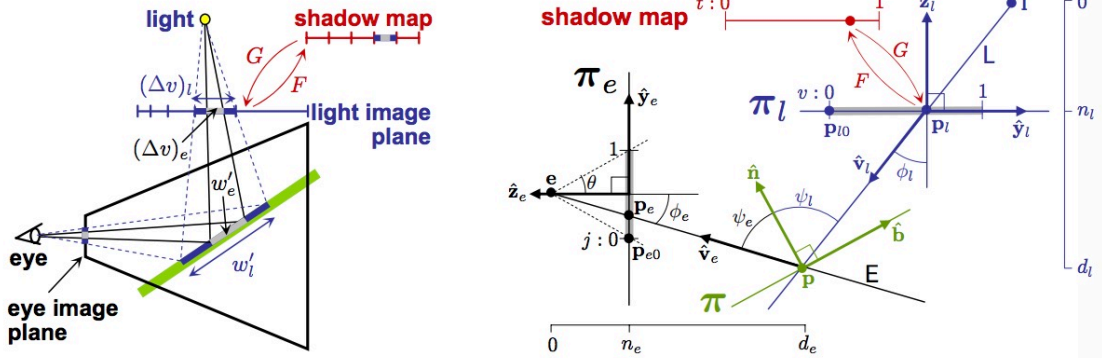


图9

## 准确的采样误差分析

Brandon Lloyd提出了一种准确的采样误差方法(5) (6)，但是它的推导过程以及构成原理相当之复杂，这里我们只给出结果，需要进一步了解的读者可以精读他的论文。在使用图10所示的参数化情况下，走样误差 $m$ 的计算公式为：

$$m = \underbrace{\frac{r_j}{r_t} \left( \frac{dG}{dt} \right)}_{\delta_l} \underbrace{\left( \left( \frac{W_l n_e d_l \cos \phi_l}{W_e n_l d_e \cos \phi_e} \right) \frac{\cos \psi_e}{\cos \psi_l} \right)}_{1/\delta_e}.$$



**Fig. 2: Computing aliasing error.** (Left) Spacing between locations corresponding to a pair of samples from the light (shadow map texels) and the eye (image pixels). Aliasing error can be quantified as the ratio of the spacing between the light and eye sample locations. (Right) The sample spacing is related to the derivatives of the function  $j(t)$  that maps a point  $t \in [0, 1]$  in the shadow map to  $\mathbf{p}_l$  on the light's image plane  $\pi_l$ , projects  $\mathbf{p}_l$  through the light to  $\mathbf{p}$  on a planar surface  $\pi$  in the scene, and projects  $\mathbf{p}$  through the eye to a point  $\mathbf{p}_e$  on the eye's image plane  $\pi_e$ .

$F, G$	shadow map parameterization and its inverse	$\phi_l, \phi_e$	angle between image plane normals and rays of light and eye
$(i, j)^\top$	eye image plane coordinates	$\psi_l, \psi_e$	angle between surface normal and light and eye rays
$(s, t)^\top$	shadow map coordinates	$\beta_l, \beta_e$	angle between surface normal and light and eye image planes
$(u, v)^\top$	light image plane coordinates	$d_l, d_e$	distance along $\hat{\mathbf{z}}_l$ and $\hat{\mathbf{z}}_e$ from light and eye to a point in the scene
$r_i \times r_j$	image resolution	$n_l, n_e$	distance from light and eye to respective image plane
$r_s \times r_t$	shadow map resolution	$W_l, W_e$	widths of the light and eye image planes
$\hat{\mathbf{x}}_l, \hat{\mathbf{y}}_l, \hat{\mathbf{z}}_l$	light space coordinate axes	$w_l, w_e$	widths of light and eye beams
$\hat{\mathbf{x}}_e, \hat{\mathbf{y}}_e, \hat{\mathbf{z}}_e$	eye space coordinate axes	$w'_l, w'_e$	widths of light and eye beams projected on a surface
$\hat{\mathbf{n}}$	surface normal	$w_e^\perp \hat{\mathbf{z}}_e$	width of an eye beam measured perpendicular to $\hat{\mathbf{z}}_e$
$\hat{\mathbf{b}}$	vector along the surface	$(\Delta v)_l, (\Delta v)_e$	sample spacing in $v$ on light image plane
$\pi_l, \pi_e$	light and eye image planes	$\delta_l, \delta_e$	spacing distribution function for the light and eye
$\pi$	a planar surface in the scene	$\bar{\delta}_e$	perspective factor of $\delta_e$
$\mathbf{p}_{l0}, \mathbf{p}_{e0}$	origins of coordinate systems on the light and eye image planes		
$\mathbf{p}_l, \mathbf{p}_e$	points on the light and eye image planes		
$\mathbf{p}$	point on a surface		
$\mathbf{l}, \mathbf{e}$	light and eye positions		
$\mathbf{L}, \mathbf{E}$	rays from the light and eye		
$\hat{\mathbf{v}}_l, \hat{\mathbf{v}}_e$	direction of rays from light and eye		
$\theta$	half of the frustum field of view		

图10

## 参考文献

- (1) CROW F. C.: Shadow algorithms for computer graphics. In Proceedings of the 4th annual conference on Computer graphics and interactive techniques (New York, NY, USA, July 1977), George J., (Ed.), vol. 11, ACM Press, pp. 242–248. 2

(2) WILLIAMS L.: Casting curved shadows on curved surfaces. Computer Graphics (SIGGRAPH '78 Proceedings) 12, 3 (Aug. 1978), 270–274. 2

(3) BRABEC S., ANNEN T., SEIDEL H.-P.: Shadow mapping for hemispherical and omnidirectional light sources. In Advances in Modelling, Animation and Rendering (Proceedings Computer Graphics International 2002)

(4) STAMMINGER M., DRETTAKIS G.: Perspective shadow maps. In ACM Transactions on Graphics (Proceedings of SIGGRAPH) (San Antonio, Texas, July 2002), ACM Press, pp. 557–3, 4, 6

(5) LLOYD B.: Logarithmic Perspective Shadow Maps. PhD thesis, University of North Carolina at Chapel Hill, August 2007. 4, 7, 8, 9

(6) LAURITZEN A., MCCOOL M.: Layered variance shadow maps. In GI '08: Proceedings of graphics interface 2008 (Toronto, Ont., Canada, Canada, 2008), Canadian Information Processing Society, pp. 139–146. 14

(7) Lance Williams: Casting curved shadows on curved surfaces.

(8) WIMMER M., SCHERZER D.: Robust shadow mapping with light space perspective shadow maps. In ShaderX 4 – Advanced Rendering Techniques, Engel W., (Ed.), vol. 4 of ShaderX. Charles River Media, Mar. 2006. 5

(9) WIMMER M., SCHERZER D., PURGATHOFER W.: Light space perspective shadow maps. In Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering) (Norrköping, Sweden, 2004), Springer Computer Science, Eurographics, Eurographics Association. 5, 6, 7