摘要

由于它的多功能性，速度和鲁棒性，自1978年问世以来，阴影贴图一直是快速生成硬阴影的一种流行算法，首先用于离线电影制作，后来在实时图形中也越来越多。因此，近年来与阴影图相关的出版物数量激增并不奇怪。上一次涵盖阴影映射方法但主要侧重于软阴影生成的调查可追溯到2003年[HLHS03]，而上一次针对一般阴影生成的调查可追溯到1990年[WPF90]。没有调查描述近年来硬阴影贴图生成中的所有进步。另一方面，阴影映射在游戏行业，生产中以及许多其他应用中被广泛使用，并且它是许多软阴影算法的基础。由于有关该主题的文章很多，从业人员和研究人员很难选择合适的阴影算法，因此许多应用程序都错过了最新的高质量阴影生成方法。这项调查的目的是通过对该领域进行详细概述来纠正这种情况。我们提供阴影映射错误的详细分析，并得出现有方法的全面分类。我们讨论最有影响力的算法，考虑它们的优点和缺点，从而为读者提供选择最适合其需求的阴影算法的方法。

介绍

阴影是场景中光传输的重要结果。 它们提供了视觉提示，以阐明物体之间以及物体与光源之间的几何关系。 尽管由于面光源引起的柔和阴影在诸如游戏的应用程序中变得越来越流行，但是许多应用程序仍使用由点光源或定向光引起的硬阴影。 即使应用中的某些光源使用了柔和的阴影，也可以将许多光源建模为点光源，这些点光源可以提供硬阴影或使用滤镜技术稍微柔化的阴影（例如，太阳造成的阴影） 。 我们的调查将集中在硬阴影上，因为它们是使用最广泛的阴影算法，但是由于关于该主题的大量论文，很少充分利用它们的潜力，这使得难以为特定应用选择最佳算法。

当从光源的角度看不到该点时，该点处于阴影中。阻挡光线到达该点的物体称为阴影投射器，遮挡物或阻挡器。 阴影点所在的对象称为阴影接收器（参见图1）。 存在两种主要的实时硬阴影方法：基于几何的和基于图像的。

尽管阴影算法的出现时间几乎与计算机图形本身一样长，但是强大而有效的硬阴影生成仍然是没有解决的问题。尽管基于几何的算法可产生像素完美的结果，但它们会因具有不同的观众光星座而遭受健壮性问题的困扰。由于它们的通用性，几乎所有基于几何的算法研究都集中在阴影体积上[Cro77]。该技术的主要缺点是渲染所有阴影体积所需的大量填充率。另外，必须进行轮廓检测，对于多边形多的场景，这意味着另一个性能损失。最终，只能处理多边形数据，因为需要一种简单的方法来检测和拉伸边缘。

另一方面，基于图像的算法非常快，因为其复杂度类似于标准场景渲染。阴影映射[Wil78]是一种基于图像的算法，可以处理任意的脚轮/接收器星座，可以解决自阴影问题，甚至可以处理非多边形输入。第2节介绍了基本的阴影算法。

不幸的是，阴影映射还具有许多缺点。首先，无法使用单个视锥捕获全向光。此问题在第2节中讨论。

第二个更严重的问题是由于阴影贴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出现的混叠伪影。 在第3节中，我们将分析这些混叠失真，并从该分析中得出许多不同类型的误差，而第4节则概述了可以减少采样误差的方法。

第三个问题是不正确的自阴影化，是由于每个纹理像素的阴影贴图中存储的深度信息的欠采样和不精确引起的。这就需要对深度测试进行偏置（深度偏置）以提供可靠的结果。我们将在第5节中讨论解决此问题的各种方法。

第6节介绍了过滤技术，这些技术应用采样理论来更好地重建第二遍阴影图中存储的信息。最后，第7节提供了有关如何为给定应用选择最佳算法的指南。

2 偏移

在阴影贴图中,阴影计算分两步执行：首先，渲染并存储从光源（在光空间中）看到的当前场景的深度图像（阴影图）（请参见图2，左）。 该图像包含每个纹理像素的距离光源最近的对象的深度。 这个想法是，位于这些深度之后的所有东西都不能被光源看到，因此处于阴影中。 在第二遍中，从视点（在视图空间中）渲染场景，并将每个3D片段重新投影到灯光空间中。 如果重新投影的片段深度比阴影贴图中存储的深度更远（深度测试），则片段处于阴影中并相应着色（参见图2，右）。

由于它们的球面视图，必须使用多个缓冲区来计算全向光。 没有任何一个视锥体可以反映此情况，因此必须构建大量阴影图和视锥体来划分此球形视图。 最常见的方法是使用六个视锥面（一个立方体贴图的每一面一个），这会严重影响此类灯光的性能。 通过采用抛物线映射[BAS02b]，可以实现更快的解决方案。 这样仅产生两个渲染，每个半球一个，但同时也产生了一个问题，即如何在图形硬件上有效地模拟抛物线映射（线变成曲线）。 最简单的解决方案是假定场景的细分足够精细，这样仅顶点的抛物线映射就足够了。 存在更慢且涉及更多的方法，这些方法可以直接在现代硬件上计算曲线[GHFP08]。

误差分析

渲染阴影贴图时，将创建给定场景的离散采样表示。阴影映射操作随后使用此表示来重建相对于光源的曲面的可见性条件。因此，将阴影映射视为类似于纹理映射的信号重建过程会很有帮助。信号重建包括以下步骤：

1. 最初对输入函数进行采样，即使用渲染生成阴影贴图。 由于不可能在初始采样阶段进行带宽限制来避免混叠，因此理想情况下，采样频率应高于信号的奈奎斯特频率。 请注意，评估深度比较后的实际阴影信号具有尖锐边缘，因此频率不受限制。
2. 从采样表示中重构（或内插）信号。 对于阴影贴图，将片段投影到阴影贴图空间并评估阴影贴图测试时会发生这种情况。 由于深度测试的非线性，对于直接的阴影映射，只有最近的查找是可能的。
3. 对重构信号进行带宽限制（或预滤波器），以避免在目标输出分辨率下出现太高的频率。 由于深度测试的非线性，这对于阴影映射是不可能直接实现的。
4. 在最终像素位置重新采样重建的信号。

请注意，在离散采样方案中，重建仅发生在最终像素位置，因此重建和重采样以组合方式进行。

为了通过以上步骤促进作为信号重建过程的阴影映射的心理模型，可以将阴影映射操作视为具有投影纹理贴图的纹理对象，该纹理贴图具有针对已烘焙的特定对象的阴影测试结果 这意味着纹理像素填充有1表示光线中的片段，0表示阴影中的片段。这样，可以将重建和重采样与标准阴影映射进行比较，在标准阴影映射中，重建是通过最近邻居查找或双线性插值完成的，带限制是使用mipmapping进行的，而重采样只是在最终像素位置对函数的评估。

3.1 误差类型

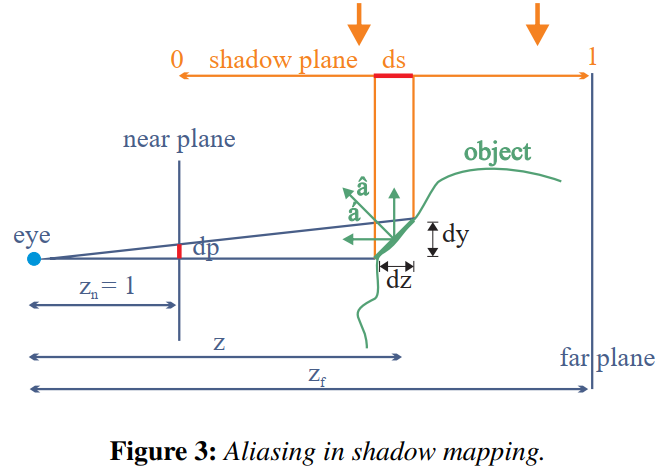
误差的主要类型是

1. 欠采样，当投影到屏幕上的阴影贴图样本的采样率低于屏幕像素时发生。这是因为初始采样频率太低。
2. 当投影到屏幕上的阴影贴图样本的采样率高于屏幕像素时，就会发生过采样。 在这种情况下，会发生纹理采样已知的经典混叠。 这可以通过调整初始采样率或上面讨论的带宽限制/预滤波步骤来解决（请参见第6节）。
3. 由于最近邻居重建而导致的重建错误或阶梯伪影。 这可以通过更好的重建滤波器来解决（请参见第6节）。
4. 如果阴影贴图的栅格化在每一帧都发生改变，则出现时间混叠或闪烁的伪影。 如果发生欠采样，这些伪像将特别出现在非最佳重建中，有关补救措施，请参见第4.6节。

特别是与信号处理或纹理映射的最重要区别是，阴影映射的过程对初始采样步骤提供了一些控制。更改初始采样可以减少所有类型错误的影响，因此，大多数有关阴影贴图的出版物都采用这种方法。因此，在3.2和3.3节中，我们将更详细地讨论采样误差，并讨论两种表征该误差类型的方法。还有一些方法专门针对阴影映射情况下的过采样和重建，我们将在第6节中进行介绍。

3.2 简化采样误差分析

减少阴影贴图采样错误的大多数算法的根源是对场景中错误分布的分析。Stamminger和Drettakis [SD02]首先针对透视阴影贴图引入了简化的误差分析，并且在许多后续方法中使用了相同的公式。 该分析假设高架定向光，并查看位于视锥的z轴上某处的表面元素。图3显示了小边缘的配置。



阴影贴图中的像素代表一束通过它的光线，并且在阴影图中的局部参数化中大小为.我们假设阴影贴图的局部参数化在观看者的近平面和远平面之间从0变为1 –这已经假设阴影贴图正确地聚焦在视锥上，而不会浪费场景不可见部分的分辨率（ 参见第4节）。 在世界空间中，以均匀阴影贴图为例，射线轴的长度为。

轴碰到一条长度为小边缘,这表示在眼睛空间中的长度为，在屏幕的投影长度为（假设近平面距离为1）.注意,我们假设小边缘可以沿z轴平移,即z是分析的自由参数.阴影贴图别名误差则为

当大于像素的大小时,或者对于高度1的近平面上的视口，当大于时,就会发生阴影贴图欠采样.如Stamminger和Drettakis [SD02]所示，这可能有两个原因：当大时透视混叠,以及大时投影混叠.

投影混叠是在几乎平行于光方向的表面上发生的局部现象（请参见图4，左）。 减少此类误差需要在此类区域中提高采样密度.只有基于场景分析在本地调整采样密度的方法才能实现此目的（第4.3.3至4.6节）。

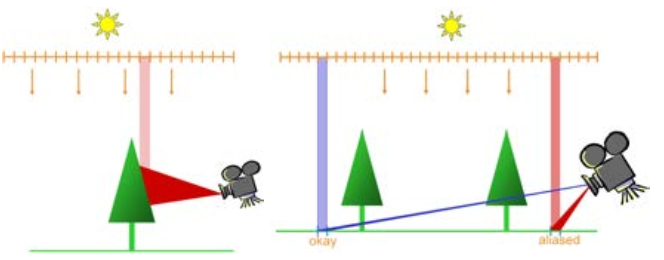


图4：在左侧的图中,投影混叠的原因是树表面的方向：它投影到阴影贴图中的一小块区域中，但是投影到相机空间中的一棵大区域中.之所以会出现右侧透视混叠，是因为阴影贴图与透视缩短无关，因此远景以及附近区域（相对于摄影机）都以相同的分辨率存储，但投影到摄影机空间的大小却非常不同.

另一方面,透视混叠是由查看器的透视投影引起的(请参见图4,右).如果透视缩短效果沿着阴影图的轴之一发生，则可能会受到阴影图参数化的影响。如果选择不同的参数化，这将导致沿着阴影图的采样密度分布不同。标准均匀参数化具有常数,因此，当大时，采样误差增大，这在靠近近平面的地方发生。这导致少量样本花费在相机附近的物体上，这使得此错误的影响非常明显。（比较图5）。为了减少透视混叠，有几种方法可以通过使用不同的参数化或通过将阴影贴图拆分为较小的部分来在查看器附近分布更多的阴影贴图样本（第4.2和4.3节）。

4 减少采样误差的策略

与通常预先确定输入图像的分辨率的纹理映射不同，在阴影映射中，对原始采样步骤有明显的控制。因此，可以调整采样，以使投影的阴影贴图样本比纯朴的阴影贴图更好地对应于屏幕空间采样率。

特别是对于放大方案，大部分负担都放在用于纹理映射的重建滤波器上，而对于阴影映射，可以通过增加采样率并从受影响的区域中删除放大率（或欠采样）来减轻这种负担，因此即使最近的邻居重建有时也可以提供良好的质量。此外，在最小化的情况下，例如对于屏幕上看起来很小的区域，可以减小初始采样率，从而避免了对昂贵的带宽限制滤波器的需求。

由于减少采样误差技术的文献很多，我们进一步将尝试将采样误差消除的方法细分为聚焦，基于扭曲，基于分区和基于不规则采样的算法。

* 1. 聚焦

可以提高采样率的最直接的方法之一是确保不将阴影图空间浪费在不可见的场景部分上。 特别是在室外场景中，如果将单个阴影贴图用于整个场景，则阴影贴图的仅一小部分实际上将与视锥面相关。 因此，首先由Brabec等人[BAS02a]引入的拟合或聚焦技术将阴影贴图视锥体拟合为包含视锥体。 几何解决方案是计算视锥面的凸包和光源位置（对于定向光源，此位置为无穷大），然后用场景边界体积和光源视锥修剪该实体（有关详细信息，请参见[WSP04，WS06] ）。 剪裁到场景边界是必要的，因为今天非常大的视锥壳很常见，并且它们经常延伸到场景边界之外。 我们称结果体为相交体B（见图7）。

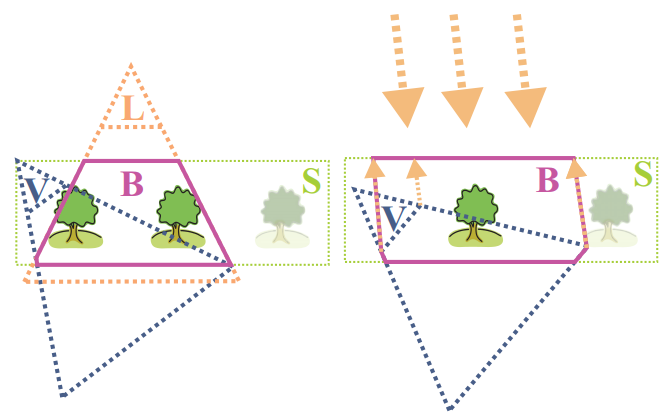


图7：通过将视锥台L，场景边界框S和视锥台V组合到边界体积B中，阴影贴图聚焦可以更好地利用可用的阴影贴图分辨率。此处的左图显示了点光源，右图显示了定向光源。

通过使用可见性算法，可以进一步减少交叉路口。 如果在创建阴影贴图之前，使用在线可见性算法（如相干层次剔除（CHC）[MBW08]）渲染了仅深度的第一遍，则可以减小远平面距离以仅覆盖最远的可见对象。

通常，拟合会导致时间混叠，因为阴影贴图的栅格化会更改每个帧。 特别是在使用可见性信息时，可能会出现强烈的时间不连续性，因此在这种情况下，使用良好的重建滤波器非常重要。

通过尝试在世界空间中保持纹理像素边界不变，还可以在某种程度上减少由于拟合而造成的时间混叠。 首先，阴影贴图需要在世界空间中保持恒定的方向，以便避免每当观众旋转时投影的阴影贴图纹理都会改变形状。 为此，阴影图需要聚焦在相交体的与轴对齐的边界框上。 为避免因在视角图视图中平视视图的平移而造成锯齿，应使用一个纹理像素边框创建阴影图，并且仅当视域平移整个视素时才进行调整。 但是，大多数观看者的移动也会导致在阴影贴图视图中缩放视锥，并且在不浪费太多阴影贴图空间的情况下更难以控制，请参见[ZZB09]。

* 1. Warping

当将视锥台投影到阴影图中时，很明显，在视点附近需要较高的采样密度，而在视点附近则需要较低的采样密度。 在某些情况下，可以在将场景投影到阴影贴图之前对其应用单个转换，以便以一种有用的方式全局更改采样密度（请参见图8）。在原始算法中，将扭曲应用于单个阴影贴图，但是后来将其与分区算法结合使用，以进一步提高采样率（请参见第4.3和4.4节）。

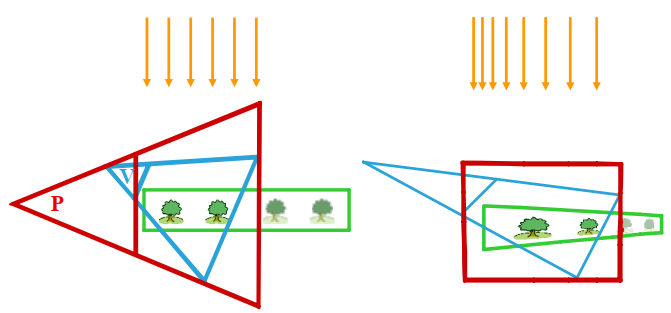


图8：光空间透视阴影贴图的示例配置，其中视锥体为V，视锥体定义了透视变换P。左：定向光，视锥体为V和透视变换为P。右：在扭曲之后，靠近视线的对象 查看器在阴影贴图中显示的更大，因此接收更多的样本。

Stamminger和Drettakis在他们的透视阴影贴图（PSM）[SD02]论文中介绍了阴影贴图变形。主要思想是在将场景渲染到阴影贴图中之前，将透视变换（即观众投影）应用于场景。因此，阴影贴图样本的分布发生了变化，使得更多的样本位于投影中心附近，而更少的样本位于投影远平面附近。这样做的好处是只使用了一个简单的透视变换，可以用4×4矩阵表示。这很好地映射到硬件，并且计算速度很快。该方法的主要问题在于，该方法的可实现质量在很大程度上取决于眼图的近平面，因为误差在可用深度范围内分布不均匀。在接近平面的情况下，大部分分辨率都用在了眼睛附近，其余阴影图的分辨率不足。作者建议分析场景以将近平面尽可能向后推，以缓解此问题。另外，观看者投影的使用可以改变光的方向或什至改变光的类型（从定向到点，反之亦然），这使实现复杂化。

通过将透视图转换与查看器分离，可以避免这些问题。这是光空间透视阴影贴图（LiSPSM）[WSP04]的主要思想，它通过光和视图对齐的变换使光空间变形。在这里，透视变换始终与视锥的轴对齐，因此，光不会改变方向或类型（请参见图8和10）。为了处理点光，首先应用点光的投影，将点光转换为定向光，然后在光的后视空间中完成LiSPSM。解耦的透视图转换还有一个额外的好处，即创建了一个自由参数，即透视图转换的近平面距离n（请参见图9）。较小的距离会导致更强的变形，而更多地聚焦在附近的物体上；较大的n会导致强度降低。请注意，以后有关变形方法的大多数工作都采用了此框架。在LiSPSM中，选择近平面距离的方式是将误差平均分配到可用深度范围内，从而产生均质的质量（请参见图11）。

Martin和Tan的梯形阴影贴图（TSM）[MT04]是一种非常相似的方法，它们使用启发式方法选择近平面距离。 在非常有见地的工作中，劳埃德（Lloyd）等人[LTYM06]证明，当同时考虑两个阴影贴图方向时，所有透视变形算法（PSM，LiSPSM，TSM）实际上会导致相同的整体误差，但是LiSPSM可以在两个阴影贴图方向之间提供最均匀的误差分布。 方向，因此是有利的。

Chong和Gortler [Cho03，CG04，CG07]通过使用多个阴影贴图来优化少量关注平面的阴影贴图质量。 他们甚至表明，可以使用计算机视觉中的“平面稳定”技术对这些平面进行完美采样（对于该平面上的每个视口像素，阴影图中的纹理像素均采样相同的几何点）。 然而，不在这些平面上的像素会表现出任意误差。

* + 1. 对数Warping