

Temporally Adaptive Shading Reuse for Real-Time Rendering and Virtual Reality

Dezeming Family

2023年8月4日

正常字体：表示论文的基本内容解释。

粗体：表示需要特别注意的内容。

红色字体：表示容易理解错误或者混淆的内容。

蓝色字体：表示额外增加的一些注释。

绿色字体：表示额外举的一些例子。

目录

| | |
|--|---|
| 一 Introduction | 1 |
| 二 RELATED WORK | 2 |
| 三 PERCEPTION OF SHADING DIFFERENCES | 2 |
| 四 TEMPORAL COHERENCE FOR SHADING REUSE | 3 |
| 4.1 Temporal Coherence of Visibility | 3 |
| 4.2 Temporal Coherence of Shading | 3 |
| 4.3 Limits of Applying Temporal Coherence | 4 |
| 五 PREDICTING SHADING CHANGES | 4 |
| 5.1 Prediction with Fixed Upsampling Rates | 4 |
| 5.2 Prediction with Shading Gradients | 4 |
| 5.3 Spatial Filtering of Temporal Gradients | 5 |
| 六 TAS FRAMEWORK | 6 |
| 6.1 Temporally Adaptive Reprojection Caching | 6 |
| 6.2 Temporally Adaptive Shading Atlas | 6 |
| 七 总结 | 7 |
| 参考文献 | 7 |

abstract

时间一致性有可能大大降低渲染中的着色成本。现有技术要么只关注空间着色重用，要么不能自适应地选择时间着色频率(shading frequencies)。

我们发现，对于大多数样本，时间着色重用在很长一段时间内是可能的，并且我们展示了用户在什么情况下感知到时间伪影。

我们的分析表明，我们可以近似着色梯度，以有效地确定何时以及多长时间可以重用着色。尽管可见性通常在90%以上的时间内从一帧到另一帧保持时间连贯，但我们发现，即使在具有高级着色的动画游戏场景中，通常也有50%以上的着色是时间连贯的。

为了利用这一潜力，我们引入了一种时间自适应着色框架，并将其应用于两种实时方法。它的应用程序节省了57%以上的着色器调用，将虚拟现实应用程序中的总体渲染时间减少了5倍，而不会显著降低视觉质量。总的来说，我们的工作表明，着色重用的潜力比目前开发的要大得多。

— Introduction

有些技术有效利用空间相关性：checkerboard rendering [El Mansouri 2016], foveated rendering [Patney et al. 2016; Swafford et al. 2016], variable rate shading [Vaidyanathan et al. 2014], and motion-adaptive shading [Yang et al. 2019]。

大多数TAA和时序重投影的使用都依赖于漫反射场景，时序着色变化不大。时序着色存在下面几个问题：

- Q1: 随着时间的推移重复使用着色样本时，时间伪影如何影响感知图像质量？
- Q2: 在具有高级着色和动画的场景中，着色重用的限制是什么？
- Q3: 如何在不实际重新加载样本的情况下提前确定着色样本何时无效？
- Q4: 给定足够的时间相关的着色样本，如何在实践中有效地重用它们？

我们的工作做出了以下的改进：

- 对着色差异的感知：我们通过进行受控用户研究来解决Q1问题，以确定由于具有高级着色和动画的场景的时间着色重用而导致的着色伪影的感知效果（论文第3节）。
- 时间连贯性：我们通过分析着色和可见性随时间变化的潜在相关性来解决Q2问题（第4节）。我们的实验表明，先前关于时间采样策略的工作所做的关于着色的时间一致性的假设在许多现实世界中可能不成立。
- 梯度：我们通过分析时间着色梯度的分析和数值一阶近似值及其预测未来着色变化幅度的能力来解决Q3问题（第5节）。此外，我们分析了时间着色梯度的空间变化，并展示了如何结合空间信息来更好地预测未来的明暗变化。
- 框架：我们用一个通用框架来解决Q4问题，该框架用于预测着色变化并在时间上重用着色（第6节）。通过将该框架集成到两种渲染方法中，我们展示了显著的性能改进，特别是对于质量差异不明显的高通量VR渲染（Fig.1）。为了评估该框架，我们进行了两项用户研究，以展示我们的方法的适用性，并详细分析了其在各种场景的着色预测成本、着色成本节约和整体帧率增益方面的增益（第7节）。

总体而言，我们的框架将总着色负载降低了57%到90%，具体取决于渲染序列的动态。如果可能的话，它会在几秒钟内保留着色结果，而不会出现任何质量下降。这些节省使我们的方法成为在当前未修改的图形硬件上运行的最有效的时间自适应着色(TAS)方法，特别是对于VR应用。



Fig 1: 使用我们提出的方法将实用的时间自适应着色应用于三个不同的场景。(a)Sponza测试场景包含物理动画巨石和移动聚光灯。(b)我们在虚幻引擎中实现的技术应用于Showdown VR演示序列，其中包含大量与视图相关的效果，包括镜面高光和闪亮表面上的反射。(c)即使是灵魂之城场景中的大量的反射表面以及动画、潮湿材质，也可以通过我们的方法进行处理，而无需手动微调。这些也显示了我们方法的基本步骤。(a)我们依靠着色梯度来估计着色的变化和着色的潜在重用。(b)使用经过最大滤波(maximum filtered)的绝对着色差异，无论是在每个基元还是在图像空间邻域中，在实践中都能很好地工作。(c)最终的着色决策基于之前计算的着色差异，根据动力学将着色减少57%到90%，同时在视觉上与每帧中的完全着色无法区分。

二 RELATED WORK

着色缓存[Tele等人2002]的设计允许在路径跟踪中重用时空着色。最近，纹理空间着色方法[Baker 2016; Hillesland和Yang 2016; Tatarinov和Sathe 2018]已经普及，甚至已用于VR客户端上的时间上采样[Mueller等人2018]。然而，所有这些方法都只允许固定的时间上采样率。相反，我们的方法显示了如何在可变的、可能非常长的时间内保留样本。

Numerous image quality metrics try to model the human perception of images. Although the peak signal-to-noise ratio (PSNR) provides an objective and easy to compute metric, it fails to capture human perception. The popular structural similarity index measure (SSIM) [Wang et al. 2004], which has been designed to more closely resemble perception, is still the prevalent image quality metric. Various improvements, such as IW-SSIM [Wang and Li 2011], have been made to SSIM to enhance the predictions of the metric. Other metrics, such as HDR-VDP-2 [Mantiuk et al. 2011], even try to model the visual system to some extent. Swafford et al. [2016] build on HDR-VDP-2 to adapt the metric for the evaluation of foveated rendering. Another approach is the combination of different metrics, such as VMAF, which appears especially useful for video game content [Barman et al. 2018]. The novel FLIP [Andersson et al. 2020] metric is derived from the manual method of comparing images by alternating between them and provides an error map showing where differences would be perceived between the two images in comparison. A major disadvantage of all of these metrics is that they only compare images, disregarding any temporal artifacts, such as flickering. A few video quality metrics incorporate temporal aspects [den Branden Lambrecht and Verscheure 1996; Pinson and Wolf 2004; Seshadrinathan and Bovik 2010; Watson 2001; Yang et al. 2007] but focus mostly on video compression-related artifacts such as frame dropping and have not been evaluated for temporal artifacts appearing in rendered imagery. Therefore, conducting a user study remains the gold standard method for the evaluation of perceptual quality, especially when temporal effects have to be considered.

三 PERCEPTION OF SHADING DIFFERENCES

使用pairwise comparison design [Kiran Adhikarla et al. 2017; Mantiuk et al. 2012]来评估一个前向渲染获得的参考视频和一个时序重用渲染的视频。

由于剪辑的顺序是随机的，他们不知道哪个剪辑是参考视频。根据评分，我们计算出平均相对质量分数(Q)，范围从-2到+2，其中+2表示参考视频明显更好，+1略好，0表示他们被评为同等。此外，我们计算了选择参考视频而不是时序重用方法的概率 p_{ref} 。50%的 p_{ref} 表示两种方法之间没有差异；75%的 p_{ref} 被

称为1显著差异(JND)单位[Mantiuk等人, 2012]。保持在1JND以下被认为是高质量的。为了进行统计分析, 我们使用重复测量方差分析和Bonferroni调整进行事后检验。在整篇文章中, 我们使用相同的研究设计来评估各种技术和方面。有关研究设计的所有详细信息, 请参阅补充材料。

另外, 为了探测因重用样本造成的感知质量衰减, 我们使用了一个叫做temporal forward rendering(TFR)的实验。大致意思是, 我们先渲染一轮作为Reference, 然后用相同变化的场景参数进行第二轮渲染, 在渲染中, 如果当前帧像素与Reference的差距小于T, 就重用历史帧, 否则就改变为Reference的对应值。随着T设置的越来越大, 用户就能感受到越来越明显的时序伪影。注意这里的比较, 是在色调映射后0-255范围的, T从2/4/8/16依次增加, 当T=16时, 用户基本上都能感受到时序伪影。

当相机移动时, 图像的视角会发生变化, 因此更难看到着色处理伪影。相比之下, 静止的相机可以让用户将注意力集中在动画对象上, 从而更容易注意到局部着色的变化。这一假设在我们的研究后访谈中也得到了证实。“移动摄影机的场景更具挑战性。”。

类似的结果也可以从场景中得出: 机器人实验室的反射和突然移动最少。阴影边界缓慢而稳定地移动; 移动物体之间几乎没有重叠, 因此很容易发现错误。Sponza包含了很多运动, 因为移动的聚光灯和落下的巨石。空间是闪亮的, 大的亮度梯度引起了人们的注意。

结合深入分析的结果表明, 像机器人实验室这样的场景可能具有最高的重用潜力, 但用户可能会更早发现质量问题。更动态的场景具有更低的重用潜力, 但发现着色缺陷也更困难。因此, 在需要更多重载的情况下, 可以使用更积极的重用设置, 从而实现跨场景的总体均衡重用潜力。应该记住, 研究参与者正在积极寻找着色质量的差异。分心的观察者, 比如参与游戏的用户, 可能会容忍更高的阈值。

四 TEMPORAL COHERENCE FOR SHADING REUSE

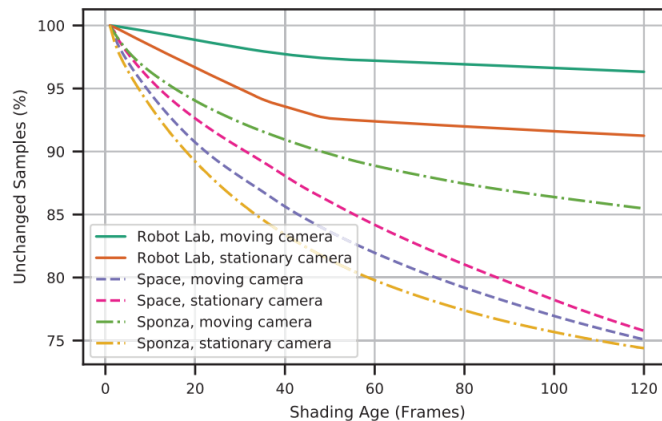
许多渲染加速技术利用了时间重用, 但除了非正式比较之外, 并没有提供方法质量的证明。这使得在实践中可以在多大程度上或多长时间重用着色结果的问题没有得到解答(Q2)。

4.1 Temporal Coherence of Visibility

根据Nehab等人[2007], 超过90%的样本在帧之间保持可见。最显著的能见度干扰是由大幅度相机移动或快速移动的物体引起的。

4.2 Temporal Coherence of Shading

为了分析着色样本的潜在理想重用, 我们依赖于第一次用户研究的结果, 并使用阈值为T=8的TFR分析所需的片段着色器调用次数。图4显示, 在大多数情况下, 即使对于Space这样的复杂场景, 超过74%的着色结果平均在120帧以上保持有效(在T=8下)。作为运动最少的场景, Robot Lab在测量的时间范围内保持在90%以上。Sponza的固定摄像机聚焦在坠落的巨石上, 因此实际上比更复杂的场景Space产生了更多的样本变化。包含其他阈值详细数据的报告可在补充材料中找到。



4.3 Limits of Applying Temporal Coherence

可见性的时间连贯性和着色的时间连贯都显示出在许多帧上重用着色的非常高的潜力。然而，实际实现还需要考虑着色的空间采样，即着色样本的漂移、它们的重投影误差和所需的滤波。

为了根据这些问题评估着色重用的潜力，我们评估了两种实用的着色重用渲染方法。第一种方法是反向重投影缓存(RRC)，由Nehab等人[2007]提出，该方法已在可见性的时间一致性实验中使用。RRC将样本从前一帧重新投影到当前帧，从而可能累积空间采样误差。我们的实现分为两个过程运行，一个是深度预过程(depth pre-pass)，另一个是使用缓存或重新渐变的前向渲染过程(forward rendering pass)。为了避免这些误差的累积，可以在时间不变的空间（如对象空间或纹理空间）中收集着色样本。

因此，我们的第二种方法，着色图谱(SA)，结合了Mueller等人[2018]提出的着色图谱与纹素着色的渲染管线[Hillesland和Yang 2016]。简而言之，这种方法对在单个纹理(着色图集, shading atlas)中动态分配的矩形块中的三角形对进行着色。着色采样在atlas中的位置保持不变，直到三角形的可见性发生变化，或者由于细节级别的变化而导致其分辨率发生变化，在这种情况下，将重新计算着色，即不重复使用这些采样。

使用这两种渲染方法，我们可以评估在不同帧之间可以重用的样本数量。对于RRC，我们将可能的重用计算为屏幕分辨率的分数，而对于SA，我们考虑每帧中图集中总分配空间的分数。

论文中的图5的结果表明，在大多数情况下可以实现70%以上的重用。随着空间重投影误差的积累，RRC在相机移动时具有更差的重用潜力。SA只在整个块可重用时才考虑样本的可重用性，从而导致总体重用性稍差。Robot Lab和Sponza中固定摄像机SA的重复使用较少，这与着色变化的时间一致性有关（见图4），因为摄像机的姿态集中在场景的动态部分。

五 PREDICTING SHADING CHANGES

在上一节中，我们已经确定了时间着色重用具有令人信服的潜力。现有的几种方法使我们能够通过图像空间重投影或在时间上不受影响的空间（如对象空间或纹理空间）中进行着色，将着色样本从一帧映射到下一帧。然而，我们需要对将来着色采样将变得无效的点进行有效预测，以了解着色可以重复使用多长时间。在本节中，我们将讨论哪种机制最适合提供这样的预测。

5.1 Prediction with Fixed Upsampling Rates

以前的策略依赖于均匀的时间上采样，即每N帧对采样进行着色。我们从RRC[Nehab et al. 2007]和shading atlas [Mueller et al. 2018] (SAU) 实现均匀的时间上采样。RRC使用一个常量刷新率来更新 16×16 像素的tiles，在每一帧中，所有tiles的恒定部分都会更新，从而为每个tile带来固定的生存期。对于具有上采样(SAU)的着色图集，缓存项的生存期也是恒定的，但每个缓存项都有一个单独的剩余生存时间，具体取决于它何时可见。RRC和SAU都将着色负载分布在多个帧上，而不会出现明显的全屏着色刷新。

为了评估固定速率的时间上采样如何影响渲染质量，我们使用第3节中描述的相同用户研究设计来评估RRC和SAU。如论文图6所示，即使只重用着色一次（ $2 \times$ 上采样），均匀上采样也无法很好地利用着色重用的潜力。在82%的例子中，2倍时间上采样的RRC导致明显差异，平均而言，据报道质量“略差”。对于更高的上采样率（4和8），所有参与者总是注意到差异，并报告图像质量接近“明显更差”。通过使用更好的滤波器进行时间上采样，可以避免一些质量恶化——我们使用了线性插值。上采样因子为 $2 \times$ 的SAU保持在1 JND以下。然而，使用上采样因子大于2倍的SAU被认为“明显更糟”，100%的参与者注意到了这种差异。总的来说，我们得出的结论是，均匀的上采样频率不足以实现更长的着色重用。

5.2 Prediction with Shading Gradients

在排除了固定速率上采样以释放时间着色重用的潜力之后，我们考虑着色数据的数学分析是否可以用作预测因子。可以说，最佳方法是阴影的频率分析[Durand等人, 2005]，但这也是最昂贵（也有点笨拙）的选择。Ramamourthi等人[2007]表明，一阶梯度分析在空间域中通常是足够的。因此，我们建议将

其中一些发现转移到时域中。我们认为着色函数 s 的泰勒近似是预测的一个明显选择。一个简单的线性预测器可以公式化为从时间 t_0 到 t 的一阶泰勒展开 ($e(t)$ 是残余项):

$$s(t) = s(t_0) + s'(t_0) \cdot (t - t_0) + e(t) \quad (五.1)$$

基于颜色阈值,我们就可以基于每通道最大颜色梯度 $s'(t_0)$ 计算出要去更新着色的deadline:

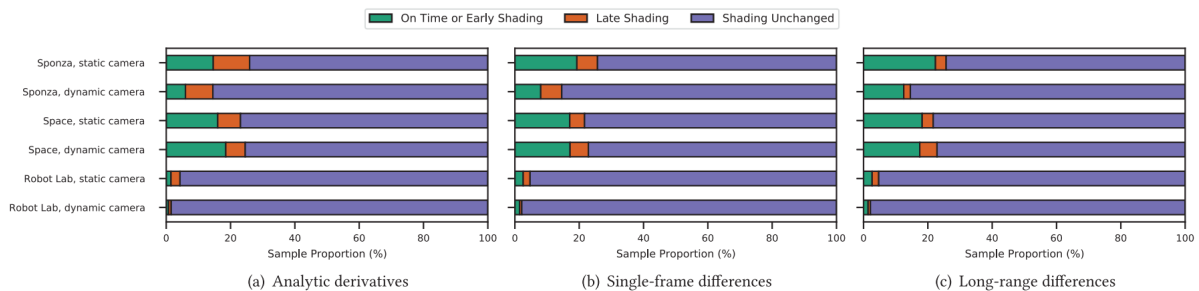
$$d = t - t_0 = \frac{T}{s'(t_0)} \quad (五.2)$$

求梯度有解析法和有限差分法。

对于有限差分法,在最简单的情况下,我们取两个着色结果之间的向后差。这种后向差异可以在连续帧中的着色之间或者在时间上更远的帧之间计算。前者的优点是它更接近有限差的极限。但是,它总是需要连续着色两次。在更长的时间间隔上计算有限差更经济,并且对伪阴影变化具有潜在的有益的低滤波效果。当然,第一次着色采样必须始终进行两次。通过这种方式,从初始梯度推断出重新加载的第一个deadline。从那时起,每当发生重新加载时,都会估计Long-range差分(更长时间间隔的差分)。

之后我们使用TFR实验来验证。使用TFR,我们渲染着色时间不断增加的多个帧,导致帧具有相同的采样位置,但着色越来越过时。我们使用这些数据来回顾性(retrospectively)地获得理想的deadline。从包含正确着色结果的当前帧开始,我们确定着色差异超过阈值 T 的过去的确切帧。在过去的这一点上,解析求导和有限差分都被用来直接预测未来的deadline。对于Long-range差分,我们重复该过程,以找到过去超过阈值的下一帧。我们将搜索过程限制在过去的119帧,有效地将deadline限制在1到118帧的范围内。

下图表示我们使用TFR来评估计算着色处理的时间梯度的不同方法,以及它们对着色处理未来变化的预测效果。着色太迟会在最终图像输出中造成伪影,因此应避免。过早着色可能会影响性能,但不会降低质量。理想情况下,一种技术可以完全避免后期阴影,同时尽可能降低早期阴影。根据我们对着色处理的时间一致性的测试,大多数样本在120帧以上保持不变。请注意,简单的Long-range差分(两个着色点之间)显示了最少的延迟着色,而过早着色仅略有增加。

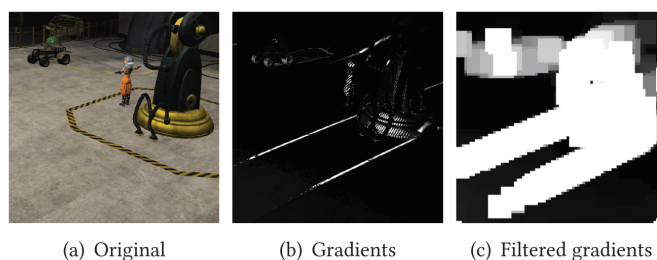


也就是说,这种根据有限差分来近似梯度然后预测着色变化的方法还是比较合理的。

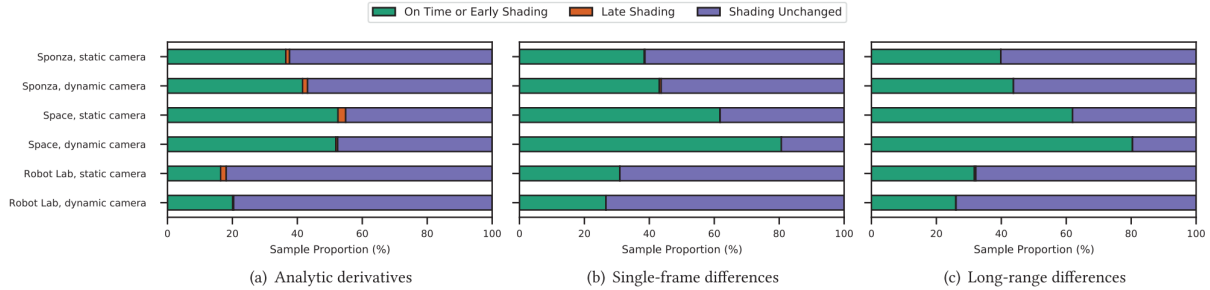
5.3 Spatial Filtering of Temporal Gradients

对着色梯度预测失误原因的实证检验表明,主要原因是在空间中相关移动,但突然改变单个样本的着色的现象。例如,由移动灯光或阴影边界引起的边界。

这表明,我们必须捕捉空间连贯、时变变化的影响。因此,我们在图像空间中提出了一个简单的最大值滤波器。这种方法的灵感来自渲染缓存[Walter等人1999]和着色缓存[Tole等人2002],它们根据相邻样本的估计时间梯度做出着色决策。这种过滤器对移动阴影边界的着色梯度的影响如图8所示。显然,梯度滤波将阴影边界的高度局部化的梯度分布到周围环境。



我们使用TFR来评估该滤波器，以隔离空间滤波的影响，同时避免其他预测错误的来源，如重投影错误。我们使用 8×8 的下采样因子，然后使用矩形内核大小为 9×9 （总像素数为 72×72 ）的卷积。我们根据经验确定这些参数是足够的，因为以较低的着色重用为代价，我们总体上避免了后期着色，我们认为这对于避免伪影更重要。当将前面描述的时间梯度估计方法与空间最大滤波相结合时，我们得到了如图9所示的结果。在应用滤波器时，解析导数、单帧差分 and Long-range差分令人信服地避免了着色过迟，代价是在120帧内进行更多早期着色。考虑到这些方法的相似性质，Long-range差分最具吸引力，因为它们是最有效的计算方法。



六 TAS FRAMEWORK

在我们之前的发现的基础上，我们设计了时间自适应着色（TAS）框架，它可靠地避免了重复的着色计算，同时对着色发生快速变化的区域做出即时响应。为了使框架在很大程度上独立于它所应用的渲染算法，我们引入了重用单元（RU）的概念。RU是一组样本，对其进行统一的决策，决定是重新对样本进行着色还是重复使用着色。这些单元的样本一起着色，因此必须一起存储在缓存数据结构中。渲染器独立地确定每个单元的可见性。例如，前向或延迟渲染器使用深度缓冲区来确定每像素的可见性，而其他渲染器类型可以确定每基元的可见性。因此，在反向重投影缓存的情况下，RU可以是单个像素[Nehab等人, 2007]，也可以是着色图谱(shading atlas)中的整个块[Mueller等人, 2018]。考虑到这些因素，我们简要地指定了TAS的三步算法：

- 将来自最后一帧的空间滤波的着色梯度与自每个RU的最后一次着色以来经过的时间相乘，并与阈值（T）进行比较，以决定是否需要重新加载。对于两个连续帧，总是对新可见的单元进行着色，以根据有限差确定梯度。
- 着色将被重复使用，或者单元将被重新加载。在后一种情况下，将为每个样本计算与先前着色结果的新着色差分。
- 基于着色差分来估计着色梯度，着色差分按它们之间的时间差缩放，并且应用空间滤波器来分布着色梯度信息。

该框架足够通用，几乎可以应用于任何能够参考先前着色结果的渲染架构。

6.1 Temporally Adaptive Reprojection Caching

在我们的第一个参考实现中，时间自适应重投影缓存（TARC），扩展了Nehab等人[2007]的方法，图像空间像素用作RU。我们用框架的前两个步骤替换了反向重投影缓存着色器的定期刷新。此外，我们将每个单位和每个样本的变量存储在双缓冲G缓冲区中。输入缓冲器被重新投影，并且潜在改变的值被存储在输出缓冲器中。在深度预处理过程中，我们还将估计的着色梯度存储在G缓冲区中。我们通过使用具有重叠平方核的最大滤波器对梯度缓冲区进行下采样来实现空间最大滤波。与标准的反向重投影缓存相比，TARC需要额外的内存用于屏幕大小缓冲区，以存储着色差异和自上次着色以来的时间。

6.2 Temporally Adaptive Shading Atlas

第二种实现的框架是temporally adaptive shading atlas (TASA)，因为我没有了解过论文《Shading atlas streaming》，所以暂时不研究这个内容。

七 总结

实验验证也都很简单，所以不再赘述。该论文虽然理论比较简单有趣，但工作量确实不小。主要研究的创新点就是，首先通过用户实验来确定阈值 T ，然后基于 T 和空间滤波后的有限差分来估计更新deadline。整体算是一个挺不错的工程实践类论文。

参考文献

- [1] Igouchkine O , Zhang Y , Ma K L .Multi-Material Volume Rendering with a Physically-Based Surface Reflection Model[J].IEEE Transactions on Visualization & Computer Graphics, 2017, PP(99):1-1.DOI:10.1109/TVCG.2017.2784830.